Adaptive Multi-Threshold Encoding for Energy-Efficient ECG Classification Architecture Using Spiking Neural Network

Diware, Sumit; Dong, Yingzhou; Yaldagard, Mohammad Amin; Hamdioui, Said; Bishnoi, Rajendra

DOI
10.23919/date64628.2025.10993026

Publication date
2025

Document Version
Final published version

Published in
2025 Design, Automation and Test in Europe Conference, DATE 2025 - Proceedings

Adaptive Multi-Threshold Encoding for Energy-Efficient ECG Classification Architecture using Spiking Neural Network

Sumit Diware Yingzhou Dong Mohammad Amin Yaldagard Said Hamdioui Rajendra Bishnoi

Computer Engineering Lab, Delft University of Technology, Delft, The Netherlands.
Email: {S.S.Diware, M.A.Yaldagard, S.Hamdioui, R.K.Bishnoi}@tudelft.nl

***Abstract*—Timely identification of cardiac arrhythmia (abnormal heartbeats) is vital for early diagnosis of cardiovascular diseases. Wearable healthcare devices facilitate this process by recording heartbeats through electrocardiogram (ECG) signals and using AI-driven hardware to classify them into arrhythmia classes. Spiking neural networks (SNNs) are well-suited for such hardware as they consume low energy due to event-driven operation. However, their energy-efficiency and accuracy are constrained by encoding methods that translate real-valued ECG data into spikes. In this paper, we present an SNN-based ECG classification architecture featuring a new adaptive multi-threshold spike encoding scheme. This scheme adjusts encoding window and granularity based on the importance of ECG data samples, to capture essential information with fewer spikes. We develop a high-accuracy SNN model for such spike representation, by proposing a technique specifically tailored to our encoding. We design a hardware architecture for this model, which incorporates optimized layer post-processing for energy-efficient data-flow and employs fixed-point quantization for computational efficiency. Moreover, we integrate this architecture with our encoding scheme into a system-on-chip implementation using TSMC 40 nm technology. Our approach provides up to 5.1× energy-efficiency compared to state-of-the-art SNN-based ECG classifiers, with high accuracy.**

I. Introduction

Cardiovascular diseases (CVDs) are a group of disorders involving heart and blood vessels [1]. They are a leading cause of global deaths, with the World Health Organization projecting around 23 million deaths by 2030 [2], [3]. Given such high mortality impact, early diagnosis of CVDs is essential for timely medical intervention. This can be achieved by identifying abnormal heart activity known as arrhythmia, which is prevalent among CVD patients. There exist several types of arrhythmia based on its underlying cause and classifying them is crucial for CVD diagnosis [4]. Wearable healthcare devices offer the most convenient solution for this task [5]. They contain sensors to record the heart activity as electrocardiogram (ECG) signal, which is then processed by a classifier hardware to identify the arrhythmia class. This classifier hardware typically employs neural networks as they inherently excel at classification.

Artificial neural network (ANN) based ECG classifier hardware [6]–[9] consumes high energy due to its continuous real-valued operation. In contrast, SNN-based ECG classifier hardware achieves energy-efficiency because of its event-based operation. Its performance mainly depends on the encoding scheme used to convert real-valued inputs into spikes. This scheme affects energy consumption by determining the number of spikes. It also controls accuracy by governing the information quality in the spikes. The encoding schemes in existing SNN-based ECG classifier hardware limit either its energy-efficiency or its accuracy. For instance, Poisson encoding [10] generates long spike sequences, consuming higher energy. Level-crossing encoding [11], [12] produces fewer spikes than Poisson encoding, but still consumes significant energy due to a high enough spike count. Dual-purpose binary encoding [13] further reduces the spike count for better energy efficiency. However, its aggressive data discarding causes information loss and limits the classification accuracy on complex data with finer important features (e.g. diverse patient groups). Hence, there is a need for an SNN-based ECG classifier hardware that is equipped with energy-efficient and accurate spike encoding.

In this paper, we present an SNN-based ECG classification architecture that incorporates a new adaptive multi-threshold spike encoding scheme. This encoding refines both its focus region and granularity as per the significance of heartbeat data samples, acquiring high quality information using fewer spikes for robust classification. We create a high-accuracy SNN compatible with this spike format, using a model development technique introduced specifically for our encoding. We design a hardware architecture for this SNN model, which includes optimized layer post-processing for energy-efficient data-flow and fixed-point quantization for computational efficiency. Finally, we combine this SNN architecture with the encoding scheme to implement an ECG classification system-on-chip using TSMC 40 nm technology. Our key contributions are as follows:

- Propose an adaptive multi-threshold spike encoding to capture high quality information with fewer spikes.
- Introduce a technique to develop high-accuracy SNN model customized for our encoding.
- Implement a system-on-chip optimized for our SNN-based ECG classifier architecture with on-chip encoding.

Our proposed design achieves up to 5.1× energy-efficiency compared to state-of-the-art SNN-based ECG classification

This work is partially funded by the European Union, DAIS (Grant No. 101007273), NEUROKIT2E (Grant No. 101112268), and is also supported by the TU Delft AI labs program.

978-3-9826741-0-0/DATE25/© 2025 EDAA

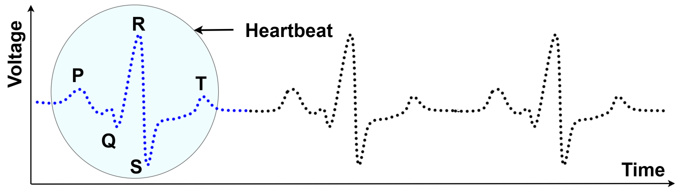

Fig. 1. ECG signal with ‘PQRST’ cycle annotated for one of its heartbeats.

hardware, while also ensuring high classification accuracy.

The rest of the paper is organized as follows: Section II presents the basics of arrhythmia classification and spiking neural networks. Section III provides details of the proposed encoding and ECG classifier design, followed by simulation results in Section IV. Finally, Section V concludes the paper.

II. Background

A. Arrhythmia and its Classification

Electrocardiogram (ECG) signal shown in Fig. 1 records the heart activity as voltage changes over time. It captures multiple contraction-relaxation cycles of the four heart chambers, known as heartbeats. A heartbeat begins with contraction of the upper two heart chambers called atria, marked as ‘P’. This is followed by atrial relaxation and contraction of the lower two heart chambers called ventricles. The three phases of ventricular contraction are indicated as ‘Q’, ‘R’, and ‘S’. Finally, the heartbeat ends with relaxation of ventricles denoted as ‘T’.

Arrhythmia refers to the deviation of a heartbeat from its expected (normal) pattern. There exist various arrhythmia classes based on the nature of such deviation. Thus, arrhythmia classification can help in identifying underlying heart issues and aid in the diagnosis of cardiovascular diseases. The most widely used dataset in arrhythmia classification research is the MIT-BIH arrhythmia dataset [14] with 15 arrhythmia classes. These classes are further grouped into five super-classes by Association for the Advancement of Medical Instrumentation (AAMI) [15] shown in Fig. 2. However, AAMI grouping does not account for severity impact of these classes. This is addressed by severity-based super-class organization in [9] depicted in Fig. 2. As the severity-based grouping offers better practical utility for both the patients and the medical professionals, we adopt it in the context of this work.

B. Spiking Neural Networks

Spiking neural networks (SNNs) perform computations only during discrete events called spikes and remain idle otherwise, unlike conventional artificial neural networks (ANNs) which operate continuously [16]–[19]. This event-driven nature saves energy, making SNNs suitable for energy-constrained healthcare devices like wearables [20]. SNNs use spiking neurons as their building blocks, with the integrate-and-fire (IF) neuron variant being the most widely used [21] as shown in Fig. 3. It accumulates the input spikes through weighted connections. When this accumulation exceeds a pre-defined threshold, it fires an output spike. These neurons can be organized into layers and joined by weighted connections to create an SNN as shown in Fig. 3. Input layer feeds the input spike patterns to subsequent layers. The hidden and output layers process these patterns while communicating via spikes. Once all the input spikes are processed, output layer determines the predicted class.

AAMI class	Sub-classes
Normal beat	N, L, R, e, j
Supraventricular ectopic beat	A, a, J, S
Ventricular ectopic beat	V, E
Fusion beat	F
Unknown Beat	Q, f, /

Severity class	Sub-classes
Normal beat	N, /
Mild beat	L, R, e, j
Moderate beat	A, a, J, S, E
Severe beat	F, V, f, Q

Fig. 2. AAMI and severity groupings for 15 MIT-BIH classes in [14].

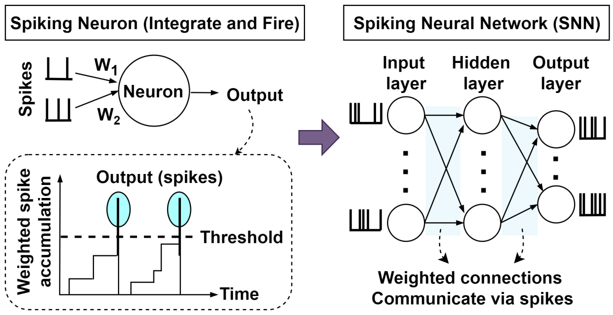

Fig. 3. Fundamentals of spiking neural networks.

III. Proposed SNN-based ECG Classification

A. Overview

SNN-based ECG classification, as shown in Fig. 4, involves two key steps: (i) encoding the ECG data into spikes, and (ii) performing classification with SNN hardware. Conventional encoding methods like [10], [11] generate long encoded spike sequences. This leads to high energy consumption and reduced battery life in wearable healthcare devices. In contrast, methods like [13] reduce the spike count by encoding very few samples of the original data. This degrades the information quality in the spikes and limits classification accuracy on complex data (e.g. varied patient demographics) that requires attention to finer details. Additionally, conventional SNN hardware cannot overcome these issues, leading to suboptimal ECG classification.

We propose an adaptive multi-threshold spike encoding to overcome these challenges. It achieves this by adjusting the region of interest within the heartbeat and the encoding granularity. First, it performs fine-grained encoding on a broad important subset of the heartbeat. This captures subtle data variations as high quality information and needs fewer spikes due to using a subset of the total data. Next, it re-encodes a smaller and more important region within the broad subset in a coarse-grained manner. This adds new critical information

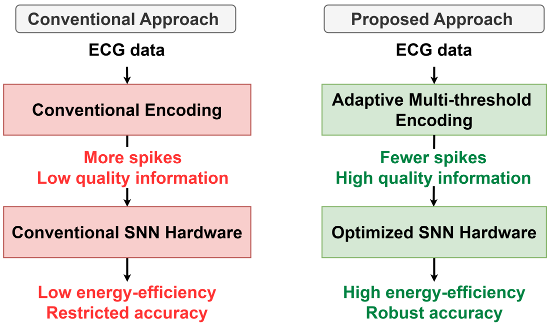

Fig. 4. Overview of conventional and proposed SNN-based ECG classification.

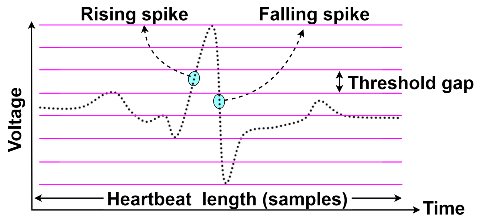

Fig. 5. A basic example of threshold-based encoding.

and reduces noise through regularization, further improving information quality. Moreover, coarse-grained approach uses less spikes to encode this new information. Thus, we represent the heartbeat using fewer spikes and high quality information. This provides high energy-efficiency and robust accuracy against data complexity/diversity. Moreover, we optimize the SNN hardware for this encoding to further enhance the energy-efficiency, resulting in optimal ECG classification. We now describe our encoding scheme in the next subsection.

B. Adaptive Multi-threshold Spike Encoding

As our encoding uses a threshold-based approach, we start with a basic example of thresholding in Fig. 5. The horizontal lines denote thresholds separated by a consistent gap. The signal generates a spike when it crosses any of the thresholds. They are categorized as rising or falling spikes, based on whether they occur during a rising or a falling signal transition respectively. This approach exhibits a trade-off between energy and information, influenced by two main factors. The first factor is the number of data samples used for spike conversion. More samples preserve the information, but cause more spikes and higher energy consumption. Fewer samples lead to less spikes and save energy, but lose information about subtle details in the data. The second factor is the threshold gap. A larger gap produces less spikes and conserves energy, but misses out on subtle data variations. A smaller gap captures these finer variations, but generates more spikes causing high energy consumption. Thus, balancing energy-efficiency with information quality is a challenge in threshold-based approach.

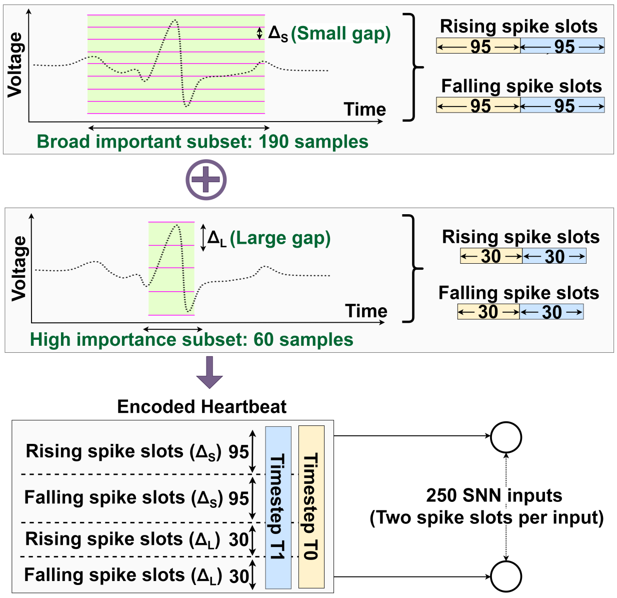
Voltage
Δ_S (Small gap)
Time
Broad important subset: 190 samples
Rising spike slots
95
95
Falling spike slots
95
95
Voltage
Δ_L (Large gap)
Time
High importance subset: 60 samples
Rising spike slots
30
30
Falling spike slots
30
30
Encoded Heartbeat
Rising spike slots (Δ_S) 95
Falling spike slots (Δ_S) 95
Rising spike slots (Δ_L) 30
Falling spike slots (Δ_L) 30
Timestep T1
Timestep T0
250 SNN inputs
(Two spike slots per input)

Fig. 6. Proposed adaptive multi-threshold spike encoding scheme. It encodes a broad region of interest with Δ_S and re-encodes its critical subset with Δ_L. The resulting spike slots are merged into two timesteps for SNN inputs.

Algorithm 1: Adaptive multi-threshold spike encoding.

```
input : Heartbeat data samples ($H$), small threshold
        gap ($\Delta_S$), large threshold gap ($\Delta_L$)
output: Spike encoded heartbeat $H_S$
$T_S$, $T_L$ ← threshold_sets($H$, $\Delta_S$, $\Delta_L$);
$S_{broad}$, $S_{crit}$ ← important_subsets($H$);
RS, FS, RL, FL ← $\emptyset$ $\emptyset$ $\emptyset$ $\emptyset$;
foreach sample in $S_{broad}$ do
    $S_{rise}$, $S_{fall}$ ← gen_spikes(sample, $T_S$);
    RS.insert($S_{rise}$), FS.insert($S_{fall}$);
    if sample $\in S_{crit}$ then
        $L_{rise}$, $L_{fall}$ ← gen_spikes(sample, $T_L$);
        RL.insert($L_{rise}$), FL.insert($L_{fall}$);
    end
end
$H_S$ ← split_and_concat(RS, FS, RL, FL);
return $H_S$;
```

Our proposed spike encoding shown in Fig.6 addresses this challenge, by adapting to the relative importance of heartbeat data samples. We begin by analyzing the heartbeats in our ECG dataset and make two key observations. First, 60 tail-end samples (total at both ends) have minimal impact on classification. Second, 60 samples in QRS region are most critical for classification. Leveraging these insights, we develop a two-pass adaptive encoding process. The first pass maximizes information coverage, while maintaining high information quality and lower spike count. It discards the 60 tail-end samples and encodes the remaining 190 samples using a small threshold gap (Δ_S). Thus, it covers a broad important subset and capture subtle data variations as high quality information. It produces fewer spikes despite a small Δ_S by encoding just a subset of the heartbeat. The second pass extracts additional critical information with fewer spikes. It revisits the 60 more important samples in QRS region and re-encodes them with a larger threshold gap (Δ_L). This captures broad variations in this critical region and also reduces noise impact through regularization effect. Thus, we obtain new critical information of high quality. The number of spikes still stays small due to high Δ_L. Thus, our encoding captures high quality information from the heartbeat using fewer spikes. We use two spike slots per sample to cover both rising and falling spike possibilities, leading to 500 spike slots per heartbeat (190×2 + 60×2). We split them into two timesteps of 250 spike slots each as shown in Fig. 6 for feeding to our SNN model in Section III-C. The encoding process is summarized in Algorithm 1 and the next subsection describes SNN model development using this encoded data.

C. Encoding-Aware SNN Model Development

We develop our SNN model using ANN-to-SNN conversion approach [22], [23] as it achieves (i) higher accuracy than unsupervised SNN learning [24], [25], and (ii) comparable accuracy to supervised SNN learning without the associated

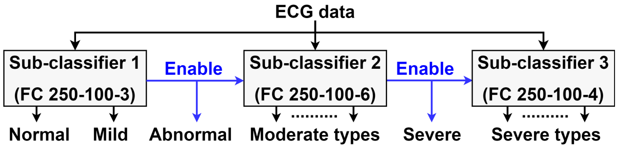

Fig. 7. Three-stage SNN-based ECG classifier model (ANN baseline in [9]).

training complexity [26], [27]. We select ANN model from [9] in Fig. 7 as our baseline for SNN conversion. It uses severity-based arrhythmia classes in Fig. 2, to help users with timely medical attention and doctors with faster diagnosis. Moreover, it saves energy by deactivating sub-classifiers 2 and 3 for prolonged periods as more severe heartbeats occur infrequently.

We use quantization clip-floor-shift (QCFS) approach [23] to obtain equivalent SNNs for the three baseline ANN sub-classifiers. It involves training an ANN with QCFS activation and replacing QCFS with integrate-and-fire (IF) neurons to create the SNN. However, this presents two challenges shown in Fig. 8 (red). First, it only supports using real-valued data directly as SNN input. This requires full-precision multipliers instead simple gating logic. Moreover, it increases bit sizes for membrane potentials and interim calculations. This increases energy consumption due to higher computation and memory demands. Second, it has no provisions to handle class imbalance in ECG data. Hence, loss function biases the classification towards majority classes (less severe heartbeats) with subpar performance on minority classes (more severe heartbeats).

We address these challenges through encoding-aware ANN-to-SNN conversion in Fig. 8 (green). It involves training QCFS-based ANN by treating input spikes as real values. Spikes cannot be considered as two-bit integers due to lack of positional weightage. So, we treat them as two binary inputs and add their outputs before backpropagation. This is consistent with IF neurons which accumulate inputs into the same membrane potential. We also use a weighted loss function, to penalize mistakes on minority classes more heavily. This ensures fair classification with good performance on minority (more severe) classes. After the training, we replace QCFS with IF neurons to obtain the SNN. This SNN seamlessly integrates our encoding scheme, with high accuracy. After obtaining SNNs for three sub-classifiers, we optimize them for hardware design next.

D. Optimizations for Hardware Design

1) SNN Dataflow: We present each heartbeat (T0 and T1) twice to the SNN for improved accuracy as shown in Fig. 9, with further repetitions offering no benefit. However, these repeated computations increase energy consumption. We propose an optimized SNN dataflow to address this. It starts with swapping T0 and T1 between the two input instances. This does not affect accuracy if the firing pattern of neurons remains consistent. We confirm this by evaluating post-swap accuracy. Next, we account for these post-swap pairs of identical inputs using layer post-processing. Our SNNs (input→FC1→FC2) have two layers: FC1 and FC2. We start with doubling the weighted sum of spikes in FC1, needing a simple left-shift in hardware. As FC1 neurons can spike up to twice over two timesteps, we compare the new membrane potentials with both the threshold and 2×threshold for the same effect. We extend this to FC2, which multiplies weights with spike count (0, 1, or 2) from FC1. The resulting FC2 output has three possible values: zero, original weight, or double the weight. This needs simple logic for inhibiting, passing, or left-shifting the weights. Thus, we achieve the same outputs as sending the heartbeat twice while sending it just once, with minimal hardware and no extra design complexity. We also use FC2 neuron membrane potentials directly for output classification, eliminating firing and resetting operations to save hardware resources.

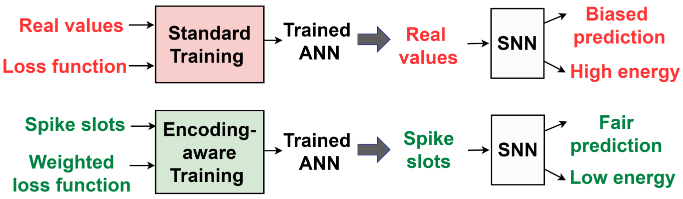

Fig. 8. Standard (red) vs proposed (green) ANN-to-SNN conversion methods.

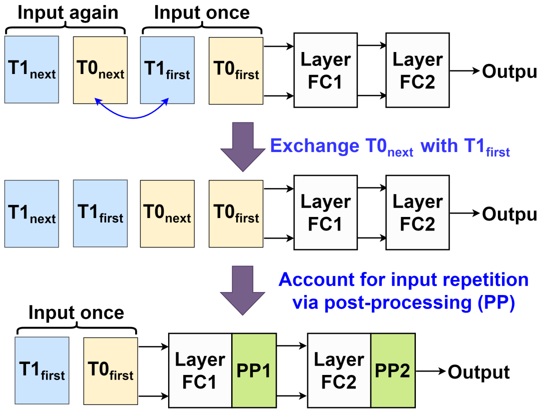

Fig. 9. SNN dataflow optimization to improve accuracy and preserve energy-efficiency. Both T0's hold identical data and both T1's contain same data.

2) Fixed-point Quantization: Fixed-point format provides energy-efficiency by allowing real number processing with integer arithmetic. It expresses a real number as binary value with an implicit radix point. As SNN inputs are spikes, we just quantize the weights to fixed-point format. However, information loss due to aggressive quantization can degrade the accuracy. To mitigate this, we perform design space exploration to minimize the weight bit sizes while maintaining high accuracy.

E. System Architecture Design

Our SNN-based ECG classification architecture in Fig. 10 consists of an encoding module, an SNN inference module, and a global control logic to manage on-chip as well as off-chip module interactions. ECG data is first sent to the encoding module. It implements our adaptive multi-threshold encoding using two sliding threshold windows (Δ_S and Δ_L). Each window internally employs two thresholds to detect both rising and falling spikes. Hence, we use four comparators as spike detectors using these four thresholds. The thresholding logic adjusts these thresholds based on spiking event history reported by spike collection unit. The spike collection unit also inserts each spike in the heartbeat register of the inference module, as per the encoding format in Fig. 6.

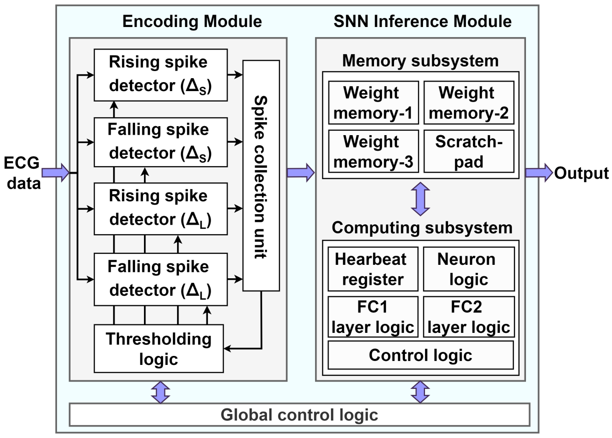

Fig. 10. Proposed ECG classification system architecture.

The SNN inference module includes a memory subsystem, a computing subsystem, and control logic. The control logic manages interactions between these subsystems and sequences relevant sub-classifier operation as per classification history. The memory subsystem has four SRAMs: three for each sub-classifier's weights and one as scratchpad for interim calculations. The computing unit is shared by all sub-classifiers. It features a heartbeat register, integrate-and-fire neuron logic, and layer state machines with optimized dataflow in Section III-D1. It first scans the encoded heartbeat to determine which data slots contain spikes. It then retrieves the relevant sub-classifier weights only for these slots from the memory subsystem. The computing subsystem then processes these spikes through the sub-classifier layers. Thus, we perform event-driven processing to execute only the necessary computations. After this, layer-2 neuron with the highest membrane potential determines the output class. Once the appropriate sequence of sub-classifiers is executed for the current heartbeat, the classifier module becomes ready to process the next one. We further implement this architecture into a system-on-chip in Section IV-C.

IV. Simulation Results

A. Setup

We use MIT-BIH ECG database [14] with severity-based class grouping in Fig 2. We divide it as 60% for training, 20% for hyperparameter tuning, and 20% for testing the final model. We integrate the baseline ANN [9], our encoding scheme, and our ANN-to-SNN conversion method into the framework from [23] to obtain the SNN models. We quantize the models to 8-bit fixed-point and implement via RTL (Register Transfer Level) to GDSII (Graphic Data System II) flow using TSMC 40 nm technology. We use Verilog for RTL design, Cadence Genus for RTL synthesis and Cadence Innovus for physical design. We verify the design through static timing analysis at 100 MHz clock frequency in Cadence Innovus. It covers global process variability by multi-corner analysis and local intra-process variability via on-chip variation (OCV) analysis.

B. SNN Model Development

We assess the performance of ANN (floating-point) and SNN (quantized) versions of the sub-classifiers from Fig.7, using accuracy and critical accuracy as shown in Fig 11. Here, accuracy measures the percentage of correctly identified heartbeats out of the total heartbeats. Critical accuracy refers to the accuracy of the most critical class within a sub-classifier, which triggers the next stage (blue lines in Fig.7). This critical classes in sub-classifier 1 and sub-classifier 2 are the abnormal class and the severe class respectively. As sub-classifier-3 is the final stage, it does not have a critical class and critical accuracy.

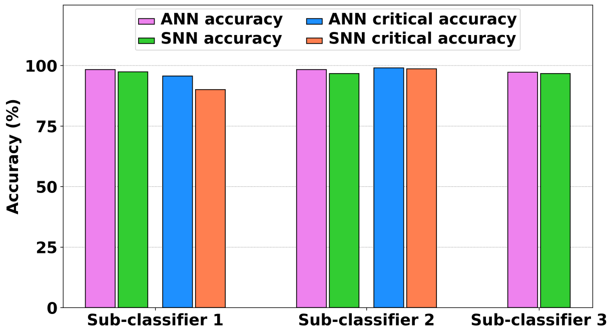

Fig. 11. Accuracy comparison between baseline ANNs and converted SNNs.

The quantized SNNs achieve comparable performance to the baseline ANNs, with an average accuracy loss of 1% across the three sub-classifiers. This is due to information loss during spike encoding, activation-to-IF neuron replacement, and weight quantization. The quantized SNN matches the critical accuracy of the baseline ANN for sub-classifier 2, but incurs a 5.6% loss in critical accuracy for sub-classifier 1. However, this loss is acceptable as abnormal heartbeats occur in sequences rather than isolated events. Thus, detecting 94 abnormal heartbeats out of every 100 (∼6% miss rate) is sufficient to provide timely feedback to the user.

C. Hardware Design

The system-on-chip implementation for our architecture is shown in Fig. 12, with core area of 0.36 mm^2. SRAMs account for 99.30% of this area, due to slower technology scaling of memory compared to logic. Classification module and encoding module occupy just 0.54% and 0.16% of the area respectively. This highlights the area-efficiency of our encoding scheme.

Fig. 13 shows that higher severity heartbeats require more inference energy, as they activate additional sub-classifiers. On average, our design consumes 59.1 nJ per inference. SRAM macros contribute to 69.02% of this energy reflecting the memory wall, where data transfers consume more energy than computations [28]. The classification module consumes 26.87% of the energy, while the encoding module just needs

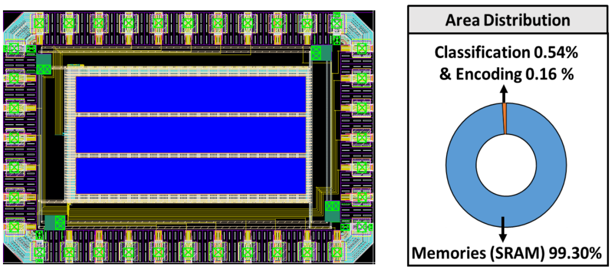

Fig. 12. System-on-chip in TSMC 40 nm technology, with its area distribution.

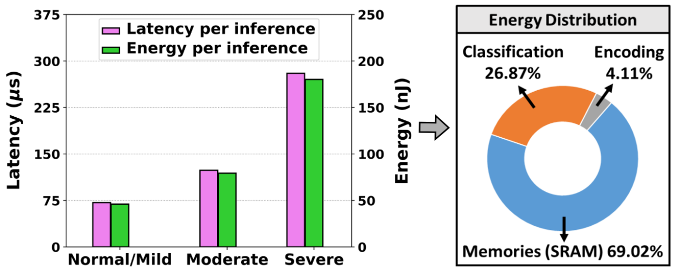

Fig. 13. Average inference latency and average energy consumption (with its distribution) for various classes of heartbeats.

4.11% of the energy. This demonstrates the energy-efficiency of our encoding scheme. Figure 13 also shows that latency per inference increases with severity of heartbeats, due to usage of more sub-classifiers. We achieve average latency of 90 μs, with a maximum latency of 680 μs for a severe heartbeat that uses all three sub-classifiers. With studies recording up to 600 heartbeats per minute [29], the maximum inference latency can be 100 ms. As our maximum latency is much smaller than this value, our design is suitable for real-time ECG classification.

D. Comparison with State-of-the-art

Performance comparison of our SNN-based ECG classifier design with state-of-the-art is shown in Table I. Section IV-B already covers accuracy comparison with [9] which served as baseline for developing our SNN models. We cannot directly compare our accuracy with the rest of AAMI-based works. This is because MIT-BIH sub-classes in the same AAMI group can map to different severity-based classes (Fig. 2), leading to different classification boundaries. Instead, the focus should be on the high classification accuracy achieved by our design.

We reduce energy per inference by nearly 2× compared to the most energy-efficient ANN [9] in Table I. This shows the energy-efficiency superiority of SNNs. Notably, we achieve this despite using SRAMs which consume more energy than memristors in [9]. SRAM's lower scalability than memristors however leads to higher area than [9]. Despite the benefits of memristor-based design, challenges with their fabrication and integration make our digital CMOS design the optimal choice.

We achieve lower energy consumption than works with poisson [10] and level-crossing [11], [12] encodings, due to fewer encoded spikes per heartbeat. The work in [13] achieves a spike count similar to ours, by encoding just 74 samples in QRS region. However, this degrades information quality by missing useful data outside this region. In contrast, we encode 190 samples covering this important missing data, with similar spike count. This makes our design robust to input variability and noise. It also facilitates generalization to more complex inputs (e.g. new patient demographics) that require finer details for accurate classification. Moreover, their design consumes higher energy despite the similar spike count due to several reasons. First, their encoding scheme is more complex and requires more energy. Second, they need higher bit precision for weights. This increases the energy required for data storage, data transfer, and computations. Third, they perform memory accesses even for layer inputs containing no spikes, leading to higher energy. Last, their on-chip learning involves sporadic weight updates using manual labels from doctors. Thus, it just consumes extra energy and is better managed off-chip. We address these issues through simpler encoding, lower weight precision and selective memory accesses for spike-containing inputs. This leads 5.1× less energy than [13] despite using an older technology node. Thus, our implementation outperforms state-of-the-art ECG classification in terms of energy-efficiency.

V. Conclusions

We presented an SNN-based ECG classification architecture based on a new adaptive multi-threshold spike encoding method. This encoding captured key information using fewer spikes, by adjusting its focus region and granularity. We developed an SNN model compatible with this encoding through our ANN-to-SNN conversion method. We created an optimized system architecture integrating the SNN model with encoding scheme, and implemented it using TSMC 40 nm technology. Our design provides up to 5.1× energy-efficiency compared to state-of-the-art SNN-based ECG classification. This highlights its potential for AI-driven healthcare targeting resource-constrained edge hardware.

TABLE I
Comparison of our proposed ECG classification with state-of-the-art. Here, 'behavioral' refers to works without a physical layout. We use '–' to denote unavailable area values and also to indicate that spike encoding is not applicable to ANNs.

Performance Metric	Wu [6] Access-2019	Lu [7] TCAS-I-2022	Diware [9] TBioCAS-2023	Liu [8] CICC-2024	Amirshahi [10] TBioCAS-2019	Chu [11] TBioCAS-2022	Mao [13] TBioCAS-2022	Tian [12] ESSCIRC-2023	This Work
Implementation	Digital ASIC	Digital ASIC	Mixed-signal ASIC (Behavioral)	Digital ASIC	Analog ASIC (Behavioral)	Digital ASIC	Digital ASIC	Analog ASIC	Digital ASIC
Technology	40 nm	40 nm	32 nm	55 nm	28 nm	40 nm	28 nm	40 nm	40 nm
Network Type	ANN	ANN	ANN	ANN	SNN	SNN	SNN	SNN	SNN
Spike Encoding	–	–	–	–	Poisson encoding	Level crossing encoding	Dual-purpose binary encoding	Level crossing encoding	Multi-threshold encoding
No. of classes	5	5	11	5	4	5	5	5	11
Accuracy (%)	96.06	98.99	98.29	98.70	97.90	98.22	98.60	95.31	97.42
Area (mm^2)	1.40	2.04	0.11	1.40	–	0.33	0.54	1.12	0.36
Energy per inference (μJ)	2.78	3.93	0.11	0.18	1.78	0.75	0.30	0.48	0.059

References

[1] V. Fuster, R. A. Harrington, J. Narula, and Z. J. Eapen, *Hurst's The Heart*. McGraw-Hill Education, 2017.

[2] L. D. Colantonio and P. Muntner, "It Is Time for Reducing Global Cardiovascular Mortality," *Circulation*, vol. 140, no. 9, pp. 726–728, 2019.

[3] D. N. Tran, B. Njuguna, T. Mercer, I. Manji, L. Fischer, M. Lieberman, and S. D. Pastakia, "Ensuring Patient-Centered Access to Cardiovascular Disease Medicines in Low-Income and Middle-Income Countries Through Health-System Strengthening," *Cardiology Clinics*, vol. 35, no. 1, pp. 125–134, 2017.

[4] American Heart Association. (2024) What is cardiovascular disease? [Online]. Available: https://www.heart.org/en/health-topics/consumer-healthcare/what-is-cardiovascular-disease

[5] Sarah Handzel. (2022) Wearable ECG Devices: Considerations for Cardiologists. [Online]. Available: https://www.gehealthcare.com/insights/article/wearable-ecg-devices-considerations-for-cardiologists

[6] J. Wu, F. Li, Z. Chen, Y. Pu, and M. Zhan, "A Neural Network-Based ECG Classification Processor With Exploitation of Heartbeat Similarity," *IEEE Access*, vol. 7, pp. 172 774–172 782, 2019.

[7] J. Lu, D. Liu, X. Cheng, L. Wei, A. Hu, and X. Zou, "An Efficient Unstructured Sparse Convolutional Neural Network Accelerator for Wearable ECG Classification Device," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 69, no. 11, pp. 4572–4582, 2022.

[8] J. Liu, Z. Xie, X. Wang, X. Liu, X. Qiao, J. Fan, H. Qin, C. Guo, J. Xiao, S. Lin, and J. Zhou, "BioWAP: A Reconfigurable Biomedical AI Processor with Adaptive Processing for Co-Optimized Accuracy and Energy Efficiency," in *IEEE Custom Integrated Circuits Conference*, 2024, pp. 1–8.

[9] S. Diware, S. Dash, A. Gebregiorgis, R. V. Joshi, C. Strydis, S. Hamdioui, and R. Bishnoi, "Severity-Based Hierarchical ECG Classification Using Neural Networks," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 17, no. 1, pp. 77–91, 2023.

[10] A. Amirshahi and M. Hashemi, "ECG Classification Algorithm Based on STDP and R-STDP Neural Networks for Real-Time Monitoring on Ultra Low-Power Personal Wearable Devices," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 13, no. 6, pp. 1483–1493, 2019.

[11] H. Chu, Y. Yan, L. Gan, H. Jia, L. Qian, Y. Huan, L. Zheng, and Z. Zou, "A Neuromorphic Processing System With Spike-Driven SNN Processor for Wearable ECG Classification," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 16, no. 4, pp. 511–523, 2022.

[12] F. Tian, X. Wang, J. Chen, J. Zheng, H. Wu, X. Liu, F. Tu, J. Yang, M. Sawan, C.-Y. Tsui, and K.-T. T. Cheng, "BIOS: A 40nm Bionic Sensor-defined 0.47pJ/SOP, 268.7TSOPs/W Configurable Spiking Neuron-in-Memory Processor for Wearable Healthcare," in *IEEE European Solid State Circuits Conference*, 2023, pp. 225–228.

[13] R. Mao, S. Li, Z. Zhang, Z. Xia, J. Xiao, Z. Zhu, J. Liu, W. Shan, L. Chang, and J. Zhou, "An Ultra-Energy-Efficient and High Accuracy ECG Classification Processor With SNN Inference Assisted by On-Chip ANN Learning," *IEEE Transactions on Biomedical Circuits and Systems*, vol. 16, no. 5, pp. 832–841, 2022.

[14] G. Moody and R. Mark, "The impact of the MIT-BIH Arrhythmia Database," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 3, pp. 45–50, 2001.

[15] Association for the Advancement of Medical Instrumentation, "AAMI Recommended Practice: Testing and Reporting Performance Results of Ventricular Arrhythmia Detection Algorithms," 1987.

[16] S. Diware, A. Singh, A. Gebregiorgis, R. V. Joshi, S. Hamdioui, and R. Bishnoi, "Accurate and energy-efficient bit-slicing for rram-based neural networks," *IEEE Transactions on Emerging Topics in Computational Intelligence*, vol. 7, no. 1, pp. 164–177, 2023.

[17] M. Bouvier, A. Valentian, T. Mesquida, F. Rummens, M. Reyboz, E. Vianello, and E. Beigne, "Spiking Neural Networks Hardware Implementations and Challenges: A Survey," *ACM Journal on Emerging Technologies in Computing Systems*, vol. 15, no. 2, 2019.

[18] A. Singh, M. A. Lebdeh, A. Gebregiorgis, R. Bishnoi, R. V. Joshi, and S. Hamdioui, "SRIF: Scalable and Reliable Integrate and Fire Circuit ADC for Memristor-Based CIM Architectures," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 68, no. 5, pp. 1917–1930, 2021.

[19] S. Diware, A. Gebregiorgis, R. V. Joshi, S. Hamdioui, and R. Bishnoi, "Mapping-aware biased training for accurate memristor-based neural networks," in *IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, 2023, pp. 1–5.

[20] L. Huijbregts, H.-H. Liu, P. Detterer, S. Hamdioui, A. Yousefzadeh, and R. Bishnoi, "Energy-efficient SNN Architecture using 3nm FinFET Multiport SRAM-based CIM with Online Learning," in *ACM/IEEE Design Automation Conference*, 2024, pp. 1–6.

[21] L. Deng, Y. Wu, X. Hu, L. Liang, Y. Ding, G. Li, G. Zhao, P. Li, and Y. Xie, "Rethinking the performance comparison between SNNS and ANNS," *Neural Networks*, vol. 121, pp. 294–307, 2020.

[22] B. Rueckauer, I.-A. Lungu, Y. Hu, M. Pfeiffer, and S.-C. Liu, "Conversion of Continuous-Valued Deep Networks to Efficient Event-Driven Networks for Image Classification," *Frontiers in Neuroscience*, vol. 11, 2017.

[23] T. Bu, W. Fang, J. Ding, P. Dai, Z. Yu, and T. Huang, "Optimal ANN-SNN Conversion for High-accuracy and Ultra-low-latency Spiking Neural Networks," in *International Conference on Learning Representations*, 2023.

[24] Y. Dong, D. Zhao, Y. Li, and Y. Zeng, "An unsupervised STDP-based spiking neural network inspired by biologically plausible learning rules and connections," *Neural Networks*, vol. 165, pp. 799–808, 2023.

[25] Y. Hao, X. Huang, M. Dong, and B. Xu, "A biologically plausible supervised learning method for spiking neural networks using the symmetric STDP rule," *Neural Networks*, vol. 121, pp. 387–395, 2020.

[26] C. Lee, S. S. Sarwar, P. Panda, G. Srinivasan, and K. Roy, "Enabling Spike-Based Backpropagation for Training Deep Neural Network Architectures," *Frontiers in Neuroscience*, vol. 14, 2020.

[27] W. Fang, Z. Yu, Y. Chen, T. Huang, T. Masquelier, and Y. Tian, "Deep Residual Learning in Spiking Neural Networks," in *Advances in Neural Information Processing Systems*, vol. 34, 2021, pp. 21 056–21 069.

[28] A. Sebastian *et al.*, "Memory devices and applications for in-memory computing," *Nature Nanotechnology*, 2020.

[29] L. Chhabra, N. Goel, L. Prajapat, D. H. Spodick, and S. Goyal, "Mouse heart rate in a human: diagnostic mystery of an extreme tachyarrhythmia," *Indian Pacing and Electrophysiology Journal*, vol. 12, no. 1, pp. 32–5, 2012.