



Hybrid Bi-LSTM-Spiking Neural Network for Energy-Efficient ECG Arrhythmia Classification

Nor Amalia Dayana Mohamad Noor^{1,2}, Wong Yan Chiew^{1*}, Zarina Mohd Noh¹ and Ranjit Singh Sarban Singh², Cheng Xiang³

¹Faculty of Electronics and Computer Technology and Engineering, Universiti Teknikal Malaysia Melaka, Durian Tunggal, Melaka, Malaysia.

²School of Engineering and Technology, Sunway University, Selangor, Malaysia.

³School of Computer Science and Intelligence Education, Lingnan Normal University, China.

Article Info

Article history:

Received July 6th, 2026

Revised Aug 12th, 2026

Accepted Aug 16th, 2026

Published Sep 30th, 2026

Index Terms:

Electrocardiogram classification

Long Short-Term Memory

Spiking neural networks

Surrogate gradients

Abstract

Wearable ECG monitoring requires a classifier that can operate within a limited power budget. This study therefore combined a bidirectional long short-term memory (Bi-LSTM) layer, which processed the waveform in both temporal directions, with two leaky integrate-and-fire (LIF) spiking layers. The recurrent output was converted into sparse events, while a fast-sigmoid surrogate gradient enabled training through the nondifferentiable spike function using backpropagation through time. Twenty hyperparameter configurations were evaluated using 60-40, 70-30, and 80-20 training-testing partitions. The best configuration, Configuration 3 using the 80-20 partition, comprised one 64-unit Bi-LSTM layer, a membrane decay parameter of 0.85, and a learning rate of 0.00020. It achieved a training accuracy of 97.64% and a test accuracy 96.88%, exceeding the reporting accuracy of the standalone SNN by 5.21 percentage points. Thus, the model maintained high ECG classification performance while shifting its later-stage processing to event-driven spike activity.

This is an open access article under the [CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/) license.



*Corresponding Author: ycwong@utem.edu.my

I. INTRODUCTION

Cardiovascular diseases (CVDs) remains the leading cause of death worldwide. For patients requiring prolonged observation, ambulatory electrocardiogram (ECG) recording can facilitate the early detection of arrhythmias and support timely intervention to reduce the risk of sudden cardiac arrest [1]-[5]. The main challenge lies in reviewing the large volume of resulting ECG traces: expert interpretation is time-consuming, difficult to scale, and increasingly prone to error as the workload grows [6], [7]. Earlier automated systems relied on handcrafted morphological descriptors [8]-[10]; however, their performance deteriorated when recordings were noisy or ECG characteristics varied among patients [11].

Deep learning (DL) has reduced reliance on hand-engineered descriptors by learning features directly from raw ECG samples [12], [13]. Convolutional neural networks (CNNs) are well suited to capturing local waveform patterns [14], [15]. Recurrent neural networks (RNNs) and long short-term memory (LSTM) networks, in contrast, can capture temporal dependencies across successive heartbeats [16]-[19]. Recent ECG classifiers combine these capabilities in various forms, including squeeze-and-excitation CNN-LSTM networks [20], frequency-channel attention mechanisms [21], bidirectional LSTM (Bi-LSTM) networks [22], and shallow CNNs [23].

High classification accuracy alone provides limited insight into whether a model can operate effectively on a wearable sensor. Dense neural networks require considerable parameter storage and computational resources, both of which place a burden on battery-powered Internet of Medical Things (IoMT) devices [24], [25]. Sending each recording to the cloud reduces local processing requirements, but introduces communication delays, dependence on network connection, and potential privacy and security risks during real-time operation [26]. A spiking neural network (SNN) provides an alternative because leaky integrated-and-fire (LIF) neurons communicate through sparse binary events; when a neuron does not generate a spike, little computation is required. However, a conventional SNN may require many simulation time steps and is generally less effective than a recurrent network in capturing long-range sequential information.

Our design places a compact Bi-LSTM network before two LIF-based spiking layers. The recurrent stage captures the temporal context of the ECG signal, while an interface converts its hidden representation into spike trains. The proposed hybrid classifier was then evaluated using the MIT-BIH Arrhythmia Database and compared with previously published models.

II. LITERATURE REVIEW

A. Clinical Significance of ECG and Wearable Telemetry

An ECG measures the heart's electrical activity non-invasively and is therefore widely used in clinical and remote-monitoring settings [1]. However, ECG interpretation is not straightforward because waveform morphology changes over time, while noise and baseline drift can obscure diagnostically useful patterns [2]. CNN classifiers can improve arrhythmia recognition; however, their memory requirements and computational costs make continuous operation on small, battery-powered devices challenging.

SNNs represent information through the timing of relatively sparse spikes. Because computation is triggered when these events occur, neuromorphic hardware can remain inactive between events. This property has motivated research on continuous ECG classification [3] and cardiac monitoring performed directly on wearable devices [4].

The global burden of cardiovascular disease [5] makes reliable long-term monitoring clinically important. Previous wearable studies have shown that continuous ECG acquisition can facilitate the detection of cardiac abnormalities [6]. In the proposed architecture, the LSTM network learns the temporal characteristics of the ECG waveform, after which the SNN processes the learned representation as sparse events [7]. This study investigates whether such a division of work can preserve classification performance while employing a computational pattern that is more compatible with wearable hardware.

B. Electrocardiogram Waveforms and Wearable Diagnostics

Electrocardiography records the electrical activity associated with depolarization and repolarization at the skin surface. A normal trace contains the P wave, QRS complex, and T wave [8]. Their shapes, amplitudes, and timing intervals provide the information required to distinguish arrhythmias and other cardiac conditions [9]. A representative normal heartbeat is shown in Figure 1.

Recording ECGs outside clinic settings enable longer observation periods and may allow abnormal heart rhythms to be detected earlier [10]. If the classifier is embedded directly in the wearable device, the time between signal acquisition and classification can be reduced [11]. Such a classifier must retain the temporal order of waveform features rather than consider only their local shape. Placing an LSTM network before an SNN provides this temporal context while reserving the later proceeding stages for sparse computation.

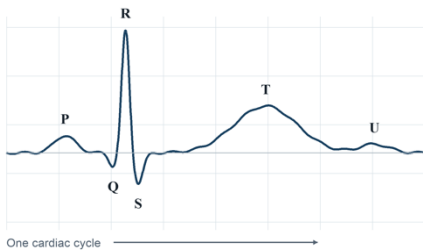


Figure 1. Normal beat: Standard ECG Signal Pattern [9]

C. Related Work

Early ECG classifiers used decision rules and compact statistical features. However, these approaches often struggled when waveform morphology varied across patients

or recording sessions. Machine-learning and deep-learning approaches were subsequently introduced to learn a broader range of patterns directly from recordings.

CNNs can recognize arrhythmias effectively because they learn local waveform structure. However, they do not explicitly retain long-term sequential information and may be computationally costly for continuous operation. LSTM networks provide this temporal memory, although at the cost of recurrent computation and additional storage. CNNs have also been applied to lung-disease classification using chest images [12]; however, their success in that setting does not address the resource constraints associated with always-on wearable ECG devices.

An SNN represents neuronal activity using discrete spikes and can perform event-triggered operations in real time. Earlier ECG research [13] has demonstrated the potential of SNNs; however, directly converting ECG signals into spikes may discard important temporal information. In the proposed architecture, the LSTM network learns that structure before its representation is passed to the spiking layers, thereby combining recurrent feature extraction with sparse downstream processing.

D. Heart Anatomy

The sinoatrial (SA) node initiates the electrical impulse that produces cardiac activity recorded by an ECG, after which the impulse travels through the heart's conduction system. Disturbances in the conduction pathway can alter the ECG waveform and may indicate an abnormal heart rhythm [10]. Therefore, ECG classification must account for the ordered P-Q-R-S-T pattern rather than evaluate each signal sample independently.

E. Spiking Neural Network (SNN) Architecture

SNNs operate using discrete events, and the timing of these events can carry information. One learning mechanism used in SNNs is spike-timing-dependent plasticity (STDP), in which the time interval between presynaptic and postsynaptic spikes determines how the corresponding synaptic weight is modified [14].

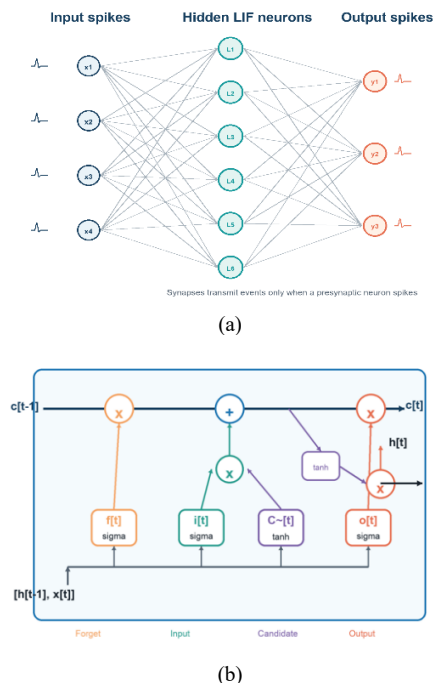


Figure 2. The architecture of: (a) SNN [9] and (b) LSTM cell [15]

F. Long Short-Term Memory (LSTM) Architecture

An LSTM augments a recurrent unit with input, forget, and output gates. Together, the gates determine which information enters the cell state, how much the existing state is retained, and what information is provided as the hidden state. This controlled memory allows the network to capture ECG context over a longer interval. In the proposed model, temporal feature extraction is performed by the LSTM network before spike generation. Figure 2(b) illustrates the LSTM cell, while [15] describes its application to the

detection of subtle ECG abnormalities.

G. The Evolution of Deep Learning Approaches in Electrocardiogram (ECG) Classification

Table 1 summarizes earlier CNN-, LSTM-, and hybrid-based studies using ECG datasets. The reported accuracy and F1-scores provide a broad overview of the progress achieved. However, these values should not be interpreted as a strict ranking because the studies used different datasets, class definitions, preprocessing methods, and evaluation protocols.

Table 1
Past Research comparison table on the ECG dataset

Reference	Model	Dataset	Classes	Accuracy (%)	Precision	Recall	F1-Score
[16]	EGCNet (GCN)	PhysioNet	24	97.98	0.756	—	0.729
[17]	ResNet CNN	MIT-BIH	5	99.2	0.992	0.992	—
[18]	Dilated CNN + BiGRU-BiLSTM + Attention	PhysioNet	Multi	99.6	0.9766	0.996	0.9821
[19]	CNN-LSTM	MIT-BIH	9	98.24	—	—	—
[20]	CNN-LSTM-SE	MIT-BIH	5	98.5	>0.97	>0.98	>0.98
[21]	CNN + FCA	MIT-BIH	5	99.6	—	—	—
[22]	Bi-LSTM	MIT-BIH	5	98.25	—	—	—
[23]	CNN	MIT-BIH	2	97.4	0.957	0.961	—
[24]	2D CNN + RR	MIT-BIH	2	98.74	0.7075	0.6747	0.6876
[25]	CNN-BiLSTM	MIT-BIH	3	—	—	—	0.82
[26]	SNN	PhysioNet	2	91.6667	—	—	—

III. METHODS

The evaluation followed a consistent sequence. First, the ECG records were prepared, hybrid network was specified, and the candidate hyperparameters were evaluated. The selected configuration was subsequently retrained and assessed using data held out from model fitting.

A. Dataset and Preprocessing

The study sample comprised 200 digitized ECG records selected from healthy volunteers and patients with cardiac conditions. Class 0 contained 80 normal records, whereas Class 1 contained 120 abnormal records.

All records were obtained from version 1.0.0 of the PTB Diagnostic ECG Database available through PhysioNet [27]. The non-commercial PTB prototype recorder has 16 input channels: 14 ECG channels, one respiration channel, and one line-voltage channel. The signals in the database were synchronously sampled at 1000 Hz with 16-bit resolution (0.5 $\mu\text{V}/\text{LSB}$), a bandwidth of 0-1 kHz, and a digitalization range of ± 16.384 mV.

Each ECG signal was rescaled to the range [0, 1] using min-max normalization to reduce amplitude differences among records. The continuous signals were then segmented into equal-length sequences while preserving the heartbeat morphology. The evaluation was repeated using training-testing partitions of 60-40%, 70-30%, and 80-20% to examine the effect of training-set size.

Figure 3 illustrates the complete processing pipeline from ECG preprocessing and recurrent feature extraction to classification by the final spiking layers.

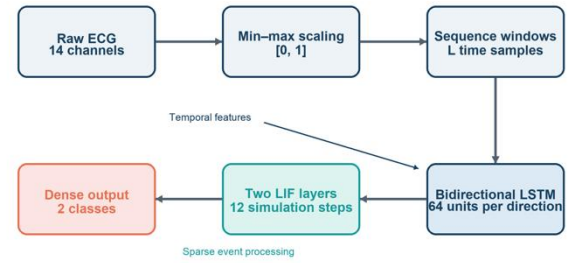


Figure 3. High-level workflow flowchart illustrating the raw ECG preprocessing and hybrid network sequence tracking stages

B. Hybrid LSTM-SNN Model Architecture

The proposed network begins with a Bi-LSTM layer and ends with two LIF-based spiking layers. The Bi-LSTM processes the ECG sequence in both temporal directions. Its context-aware output is then passed to the event-based SNN.

Figure 4 presents the detailed layer sequence, including the forward and backward recurrent paths, their concatenation, and the two LIF layers.

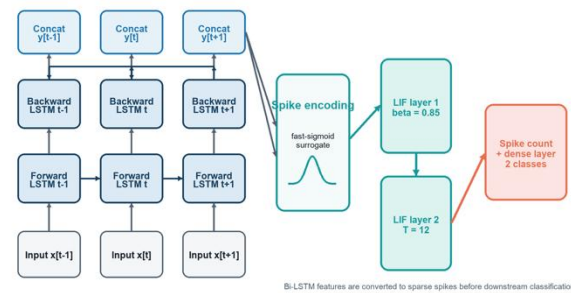


Figure 4. Detailed low-level neural architecture schematic showing the forward and backward LSTM gating sequences, temporal concatenation layers, surrogate fast-sigmoid gradient approximation, and two-layer spiking LIF neurons

1. Input encoding and rasterization: A preprocessed ECG sequence is presented over (T) discrete time steps. During SNN processing, the spike raster records the resulting neuronal activity, allowing variations in spike density to be examined over time.
2. Temporal feature extraction using the Bi-LSTM: The forward and backward recurrent paths process the ECG sequence in opposite temporal directions. The hidden state of each path contains contextual information from its respective direction. Within a cell, the input, forget, and output gates update the cell state C_t and hidden state h_t . Equations (1)-(6) specify, in sequence, the forget gate, input gate, candidate state, updated state, output gate, and final hidden state:

$$ft = \sigma(Wf \cdot [ht - 1, xt] + bf) \quad (1)$$

$$it = \sigma(Wi \cdot [ht - 1, xt] + bi) \quad (2)$$

$$\tilde{C}t = \tanh(WC \cdot [ht - 1, xt] + bC) \quad (3)$$

$$Ct = ft \odot Ct - 1 + it \odot \tilde{C}t \quad (4)$$

$$ot = \sigma(Wo \cdot [ht - 1, xt] + bo) \quad (5)$$

$$ht = ot \odot \tanh(Ct) \quad (6)$$

Here, xt is the input vector at time step t ; W and b are the weight matrices and bias vectors; σ and $\tanh$ are the activation functions; and $\odot$ denotes element-wise multiplication.

When multiple LSTM layer are used, dropout is applied between successive layers to reduce feature co-adaptation.

3. Spiking neural network processing: The forward and backward hidden states produced by the Bi-LSTM are concatenated and supplied to two spiking layers implemented with `snn.Leaky` neurons in `snnTorch`. The membrane-decay parameter β determines how rapidly the stored membrane potential decays between time steps. The discrete membrane potential update is given in Equation (7). ($Yt = [\vec{h}_t || h_t^-]$)

$$U[t + 1] = \beta U[t] + I[t + 1] - S[t]\theta \quad (7)$$

In (7), $I[t+1]$ is the incoming synaptic current, θ is the firing threshold, and $S[t]$ is the binary spike output.
4. Surrogate gradient: A hard spike threshold represents by a step function whose derivative is zero almost everywhere and is therefore unsuitable for conventional backpropagation. During the backward pass, a fast-sigmoid surrogate ingredient is used to approximate this derivative. Equation (8) defines the approximation used during training.

$$\tilde{f}'(U) = \frac{\partial s}{\partial U} \approx \frac{1}{(1 + \alpha|U - \theta|)^2} \quad (8)$$

In (8), U denotes the membrane potential relative to the firing threshold, while the remaining parameters controls the slope of the surrogate-gradient function. This approximation allows gradients to propagate through the spiking layers during backpropagation through time (BPTT).

5. Classification output: The spikes emitted by the second LIF layer are counted over all T simulation steps. Following dropout, a fully connected linear layer maps the spike counts to the corresponding class scores.

The model was implemented in PyTorch using `snnTorch` and `scikit-learn`.

C. Hyperparameter Optimization Strategy

The search was conducted for each train-test partition by varying the recurrent width and depth, dropout rate, membrane-decay parameter, learning rate, simulation duration, and batch size. Within each training partition a validation subset was used for hyperparameter selection.

- LSTM hidden size (H): 32, 64, or 128 units.
- Bidirectional LSTM depth (L): one, two, or three layers.
- LSTM and SNN dropout rate: 0.5, 0.6, or 0.7.
- Membrane decay constant (β): 0.75, 0.80, or 0.85.
- Learning rate (η): 0.00002, 0.00005, 0.00020, or 0.00030 with AdamW.
- Simulation length (T): 8, 12, or 15 steps.
- Batch size (B): 32 or 64 samples.

A total of 20 candidate configurations were trained. Model selection considered validation accuracy together with the training-validation performance gap. When two configurations produced similar results, the smaller network was preferred.

D. Model Training and Evaluation Protocol

Each configuration was trained end to end using BPTT, the AdamW optimizer, and cross-entropy loss. As shown in Equation 9, the recurrent gradients were clipped to a maximum norm of 1.0.

$$g \leftarrow \min(1.0, \frac{1.0}{\|g\|_2})g \quad (9)$$

Training and validation accuracy was recorded for each run. After hyperparameter selection, the chosen network was retrained using the entire training partition and evaluated once on the held-out test partition. For the selected network, class-wise precision, recall, and F1 score were also calculated, and the confusion matrix was examined. The trained model was then exported for deployment.

IV. RESULTS AND DISCUSSION

Across the three data partitions, 20 Bi-LSTM-SNN configurations were evaluated. Each configuration used a fast-sigmoid surrogate gradients, dropout, gradient clipping, and stratified five-fold cross-validation within the training data. The resulting classification scores and the spike patterns

produced by the selected network are reported.

A. Grid Search Exploration

The training and test accuracies obtained for the 20 configurations across the three training-testing partitions are presented in Table 2.

Configuration 3 achieved the highest test accuracy for the 80-20 partition. It used 64 hidden units, one LSTM layer, 0.5 dropout in both the recurrent and spiking stages, $\beta = 0.85$, a 0.00020 learning rate, 12 simulation steps, and a batch size of 64. It achieved a training accuracy of 97.64%, and a test accuracy of 96.88%.

Table 2
Performance Evaluation of Hyperparameter Tuning Configurations

No.	Dataset Split	Combo ID	Hidden Size	LSTM Layers	LSTM Dropout	SNN Dropout	Beta (β)	Learning Rate	Num Steps (T)	Batch Size	Train Accuracy (%)	Test Accuracy (%)
1	80-20	3	64	1	0.5	0.5	0.85	0.00020	12	64	97.6378	96.8750
2	60-40	1	32	1	0.5	0.5	0.80	0.00030	8	32	92.6316	95.3125
3	70-30	3	64	1	0.5	0.5	0.85	0.00020	12	64	98.1982	93.7500
4	80-20	5	128	2	0.5	0.5	0.85	0.00005	15	64	90.5512	93.7500
5	60-40	3	64	1	0.5	0.5	0.85	0.00020	12	64	98.9474	92.1875
6	70-30	1	32	1	0.5	0.5	0.80	0.00030	8	32	95.4955	91.6667
7	70-30	7	64	2	0.5	0.5	0.80	0.00030	8	64	79.2793	91.6667
8	70-30	4	64	2	0.7	0.7	0.85	0.00020	12	64	83.7838	91.6660
9	80-20	1	32	1	0.5	0.5	0.80	0.00030	8	32	96.8504	90.6250
10	80-20	2	32	2	0.6	0.6	0.75	0.00030	8	32	81.8898	90.6250
11	80-20	4	64	2	0.7	0.7	0.85	0.00020	12	64	85.8268	90.6250
12	60-40	4	64	2	0.7	0.7	0.85	0.00020	12	64	83.1579	87.5000
13	70-30	5	128	2	0.5	0.5	0.85	0.00005	15	64	90.9910	85.4167
14	60-40	2	32	2	0.6	0.6	0.75	0.00030	8	32	74.7368	84.3750
15	70-30	2	32	2	0.6	0.6	0.75	0.00030	8	32	71.1712	83.3333
16	60-40	5	128	2	0.5	0.5	0.85	0.00005	15	64	88.4211	78.1250
17	60-40	7	64	2	0.5	0.5	0.80	0.00030	8	64	73.6842	71.8750
18	80-20	7	64	2	0.5	0.5	0.80	0.00030	8	64	76.3780	65.6250
19	60-40	6	128	3	0.6	0.6	0.80	0.00005	15	64	60.0000	60.9375
20	70-30	8	128	2	0.7	0.7	0.80	0.00002	15	64	65.7658	60.4167

Increasing the recurrent width and depth did not improve performance on this dataset. Configuration 6 for the 60-40 partition and Configuration 8 for the 70-30 partition used 128 hidden units with three and two LSTM layers, respectively. However, their test accuracies decreased to 60.94% and 60.42%. Given the small sample size, this pattern is consistent with excessive recurrent complexity optimization difficulties.

Figure 5 compares the best-performing configuration from each partition side by side. The configuration evaluated using the 80-20 partition achieved the highest test accuracy. Its training-test performance gap also differed from those obtained using smaller training partitions.

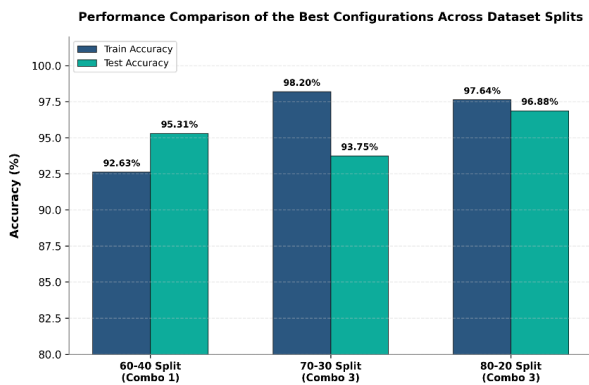


Figure 5: Performance comparison of the best configurations across dataset splits

B. Classification Results and Performance Metrics

The class-level performance measures for Configuration 3 using the 80-20 test partition are presented in Table 3.

Table 3
Classification Report of the Optimal Hybrid LSTM-SNN Model

Class	Precision	Recall	F1-Score	Support
0 (Normal)	0.93	1.00	0.96	28
1 (Abnormal)	1.00	0.95	0.97	36
Macro Average	0.97	0.97	0.97	64
Weighted Average	0.97	0.97	0.97	64
Overall Accuracy	96.88%			64

Of the 64 held-out records, the network correctly classified 96.88%. For the 28 normal records, recall was 1.00 and precision was 0.93, indicating that no normal record was misclassified in this test set. For the 36 abnormal records, precision reached 1.00 and recall 0.95; therefore, every record predicted as abnormal was an abnormal record.

Figure 6 shows the training and test accuracy trajectories over 25 epochs. Both series varied during the early stages of training because of mini-batch reshuffling before approaching their final levels.

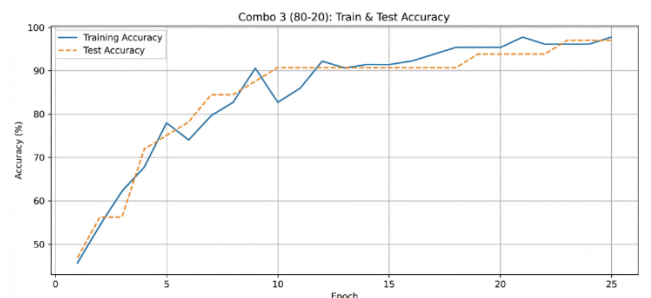


Figure 6: Empirical training and testing accuracy trajectories of the optimal Combo 3 configuration over 25 epochs under an 80-20 dataset split

The fast-sigmoid surrogate provided a non-zero training signal near the spike threshold; without this approximation, the derivative of the step function would prevent gradients from propagating through the spiking layers. A gradient-norm limit of 1.0 also constrained unusually large recurrent updates during BPTT.

The shallow Bi-LSTM produced continuous temporal features, which the LIF layers converted into spike activity. Periods without spikes require fewer downstream operations. Nevertheless, the network retained a test accuracy of 96.88%. However, actual energy savings cannot be established until the model is evaluated on hardware.

Figure 7 shows the spike counts produced by the second spiking layer over 12 simulation steps for eight test representation test samples.

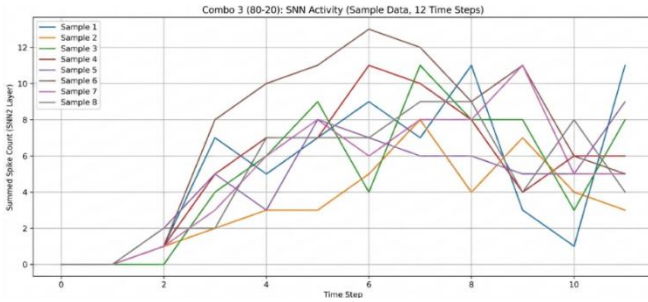


Figure 7. Summed spike counts of the second Spiking Neural Network layer (SNN2) across twelve simulation time steps for eight representative samples under the optimal Combo 3 configuration

No spikes appeared at the beginning of the sequence. The spike counts then increased before following a less regular recovery pattern.

1. Initial latency ($t = 0$ to $t = 1$): All eight samples remained silent during the first two simulation steps. The synaptic input generated from the Bi-LSTM representation first had to accumulate and raise the downstream LIF membrane potential $U[t]$ above the firing threshold ($= 1.0$).
2. Spike accumulation ($t = 2$ to $t = 6$): Spiking began at $t = 2$ as the recurrent representation supplied current to the LIF layers. Activity was highest around $t = 5$ to $t = 7$. At $t = 6$, Sample 6 produced 13 spikes whereas Sample 4 produced 11 spikes. One plausible explanation is that this window overlaps with waveform regions containing pronounced local structures, including the QRS complex.
3. Recovery ($t = 7$ to $t = 11$): After the peak, the samples no longer followed a uniform pattern. Sample 6 decreased from 13 spikes at $t = 6$ to 5 at $t = 11$; Sample 1 fell to one spike at $t = 10$ but rebounded to 11 at the next step. This behavior resulted from the $\beta = 0.85$ membrane decay and subtractive reset: once a spike occurs, the current must accumulate again before the membrane crosses threshold.

C. Benchmarking Table

Table 4 compares the proposed results with those of published ECG classifiers using accuracy, precision, recall, and F1 score.

The hybrid network combines recurrent feature learning with event-driven spiking. Its accuracy of 96.88% is lower than the highest CNN accuracies reported in Table 4, although its

later-stage computation are represented by sparse spike events.

Table 4
Comparative Benchmarking against Existing Literature

Reference	Accuracy (%)	Precision	Recall	F1-Score
[16]	97.98	0.756	–	0.729
[17]	99.2	0.992	0.992	–
[18]	99.6	0.9766	0.996	0.9821
[19]	98.24	–	–	–
[20]	98.5	>0.97	>0.98	>0.98
[21]	99.6	–	–	–
[22]	98.25	–	–	–
[23]	97.4	0.957	0.961	–
[24]	98.74	0.7075	0.6747	0.6876
[25]	–	–	–	0.82
[26]	91.6667	–	–	–
Proposed Model Class 0	96.88	0.93	1.00	0.96
Proposed Model Class 1	96.88	1.00	0.95	0.97

The CNN-based studies in [18] and [21] reported accuracies of up to 99.60%. These models use continuous activations and may require more computation than a spiking implementation; however, Table 4 provides no common hardware or energy measurements for a direct comparison. The standalone SNN in [26] achieved 91.67% and lacked a recurrent stage for long-range context.

With an accuracy of 96.88%, the hybrid model performed between the standalone SNN, which achieved 91.67%, and the highest-performing CNNs. Recall for Class 0 and precision for Class 1 both reached 1.00. Whether the architecture also reduced energy consumption must be evaluated on its intended edge or neuromorphic platform.

D. Discussion on Hybrid vs. Standalone Spiking Architectures

For a direct baseline comparison, the standalone SNN reported in [26] was used. Table 5 shows the architectural and performance differences relevant to this comparison.

The baseline in [26] converted segmented ECG amplitudes directly into spike trains for LIF processing. This approach was computationally lightweight, but its memory of previous inputs was governed by membrane decay. This limitation is important because an arrhythmia may be represented by relationships across the complete P-Q-R-S-T sequence. As the LIF state decayed, earlier parts of the heartbeat contributed less to subsequent processing; the reported baseline accuracy was 91.67%.

The hybrid architecture added a Bi-LSTM before the SNN. Its gates preserved heartbeat context in both temporal directions, and the concatenated recurrent representation was then supplied to the LIF layers.

To make the computational trade-off explicit, an arithmetic cost using a common scale was estimated. For an input width of $I = 14$ and bidirectional hidden size of $H = 64$, the recurrent front end performed $2 \times 4H(I + H) = 39,936$ multiply-accumulate operations (MACs) per ECG time step, equivalent to $0.039936L$ MMACs for a sequence of L samples. Using the 45-nm FP32 estimates reported in [28] (3.7 pJ per multiplication and 0.9 pJ per addition), this recurrent computation required approximately 0.184 μ J per time step, excluding memory access and data movement. Applying the same arithmetic-only calculation to the CNN-

LSTM-SE dimensions in [20] produced approximately 2.19 MMACs, or 10.1 μ J, for one 360-sample heartbeat. The two operation counts were approximately equal at $L = 55$; beyond this sequence length, the recurrent front end performed more arithmetic. Its downstream SNN avoided dense MACs when the neurons were silent, but the resulting benefit depended on spike activity and fan-out. These values are estimates rather than measurements because the studies used different input lengths and did not report energy consumption on the same hardware.

Table 5
Comparative Analysis of Standalone SNN versus Hybrid LSTM-SNN

Feature / Dimension	Standalone SNN (Baseline Model [26])	Proposed Hybrid LSTM-SNN Model
Architectural Pipeline	Raw Segmented Input → Encoding → LIF Spiking Neurons	Raw Input → Bidirectional LSTM → LIF Spiking Layers → Dense Output
Temporal Context Memory	Short-term temporal decay limited by LIF membrane potential decay (β)	Long-range bidirectional context captured via recurrent LSTM gates
Feature Extraction	Static spatial/rate-based spike encoding	Deep, highly expressive dynamic sequence features
Optimal Split Accuracy	91.67%	96.88% (Combo 3, 80-20 Split)
Class-Specific Precision	Standardized, moderate class boundary separation	Class 0 (Normal): 0.93 Class 1 (Abnormal): 1.00
Class-Specific Recall	Variable; higher false-negative rates on complex signals	Class 0 (Normal): 1.00 Class 1 (Abnormal): 0.95
Primary Advantage	Ultra-low computational overhead	High accuracy with sparse, energy-efficient operations
Deployment Target	Basic microcontrollers with restricted memory	Intelligent, ultra-low-power wearable ECG patches

The memory requirements were estimated more directly. With $2 \times 4H(I + H + 1) = 40,448$ parameters, the Bi-LSTM weights required 158.0 KiB in FP32, 79.0 KiB in FP16, or 39.5 KiB in int8. The two directions jointly maintained 256 hidden- and cell-state values, corresponding to 1.0 KiB in FP32 or 256 bytes in int8. Retaining every 128-dimensional bidirectional output would require an additional FP32 buffer of 512L bytes, which may become the main constraint on a small microcontroller. Streaming or chunked inference could reduce the need to store the entire sequence. These totals exclude the SNN and classifier weights; therefore, a target-specific assessment should be conducted after quantization and operator fusion.

1. Accuracy: Moving from the standalone SNN to the hybrid model increased test accuracy from 91.67% to 96.88%, representing an absolute gain of 5.21 percentage points. On the 64-record test set, recall for Class 0 and precision for Class 1 both reached 1.00.
2. Deployment implication: The estimates demonstrated the likely arithmetic and memory costs, but they did not establish an end-to-end reduction in energy consumption. Latency, memory traffic, and power consumption must be measured using the intended microcontroller or neuromorphic processor.

V. CONCLUSION

This work examined a Bi-LSTM-SNN for binary ECG classification. The recurrent front end captured sequential context, while the two LIF layers operated on the resulting spike-based representation.

20 configurations across three train-testing partitions were evaluated. Configuration 3 for the 80-20 partition used a single 64-unit bidirectional LSTM, $\beta = 0.85$, a 0.00020 learning rate, 12 simulation steps, and a batch size of 64. The training accuracy reached 97.64%, while test accuracy reached 96.88%. Within the test set, recall was 1.00 for Class 0 and precision was 1.00 for Class 1.

The gated recurrent path preserved information across the ECG sequence, while the fast-sigmoid surrogate enabled gradients to propagate through the discrete spiking layers during BPTT. Compared with the standalone SNN reported in [26], the test accuracy increased by 5.21 percentage points.

These results warrant hardware evaluation, but not yet support a claim of lower energy consumption. Future work will measure latency, memory consumption, and energy use on the intended edge or neuromorphic platform and repeat the analysis with a larger, patient-independent cohort.

ACKNOWLEDGMENT AND FUNDING

The authors acknowledge support from the Ministry of Higher Education, Malaysia, through grant FRGS/1/2024/ICT02/UTEM/02/3 (FRGS/1/2024/FTKEK/F00566), and from Universiti Teknikal Malaysia Melaka (UTeM).

CONFLICT OF INTEREST

The authors declare no conflict of interest related to this paper.

AUTHOR CONTRIBUTION

The authors contributed as follows:

Study conception and design: Wong Yan Chiew and Nor Amalia Dayana Mohamad Noor; data collection: Nor Amalia Dayana Mohamad Noor; analysis and interpretation: Nor Amalia Dayana Mohamad Noor and Wong Yan Chiew; manuscript drafting: Nor Amalia Dayana Mohamad Noor, Zarina Mohd Noh, Ranjit Singh Sarban Singh, and Chen Xiao. All authors reviewed the findings and approved the final manuscript.

REFERENCES

- [1] H. Wu et al., "A fully-automated paper ECG digitisation algorithm using deep learning," *Sci. Rep.*, vol. 12, no. 1, pp. 1–12, 2022, <https://doi.org/10.1038/s41598-022-25284-1>.
- [2] F. Khan, X. Yu, Z. Yuan, and A. U. Rehman, "ECG classification using 1-D convolutional deep residual neural network," *PLoS One*, vol. 18, no. 4, 2023, <https://doi.org/10.1371/journal.pone.0284791>.
- [3] N. Rafie, A. H. Kashou, and P. A. Noseworthy, "ECG interpretation: clinical relevance, challenges, and advances," *Hearts*, vol. 2, no. 4, pp. 505–513, 2021, <https://doi.org/10.3390/hearts2040039>.
- [4] H. Zheng et al., "Temporal dendritic heterogeneity incorporated with spiking neural networks for learning multi-timescale dynamics," *Nat. Commun.*, vol. 15, 2024, <https://doi.org/10.1038/s41467-023-44614-z>.

- [5] Z. Yan, J. Zhou, and W. F. Wong, "Energy efficient ECG classification with spiking neural network," *Biomed. Signal Process. Control*, vol. 63, 2021, <https://doi.org/10.1016/j.bspc.2020.102170>.
- [6] L. Xiaoxue et al., "Review of medical data analysis based on spiking neural networks," *Procedia Comput. Sci.*, vol. 221, pp. 1527–1538, 2023, <https://doi.org/10.1016/j.procs.2023.08.138>.
- [7] Y. Guo et al., "Real spike: learning real-valued spikes for spiking neural networks," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13672 LNCS, pp. 52–68, 2022, https://doi.org/10.1007/978-3-031-19775-8_4.
- [8] M. U. Khan, S. Aziz, S. Z. H. Naqvi, and A. Rehman, "Classification of coronary artery diseases using electrocardiogram signals," in *Proc. of 2020 International Conference on Emerging Trends in Smart Technologies (ICETST)*, 2020, <https://doi.org/10.1109/ICETST49965.2020.9080694>.
- [9] A. Rana, and K. K. Kim, "A novel spiking neural network for ECG signal classification," *Journal of Sensor Science and Technology*, vol. 30, no. 1, pp. 20–24, 2021, <https://doi.org/10.46670/JSST.2021.30.1.20>.
- [10] S. Aziz, M. U. Khan, M. Alhaisoni, T. Akram, and M. Altaf, "Phonocardiogram signal processing for automatic diagnosis of congenital heart disorders through fusion of temporal and cepstral features," *Sensors*, vol. 20, no. 13, pp. 1–20, 2020, <https://doi.org/10.3390/s20133790>.
- [11] Q. Liu, and Z. Zhang, "Ultralow power always-on intelligent and connected SNN-based system for multimedia IoT-enabled applications," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15570–15577, 2022, <https://doi.org/10.1109/JIOT.2022.3150307>.
- [12] I. P. Kamila, C. A. Sari, E. H. Rachmawanto, and N. R. D. Cahyo, "A good evaluation based on confusion matrix for lung diseases classification using convolutional neural networks," *Advance Sustainable Science, Engineering and Technology*, vol. 6, no. 1, pp. 1–8, 2024, <https://doi.org/10.26877/asset.v6i1.17330>.
- [13] N. Rathi, G. Srinivasan, P. Panda, and K. Roy, "Enabling deep spiking neural networks with hybrid conversion and spike timing dependent backpropagation," in *Proc. of International Conference on Learning Representations (ICLR)*, 2020.
- [14] F. Liu, W. Zhao, Y. Chen, Z. Wang, T. Yang, and L. Jiang, "SSTDTP: supervised spike timing dependent plasticity for efficient spiking neural network training," *Front. Neurosci.*, vol. 15, 2021, <https://doi.org/10.3389/fnins.2021.756876>.
- [15] I. K. Ihianle, A. O. Nwajana, S. H. Ebebuwa, R. I. Otuka, K. Owa, and M. O. Orisatoki, "A deep learning approach for human activities recognition from multimodal sensing devices," *IEEE Access*, vol. 8, pp. 179028–179038, 2020, <https://doi.org/10.1109/ACCESS.2020.3027979>.
- [16] C. J. Zhang, Yuan-Lu, F. Q. Tang, H. P. Cai, Y. F. Qian, and Chao-Wang, "Heart failure classification using deep learning to extract spatiotemporal features from ECG," *BMC Med. Inform. Decis. Mak.*, vol. 24, 2024, <https://doi.org/10.1186/s12911-024-02415-4>.
- [17] H. Zhang, W. Zhao, and S. Liu, "SE-ECGNet: A multi-scale deep residual network with squeeze-and-excitation module for ECG signal classification," in *Proc. of 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2020, pp. 2685–2691, <https://doi.org/10.1109/BIBM49941.2020.9313548>.
- [18] M. S. Islam et al., "HARDC: A novel ECG-based heartbeat classification method to detect arrhythmia using hierarchical attention based dual structured RNN with dilated CNN," *Neural Networks*, vol. 162, pp. 271–287, 2023, <https://doi.org/10.1016/j.neunet.2023.03.004>.
- [19] N. Alamatsaz, L. Tabatabaei, M. Yazdchi, H. Payan, N. Alamatsaz, and F. Nasimi, "A lightweight hybrid CNN-LSTM model for ECG-based arrhythmia detection," *Biomed. Signal Process. Control*, vol. 90, pp. 1–7, 2022, <https://doi.org/10.1016/j.bspc.2023.105884>.
- [20] A. Sun, W. Hong, J. Li, and J. Mao, "An arrhythmia classification model based on a CNN-LSTM-SE algorithm," *Sensors*, vol. 24, no. 19, 2024, <https://doi.org/10.3390/s24196306>.
- [21] F. Zhou and D. Fang, "Multimodal ECG heartbeat classification method based on a convolutional neural network embedded with FCA," *Sci. Rep.*, vol. 14, pp. 1–10, 2024, <https://doi.org/10.1038/s41598-024-59311-0>.
- [22] S. M. H. Irid, D. Moussaoui, M. Hadjila, and O. Azzoug, "Classification of ECG signals based on MIT-BIH dataset using bi-LSTM model for assisting cardiologists diagnosis," *Traitement du Signal*, vol. 41, no. 6, pp. 3245–3253, 2024.
- [23] T. Rasti, and Q. Zhu, "ECG-based arrhythmia detection by a shallow CNN model," *Research Square*, 2024, [Preprint.] Available: <https://doi.org/10.21203/rs.3.rs-3407809/v1>.
- [24] T. Wang, C. Lu, Y. Sun, M. Yang, C. Liu, and C. Ou, "Automatic ECG classification using continuous wavelet transform and convolutional neural network," *Entropy*, vol. 23, no. 1, pp. 1–13, 2021, <https://doi.org/10.3390/e23010119>.
- [25] J. Wang and W. Li, "Atrial fibrillation detection and ECG classification based on CNN-BiLSTM," *arXiv*, 2020, <https://doi.org/10.48550/arXiv.2011.06187>.
- [26] N. A. D. B. M. Noor, W. Y. Chiew, Z. M. Noh, and R. S. S. Singh, "Neurons to heartbeats: spiking neural networks for electrocardiogram pattern recognition," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 36, no. 2, pp. 863–871, 2024, <https://doi.org/10.11591/ijeecs.v36.i2.pp863-871>.
- [27] R. Bousseljot, D. Kreiseler, and A. Schnabel, "Nutzung der EKG-Signaldatenbank CARDIODAT der PTB über das Internet," *Biomedical Engineering*, vol. 40, suppl. 1, pp. 317–318, 1995, <https://doi.org/10.13026/C28C71>.
- [28] M. Horowitz, "1.1 Computing's energy problem (and what we can do about it)," in *Proc. of 2014 IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers (ISSCC)*, 2014, pp. 10–14, <https://doi.org/10.1109/ISSCC.2014.6757323>.