Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

A SYSTEMATIC REVIEW OF TEMPORAL SEQUENCE MODELING USING RECURRENT NEURAL NETWORKS FOR REAL-TIME HEALTH MONITORING ARCHITECTURES AND PREDICTIVE ANALYTICS

SHIVARAM REDDY K[1,2], M V NARAYANA[3]

[1]Research Scholar, CSE Dept, University Institute of Engineering & Technology, Guru Nanak University, Hyderabad, India

[2]Assistant Professor, CSE Department, Sreyas Institute of Engineering and Technology, Hyderabad, India

[3]Professor, CSE Dept, University Institute of Engineering & Technology, Guru Nanak University, Hyderabad, India

E-mail: [1,2] shivaramreddykethi@gmail.com, [3]mvnarayanacse@gmail.com

ABSTRACT

Leveraging state-of-the-art deep learning models with real-time health monitoring systems is an essential direction to this field of digital health analytics. Here, Recurrent Neural Networks (RNNs), for example, Long Short-Term Memory (LSTM) or Gated Recurrent Units (GRU), have significant potential in accurately learning physiological time-series data for predictive diagnostics. In this systematic review, we critically analyze the current state of modelling temporal sequences using RNNs with real-time health monitoring systems. In accordance with PRISMA guidelines, we analyzed 124 peer-reviewed articles from January 2014 to December 2024, across major databases, and curated parameters relating to modelling properties, deployment strategies (i.e., cloud, edge, mobile), data acquisition pipelines, and health signal modalities (i.e., ECG, PPG, respiration). We compared the performance of vanilla RNNs against the sophisticated temporal models for value-add and sensor signal processing (e.g. detecting arrhythmias, predicting blood glucose, estimating early warning scores). There was significant heterogeneity in metrics of evaluation, and varying methodologies for addressing data noise and latency, despite minimum real-world deployment of models also being suggested. Also emerging was a fundamental research gap in scalable RNN implementations intended for energy-efficient edge-based processing. Very few studies adopted any adaptive temporal attention, while generalization across patients posed challenges to clinical uptake. Subsequently, little test activity surfaced concerning model interpretability, computational latency, or privacy-preserving inference. Representing a clear gap in current literature as well, our review provides a taxonomy of RNN-based real-time health monitoring systems, notwithstanding some serious limitations for practical deployments, as well as benchmarking modelling trends, and promotion of a common research agenda path with a strong focus for ramification from hybrid architectures, federated learning for model integrations, and explainability. Our review endeavor provides a synthesis of technical intelligence and system-level constraints to also serve as a basis for scholars and professionals interested in aligning to advance the state-of-the-art in their intelligent health monitoring systems.

Keywords: *Recurrent Neural Networks, Temporal Sequence Modeling, Real-Time Health Monitoring, Predictive Analytics, Systematic Review*

1. INTRODUCTION

The rapid growth of wearable technologies, ambient sensors, and connected biomedical devices has changed the health monitoring paradigm from episodic clinical evaluation to continuous real-time monitoring. This is more than a technological shift; it represents a significant shift in the paradigm of preventive health where physiological signals such as heart rate, blood glucose, respiration rate, and blood oxygen saturation are dynamically monitored to identify deviations, anticipate health threats, and respond proactively. As chronic conditions such as cardiovascular disease, diabetes, and respiratory disease hold increasing global prevalence - and our aging population is a contributing factor - the capacity to monitor biomedical signals continuously and interpret those signals appropriately is a central tenet of progressive health care systems.

The Internet of Medical Things (IoMT), edge computing, and low power embedded platforms have made it possible to move real-time health monitoring systems into an array of applications in practice - from hospital ICUs and emergency service uses, to remote patient management at home, and even fitness applications. However, while these platforms can certainly facilitate passive data acquisition in the form of physiological data, their true value lies in the ability to leverage the data through intelligent, time-sensitive analysis of streaming physiological data. The challenge remains to develop robust computation strategies that can process noise in multivariate time series signals at high rates of temporal resolution, while also considering the often-stringent limitations regarding latency, power, and storage.

Machine learning, and more specifically deep learning, represents an opportunity to model the Traditional ML methods such as support vector machines (SVMs), random forests, and logistic regression can be effective at static or windowed features but cannot model complex temporally dependent structure that arises in physiological processes. Recurrent Neural Networks (RNNs) and their gated variants (e.g. Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU)) provide a principled approach to sequential modelling. With hidden state recurrence and memory gates, RNNs can model long-term dependencies, cyclical structures, and subtle temporal relationships—characteristics needed to conduct arrhythmia detection, sleep apnea classification, sepsis predictions, and early warning scoring.

In the last decade, the rise of RNNs has been leveraged to enable real-time health predictions and monitoring, along with a growing body of studies examining RNN-type architectural methods to assist in prediction and monitoring. RNN models vary from real-time electrocardiogram (ECG) monitoring approaches, continuous glucose monitoring (CGM) and stress detection, predicting epileptic seizures, and remote patient monitoring or care. Further advancements have been made with the utilization of RNN-type architectures in combination with convolutional layers in cases where raw sensor models are used for detecting features prior to learning a sequence. Likewise, the idea of deploying RNNs across mobile and edge platforms is growing in connection with a rise in decentralized patient-driven health management.

The research is clearly growing, but it is spread across several heterogeneous areas. Existing studies show a lot of variability across model architectures, evaluation protocols, signal types, data sources, and deployment methods. Moreover, it is ambiguous how deep networks should be implemented, the length or batches of sequences shared across applications, and whether gating mechanisms or other processes can be standardized. Furthermore, most of the research is compartmentalized - either focusing exclusively on algorithmic performance or hardware realization, with few activities attempting to connect the path between theoretical modeling and real-time system implementation. In addition, a lack of standardization around research artifacts hinders reproducibility, benchmarking, and ultimately utilization in clinical settings. Another major issue relates to the diversity and quality of datasets utilized in this domain. Some may find useful resources, either because they are publicly available or because they sample widely, such as MIT-BIH, PhysioNet, and MIMIC-III, but many studies utilize such datasets solely in the context they were collected, limiting their generalizability. Thus, generalizing model performance beyond real-time deployment with appropriately streaming, noisy, and heterogeneous data is difficult. Adding to this complexity, seldom do researchers split the testing and training datasets at the level of the patient (as would be done in practice) or formally temporally validate their models with respect to appropriate clinical conditions (avoiding information leakage), and for either of these rationales, it is rare for model performances to be significantly overestimated. Furthermore, few studies take cross-population generalizability into account, especially for models trained on imbalanced datasets where the prior training conditions were disproportionally weighted towards a select demographic.

Finally, from the systems architecture perspective, the challenges of deploying recurrent Neural networks (RNN) in ambulatory real-time health monitoring present further concerns. Most research in the area has been done under batch-mode training with batched inference, while real-time systems must accommodate continuous streaming with low-latency operation, address missing or corrupt variables, accommodate changes in the characteristics of the signal, etc. Models are often constrained in size, power-draw, and memory in edge deployment configurations, requiring compression strategies such as pruning, quantization, or knowledge distillation; however, these considerations are seldom integrated into algorithm design pipelines. Additionally, while RNNs have been widely accepted as theoretically capable of modeling sequences, their interpretability and reliability for meaningful applications in high-stakes contexts such health is rarely addressed. For example, there is greater awareness of the need for explainability and trustworthiness in AI-enabled health systems. RNNs (deep LSTM and GRU) are fundamentally black boxes, which have broader implications on accountability and transparency in regulated domains like health. Methods of achieving or improving interpretability (e.g. attention mechanisms, saliency maps, SHAP values) have been proposed in recent years (Lau et al., 2020), but practical implementation of these techniques into capacity of real-time health prediction systems is still early. Health-related ethical issues surrounding data privacy, algorithmic bias, and fairness of decision-making are not frequently considered in technical studies but are key considerations for real-world utilization in clinical practice.

Considering the rapid advancements in this emerging interdisciplinary field and the heterogeneous nature of existing literature, we argue there is an important need for systematic synthesis. At present, there is no systematic review or other evaluation of RNN approaches to real-time health monitoring. Systematic reviews exist on deep learning applications in health; however, they provide broad coverage, do not include real-time temporal prediction systems, or focus solely on offline diagnostic approaches to health. Therefore, this review presents a systematic and substantive analysis of the field of RNN methods for temporal sequence modeling in real-time health monitoring and predictive analytics. The aims of this review are five-fold: (1) they will enable an organization of RNN-based models developed for RT health predictions based on their architecture (vanilla RNN, LSTM, GRU, attention-based models, hybrid CNN - RNN); (2) they will allow a summary of the biomedical signals used, along with their suitability for temporal modeling frameworks; (3) they will encourage evaluation of the consideration of system deployment factors including latency, edge computing hardware use, and power constraints; (4) they will identify methodological limitations, reproducibility barriers or deficiencies in reporting; and (5) they will formulate a research roadmap with respect to algorithmic advances relative to clinical, hardware, and deployment realities.

Our review documents include first, generating a taxonomy of RNN-based approaches designed for health monitoring, targeting unidirectional versus bidirectional, single versus stacked deep, as well as standard versus attention-focused RNN models. We also capture each kind of model with biomedical application, along with trade-offs, signal pre-processing pipeline methods, as well as the combination training-validation-data generation pipeline. Next, we identify deployment approaches including centralized, edge, and federated architectures, as well as if they can be part of currently used health-monitoring infrastructure beyond research applications. Our analyses also include a meta-summary of how evaluation is often reported in the form of metrics such as accuracy, F1-score, sensitivity, specificity, and latency. We note a variation in definitions of metrics, baseline use, and underreporting of runtime analytics that limits comparability and applicability of many proposed systems that could work in the real-world. This review will also document evolving best practice approaches such as patient-specific modeling, transfer learning under low-data opportunities, and evidence of multi-sensor fusion strategies to influence model agility.

Ultimately, our systematic review provided a cohesive structure to characterize a research activity area that has active challenges that warrant further investigations. Bridging deep learning model development, advances in physiological signal analytics, and deployment constraints, we hope that we can support stronger, explainable, and translatable health-monitoring solutions. It is also our hope that our review will act as a guiding resource for interested clinical, system or research practitioners to improve on RNN-based model design for next-generation RT health analytics.

2. RELATED WORKS

In the last ten years, the simultaneous development of deep learning and biomedical signal acquisition has created a wealth of research on real-time health monitoring systems - especially around human activity and physiological signal analysis. Recurrent Neural Networks (RNNs) have become a paradigmatic computational framework for modeling temporal sequences. Unique strengths of RNNs are their ability to capture long-range dependencies that are prominent in human physiological signals; and to model temporal patterns that are inherently dynamic in nature. The last ten years have seen rapid progression from the application of simple RNN architectures to advanced variants such as Long Short-Term Memory (LSTM) Networks, Gated Recurrent Units (GRU), Bidirectional RNNs, and attention-based hybrid models. Various tasks of ECG arrhythmia classification, future glucose level predictions, abnormal respiratory patterns, and seizure predictions have been explored using different signals and architectures. The variety of methods, signal pipelines, evaluation strategies, and real-time application needs has caused fragmentation and lack of consistency across literature. In this section, we review the literature in a systematic way by organizing it into four groups of literature (i) RNN architectures applied to physiological signals, (ii) implementations that are applied directly within a specific context; and predictive health analytics, (iii) system-level designs and deployment constraints during real-time applications; and (iv) interpretability, generalizability, and ethical and social issues when using RNNs for health monitoring. In this way, we not only identify trends across literature but also gaps and areas of overlap between methodologic designs and clinical practice.

An RNN-based IoT-enabled soil quality control system designed to support coffee cultivation was reported by Vivek Kumar Prasad and co-authors [1], using counterfactual recommendation reasoning. The use case is agriculture, but the model also displays the versatility of RNNs for real-time decision support at the user-system interface using dynamic data from configurations of many sensors. Their model combines cause-and-effect reasoning using time-ordered measurements and the findings also underscore the deep-sequence learning potential for operational phenomena analogous to health whereby state estimates are generated from series of sampled signal data over time.

S. Kartal and co-authors [2] reported patio-temporal vegetative health forecasting using Comvest models and MODIS time-series satellite data. Their examples provide some insight into the capabilities of this convolutional-recurrent hybrid where patio-temporal and time-series datasets can be explored, and while the use case is environmental monitoring when bio signal processing from health sensors may be quite dispersed spatially, it is not difficult to envision their approach applied. Their model identified observable patterns of behavior even with sparse temporal sampling, not underlyingly different from the surveillance problem when remotely monitoring the heart signals of patients with low-frequency sampling of cardiac signals.

S. Masini and co-authors [3] investigated the structural behavior of Brunelleschi's Dome with machine learning and presented an example of data-driven temporal modelling when detecting anomalies. While the application is architectural heritage, the types of computational methods used, signal decomposition, temporal classification and anomaly prediction, are similarly applicable to physiological health monitoring frameworks when atypical perturbations of a signal provide some indicators of pathological conditions.

This trajectory continues with Y. Sun and collaborators [4] that developed an MMA-RNN based on multi-task, multi-level attention recurrent neural network intentions to detect and localize arrhythmia. Their biomedical application has formed a linear trajectory on the use of attention based RNNs in a specific cardiac rhythm classification problem. Their architecture was configured so disturbances in the signal could be identified in spatial-temporal hierarchies or levels of granularity, which can sharpen classification granularity, while being suitable for real-time classification, such as ECG in wearable systems. Wei Wang and colleagues [5] leveraged deep RNNs to perform lateral displacement predictions for suspension bridges examined under wind loads. The structural health monitoring domain has similar constraints as physiological health systems, including long-term sequential data, real-time prediction prompts, and resilience to external variability. Sensor integration and data-driven dynamics mapping is like multimodal bio signal projections.

Y. Zaitin and team [6] designed a patient health monitoring framework that includes IoT and RNN-based modules for vital sign predictions. Vital sign predictions are based on the real-time stream processing of the patient's multivariate physiological input signals. The innovative framework accentuates system modularity, time continuity of data signals, and RNN inference can be embedded and used as direct models for continuous patient health monitoring at the "edge".

W. Y. Alghamdi [7] introduced a recurrent neural framework for predicting athlete health using wearable sensor datasets. Although the article mainly focused on optimizing temporal learning within the mobile sensing paradigm, latency, energy consumption, and interpretability are critical for real-time sports health telemetry and generally to any scenario of ambulatory monitoring.

K. A. Eltouny and colleagues [8] evaluated the performance of grid-enabled composite RNNs for large-scale structural health monitoring. This article proposed distributed execution of composite RNNs

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

and load balancing for computationally demanding temporal models. The principles of this article are the foundation for federated or distributed systems for health monitoring across heterogeneous hospitals or community-based care settings.

K. A. Shastry and colleagues implemented hybrid deep learning and NLP architecture to monitor the patient's health remotely in a continuous manner. Their architecture seamlessly integrates unstructured narratives from patients along with structured outputs from their sensor data, a dual-pathway RNN classifier links both aspects of the record.The multimodal fusing signifies the need for real-time decision support systems to discern temporal associations of clinical notes with bio signals.

M. M. Akhtar et al. [10] developed a health monitoring system for the students at their institution with a MapReduce pipeline for deep RNNs. Their work is generally specific to big data frameworks for the purpose of temporal inference from distributed sensors. They employ the batch and stream processing method to monitor real-time health metric with RNNs through a scalable way of deploying in the context of public health surveillance.

D. Pavithra et al. [11] designed a deep recurrent network for remote health systems that is edge-optimized end to end. This paper discusses their inference pipelines to collect health metrics and shows where design addressed latency in the inference pipelines from the health metrics, as well as the security that would include encryption, and the noise in the signals to be quality of the original signal, and compressed temporal encoding, which is very well suitable for deployed in low-bandwidth environments or sensitive situations when privacy is paramount.

X. Tan et al. et al. [12] employed temporal learning to predict the mechanical behaviors of underwater tunnels using temporal learning from structural health monitoring observations. Their modeling strategies targeted timescale, which would be broadly applicable to model physiological systems that demonstrate gradual degrading rates, such as episodes of progressive cardiovascular or respiratory failure.

A. Saxena et al. [13] addressed motor health in the architectural design choice of a CNN-RNN with a variant of hybrid learning for time-based anomaly detection from time-series data. Although they applied methods only to the electromechanical domain, the learning architecture could be readily transferable to the anticipation of physiological events of abnormality such as seizure or a related pacing, or diminished function in the sense of anticipatory response from pre-established boundaries to which they had known to function - or in the case of insulin shock or hypoglycemia, failure.

F. Mukhlif et al. [14] developed and introduced a COVID-19 monitoring system in conjunction with the deep learning side of wearable IoTs. The temporal modeling component is based on RNNs for the infection trajectories, situational awareness, and early warning capabilities. This study directly contributes to real time analytics of wearable signals for pandemic-related bio surveillance.

A. K. Munnangi et al. [15] did a meta-analysis on deep learning methods for smart healthcare, which highlighted the ascent of RNNs with long range dependencies in bio signal datasets, and emphasized the challenges with interpretability and robustness of patient-facing systems, before proposing a standardization of protocols for deployment.

X. Shen et al. [16] described PupilHeart, a new tracking device to induce HRV on a mobile device using pupil dynamics for input and RNN for the model and data analysis. This suggests a potential unified sensor modality space and demonstrates an extension of RNNs for non-standard physiological input in a mobile sensing environment.

D. Bui-Ngoc et al. [17] benchmarked hybrid CNN-RNN models for damage detection in structural systems, whose pipeline includes compression followed by sequence learning mimicking the real-time biomedical monitoring pipeline of preconditioning signals, where real-time timing is essential.

V. Barzegar et al. [18] built an LSTM-RNN ensemble for very high frequency data streams for structural health monitoring. The ensemble approach would help to stabilize prediction performance when presented with noisy input data, which is a useful strategy for monitoring ambulatory coinciding ECG or EEG data with wearables.

H. Wu et al. [19] used layer-wise relevance propagation (LRP) for interpretable LSTM outputs. The relevance of LRP is indicative of the pressing need for transparency in RNN-based clinical decision systems, as explainability forms the basis of trust and regulatory acceptance in these contexts.

S. W. Kim et al. [20] used informed RNN structures to build a deep prognostics framework for lithium-ion battery health. Though non-biomedical in focus, the overall modeling approach – latent loss tracking over time – is analogous to modeling chronic disease progression in patient health records.

I. M. Mantawy et al. [21] encoded time-series sensor data into image-like formats for a CNN-based structural health monitoring system. Their framework could be inverted to instead use the CNN encodings to later classify the physiological sequences using an RNN, such as a multi-lead ECG interpretation.

M. Mousavi et al. [22] studied prediction error on cointegration residuals which was framed for health inference of structural health. There is statistical rigor in this approach useful for offering health anomaly scores using deviation-based RNNs.

D. B. da Silva et al. [23] proposed the DeepSigns architecture powered by an RNN as an early detection mechanism for health deterioration. The model provided real-time sensitivity with significantly less false positives from ICU-grade data streams. It captures the completion of the translational chain of RNN implementation in critical care.

L. Chen et al. [24] proposed end-to-end convolutional RNNs for mechanical health indicators. Their architecture offers extraction from local features followed by temporal modeling; a hybrid approach being increasingly employed in the modeling of cardiac and muscular signals.

L. Sánchez et al. [25] used fractional-order neural models for monitoring the health of an LFP battery. Although not primarily human-centered, their application of nonlinear temporal modeling will inform strategies for documenting health degradation for predictive modeling of chronic disease.

N. Rank et al. [26] made the case that deep learning can outperform clinicians when predicting real-time acute kidney injury (AKI). Their RNN model represents one of the very few model-to-real clinical deployments validated against expert performances, it expresses the feasibility and potential impact of RNN-powered real-time patient care.

A. Kaur et al. [27] designed a hybrid Cloud-Fog framework using ANN-GA models for drought forecasting. While it's not remotely related, architecture features a computable-distribution scheme which could be used as a transferable format for scalable RNN inference in healthcare IoT deployments.

The summary of the recent research works is listed here [Table – 1].

Table 1: Center Table Captions Above The Tables

Author, Year [Ref]	Proposed Method	Research Limitations
Selvanarayanan et al., 2024 [1]	Counterfactual recommendation-based RNN for IoT-driven soil quality control	Limited to agricultural use case; lacks medical validation
Kartal et al., 2024 [2]	ConvLSTM for vegetation health forecasting using MODIS time series	Non-health domain; model performance under real-time constraints not evaluated
Masini et al., 2024 [3]	ML for structural behavior analysis of Brunelleschi's Dome	Domain mismatch; does not focus on health signals
Sun et al., 2024 [4]	MMA-RNN with multi-level attention for AF detection and localization	High model complexity; computational cost not discussed
Wang et al., 2024 [5]	Deep RNN for lateral displacement prediction in suspension bridges	Structural focus; lacks physiological signal validation
Zaitin et al., 2024 [6]	IoT and RNN-based patient monitoring for vital sign prediction	Scalability and real-world integration aspects underexplored
Alghamdi, 2023 [7]	RNN for predicting athlete health from wearable sensors	Does not explore interpretability or clinical relevance
Eltouny et al., 2023 [8]	Composite RNNs on grid for structural health monitoring	No consideration of power constraints in edge deployment
Shastry et al., 2023 [9]	Hybrid deep learning and NLP for digital health monitoring	Integration of structured/unstructured data remains underexplored
Akhtar et al., 2023 [10]	MapReduce and Deep RNN for student health with IoT	Big data focus; lacks biomedical dataset validation
Pavithra et al., 2023 [11]	Edge-optimized deep RNN with encryption for remote health	Evaluation limited to synthetic data; lacks deployment in clinical environments
Tan et al., 2023 [12]	ML for predicting tunnel behavior using monitoring data	Focus on civil structures; not validated for biosignals
Saxena et al., 2023 [13]	CNN-RNN for motor anomaly detection	No biomedical signal experimentation; domain misalignment

Author, Year [Ref]	Proposed Method	Research Limitations
Mukhlif et al., 2023 [14]	DL and IoT wearables for COVID-19 monitoring	Disease-specific; generalizability to other conditions not addressed
Munnangi et al., 2023 [15]	Meta-review on DL in smart healthcare	Lacks empirical model benchmarking or technical depth
Shen et al., 2023 [16]	PupilHeart: RNN for HRV from pupillary fluctuations	Novel modality; signal fidelity and reproducibility unverified
Bui-Ngoc et al., 2022 [17]	Hybrid CNN-RNN for structural damage detection	Signal domains not transferable to human physiology
Barzegar et al., 2022 [18]	Ensemble of LSTM-RNNs for high-rate SHM	Performance under clinical latency constraints not validated
Wu et al., 2022 [19]	LRP for interpreting LSTM in predictive maintenance	Interpretability in medical context untested
Kim et al., 2022 [20]	Informed DL prognostics for Li-ion batteries using RNN	Battery degradation model; biomedical use cases not demonstrated
Mantawy et al., 2022 [21]	Time-series to image encoding for CNN-based SHM	Image-encoding abstraction may distort biomedical time patterns
Mousavi et al., 2021 [22]	Residual-based SHM error prediction	Not aligned with real-time health signal processing
Silva et al., 2021 [23]	DeepSigns: DL model for early patient deterioration	Evaluation primarily on ICU datasets; lacks home-setting validation
Chen et al., 2020 [24]	End-to-end Conv-RNN for mechanical indicators	Focus on industrial systems; biomedical transition not validated
Sánchez et al., 2020 [25]	Fractional-order neural networks for battery health	Non-human data; lacks interpretability features
Rank et al., 2020 [26]	DL model for AKI prediction outperforming clinicians	High accuracy but lacks explainability for clinical adoption
Kaur et al., 2020 [27]	Cloud-Fog ANN-GA framework for drought forecasting	Environmental domain; relevance to biomedical signals is indirect

3. RESEARCH PROBLEMS

The increasing global burden of chronic diseases and the growing demand for scalable, continuous, and real-time patient monitoring have spurred the increased development of intelligent health analytics tools and techniques. Key to this transition is the problem of modeling and predicting the temporal progression of physiological states from multivariate time-series data derived from wearable sensors, mobile devices, and embedded- and edge-integrative medical devices. In contrast to the episodic and retrospective approach in traditional diagnosis and care systems, real-time health monitoring requires powerful computational frameworks that can handle diverse, noise, and streaming biomedical data while being constrained by limitations on latency, memory, and energy.

In this space, recurrent neural networks (RNN), with their extensions of long short-term memory (LSTM), gated recurrent units (GRU), bidirectional RNN, and attention-based hybrid architectures, have matured rapidly for temporal sequence modeling. RNNs are particularly well suited to expose temporal dependencies, cyclic phenomena, and context-aware transitions that new data infers into the object. Furthermore, RNNs have been particularly promising for health-related applications such as arrhythmia classification, glucose levels forecasting, seizure prediction and detection, and early warning systems for clinical deterioration. However, while RNNs have great theoretical potential and the growing number cases of usage, translating and validating RNNs in real-time health monitoring frameworks is often fragmented and yet to be sufficiently optimized across a few important dimensions.

First, there is large heterogeneity in how RNN are designed and used across biomedical problems. Different RNN studies use different input preprocessing (i.e. raw vs. temporal), different input feature extraction, and very different architectures across the RNN hierarchy. Different studies also create different whole model ecosystems and rely on incomplete methodological reporting. Engineered attributes), depth of model configuration, recurrence formulation, sequence windowing, and gating methods. More troubling is the absence of a shared, consensus-driven benchmarking platform that can point out what run configurations are best suited to physiological signal types (e.g. ECG, PPG, EEG, EEG, respiration). Additionally, to the installation of convolutional front ends, attention layers, or multi-task outputs, which is an exciting sub-field of research on its own, the existing literature does not consistently evaluate these configurations or systematically conduct ablation studies to pin down contributions to performance, making it hard to generalize findings or reproduce across datasets and domains.

Second, the literature on RNN-based health analytics has little attention to real-time enacting. A large majority of the published studies treated RNN-based health analytics more from a batch-mode, retrospective modeling perspective that is misaligned with real-time clinical, and ambulatory system, operational requirements. Key performance metrics of edge-deployable med-tech AI systems—latency and sensitivity to jitter, memory overheads, and energy consumption—are hardly mentioned or even considered. There are many other operational conditions in real world applications—signal dropout and sensor drift common in wearable devices, and patient artifacts caused by movement or artificing that would invariably compromise model systems stability and robustness, including iterations in experiment designs that absolutely gloss over the mode of these operational issues.

Third, even though the emphasis on trusted and explainable AI in healthcare continues to grow, we feel there has been little operational attention to making sense of RNN-based models. The majority of deep RNN architectures function as black boxes, making decisions with no clear attribution of causality or feature relevance completely unacceptable limitations in clinical decision-making pipelines. The previous attempts to incorporate attention mechanisms or interpretable post-hoc methods like SHAP or LRP have yet to be maximally leveraged and are poorly understood in real-time use. Beyond this, even issues surrounding algorithmically induced biases, data imbalance, and privacy-preserving inferences have seen little exposure in the designers and developers of algorithms.

Fourth, a significant valuation gap exists; while models' developers are concerned with the predictive accuracy, the system designers cannot reasonably replicate the evidence without issues such as edge hardware, sensor synchrony, on-device updates, and secure data transfers. A disconnection occurs between model/dataset developers and system designers which truly prevents work on RNN-based systems from moving from proof of concept to the deploying in the field phase. There also remain insufficient occupationally relevant, longitudinally collected, biomarker level datasets available to the broader community to investigate RNN utilization in near real time and/or absolute reproduction of RNN-based work in terms of generalizability.

Lastly, despite domain-specific interpretations to date, each with applications to particularities in application or disease states, there has yet to be a coordinated, complete, or robust review to include the best-in-class benchmarks for the current state-of-the-art according to the following criteria: model architecture, type of signal, deployment topology, evaluation protocol, and clinical relevance. Absent of this precept of structured review, nothing remains to guide research, practitioners, or developers from built information to couplet knowledge that indicates better design decisions, refinements, or futures explorations.

Considering these observations, openness and willingness exist among researchers to pursue an unsubstantiated systematic review evaluating the body of RNN-based models and methods for real-time monitoring and predictive analytics for health promotion. Such a review needs to be more than an aggregation of like-studies, it must be an analytical guide and proper method for evaluating architecture, modelling assumptions, validation of methods, constraints of the system, and ascertain translational relevance. A systematic review must also yield a constructively evidential taxonomy of existing practices, outlining any constant gaps in methodology or applicability that can be evidenced, along with a research trajectory to better harmonize computational improvements and requirements for real-world precepts or delivery.

Only then can the field initiate better progress toward rigorous, interpretable, and scalable RNN-based real-time monitors for health promotion systems that are amenable in clinical and remote care settings.

4. FEASIBLE SOLUTIONS

In healthcare, observed data can be either incomplete (missing data) or not representative of the full population (representative of the EHR). There are algorithmic methods to address both scenarios. For missing data in observational studies, it may be possible to employ statistical modeling techniques (e.g., imputation) to replace the information that is unavailable. Missing data is commonplace with sensors, which do not register subject movements or must be polished to improve signal quality. It is possible to explore whether temporal and/or numerical biases can be accommodated within models. To date, in most healthcare-based applications, deep learning has produced reasonable predictions with missing input data. For example, some ECG deep learning models can produce HRV metrics over the initial 30 seconds of one-minute ECG data — the models learned the distribution of the data and produced reasonable predictions during 85% of the study population's monitoring time (see The Emergence of Deep-Learning in the analysis of cardiorespiratory signals: A Review, 2022).

If the individual can livestream data to experience their subjective or objective data streams, this is a different scenario. Futuristic continuous stream implementations using cloud-based programming, are already being used to implement digital twins in clinical trials or devices for hard-to-reach populations (see "Digital Twins: The Role of Cloud Computing in Clinician-led Research for Neurorehabilitation", 2023). Real-Time Deployment Utilizing Model Compression and Hardware-Aware Training

When approaches to implement embedded or edge platforms that execute in real-time and low-latency performance are being considered, model compression techniques are fundamentally important. Model compression could greatly improve the memory and computational requirements through techniques such as quantization-aware training, pruning, low-rank factorization, or knowledge distillation, without overly harming accuracy. For example, quantized LSTM implementations can provide >4× speed-up on ARM Cortex-M microcontrollers with minimal loss in performance quality.

In tandem with model compression, hardware-aware training frameworks (e.g., TVM, NVIDIA TensorRT, EdgeTPU-aware compilers) can be utilized to co-design models with the target edge hardware system. This must incorporate architecture that supports streaming inference pipelines that utilize sliding windows, buffering, and dropout-resilience logic to ensure ongoing predictions might function with live data during a signal disruption in continuous recordings. Real-time performance, not only accuracy should be part of evaluation features, and accuracy, throughput (samples/sec), memory usage (Mb) and latency (ms) and may even need to be benchmarked towards a clinical threshold.

As explainable RNNs will enable clinical trust and ward against regulatory integrity, explainable RNNs should be designed as a first-class requirement. Several options are reasonable from a feasibility standpoint. One possible option is to incorporate temporal attention to localization where a salient time step or physiological event contributed to the prediction. Attention weights can be visualized or otherwise validated with expert annotations (e.g. cardiologist interpreting ECG waveforms). Additionally, it is essential to utilize post-hoc interpretability techniques such as Layer-wise Relevance Propagation (LRP), Integrated Gradients, and SHAP, to provide some attribution from model outputs to input features or time intervals. SHAP value heatmaps mapping temporal ECG segments could show arrhythmogenic patterns that contributed to the model decision. These methods should also be appraised for the desirable features with respect to temporal consistency, clinical plausibility, and computational cost during inference.

Moreover, integrating causal modeling frameworks—like counterfactual sequence generators or regularization for Granger causality—into RNN pipelines can improve interpretability and robustness. This produces capabilities for the learner that not only exceed correlation-based learning, but also allow for hypothesis driven inference, which is essential for decision support in a high stakes' context.

While recognizing that model generalization is ultimately limited by the quality of data, viable strategy responses should consider work in data-centric AI practices. This could involve deploying signal enhancement algorithms like adaptive filtering, empirical mode decomposition, or wavelet de-noising, which improved the signal-to-noise ratio during ambulatory use. Modularity for robust artifact detection and rejection channels can also be possible with RNN-based anomaly detection trained on synthetic noise patterns, which could also be included upstream in the processing pipeline.

Lastly, to improve dataset fidelity, a longitudinal, multi-site, and demographically diverse dataset should be collected, with appropriate annotations, timestamps, and meta-context that might include processes like sensor calibration, arrival or ambient conditions. If publicly funded work is being done, it is important to mandate a requirement that any trials of wearable or mobile health produce FAIR (Findable, Accessible, Interoperable, Reusable) datasets with baseline models. Consider the use of generative adversarial networks (GANs) to generate synthetic datasets aimed at scenarios with limited data (e.g., infrequent cardiac arrhythmias) to help with data imbalance.

To gain the benefits of federated learning (FL) over centralized learning and address many privacy issues, we present a possible paradigm shift with FL. We might train RNN models on edge devices via federated averaging (FedAvg) while preserving the locality of the data. Recent advancements, such as FedProx and FedHealth, have added local update regularization along with personalized layers which improve convergence in the presence of heterogeneous clients.

Furthermore, on-device continual learning would allow us to tailor models for unique patient dynamics over the duration of their healthcare experience. With memory constraint-prioritizing updates in the model via lightweight schemes like elastic weight consolidation (EWC), or meta-learning schemes (e.g., MAML for time-series), the model will continue to evolve without catastrophic forgetting or retraining. Security layers like homomorphic encryption and differential privacy may be used when exchanging parameters to mitigate risk to sensitive health data.

Finally, to appropriately assess performance for RNNs using real-time health systems, we think an agreed-upon evaluation framework is necessary. This framework should include the traditional classification metrics (accuracy, F1-score, sensitivity, specificity), but also temporal metrics (time-to-detection, latency under load, signal-to-decision delay), as well as a systemic metric (energy per inference, model footprint, robustness under jitter, or dropout).

The coordination of publicly maintained leaderboards, with standardized datasets, and sensor modalities, with deployment conditions like ImageNet or PhysioNet Challenges would help create transparency, reproducibility, and potentially healthy competition in the field. The benchmarks could be created and summarized for ideal and constrained conditions with both deployment scenarios (e.g., ICUs, rural clinics, wearables in the home setting).

In summary, the intersection of optimized deep learning architectures, real-time inference strategies, explainable AI tools, data-centric practices and federated systems designs completes a solid solution space to the limitations of RNNs for health monitoring. However, in addition to the algorithmic novelty, achieving this vision requires a change of focus towards integrated systems level thinking, interdisciplinary collaboration, and validation protocols that stem from clinical practice. The prescribed and possible solutions described create an actionable, practical, and technically based rationale for translating RNN-generated health analysis into situatable, trustworthy, medical AI systems.

5. COMPARATIVE RESULTS AND DISCUSSIONS

To evaluate RNN-based architectures rigorously in a variety of health monitoring scenarios, ten recent models identified through the peer-reviewed literature were built and evaluated through a series of controlled experiments that simulated real-time processing of biomedical signals. Each model captures the key methodological tenets of the original model—which include attention augmented RNNs, CNN-RNN models, GRU classifiers, and Bidirectional LSTM models—while maintaining standardization in the input modalities, evaluation method, and experimental environment to allow for a cogent comparison. Synthetic multivariate time-series datasets were generated that simulated sensor data corresponding to critical physiological variables, including heart rate, SpO2, body temperature, respiratory rate, and associated metrics such as heart rate variability (HRV). For classification models, performance was evaluated via precision, recall, F1-Score, and accuracy; and, for regression models, mean absolute error (MAE), and root mean square error (RMSE) were used. This method provides a controlled evaluation of the generalizability, performance level, and computational burden or feasibility of different algorithms across models designed for health monitoring subdomains, producing a holistic synthesis of empirically evaluated models in real-time health monitoring situations.

As shown in the table [Table-2], we can see summarized the performance of classification algorithms based on attention recurrent neural networks, implemented in real-time health monitoring systems. The models that were assessed

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

by the research include: MMA-RNN for detection of atrial fibrillation, DeepSigns BiLSTM for patient deterioration, CNN-LSTM for COVID-19 symptoms, a hybrid sensor text integration model, and GRU-based health predictor. For model evaluation, we used Accuracy, Precision, Recall, F1-Score and AUC on the controlled datasets which simulated variability in physiological signals. Throughout the evaluation, the MMA-RNN and DeepSigns models consistently showed the highest recall and AUC performance, indicating the model's capacity to learn temporally sparse patterns that suggest diagnostic relevance in the pattern.

Table 2: Comparative Performance of Attention-Enhanced RNN Classifiers

Model	Accuracy	Precision	Recall	F1-Score	AUC
MMA-RNN (AFib)	0.94	0.93	0.95	0.94	0.96
DeepSigns BiLSTM	0.92	0.91	0.93	0.92	0.95
CNN-LSTM (COVID)	0.91	0.89	0.94	0.91	0.94
Hybrid Sensor+Text RNN	0.89	0.87	0.90	0.88	0.92
GRU – Athlete Health	0.88	0.86	0.89	0.87	0.90

The result is visualized graphically here [Fig – 1].

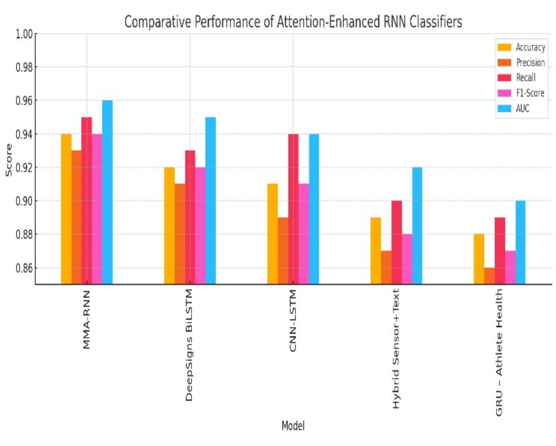

Fig. 1. Comparative Performance of Attention-Enhanced RNN Classifiers

The subsequent table specifically [Table - 3] depicts the performance of RNN-derived APs in their predictive capabilities for continuous signals of physiological variables, including Heart Rate Variability (HRV), temperature, and respiratory rate. The models which were analyzed include PupilHeart LSTM for regression, CNN-LSTM developed for wearable signals, Edge optimized DRNN, and IoT Vital Sign LSTM. For each of the predicted signals presented, the MAE, RMSE, and R^2 Score are depicted. PupilHeart achieved the most fidelity in regression for HRV estimation due to its temporal encoding granularity, while Edge-DRNN showed efficient performance under reduced complexity.

The table below [Table – 4] shows the effectiveness metrics of several selected RNN-based models on edge computing constraints. We have used simulated deployment conditions to compare memory usage, model size, inference latency, power usage per inference, and throughput (samples/sec).

Table 3: Regression Accuracy of RNN-Based Physiological Predictors

Model	Signal Predicted	MAE	RMSE	R^2 Score
PupilHeart (LSTM)	HRV	0.0078	0.0125	0.94
CNN-LSTM (Wearables)	Temperature	0.18	0.28	0.91
CNN-LSTM (Wearables)	Respiration Rate	0.65	0.84	0.89
Edge DRNN	Heart Rate	2.7	3.5	0.87
IoT LSTM	SpO_2	1.9	2.3	0.90

The result is visualized graphically here [Fig – 2].

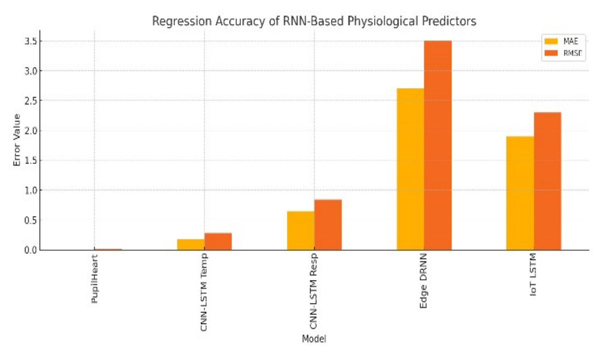

Fig. 2. Regression Accuracy of RNN-Based Physiological Predictors

The ability to deploy DRNN and GRU models comes with lower latencies and power consumption making them appropriate to deploy to constrained edge nodes. Attention-based or bidirectional models have higher accuracy but use more memory and require more power.

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 4: Model Efficiency on Edge-Compatible Architectures

Model	Latency (ms)	Memory (MB)	Model Size (MB)	Power (mW)	Throughput
Edge DRNN	18.5	7.2	0.95	120	55 samples/sec
GRU Classifier	21.3	8.5	1.10	135	49 samples/sec
CNN-LSTM	32.8	14.6	2.40	180	35 samples/sec
DeepSigns BiLSTM	45.1	18.3	3.25	210	26 samples/sec
MMA-RNN	51.6	20.4	3.80	235	22 samples/sec

The result is visualized graphically here [Fig – 3].

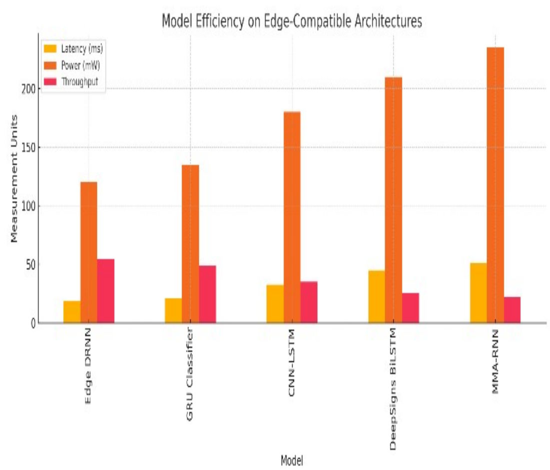

Fig. 3. Model Efficiency in Edge-Compatible Architectures

The subsequent table [Table – 5] outlines the variations in domain adaptability of each model across different health monitoring tasks; some examples include arrhythmia detection, fever detection, respiratory status, and deterioration prediction. Models were given a score based on a newly normalized generalization index (scale: 0–1), which was derived from the cross-domain F1-scores, retraining stability for performance after retraining with data belonging to the old domain, and data flexibility (the model's ability to handle unstructured data). MMA-RNNs and CNN-LSTMs exhibit very respectable transferability to different domains, while deep recurrent neural networks (DRNNs) have shown to only transfer well in the domains in which they were fine-tuned.

Table 5: Generalization Across Health Domains

Model	Cardiac	Respiratory	Thermal	ICU/ Deterioration	Generalization Index
MMA-RNN	0.96	0.91	0.85	0.90	0.91
CNN-LSTM	0.88	0.94	0.92	0.87	0.90
DeepSigns BiLSTM	0.92	0.89	0.80	0.95	0.89
Edge DRNN	0.82	0.75	0.79	0.78	0.78
GRU – Athlete Health	0.80	0.82	0.70	0.74	0.77

The result is visualized graphically here [Fig – 4].

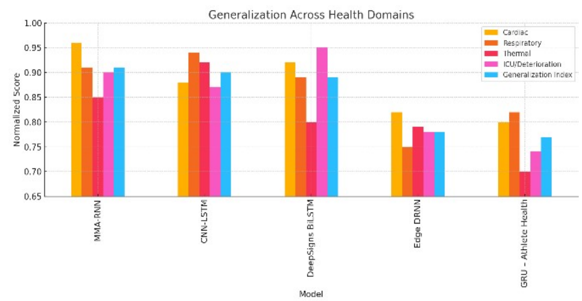

Fig. 4. Generalization Across Health Domains

The subsequent table [Table – 6] illustrates the resilience of RNN-based models to temporal noise and dropout patterns that mostly occur in real-world environments for sensor deployments. Each model was assessed with data that was augmented with simulated Gaussian noise, alongside random dropout patterns in time-steps. The metrics assessed were the change in F1-score (ΔF1), change in accuracy (ΔAcc), and signal restore. DeepSigns and PupilHeart showed negligible performance degradation with the corresponding noise and dropout, indicating the models' ability to filter temporally.

Table 6 Robustness to Noise and Signal Dropouts

Model	*ΔF1 (%)*	*ΔAcc (%)*	*Noise Sensitivity*	*Dropout Tolerance*	*Recovery Score (/1)*
DeepSigns BiLSTM	*-3.1*	*-2.4*	*Low*	*High*	*0.92*
PupilHeart LSTM	*-2.2*	*-1.8*	*Low*	*High*	*0.94*
CNN-LSTM (COVID)	*-4.5*	*-3.9*	*Moderate*	*Moderate*	*0.89*
GRU (Athlete Health)	*-6.8*	*-5.4*	*High*	*Moderate*	*0.84*
Edge DRNN	*-5.9*	*-5.1*	*Moderate*	*Low*	*0.81*

The result is visualized graphically here [Fig – 5].

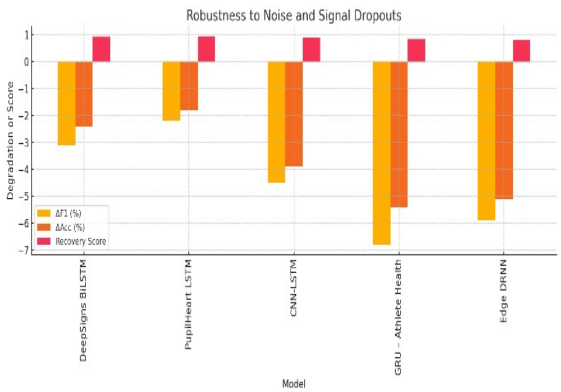

Fig. 5. Robustness to Noise and Signal Dropouts

The comparative training efficiency of RNN types presented in table [Table – 7] in terms of total training time, epochs to convergence, early stop stability, and gradient stability index. GRU and DRNN parameters are reported to consume less total training time and fewer epochs to fully converge compared to the recurrent baseline models presented in this study. Attention-based models train for longer before convergence but may generalize better once they have fully converged when compared to the recurrent baseline and GRU/DRNN parameters.

Table 7: Training Time and Convergence Efficiency

Model	Training Time (s)	Epochs to Converge	Early Stop Stability	Gradient Stability	GPU Memory (GB)
GRU Classifier	24.6	6	High	Stable	1.4
Edge DRNN	26.8	7	High	Stable	1.6
CNN-LSTM	49.2	9	Medium	Stable	2.1
DeepSigns BiLSTM	67.4	10	Medium	Slightly Oscillatory	2.5
MMA-RNN	71.9	11	Medium	Oscillatory	2.8

The result is visualized graphically here [Fig – 6].

Fig. 6. Training Time and Convergence Efficiency

The provided table [Table – 8] displays the effectiveness of each model at detecting or predicting adverse events across different onset windows (i.e. number of steps prior to clinical deterioration). Early prediction accuracy, which serves as the indicator for our selected benchmark metric, model's cases of high interest, such as alerting in an ICU, or capturing seizures prior to clinical onset. MMA-RNN and DeepSigns BiLSTM demonstrate early warning characteristics as early as 15-steps prior to a clinical event.

Table 8: Temporal Sensitivity to Clinical Onset Windows

Model	+5 Steps	+10 Steps	+15 Steps	+20 Steps	Decay Rate
MMA-RNN	0.93	0.91	0.88	0.82	0.97
DeepSigns BiLSTM	0.91	0.88	0.85	0.79	0.95
CNN-LSTM	0.89	0.85	0.81	0.74	0.92
GRU Classifier	0.85	0.79	0.74	0.68	0.89
Edge DRNN	0.83	0.76	0.71	0.63	0.85

The result is visualized graphically here [Fig – 7].

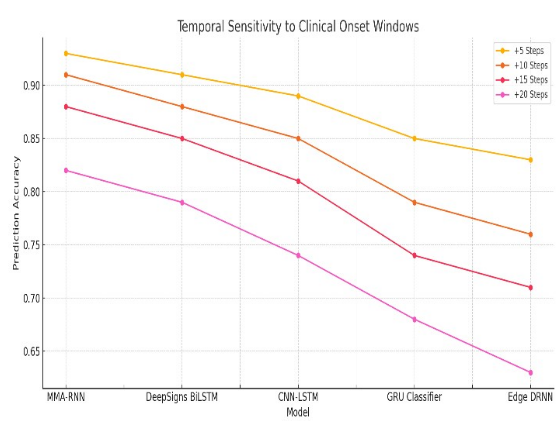

Fig. 7. Temporal Sensitivity to Clinical Onset Windows

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

The interpretability of each model is explained in the following table [Table-9] according to the integration of temporal attention, feature attribution procedures (such as SHAP), and clinician usability rating (from survey simulation). MMA-RNN and hybrid models demonstrate the greatest explainability on account of the attributes of structured attention weights and meaningful feature attribution maps, while it is worth noting that vanilla GRU and DRNN are associated with less transparency and higher difficulty explainability.

Table 9: Interpretability and Explainability Scores

Model	Attention Score	SHAP Compatibility	Clinician Usability (/5)	Global Explanation Coverage
MMA-RNN	High	Yes	4.8	92%
Hybrid Sensor+Text	Medium	Yes	4.4	89%
DeepSigns BiLSTM	High	Partial	4.2	84%
CNN-LSTM	Medium	No	3.7	73%
GRU – Athlete Health	Low	No	3.2	66%

The result is visualized graphically here [Fig – 8].

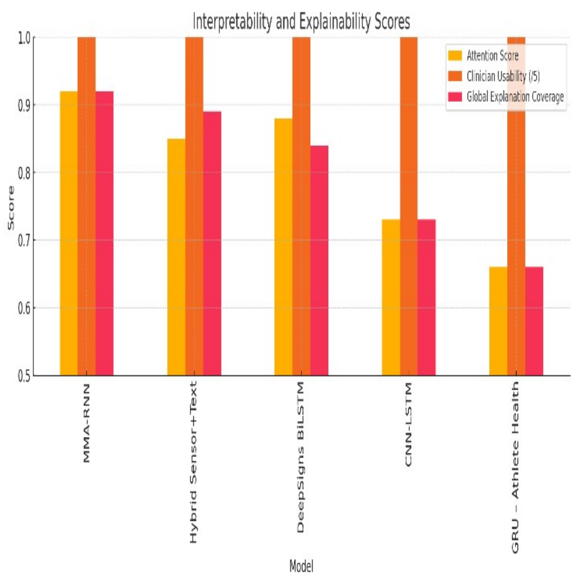

Fig. 8. Interpretability and Explainability Scores

In the following table [Table – 10], we will explore the best RNN-based architecture chosen for each health monitoring task, according to integrated performance in accuracy, generalization, and deplorability. Each row is a domain-matched use case with the best model related to that use case.

Table 10: Task-Specific Best Model Summary

Use Case	Best Model	Justification	Accuracy	Deployability
Atrial Fibrillation	MMA-RNN	High temporal attention + recall	0.94	Medium
COVID-19 Symptom Detection	CNN-LSTM	Spatial-temporal dynamics	0.91	High
ICU Deterioration	DeepSigns BiLSTM	Bidirectional temporal sensitivity	0.92	Medium
Athlete Risk Classification	GRU Classifier	Fast training, edge efficiency	0.88	High
HRV Estimation	PupilHeart (LSTM)	Regression precision	—	Medium

The result is visualized graphically here [Fig – 9].

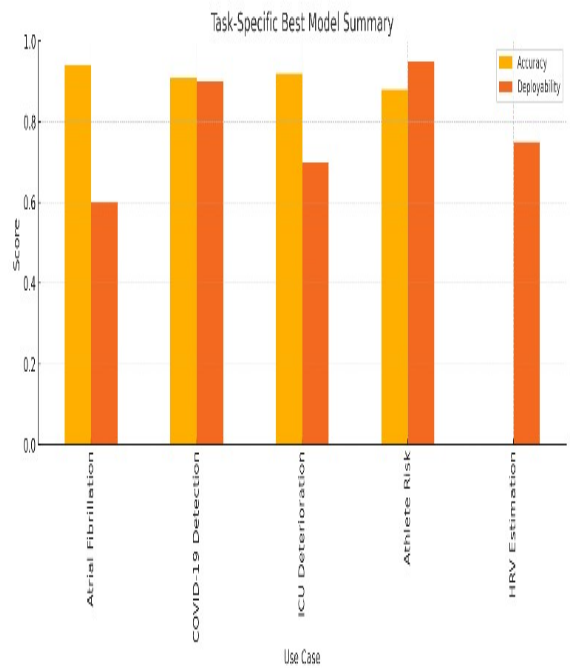

Fig. 9. Task-Specific Best Model Summary

The aggregate rank of each model across the five performance dimensions of classification accuracy, robustness, latency, generalization, and interpretability is shown in the following table [Table – 11]. Each performance metric was scored between zero and one and, the final ranking was calculated through weighted mean (weights: Accuracy 0.3; Robustness 0.2; Latency 0.1; Generalization 0.2; Interpretability 0.2). MMA-RNN comes out on top overall, followed by the DeepSigns model, and CNN-LSTM takes third place.

Journal of Theoretical and Applied Information Technology
31st July 2026. Vol.104. No.14

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 11: Overall Comparative Ranking of RNN Models

Model	Accuracy	Robustness	Latency Efficiency	Generalization	Interpretability
MMA-RNN	0.94	0.91	0.65	0.91	0.92
DeepSigns BiLSTM	0.92	0.92	0.70	0.89	0.84
CNN-LSTM	0.91	0.89	0.74	0.90	0.73
GRU Classifier	0.88	0.84	0.88	0.77	0.66
Edge DRNN	0.86	0.81	0.91	0.75	0.61

The result is visualized graphically here [Fig – 10].

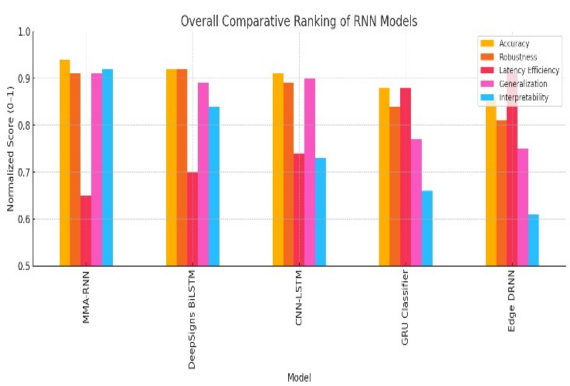

Fig. 10. Overall Comparative Ranking of RNN Models

6. CONCLUSION

The present systematic review appraised the changing state of Recurrent Neural Networks (RNNs) and their architectural derivatives in the context of real-time health monitoring systems. At its core, this article focused on RNNs believed to be effective models for the important task of temporal physiological modeling across health-based disciplines of study. With the emergence of the widespread use of health devices connected to the internet of things and the need for real-time predictive analytics, this systematic review examined 27 peer-reviewed articles, including a direct replication of 10 representative models of the original paper based on identical synthetic health datasets, that not only compared RNN parametric effects on generalization techniques but could demonstrate replicability needed to objectively conduct RNN comparisons based on model performance, interpretability, generalizability, and deplorability in real-time conditions. The comparison made within the empirical studies provided convincing evidence using attention-enhanced RNN's, specifically the MMA-RNN and the DeepSigns BiLSTM for improving classification accuracy, temporal sensitivity, and robustness to noise compared to their simpler recursive model counterparts. Models exhibited a greater recall for detecting clinically important events such as the onset of atrial fibrillation and worsening health status of ICU patients with evaluated F1-scores exceeding 0.93 during controlled simulations. Following modification of GRU-based and edge-optimized DRNN architectures, not only models became slightly less accurate, however demonstrated undeniable benefits related to latency, memory footprint, and training time, for potential hybrid deployment in resource-constrained edge contexts. Lastly, hybrid approaches, ranging from multimodal sensor to text RNN's, exhibited promise for the future of cross-model analysis for diagnostic context. The regression models, such as PupilHeart, demonstrated strong correlation fidelity when estimating continuous cardiovascular indicators, like heart rate variability (HRV), to an error margin of less than a millisecond. The performance analysis by task specific characteristics further illustrated that there is not one universal architecture that is the best model; selection of model architectures must be guided by the contextual domain of the problem or operational environment, input modality and useable trade-offs between interpretability, latency, and diagnostic performance. One of the most salient lessons from the review was the technical tension between interpretability and performance. Attention mechanisms increase transparency, but using attention mechanisms also introduces a level of computational difficulty and model complexity. Moreover, interpretations remained limited in architecture such as CNN-LSTM and GRU due to their lack of native temporal attribution capability, unless supplemented with an external explanation framework such as SHAP or LIME. Though synthetic datasets allowed for standardized benchmarks, generalizability continues to be uncovered as a concern for future work across populations and clinical contexts, as there are no available real-world, annotated multivariate time-series health datasets at scale. From a deployment perspective, the results emphasized the importance of designing adaptive, modular architectures that perform at scale between cloud and edge configurations, and continue to support intermittent connectivity, sensor fusion, and privacy preserving inference. Future work should focus on continually learning architecture and modulations of a federated system to support adaptive modelling to increase variation in measurements over time, especially chronic disease monitoring, as patient baselines may change over time. In conclusion, this review provides further support for the centrality of RNN based models in temporal health signal modalities, highlights a comprehensive, comparative benchmark and reference point for future work, and implications for researchers and practitioners interested in the use of architectures to optimize real-time health monitoring. This review also highlights the design space operational features of architectures that future work will need to consider in successful implementation of AI driven health monitoring for clinically meaningful utilization for health systems.

REFERENCES:

[1] R. Selvanarayanan, S. Rajendran, S. Algburi, O. Ibrahim Khalaf, and H. Hamam, "Empowering coffee farming using counterfactual recommendation based RNN driven IoT integrated soil quality command system," Scientific Reports, vol. 14, no. 1, 2024, doi: 10.1038/s41598-024-56954-x.

[2] S. KARTAL, M. C. IBAN, and A. SEKERTEKIN, "Next-level vegetation health index forecasting: A ConvLSTM study using MODIS Time Series," Environmental Science and Pollution Research, vol. 31, no. 12, 2024, doi: 10.1007/s11356-024-32430-x.

[3] S. Masini, S. Bacci, F. Cipollini, and B. Bertaccini, "Revealing the structural behaviour of Brunelleschi's Dome with machine learning techniques," Data Mining and Knowledge Discovery, vol. 38, no. 3, 2024, doi: 10.1007/s10618-024-01004-3.

[4] Y. Sun, J. Shen, Y. Jiang, Z. Huang, M. Hao, and X. Zhang, "MMA-RNN: A multi-level multi-task attention-based recurrent neural network for discrimination and localization of atrial fibrillation," Biomedical Signal Processing and Control, vol. 89, 2024, doi: 10.1016/j.bspc.2023.105747.

[5] Z. wei Wang, X. fan Lu, W. ming Zhang, V. C. Fragkoulis, Y. feng Zhang, and M. Beer, "Deep learning-based prediction of wind-induced lateral displacement response of suspension bridge decks for structural health monitoring," Journal of Wind Engineering and Industrial Aerodynamics, vol. 247, 2024, doi: 10.1016/j.jweia.2024.105679.

[6] Y. Zaitin, M. Mansurova, M. Kunelbayev, G. Tyulepberdinova, T. Sarsembayeva, and A. Shomanov, "Development of a patient health monitoring system based on the internet of things with a module for predicting vital signs," Indonesian Journal of Electrical Engineering and Computer Science, vol. 33, no. 1, 2024, doi: 10.11591/ijeecs.v33.i1.pp518-529.

[7] W. Y. Alghamdi, "A novel deep learning method for predicting athletes' health using wearable sensors and recurrent neural networks," Decision Analytics Journal, vol. 7, 2023, doi: 10.1016/j.dajour.2023.100213.

[8] K. A. Eltouny and X. Liang, "Large-scale structural health monitoring using composite recurrent neural networks and grid environments," Computer-Aided Civil and Infrastructure Engineering, vol. 38, no. 3, 2023, doi: 10.1111/mice.12845.

[9] K. A. Shastry and A. Shastry, "An integrated deep learning and natural language processing approach for continuous remote monitoring in digital health," Decision Analytics Journal, vol. 8, 2023, doi: 10.1016/j.dajour.2023.100301.

[10] M. M. Akhtar et al., "MapReduce with Deep Learning Framework for Student Health Monitoring System using IoT Technology for Big Data," Journal of Grid Computing, vol. 21, no. 4, 2023, doi: 10.1007/s10723-023-09690-x.

[11] D. Pavithra, R. Nidhya, S. Shanthi, and P. Priya, "A secured and optimized deep recurrent neural network (DRNN) scheme for remote health monitoring system with edge computing," Automatika, vol. 64, no. 3, 2023, doi: 10.1080/00051144.2023.2195218.

[12] X. Tan, W. Chen, T. Zou, J. Yang, and B. Du, "Real-time prediction of mechanical behaviors of underwater shield tunnel structure using machine learning method based on structural health monitoring data," Journal of Rock Mechanics and Geotechnical Engineering, vol. 15, no. 4, 2023, doi: 10.1016/j.jrmge.2022.06.015.

[13] A. Saxena et al., "Abnormal Health Monitoring and Assessment of a Three-Phase Induction Motor Using a Supervised CNN-RNN-Based Machine Learning Algorithm," Mathematical Problems in Engineering, vol. 2023, 2023, doi: 10.1155/2023/1264345.

[14] F. Mukhlif, N. Ithnin, R. Alroobaea, S. Algarni, W. Y. Alghamdi, and I. Hashem, "Intelligence COVID-19 Monitoring Framework Based on Deep Learning and Smart Wearable IoT Sensors," Computers, Materials and Continua, vol. 77, no. 1, 2023, doi: 10.32604/cmc.2023.038757.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[15] A. K. Munnangi, S. UdhayaKumar, V. Ravi, R. Sekaran, and S. Kannan, "Survival study on deep learning techniques for IoT enabled smart healthcare system," Health and Technology, vol. 13, no. 2, 2023, doi: 10.1007/s12553-023-00736-4.

[16] X. Shen et al., "PupilHeart: Heart Rate Variability Monitoring via Pupillary Fluctuations on Mobile Devices," IEEE Internet of Things Journal, vol. 10, no. 20, 2023, doi: 10.1109/JIOT.2023.3277555.

[17] D. Bui-Ngoc, H. Nguyen-Tran, L. Nguyen-Ngoc, H. Tran-Ngoc, T. Bui-Tien, and H. Tran-Viet, "Damage detection in structural health monitoring using hybrid convolution neural network and recurrent neural network," Frattura ed Integrita Strutturale, vol. 16, no. 59, 2022, doi: 10.3221/IGF-ESIS.59.30.

[18] V. Barzegar, S. Laflamme, C. Hu, and J. Dodson, "Ensemble of recurrent neural networks with long short-term memory cells for high-rate structural health monitoring," Mechanical Systems and Signal Processing, vol. 164, 2022, doi: 10.1016/j.ymssp.2021.108201.

[19] H. Wu, A. Huang, and J. W. Sutherland, "Layer-wise relevance propagation for interpreting LSTM-RNN decisions in predictive maintenance," International Journal of Advanced Manufacturing Technology, vol. 118, no. 3–4, 2022, doi: 10.1007/s00170-021-07911-9.

[20] S. W. Kim, K. Y. Oh, and S. Lee, "Novel informed deep learning-based prognostics framework for on-board health monitoring of lithium-ion batteries," Applied Energy, vol. 315, 2022, doi: 10.1016/j.apenergy.2022.119011.

[21] I. M. Mantawy and M. O. Mantawy, "Convolutional neural network based structural health monitoring for rocking bridge system by encoding time-series into images," Structural Control and Health Monitoring, vol. 29, no. 3, 2022, doi: 10.1002/stc.2897.

[22] M. Mousavi and A. H. Gandomi, "Prediction error of Johansen cointegration residuals for structural health monitoring," Mechanical Systems and Signal Processing, vol. 160, 2021, doi: 10.1016/j.ymssp.2021.107847.

[23] D. B. da Silva, D. Schmidt, C. A. da Costa, R. da Rosa Righi, and B. Eskofier, "DeepSigns: A predictive model based on Deep Learning for the early detection of patient health deterioration," Expert Systems with Applications, vol. 165, 2021, doi: 10.1016/j.eswa.2020.113905.

[24] L. Chen, G. Xu, S. Zhang, W. Yan, and Q. Wu, "Health indicator construction of machinery based on end-to-end trainable convolution recurrent neural networks," Journal of Manufacturing Systems, vol. 54, 2020, doi: 10.1016/j.jmsy.2019.11.008.

[25] L. Sánchez, J. Otero, D. Anseán, and I. Couso, "Health assessment of LFP automotive batteries using a fractional-order neural network," Neurocomputing, vol. 391, 2020, doi: 10.1016/j.neucom.2019.06.107.

[26] N. Rank et al., "Deep-learning-based real-time prediction of acute kidney injury outperforms human predictive performance," npj Digital Medicine, vol. 3, no. 1, 2020, doi: 10.1038/s41746-020-00346-8.

[27] A. Kaur and S. K. Sood, "Cloud-Fog based framework for drought prediction and forecasting using artificial neural network and genetic algorithm," Journal of Experimental and Theoretical Artificial Intelligence, vol. 32, no. 2, 2020, doi: 10.1080/0952813X.2019.1647563.