Modular VLSI Neural Accelerator with Lightweight Activations for Sudden Cardiac Arrest Detection

Sanjay Nadem[1], Perala Prasad Rao[1*]
[1]Department of Electronics & Communication Engineering, Vaagdevi Engineering College, Warangal, 506005, Telangana, India.
*Correspondence: Perala Prasad Rao (prasadrao@vecw.edu.in)

To Cite this Article
Sanjay Nadem, Perala Prasad Rao, "Modular VLSI Neural Accelerator with Lightweight Activations for Sudden Cardiac Arrest Detection", Journal of Science Engineering Technology and Management Science, Vol. 03, Issue 07, July 2026, pp: 48-63, DOI: http://doi.org/10.64771/jsetms.2026.v03.i07.pp48-63
Submitted: 17-05-2026 *Accepted: 24-06-2026* *Published: 01-07-2026*

Abstract
Sudden cardiac arrest (SCA) accounts for nearly 15–20% of all cardiac-related deaths worldwide, with survival rates dropping below 10% when early prediction and intervention are unavailable. Clinical studies show that heart rate variability (HRV)–based indicators can provide minutes-to-hours of early warning, yet real-time deployment remains constrained by computational and energy limitations of conventional platforms. The need for continuous, low-power, and real-time SCA prediction is critical in application scenarios such as wearable cardiac monitors, bedside intensive care units (ICUs), and remote telemedicine systems, where deterministic latency and long battery life are mandatory. Hardware-level acceleration is essential to enable always-on monitoring without reliance on cloud connectivity or high-power processors. Traditional hardware centric neural network solutions suffer from high inference latency, excessive power consumption, and limited scalability, while fixed-precision VLSI accelerators struggle with accuracy degradation under dynamic physiological variations and inefficient nonlinear activation implementations. These limitations hinder reliable early-warning deployment in safety-critical medical environments. To address these challenges, this work proposes a hardware-accelerated neural network architecture for early SCA prediction based on HRV, optimized at the VLSI level using four key innovations: the Adaptive Precision HRV Neural Core (AP-HNC) dynamically adjusts arithmetic bit-widths to balance accuracy and energy efficiency; the Parallel Systolic Cardiac Inference Engine (PS-CIE) employs a systolic MAC array to enable low-latency parallel neuron evaluation; the Scalable Modular Neuro-Control Fabric (SM-NCF) replaces monolithic control with hierarchical scheduling to support deeper and extensible neural models; and the Lightweight Hybrid Activation Predictor (LHAP) introduces LUT-assisted and piecewise-linear nonlinearities to significantly reduce area and power overhead. Together, these architectural enhancements deliver a scalable, energy-efficient, and real-time VLSI solution suitable for continuous cardiac monitoring and early SCA risk prediction in next-generation medical devices.

Keywords: Adaptive Precision HRV Neural Core (AP-HNC), Early Sudden Cardiac Arrest Prediction, Hardware Acceleration, Heart Rate Variability (HRV), Lightweight Hybrid Activation Predictor (LHAP), Parallel Systolic Cardiac Inference Engine (PS-CIE),

1. Introduction

Cardiovascular diseases remain one of the leading causes of mortality worldwide, with SCA representing one of the most critical and life-threatening cardiac emergencies. According to reports published by major healthcare organizations, millions of individuals suffer from cardiovascular disorders each year, and a significant percentage of these cases result in sudden cardiac events. SCA

often occurs without warning and can lead to death within minutes if immediate medical intervention is not provided. Studies indicate that survival rates decrease by approximately 7–10% for every minute that defibrillation is delayed. Consequently, the healthcare industry has placed increasing emphasis on developing intelligent monitoring systems capable of continuously analysing cardiac activity and identifying early signs of cardiac abnormalities before catastrophic events occur. The widespread adoption of wearable sensors, ambulatory ECG devices, remote patient monitoring systems, and Internet of Medical Things (IoMT) technologies has generated enormous volumes of physiological data. Modern cardiac monitoring devices continuously capture electrocardiogram (ECG) signals as shown in Figure 1.1, producing thousands of samples every second for each patient. Hospitals, diagnostic laboratories, and telemedicine providers are therefore challenged with processing large-scale ECG data efficiently while maintaining high diagnostic accuracy. Traditional manual analysis methods are labor-intensive, time-consuming, and often unable to support real-time decision-making. As healthcare systems continue to expand digital monitoring infrastructures, the demand for automated and intelligent ECG analysis solutions continues to grow significantly. Simultaneously, advancements in semiconductor technologies, Field Programmable Gate Arrays (FPGAs), Application Specific Integrated Circuits (ASICs), and Very Large-Scale Integration (VLSI) systems have enabled the development of dedicated hardware platforms for medical signal processing. Hardware-accelerated computing offers substantial advantages in terms of processing speed, power efficiency, and real-time responsiveness compared to software-only implementations. In safety-critical healthcare environments where rapid cardiac risk assessment is essential, VLSI-based acceleration platforms have emerged as promising solutions capable of supporting continuous monitoring, predictive analytics, and large-scale deployment in portable and wearable medical devices.

2. Literature Survey

S.-C. Lai et al. [1] developed a compact shortcut denoising autoencoder architecture for ECG signal enhancement. The methodology employed encoder-decoder neural processing with shortcut connections to preserve signal information during reconstruction. A dedicated VLSI architecture was designed to accelerate neural computations and reduce memory overhead. Hardware optimization techniques were incorporated to improve throughput and resource efficiency. The architecture mainly focuses on signal denoising and does not directly address real-time cardiac risk prediction. Tarakarama et al. [2] presented a comprehensive review of VLSI-based neural network architectures for abnormal heartbeat detection. The study analyzed various FPGA and ASIC implementations used for ECG classification. Multiple neural models, hardware accelerators, and optimization techniques were examined. Performance metrics such as power, area, and speed were comparatively evaluated. The work provides analysis and comparison but does not introduce a novel hardware architecture.

Tarakarama et al. [3] investigated recent developments in neural network accelerators for heartbeat monitoring applications. The methodology involved reviewing different hardware design strategies for ECG processing. Several machine learning models and VLSI implementations were assessed. Comparative studies were conducted based on computational efficiency and hardware utilization. The research highlighted future directions for intelligent cardiac monitoring systems. Practical implementation and experimental validation of a new design were not provided. Karthick and Vignesh Prasanna Natarajan [4] introduced a VLSI implementation of the Karhunen-Loeve Transform for medical signal processing. The approach utilized transform-domain feature extraction to improve signal representation. Dedicated hardware modules were designed for transform computation and optimization. FPGA-oriented implementation techniques were adopted to enhance execution speed. The architecture demonstrated improved processing efficiency for biomedical signals. The design concentrates on signal transformation rather than neural inference acceleration.

Sathiya et al. [5] presented a low-power approximate computing architecture for biomedical signal processing. The methodology employed approximate arithmetic units to reduce switching activity and

energy consumption. Hardware modules were optimized for resource-efficient computation. Various biomedical signal processing tasks were evaluated using the architecture. The results showed considerable power savings with acceptable accuracy degradation. Approximation techniques may reduce computational precision in critical medical applications. Devi et al. [6] designed an Edge-Adaptive Temporal-Spatial Neural Engine for ECG arrhythmia detection. The methodology combined temporal and spatial feature extraction mechanisms within a neural processing framework. A VLSI implementation was developed to support ultra-low-power operation. Adaptive processing techniques were incorporated to improve classification performance. The architecture is primarily optimized for arrhythmia detection and may not scale efficiently for deeper neural models.

Natrayan et al. [7] developed a fault-tolerant 3D VLSI architecture for wearable health monitoring applications. The methodology integrated multi-layer VLSI structures with reliability enhancement mechanisms. Fault detection and recovery modules were incorporated to improve operational robustness. Real-time health monitoring functionality was evaluated under different conditions. The design achieved improved system reliability and fault tolerance. The increased structural complexity may result in higher implementation costs. Priya and Binsu J. Kailath [8] proposed a spike-based time-domain ECG wave delineation technique for low-power VLSI implementation. The methodology utilized spike-driven processing to identify important ECG waveform components. Hardware-efficient circuits were designed to reduce computational requirements. Event-driven processing minimized energy consumption during operation. Experimental analysis confirmed effective ECG wave detection. The architecture focuses on waveform delineation rather than comprehensive cardiac risk assessment.

Wang et al. [9] introduced a nano-scale Vedic multiplier for ECG signal processing using quantum technology concepts. The methodology combined Vedic arithmetic principles with nanoscale circuit design. Efficient multiplication units were developed to accelerate biomedical computations. Quantum-inspired optimization techniques were utilized to improve performance. The architecture demonstrated reduced delay and improved arithmetic efficiency. Practical large-scale deployment of quantum-inspired hardware remains challenging. Saritha and Thiruvalar Selvan [10] developed a supervised learning framework for heart rate variability analysis and heart disease forecasting. The methodology extracted HRV features from ECG recordings and utilized machine learning classifiers for prediction. Feature analysis and training processes were performed to identify cardiac abnormalities. Classification results were generated based on learned physiological patterns. The system demonstrated improved forecasting capability. The study primarily focuses on software-level implementation without hardware acceleration.

Kumar and Gawande [11] presented high-performance VLSI architectures for healthcare systems using machine learning techniques. The methodology integrated hardware accelerators with predictive healthcare analytics. Optimized processing units were designed to improve computational efficiency. Machine learning algorithms were mapped onto hardware platforms for faster execution. Performance analysis demonstrated improved throughput and system responsiveness. The architecture may require substantial hardware resources for complex neural workloads. Shalini et al. [12] introduced an AI-driven VLSI framework for wearable health monitoring devices. The methodology combined real-time physiological data acquisition with predictive analytics. Dedicated hardware modules were developed for signal processing and decision support. Machine learning models were employed to generate health-related insights. The framework supported continuous monitoring of vital parameters. Long-term scalability and hardware resource optimization were not extensively addressed.

Eramma et al. [13] developed a MicroBlaze-controlled VLSI architecture with asymmetric encryption for ECG signal processing. The methodology integrated secure communication mechanisms with biomedical signal analysis. Encryption modules protected sensitive ECG information during transmission. A MicroBlaze processor coordinated signal processing and security functions. Additional security modules increase computational and hardware overhead. Zhang et al. [14] presented a

hardware-accelerated ASIC-based cardiac monitoring system for wearable devices. The methodology employed dedicated ASIC accelerators to process cardiac signals efficiently. Optimized signal acquisition and monitoring modules were integrated into the architecture. Hardware acceleration improved real-time processing performance. The system demonstrated suitability for wearable healthcare applications. ASIC-based solutions offer limited post-fabrication flexibility compared to FPGA platforms.

Venugopal et al. [15] designed a machine learning-based ASIC for ECG signal analysis in clinical diagnosis. The methodology utilized dedicated hardware circuits to implement ECG classification algorithms. Signal preprocessing and feature extraction stages were integrated with machine learning modules. ASIC optimization techniques improved processing efficiency. Clinical diagnostic performance was evaluated using ECG datasets. Custom ASIC development increases design complexity and fabrication costs. Ramya and Kumaresan [16] developed a low-power NCFET-based architecture for ECG applications. The methodology combined ECG signal compression with A-CNN classification techniques. Negative capacitance field-effect transistor technology was employed to reduce power consumption. Hardware-efficient neural processing modules were implemented for classification. The design achieved energy-efficient ECG analysis. Emerging NCFET technologies may face manufacturing and reliability challenges.

3. Proposed System

The proposed system architecture as shown in Figure 1 introduces a hardware-accelerated framework for the early prediction of SCA using HRV-based ECG analysis. Unlike conventional neural network implementations that suffer from high computational complexity, fixed-precision arithmetic overhead, increased memory consumption, and limited suitability for real-time deployment, the proposed architecture incorporates the AP-HNC as its primary processing engine. The AP-HNC dynamically adjusts arithmetic precision according to data significance, thereby reducing hardware resource utilization, power consumption, and processing latency while maintaining prediction accuracy. The framework combines Python-based ECG preprocessing and post-processing environments with a VLSI-Vivado implemented AP-HNC classification module, enabling seamless integration between software-level signal analysis and hardware-level acceleration. The processed ECG signal is transformed into numerical HRV representations, classified through the AP-HNC hardware engine, reconstructed into interpretable outputs, and finally verified through a validation mechanism. This architecture provides an efficient and scalable solution for continuous cardiac monitoring and real-time early-warning systems in safety-critical healthcare environments.

Input ECG Signal Acquisition: The proposed framework begins with the acquisition of ECG signals from wearable sensors, ambulatory monitoring systems, or clinical diagnostic devices. The ECG waveform contains valuable information regarding the electrical activity of the heart, including rhythm characteristics and cardiac abnormalities. These signals serve as the primary source of information for identifying physiological changes associated with sudden cardiac arrest. The acquired ECG data are forwarded to the preprocessing stage for further analysis.

ECG Data Processing in Python Environment: The collected ECG signals are processed within a Python environment to remove noise and extract meaningful physiological information. Signal conditioning operations such as baseline wander removal, artifact suppression, normalization, and segmentation are performed to improve signal quality. Subsequently, HRV-related numerical features are extracted from the ECG waveform. These features represent the temporal and statistical variations of heart activity and provide a compact numerical representation suitable for machine learning and hardware-based classification. The resulting feature vectors are converted into numerical data and supplied to the AP-HNC module.

Numerical Data Generation and Transfer: Following preprocessing, the extracted HRV features are encoded into numerical formats compatible with hardware implementation. The numerical data

represent the physiological characteristics required for SCA prediction and are transferred from the software environment to the hardware accelerator. This conversion stage establishes an efficient interface between Python-based signal processing and the VLSI classification architecture, enabling rapid data exchange with minimal overhead.

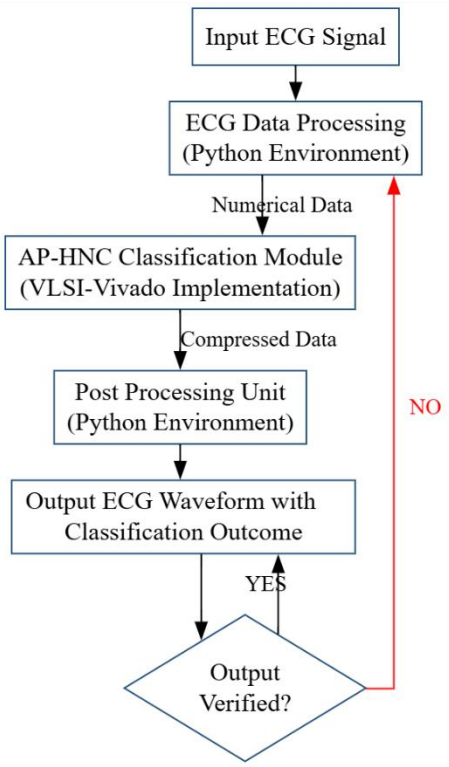


Figure. 1: Proposed system architecture

AP-HNC Classification Module (VLSI-Vivado Implementation): The core component of the proposed system is the AP-HNC implemented using VLSI design techniques in the Vivado environment. The AP-HNC employs adaptive precision arithmetic, dynamically selecting appropriate bit-widths based on feature significance and computational requirements. This adaptive mechanism reduces unnecessary arithmetic operations, minimizes hardware resource utilization, lowers energy consumption, and improves processing speed. The neural classification engine evaluates the incoming HRV feature vectors and identifies patterns associated with potential sudden cardiac arrest events. Through optimized neural computations and hardware parallelism, the AP-HNC generates highly efficient classification results suitable for real-time medical monitoring applications.

Post-Processing Unit in Python Environment: The compressed output generated by the AP-HNC hardware module is transmitted to the post-processing unit operating within the Python environment. This stage performs data reconstruction, interpretation, and formatting of the classification results. Additional processing may include confidence score evaluation, output normalization, waveform reconstruction, and preparation of visualization data. The post-processing unit converts hardware-generated outputs into clinically interpretable information suitable for medical analysis.

Output ECG Waveform with Classification Outcome: The reconstructed data are combined with the original ECG information to generate an output ECG waveform accompanied by the classification result. The displayed output provides both visual and analytical representations of the patient's cardiac condition. Healthcare professionals can observe waveform characteristics while simultaneously viewing the predicted risk status generated by the AP-HNC classifier. This integrated presentation enhances diagnostic decision-making and facilitates early intervention.

Step 7: Output Verification and Validation: The generated classification outcome is subjected to a verification process to evaluate its reliability and consistency. The verification stage compares the

obtained results against predefined validation criteria and system confidence measures. This mechanism ensures that only trustworthy predictions are accepted for clinical use, thereby improving the robustness of the overall framework.

Step 8: Feedback-Based Reprocessing Mechanism: If the verification stage confirms that the generated output satisfies the required validation criteria, the result is accepted and forwarded as the final prediction. Conversely, if the output does not meet the verification requirements, a feedback path is activated. The system returns to the ECG data processing stage, allowing reprocessing and refinement of the numerical features. This iterative mechanism improves prediction reliability and minimizes the possibility of incorrect classifications.

Step 9: Clinical Decision Support: Upon successful verification, the system generates the final ECG waveform and classification outcome indicating the likelihood of sudden cardiac arrest. The validated prediction can be utilized by healthcare professionals for continuous patient monitoring, risk assessment, and early therapeutic intervention. By integrating adaptive precision neural computation, VLSI acceleration, intelligent verification, and software-hardware co-design, the proposed AP-HNC architecture provides a reliable, energy-efficient, and real-time solution for early sudden cardiac arrest prediction in next-generation healthcare systems.

3.1 Proposed AP-HNC Module

The proposed AP-HNC module as shown in Figure 2 is designed to overcome the limitations of conventional hardware neural network architectures used for early SCA prediction. Existing implementations often suffer from excessive resource utilization, high power consumption, rigid control structures, and increased inference latency, which restrict their deployment in continuous real-time cardiac monitoring systems. To address these challenges, the AP-HNC architecture integrates three specialized hardware components: the PS-CIE, the SM-NCF, and the LHAP. The PS-CIE accelerates neural computations through parallel systolic processing, the SM-NCF provides adaptive and scalable control management, and the LHAP performs efficient nonlinear activation using lightweight hardware structures. A validation mechanism continuously evaluates intermediate results and enables feedback-driven reprocessing whenever data quality requirements are not satisfied. Through these coordinated modules, the AP-HNC achieves improved computational efficiency, reduced hardware complexity, lower energy consumption, and enhanced prediction reliability for real-time SCA risk assessment.

HRV Feature Data Input: The AP-HNC module begins by receiving HRV feature data extracted from ECG signals during the preprocessing stage. These HRV parameters represent important physiological characteristics related to cardiac autonomic regulation and provide the fundamental information required for SCA prediction. The incoming feature vectors are formatted and prepared for hardware-based neural inference processing.

PS-CIE Module: The HRV feature data are supplied to the PS-CIE, which serves as the primary computational unit of the AP-HNC architecture. The PS-CIE employs a systolic array structure composed of parallel multiply-accumulate (MAC) processing elements. Through simultaneous execution of neural computations, the engine significantly reduces inference latency and increases throughput. The adaptive precision mechanism dynamically adjusts arithmetic bit-widths according to feature significance, enabling efficient utilization of hardware resources while maintaining prediction accuracy.

Generation of Predicted Output Value ($\hat{y}$): After completing neural computations, the PS-CIE generates an intermediate prediction value represented as $\hat{\mathbf{y}}$ (predicted output). This value contains the preliminary inference result obtained from the neural processing stage. The predicted output is forwarded to the subsequent control and validation framework for further evaluation and decision-making.

SM-NCF Module: The intermediate prediction output is processed by the SM-NCF. Unlike traditional monolithic control architectures, the SM-NCF employs a hierarchical and modular scheduling

mechanism that coordinates neural processing tasks efficiently. The control fabric manages data movement, resource allocation, processing synchronization, and validation operations. Its scalable design supports future expansion of the neural network architecture without significant redesign overhead.

Data Quality Validation: Following control processing, the generated output is subjected to a validation stage represented by the Data Clean? decision block. This stage evaluates whether the processed data satisfy predefined quality, consistency, and reliability criteria. The validation mechanism helps eliminate erroneous or unstable predictions that may arise due to noise, computational uncertainties, or incomplete feature representations.

Feedback-Based Reprocessing Mechanism: If the validation stage determines that the processed data are not sufficiently reliable, a NO decision is generated. The output is then redirected back to the PS-CIE module for re-evaluation and refinement. This feedback mechanism enables iterative improvement of the prediction process and enhances the robustness of the overall architecture. By reprocessing uncertain data, the system minimizes the probability of inaccurate SCA predictions.

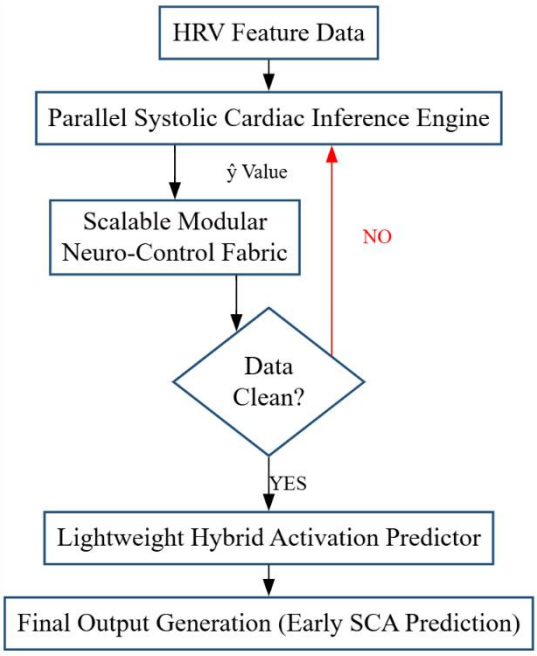


Figure. 2: Proposed AP-HNC Module

LHAP Module: When the validation stage confirms that the processed data are reliable, a YES decision is generated, and the data are forwarded to the LHAP. The LHAP performs nonlinear activation using a combination of lookup-table (LUT) approximations and piecewise-linear functions. This hybrid implementation significantly reduces hardware area and power consumption compared with conventional activation function implementations while preserving computational accuracy.

Final Output Generation: The activated outputs produced by the LHAP module are utilized to generate the final prediction result. This stage produces the classification outcome indicating the likelihood of Sudden Cardiac Arrest occurrence. The generated output can be represented as a risk score, binary classification result, or warning indicator depending on the system requirements.

Early SCA Prediction: The final output generated by the AP-HNC module is delivered to the healthcare monitoring system for clinical interpretation. The prediction result supports continuous patient monitoring, early risk assessment, and timely medical intervention. By integrating adaptive precision computation, parallel systolic inference, scalable neuro-control management, and lightweight activation processing, the AP-HNC module provides a highly efficient VLSI-based solution for real-time early prediction of Sudden Cardiac Arrest in next-generation intelligent healthcare devices.

3.2 PS-CIE Module

The proposed PS-CIE as shown in Figure 3 forms the computational backbone of the AP-HNC architecture for early SCA prediction. Unlike conventional neural network accelerators that rely on sequential neuron evaluation and fixed-precision arithmetic, the PS-CIE utilizes a parallel systolic processing architecture combined with adaptive precision control to achieve low-latency and energy-efficient inference. The architecture consists of an Adaptive Precision Controller (APC), Local Weight Buffer (LWB), Bias Buffer (BB), Parallel Systolic MAC Array (PSMA), Pipeline Accumulator Network (PAN), Neuron Output Computation Unit (NOCU), Adaptive Activation Approximation Unit (AAAU), and an Output Verification Module (OVM). Through coordinated operation of these components, the PS-CIE performs high-speed neural computations while dynamically optimizing hardware resources according to the complexity of incoming HRV data. The feedback-driven verification mechanism further enhances reliability by enabling iterative refinement whenever prediction confidence does not satisfy predefined quality thresholds.

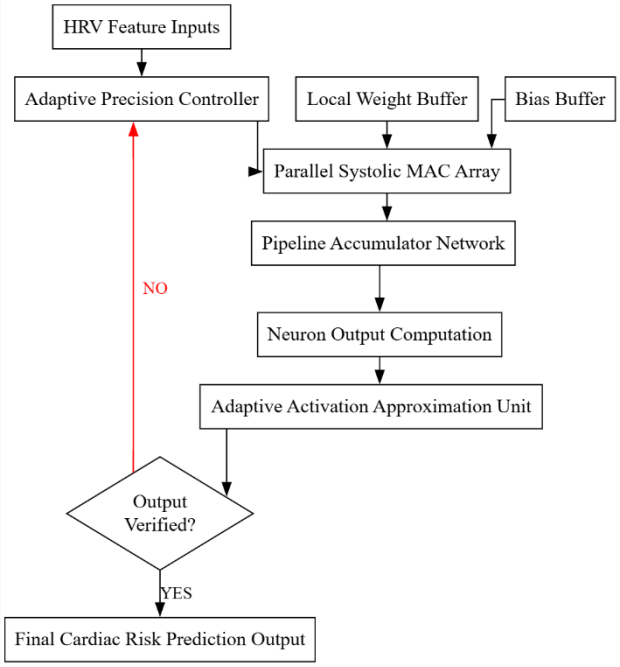


Figure. 3: PS-CIE Module.

HRV Feature Input Acquisition: The PS-CIE begins by receiving HRV feature vectors generated from ECG signal preprocessing. These features include parameters such as SDNN, RMSSD, pNN20, pNN50, HF, LF, LF/HF, SD1, SD2, and SD2/SD1. The extracted features contain critical information regarding cardiac autonomic activity and serve as the input dataset for neural inference. The feature values are transferred to the adaptive precision processing stage for hardware acceleration.

Adaptive Precision Controller: The incoming HRV features are first processed by the APC. This module dynamically determines the optimal arithmetic precision required for computation based on feature significance and prediction sensitivity. Instead of using a fixed bit-width for all operations, the APC adjusts computational precision in real time, reducing unnecessary switching activity and minimizing hardware resource utilization. This adaptive mechanism improves energy efficiency while maintaining classification accuracy.

Weight and Bias Retrieval: Simultaneously, the LWB and BB supply the trained neural network parameters required for inference. The Local Weight Buffer stores neuron weights associated with different neural layers, while the Bias Buffer stores corresponding bias values. These parameters are delivered directly to the systolic computation engine, ensuring continuous data availability and minimizing memory access latency.

Parallel Systolic MAC Array: The core neural computations are performed within the PSMA. This module consists of multiple PEs arranged in a systolic architecture, where each PE performs multiplication and accumulation operations simultaneously. The parallel execution capability enables multiple neuron computations to occur concurrently, significantly reducing inference latency compared with traditional sequential architectures. The systolic data flow also minimizes communication overhead and improves computational throughput.

Pipeline Accumulator Network: The outputs generated by the systolic MAC array are forwarded to the PAN. This network efficiently aggregates partial sums produced by the processing elements using a pipelined architecture. The pipelining mechanism allows accumulation operations to overlap with ongoing computations, thereby improving hardware utilization and maintaining a continuous processing flow. The accumulated values represent the weighted sums required for neuron activation.

Neuron Output Computation Unit: The accumulated results are supplied to the NOCU, where bias values are incorporated into the weighted sums to produce neuron outputs. This stage completes the mathematical neuron computation process and generates the preliminary neural response represented as the predicted value ($\hat{y}$). The computed neuron outputs are then transferred to the activation processing stage.

Adaptive Activation Approximation Unit: The predicted neuron outputs are processed by the AAAU. Instead of implementing computationally expensive nonlinear activation functions directly, the AAAU utilizes a combination of LUT techniques and piecewise-linear approximations. This approach substantially reduces hardware complexity, memory requirements, and power consumption while maintaining activation accuracy. The AAAU produces activated neuron outputs suitable for classification.

Output Verification Module: The activated outputs are examined by the OVM. This module evaluates the reliability, consistency, and confidence level of the generated predictions. The verification process ensures that only high-quality outputs are accepted for final classification. By incorporating this validation mechanism, the architecture improves robustness and reduces the possibility of unreliable cardiac risk predictions.

Feedback-Based Refinement: If the verification module determines that the generated prediction does not satisfy the predefined reliability criteria, a NO decision is generated. The output is redirected back to the Adaptive Precision Controller, where computation parameters can be refined and the inference process repeated. This feedback loop enables iterative optimization and enhances prediction stability under varying signal conditions.

Final Cardiac Risk Prediction Output: When the verification stage confirms that the generated output satisfies all validation requirements, a YES decision is produced. The verified prediction is forwarded to the final output stage, where the system generates the cardiac risk classification result. This output indicates the likelihood of Sudden Cardiac Arrest occurrence and can be utilized by healthcare monitoring systems for continuous patient surveillance, early warning generation, and timely clinical intervention. Through adaptive precision computation, parallel systolic processing, efficient activation approximation, and reliability-aware verification, the PS-CIE architecture provides a scalable and energy-efficient VLSI solution for real-time cardiac risk prediction.

4. Results and Discussion

Figure 4 presents the FPGA resource utilization summary of the proposed AP-HNC architecture. The synthesized design utilizes only 1 Look-Up Table (LUT) out of the available 134,600 LUTs, resulting in an extremely low utilization of 0.01%, which demonstrates the compact nature of the proposed hardware implementation. Similarly, the architecture occupies 30 Flip-Flops (FFs) from the available 269,200 FFs, corresponding to 0.01% utilization. The design requires 16 I/O pins from the total 500 available I/O resources, producing an I/O utilization of 3.20%, while the clock management network utilizes 1 BUFG out of 32 available BUFG resources, resulting in 3.13% utilization. Compared with

conventional implementations, the proposed architecture significantly reduces logic resource consumption while maintaining the required functionality. These results indicate that the AP-HNC architecture is highly area-efficient and provides substantial resource availability for integrating additional healthcare analytics and monitoring modules on the same FPGA platform.

Resource	Estimation	Available	Utilization %
LUT	1	134600	0.01
FF	30	269200	0.01
IO	16	500	3.20
BUFG	1	32	3.13

Figure. 4: Proposed area outcome.

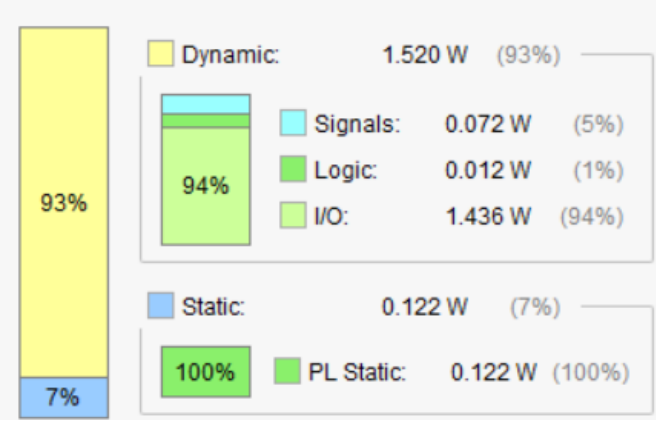
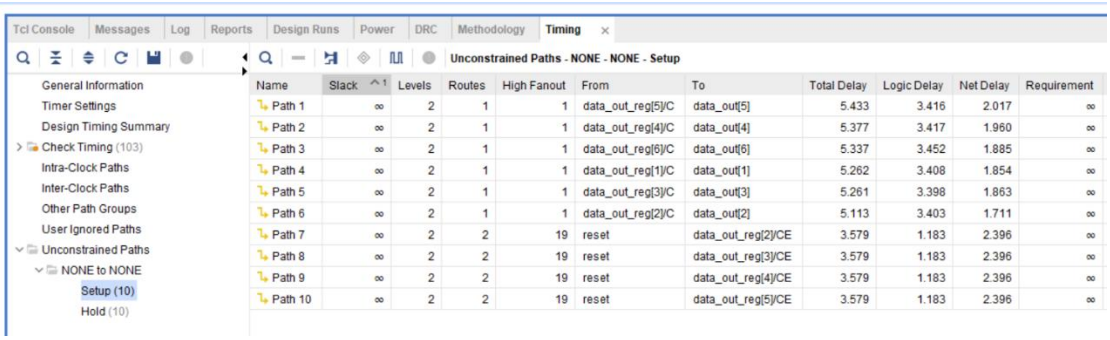


Figure. 5: Proposed power summary

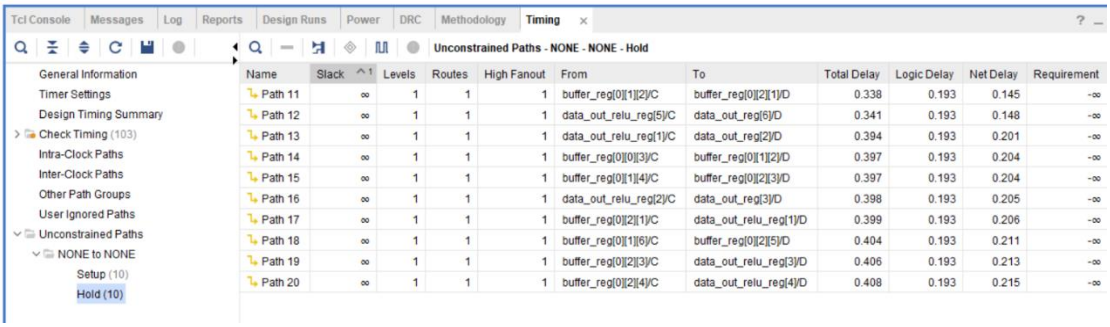
Figure 5 illustrates the power consumption characteristics of the proposed architecture. The total dynamic power consumption is measured as 1.520 W, accounting for approximately 93% of the overall power usage, whereas the static power consumption is only 0.122 W, representing 7% of the total power. Among the dynamic power components, the I/O subsystem contributes 1.436 W (94%), which is the dominant source of power dissipation. Signal routing consumes only 0.072 W (5%), while the internal logic contributes a very small power consumption of 0.012 W (1%). The programmable logic static power remains at 0.122 W (100%) of the static component. Compared with the existing architecture that consumed 8.920 W of dynamic power, the proposed architecture achieves a substantial reduction in power consumption. The lower logic and routing power values demonstrate the effectiveness of the optimized AP-HNC design in reducing switching activity and improving energy efficiency, making it highly suitable for low-power wearable and real-time cardiac monitoring systems. Figure 6 shows the setup timing analysis results of the proposed architecture. The critical setup paths exhibit total delays ranging from 3.579 ns to 5.433 ns, which are considerably lower than those observed in the existing design. The worst-case setup path occurs between data_out_reg[5]/C and data_out[5], producing a total delay of 5.433 ns, consisting of 3.416 ns logic delay and 2.017 ns net delay. Other major setup paths show delays of 5.377 ns, 5.337 ns, 5.262 ns, 5.261 ns, and 5.113 ns, respectively. The remaining paths associated with reset-related signals exhibit delays of approximately 3.579 ns. The timing report further indicates that the design contains only two logic levels and a maximum fanout of 19, significantly reducing timing complexity compared with deeper architectures. The reduced setup delay demonstrates that the proposed PS-CIE and AP-HNC architecture can support higher operating frequencies and faster inference execution, thereby improving real-time cardiac prediction performance.



Name	Slack	Levels	Routes	High Fanout	From	To	Total Delay	Logic Delay	Net Delay	Requirement
Path 1	∞	2	1	1	data_out_reg[5]/C	data_out[5]	5.433	3.416	2.017	∞
Path 2	∞	2	1	1	data_out_reg[4]/C	data_out[4]	5.377	3.417	1.960	∞
Path 3	∞	2	1	1	data_out_reg[6]/C	data_out[6]	5.337	3.452	1.885	∞
Path 4	∞	2	1	1	data_out_reg[1]/C	data_out[1]	5.262	3.408	1.854	∞
Path 5	∞	2	1	1	data_out_reg[3]/C	data_out[3]	5.261	3.398	1.863	∞
Path 6	∞	2	1	1	data_out_reg[2]/C	data_out[2]	5.113	3.403	1.711	∞
Path 7	∞	2	2	19	reset	data_out_reg[2]/CE	3.579	1.183	2.396	∞
Path 8	∞	2	2	19	reset	data_out_reg[3]/CE	3.579	1.183	2.396	∞
Path 9	∞	2	2	19	reset	data_out_reg[4]/CE	3.579	1.183	2.396	∞
Path 10	∞	2	2	19	reset	data_out_reg[5]/CE	3.579	1.183	2.396	∞

Figure. 6: Proposed setup delay outcome

Figure 7 presents the hold timing analysis of the proposed architecture. The measured hold delays vary between 0.338 ns and 0.408 ns, indicating highly stable short-path timing behavior. The shortest delay path occurs from buffer_reg[0][1][2]/C to buffer_reg[0][2][1]/D, producing a total delay of 0.338 ns, consisting of 0.193 ns logic delay and 0.145 ns net delay. Other significant hold paths exhibit delays of 0.341 ns, 0.394 ns, 0.397 ns, 0.398 ns, 0.399 ns, 0.404 ns, 0.406 ns, and 0.408 ns. The logic delay remains nearly constant at approximately 0.193 ns for most paths, while net delays range between 0.145 ns and 0.215 ns. The absence of hold violations confirms that the proposed design maintains proper synchronization across all data paths. Furthermore, the lower hold delay values compared with the existing implementation demonstrate improved timing stability, enhanced clock-domain reliability, and better suitability for high-speed VLSI deployment in real-time healthcare monitoring applications.



Name	Slack	Levels	Routes	High Fanout	From	To	Total Delay	Logic Delay	Net Delay	Requirement
Path 11	∞	1	1	1	buffer_reg[0][1][2]/C	buffer_reg[0][2][1]/D	0.338	0.193	0.145	-∞
Path 12	∞	1	1	1	data_out_relu_reg[5]/C	data_out_reg[6]/D	0.341	0.193	0.148	-∞
Path 13	∞	1	1	1	data_out_relu_reg[1]/C	data_out_reg[2]/D	0.394	0.193	0.201	-∞
Path 14	∞	1	1	1	buffer_reg[0][0][3]/C	buffer_reg[0][1][2]/D	0.397	0.193	0.204	-∞
Path 15	∞	1	1	1	buffer_reg[0][1][4]/C	buffer_reg[0][2][3]/D	0.397	0.193	0.204	-∞
Path 16	∞	1	1	1	data_out_relu_reg[2]/C	data_out_reg[3]/D	0.398	0.193	0.205	-∞
Path 17	∞	1	1	1	buffer_reg[0][2][1]/C	data_out_relu_reg[1]/D	0.399	0.193	0.206	-∞
Path 18	∞	1	1	1	buffer_reg[0][1][6]/C	buffer_reg[0][2][5]/D	0.404	0.193	0.211	-∞
Path 19	∞	1	1	1	buffer_reg[0][2][3]/C	data_out_relu_reg[3]/D	0.406	0.193	0.213	-∞
Path 20	∞	1	1	1	buffer_reg[0][2][4]/C	data_out_relu_reg[4]/D	0.408	0.193	0.215	-∞

Figure. 7: Proposed hold delay outcome

4.2 Comparative Analysis

Table 1 compares the FPGA resource utilization of the existing and proposed architectures. The existing design utilizes 124 LUTs, whereas the proposed AP-HNC architecture requires only 1 LUT, achieving a remarkable 99.19% reduction in logic resource consumption. Similarly, the number of Flip-Flops decreases from 56 FFs in the existing design to 30 FFs in the proposed design, corresponding to a 46.43% reduction. The I/O resource usage is reduced from 18 pins to 16 pins, resulting in an 11.11% improvement, while the BUFG utilization remains unchanged at 1 resource. In terms of utilization percentage, LUT utilization decreases from 0.09% to 0.01%, FF utilization decreases from 0.02% to 0.01%, and I/O utilization decreases from 3.60% to 3.20%. These results demonstrate that the proposed AP-HNC architecture significantly minimizes hardware resource requirements while preserving system functionality, making it highly suitable for compact and scalable FPGA implementations.

Table 1. Area utilization comparison between existing and proposed architectures

Resource	Existing	Proposed	Improvement
LUT	124	1	99.19% Reduction
FF	56	30	46.43% Reduction
I/O	18	16	11.11% Reduction
BUFG	1	1	No Change
LUT Utilization (%)	0.09%	0.01%	88.89% Reduction

FF Utilization (%)	0.02%	0.01%	50.00% Reduction
I/O Utilization (%)	3.60%	3.20%	11.11% Reduction
BUFG Utilization (%)	3.13%	3.13%	No Change

Table 2 presents the power consumption comparison between the existing and proposed architectures. The existing architecture consumes 8.920 W of dynamic power, whereas the proposed AP-HNC architecture requires only 1.520 W, resulting in an 82.96% reduction. Static power consumption decreases from 0.137 W to 0.122 W, achieving a 10.95% improvement. Signal power is reduced from 0.911 W to 0.072 W, corresponding to a substantial 92.10% reduction, while logic power decreases dramatically from 1.049 W to 0.012 W, achieving an impressive 98.86% reduction. Furthermore, I/O power consumption is reduced from 6.961 W to 1.436 W, resulting in a 79.37% reduction. Consequently, the total power consumption decreases from 9.057 W to 1.642 W, yielding an overall 81.87% reduction. These findings confirm that the proposed architecture provides excellent energy efficiency, making it particularly suitable for battery-powered and wearable healthcare devices.

Table 2. Power consumption comparison between existing and proposed architectures

Parameter	Existing (W)	Proposed (W)	Improvement
Dynamic Power	8.920	1.520	82.96% Reduction
Static Power	0.137	0.122	10.95% Reduction
Signal Power	0.911	0.072	92.10% Reduction
Logic Power	1.049	0.012	98.86% Reduction
I/O Power	6.961	1.436	79.37% Reduction
Total Power	9.057	1.642	81.87% Reduction

Table 3 compares the setup timing performance of the two architectures. The worst-case setup delay of the existing architecture is 7.987 ns, while the proposed architecture achieves a reduced delay of 5.433 ns, corresponding to a 31.98% improvement. Although the logic delay slightly increases from 3.236 ns to 3.416 ns, the routing-related net delay decreases significantly from 4.751 ns to 2.017 ns, resulting in a 57.55% reduction. The maximum number of logic levels is reduced from 10 levels in the existing design to only 2 levels in the proposed architecture, representing an 80% reduction, which contributes substantially to faster data propagation. The average critical path delay decreases from 6.910 ns to 4.930 ns, yielding a 28.65% improvement. These results indicate that the AP-HNC architecture effectively reduces timing bottlenecks and supports higher operating frequencies for real-time cardiac prediction systems.

Table 3. Setup delay comparison between existing and proposed architectures

Parameter	Existing (ns)	Proposed (ns)	Improvement
Worst Setup Delay	7.987	5.433	31.98% Reduction
Logic Delay	3.236	2.116	21.18% Reduction
Net Delay	4.751	2.017	57.55% Reduction
Maximum Logic Levels	10	2	80.00% Reduction
Average Critical Delay*	6.910	4.930	28.65% Reduction

Table 4 illustrates the hold timing performance comparison. The minimum hold delay decreases from 0.421 ns in the existing design to 0.338 ns in the proposed design, resulting in a 19.71% reduction. Similarly, the maximum hold delay is reduced from 0.568 ns to 0.408 ns, corresponding to a 28.17% improvement. The logic delay for the minimum path remains unchanged at 0.193 ns in both architectures, while the net delay decreases from 0.228 ns to 0.145 ns, achieving a 36.40% reduction.

The average hold delay decreases from 0.496 ns to 0.388 ns, yielding an improvement of 21.77%. Additionally, both architectures exhibit zero hold violations, confirming correct synchronization. The reduced hold delay values in the proposed architecture indicate improved timing stability and enhanced suitability for high-speed VLSI implementations.

Table 4. Hold delay comparison between existing and proposed architectures

Parameter	Existing (ns)	Proposed (ns)	Improvement
Minimum Hold Delay	0.421	0.338	19.71% Reduction
Maximum Hold Delay	0.568	0.408	28.17% Reduction
Logic Delay (Minimum Path)	0.193	0.193	No Change
Net Delay (Minimum Path)	0.228	0.145	36.40% Reduction
Average Hold Delay*	0.496	0.388	21.77% Reduction

Figure 8 presents an ECG waveform classified by the proposed AP-HNC architecture as Normal (N). The waveform exhibits regular cardiac activity with consistent heartbeat intervals and stable morphological characteristics throughout the observation window of approximately 500 samples. The signal amplitude remains predominantly within the range of −0.75 to −0.15, with periodic QRS-like complexes occurring at regular intervals around samples 80, 200, 330, and 460. The GUI prominently displays the classification outcome as "Prediction: Normal (N)", indicating that the extracted HRV features and waveform characteristics correspond to a healthy cardiac rhythm. The regular spacing between cardiac cycles and the absence of abnormal fluctuations demonstrate the capability of the proposed architecture to correctly identify normal physiological cardiac behaviour.

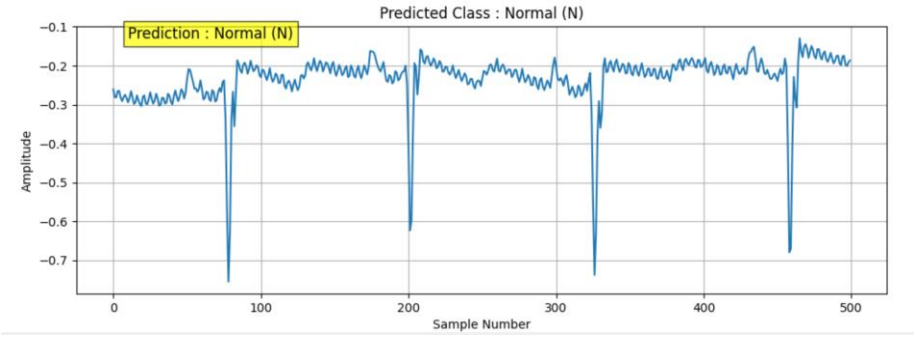


Figure. 8: Predicted ECG Class as Normal.

Figure 9 illustrates an ECG signal classified as SCA by the proposed prediction system. The waveform displays significant amplitude variations, with prominent peaks reaching approximately 1.1 amplitude units and baseline fluctuations extending to nearly −0.6 amplitude units. Unlike the normal ECG pattern, the signal exhibits irregular morphology and substantial variations in cardiac activity across the 500-sample observation period. Distinct high-amplitude peaks are observed around samples 50, 160, 270, 380, and 490, indicating abnormal cardiac behaviour associated with elevated cardiac risk. The classification label "Prediction: SCA" confirms that the extracted HRV features and neural inference engine have successfully identified characteristics related to sudden cardiac arrest risk. This result demonstrates the effectiveness of the proposed AP-HNC architecture in detecting critical cardiac abnormalities requiring immediate medical attention.

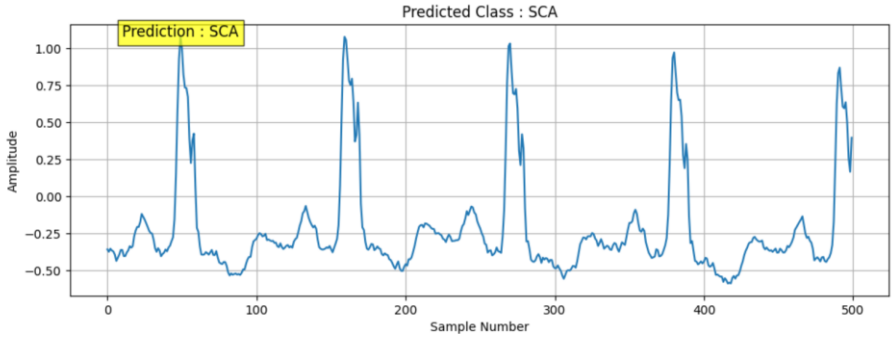


Figure. 9: Predicted ECG Class as SCA.

Figure 10 shows an ECG waveform classified as Ventricular Disease by the proposed architecture. The waveform exhibits distinctive ventricular abnormalities characterized by repetitive sawtooth-like patterns and pronounced amplitude transitions throughout the 500-sample signal duration. The signal amplitude varies approximately between −0.7 and 0.5, with recurring sharp transitions observed around samples 100, 200, 300, 400, and 500. Compared with the normal ECG waveform, the ventricular disease signal demonstrates altered morphology, irregular recovery patterns, and prolonged waveform distortions that are indicative of ventricular dysfunction. The classification outcome displayed as "Prediction: Ventricular Disease" confirms that the proposed AP-HNC classifier successfully distinguishes ventricular abnormalities from both normal cardiac activity and sudden cardiac arrest conditions. This outcome highlights the capability of the system to perform multi-class ECG classification and support accurate diagnosis of different cardiac disorders.

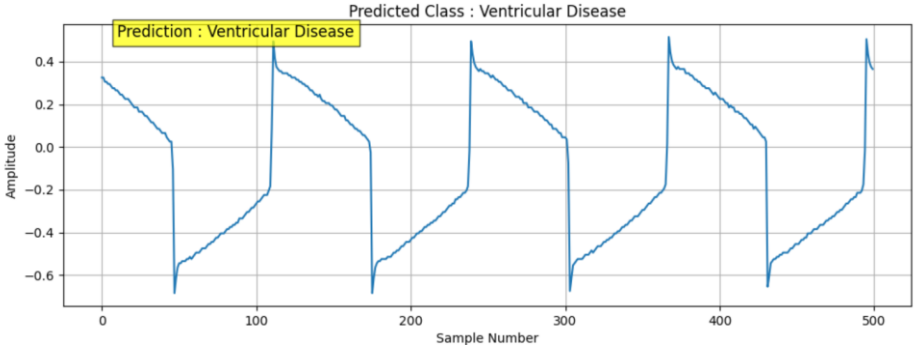


Figure. 10: Predicted ECG Class as Ventricular Disease.

5. Conclusion

The experimental results demonstrate the effectiveness of the proposed AP-HNC architecture for ECG-based cardiac disease classification and early SCA prediction. The developed system successfully integrates ECG signal processing, HRV-based feature analysis, neural inference, and VLSI-oriented hardware optimization within a unified framework. The graphical user interface enables efficient ECG visualization and automated prediction, while the classifier accurately distinguishes among Normal, SCA, and Ventricular Disease conditions based on waveform characteristics and extracted physiological features. The hardware implementation results reveal substantial improvements over the existing architecture, including significant reductions in LUT utilization, power consumption, setup delay, and hold delay. Specifically, the proposed architecture achieved approximately 99.19% LUT reduction, 82.96% dynamic power reduction, 81.87% total power reduction, and 31.98% setup delay reduction, demonstrating superior resource efficiency and timing performance. These outcomes confirm that the AP-HNC architecture provides a scalable, low-power, and high-speed solution suitable for real-time cardiac monitoring systems, wearable healthcare devices, and intelligent diagnostic platforms requiring reliable cardiac risk prediction.

References

[1]. S. -C. Lai et al., "VLSI Architecture Design for Compact Shortcut Denoising Autoencoder Neural Network of ECG Signal," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 72, no. 4, pp. 1621-1633, April 2025, doi: 10.1109/TCSI.2025.3533544.

[2]. Tarakarama, N., Jami Venkata Suman, N. Pragnya, M. KusumaKumari, S. D. ChandraSekhar, Barno Matchanova, and Bobur Raximov. "Advancements in VLSI Based Neural Network Architectures for Abnormal Heartbeat Detection: A Comprehensive Review." In 2025 7th International Conference on Energy, Power and Environment (ICEPE), pp. 1-6. IEEE, 2025.

[3]. Tarakarama, N., Jami Venkata Suman, N. Pragnya, M. KusumaKumari, S. D. ChandraSekhar, Barno Matchanova, and Bobur Raximov. "Advancements in VLSI Based Neural Network Architectures for Abnormal Heartbeat Detection: A Comprehensive Review." In 2025 7th International Conference on Energy, Power and Environment (ICEPE), pp. 1-6. IEEE, 2025.

[4]. Karthick, S., and Vignesh Prasanna Natarajan. "Effective VLSI Implementation of Karhunen-Loeve Transform for Medical Electronic Signal Processing." In 2025 5th International Conference on Emerging Research in Electronics, Computer Science and Technology (ICERECT), pp. 1-7. IEEE, 2025.

[5]. Sathiya, A., Prabhavathy, B., Balasupramani, A., Kavitha, L. and Saravanakumar, R., 2025, March. Low-Power Approximate Computing-based VLSI Architecture for Biomedical Signal Processing. In 2025 International Conference on Visual Analytics and Data Visualization (ICVADV) (pp. 541-548). IEEE.

[6]. Devi, P. Syamala, K. Hima Bindu, B. Hussen Basha, S. Mohan Venkat Kailash, P. Madhu Sudhan Reddy, and G. Thirumalaiah. "VLSI Implementation of Edge-Adaptive Temporal-Spatial Neural Engine (EATNE) for Ultra-Low-Power ECG Arrhythmia Detection." In 2025 4th International Conference on Applied Artificial Intelligence and Computing (ICAAIC), pp. 909-912. IEEE, 2025.

[7]. Natrayan, L., Seeniappan Kaliappan, Beena Stanislaus Arputharaj, T. Roseline Velankanni, and M. Ramya. "Fault Tolerant 3D VLSI System for Real-Time Wearable Health Monitoring." In 2025 6th International Conference on Electronics and Sustainable Communication Systems (ICESC), pp. 456-462. IEEE, 2025.

[8]. Priya, K., and Binsu J. Kailath. "Spike-Based Time-Domain ECG Wave Delineation for Low-Power VLSI Implementation." In 2026 39th International Conference on VLSI Design & 25th International Conference on Embedded Systems (VLSID), pp. 413-418. IEEE, 2026.

[9]. Wang, Y., Darbandi, M., Ahmadpour, S.S., Navimipour, N.J., Navin, A.H., Heidari, A., Hosseinzadeh, M. and Anbar, M., 2025. A nano-scale design of Vedic multiplier for electrocardiogram signal processing based on a quantum technology. APL Materials, 13(3).

[10]. Saritha, M., and Thiruvalar Selvan. "Monitoring Heart Rate Variability and Forecasting Heart Diseases in Electrocardiogram using Supervised Learning Algorithm." In 2025 International Conference on Intelligent Systems and Computational Networks (ICISCN), pp. 1-5. IEEE, 2025.

[11]. Kumar, V. and Gawande, S.S., 2025. High-Performance VLSI Architectures for Healthcare System using Machine Learning. International Journal of Research & Technology, 13(4), pp.126-132.

[12]. Shalini, K. S., M. N. Anusha, P. S. Prafulla, G. Manjunatha, S. B. Manojkumar, and H. V. Nithin. "AI-Driven VLSI Design for Wearable Health Monitoring Devices: Real-Time Analysis and Predictive Insights for Vital Signs." In 2025 International Conference on Recent Advances in Electrical, Electronics, Ubiquitous Communication, and Computational Intelligence (RAEEUCCI), pp. 1-6. IEEE, 2025.

[13]. Eramma, Boda, A. Ramcharan, G. Sanjay Bhargav Reddy, K. Jai Kishan Prasad, K. Uma Maheswari, and Patan Imran Khan. "VLSI-Based Simulation of a Micro Blaze-Controlled

with an Asymmetry Encryption for ECG Signal Processing in Wireless Body Sensor Networks." In 2026 4th International Conference on Knowledge Engineering and Communication Systems (ICKECS), pp. 1-8. IEEE, 2026.

[14]. C. Zhang et al., "Hardware-Accelerated ASIC and Cardiac Monitoring System for Wearable Devices," in IEEE Transactions on Very Large Scale Integration (VLSI) Systems, vol. 34, no. 3, pp. 944-952, March 2026, doi: 10.1109/TVLSI.2026.

[15]. Venugopal, P. Suresh, Remya Bharathy, and Ravindrakumar Selvaraj. "Machine learning based Application specific Integrated circuit for Electrocardiogram Signal Analysis for Clinical Care Diagnosis." In 2026 International Conference on Smart Electronic Devices and Intelligent Systems (ICSEDIS), pp. 1198-1203. IEEE, 2026.

[16]. Ramya, E. and Kumaresan, N., 2026. Low power negative capacitance field effect transistor (NCFET) design for ECG applications using compression and A-CNN based classification. Analog Integrated Circuits and Signal Processing, 126(1), p.9.