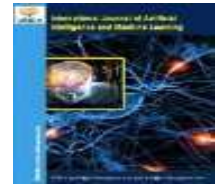


SvedbergOpen
DISSEMINATION OF KNOWLEDGEInternational Journal of Artificial
Intelligence and Machine LearningPublisher's Home Page: <https://www.svedbergopen.com/>

Research Paper

Open Access

Interpretable and Scalable Multi-Lead ECG Arrhythmia Detection Using a Selective Multi-Feature Branch Network

¹Ms. Ankita Shukla*, ²Prof. Muhamed Izharuddin¹PhD Research Scholar, Department of Computer Engineering, Zakir Husain College of Engineering & Technology (ZHCET), Aligarh Muslim University, Aligarh, 202002, U.P., India. E-mail: ankitashukla@zhcet.ac.in²Professor, Department of Computer Engineering, Zakir Husain College of Engineering & Technology (ZHCET), Aligarh Muslim University, Aligarh, 202002, U.P., India. E-mail: izharuddin@zhcet.ac.in*Corresponding mail: ankitashukla@zhcet.ac.in**Abstract**

Precise analysis of cardiac arrhythmias as influenced by multi-lead electrocardiogram (ECG) signals for early diagnosis and clinical intervention is important. However, available deep learning techniques overlook lead-dependent morphological changes or overcompensate with high cost and complexity hence less ideal for real-time and resource-constrained environments. This work introduces SMFB-Net (Selective Multi-Feature Branch Network), an efficient, accurate, and robust deep learning framework for multi-class arrhythmia classification based on 12-lead ECG signals. The architecture proposed is designed to embed an Adaptive Lead Selection Module (LSM) for prioritizing diagnostically informative leads with minimum redundant computations, which resulted in floating point operations (FLOPs) being 30.92% lower than 12-lead processing. For spatial and multi-scale features, convolution and inception blocks are employed and temporal dependencies are modeled with the help of Bidirectional Gated Recurrent Units (Bi-GRUs). The obtained representations are combined, and global average pooling is performed to classify them into nine arrhythmia subgroups. Experimental tests on the publicly available China Physiological Signal Challenge (CPSC) 2018 dataset reveal that SMFB-Net achieves a macro F1-score of 0.864. The model performed particularly well for difficult arrhythmic classes such as first-degree atrioventricular block (I-AVB) and ST-segment deviation abnormalities. Furthermore, SHAP interpretability analysis on patient and population data also confirmed clinically relevant ECG leads. In general, SMFB-Net shows a balance between classification performance, interpretability, and computational efficiency, giving it better overall performance compared to other deep learning techniques. This makes our proposed multi branch framework capable of deployment in real-time and resource-limited arrhythmia monitoring applications.

1. Introduction

Cardiovascular diseases remain the leading cause of death worldwide, killing approximately 17.9 million people annually and exacting an immense burden on health care systems worldwide¹. Among these, abnormalities in the heart's rhythm known as cardiac arrhythmias are especially high in importance, accounting for about 80% of all sudden cardiac deaths². Recognizing the arrhythmias at an early stage is essential for appropriate patient care, thereby improving patients' prognosis. The electrocardiogram (ECG), which quantifies the activity on the surface of the heart, is generally used for the evaluation of rhythm and conduction disorders. As shown in Fig.1, subfigure 1a provides the triangle generated by RA, LA, and LL electrodes that includes the three typical bipolar limb leads (I, II, III) and shows the frontal plane view of cardiac activity. Further elaborating this aspect, subfigure 1b is shown, where a 12-lead ECG sequence consists of six limb leads in the order I, II, III, aVR, aVL, and aVF and six precordial chest leads in this set (V1–V6). The limb leads record the frontal plane activity, while the precordial leads register the horizontal plane activity, and together the twelve leads give a holistic view of cardiac electrophysiology. The centrality of this system for the detection and classification of arrhythmias and for clinical decisions has been established (Geetanjali & Singh, 2025).

Automatic arrhythmia detection (AAD) has attracted much research interest to facilitate continuous monitoring, computer-aided diagnosis, and smart healthcare systems. Early AAD approaches were mainly based on hand-crafted feature extraction techniques, using morphology-based feature extraction methods, including morphological descriptors, P-QRS-T interval analysis³, wavelet transforms⁴, and conventional machine learning classifiers (e.g., SVMs)⁵, decision trees, and Bayesian models. While these methods performed reasonably well in classification work, they were heavily reliant on manual feature engineering and experienced poor performance under noisy and variable ECG input data.

Recent advances in deep learning (DL) have enabled end-to-end feature learning directly from raw ECG signals, significantly improving arrhythmia classification performance^{6,7}. Consequently, ECG

classification frameworks have evolved from conventional convolutional and recurrent architectures toward more sophisticated hybrid and multi-branch learning models capable of exploiting multi-lead spatial and temporal features^{15,16,19}. However, many existing multi-lead frameworks¹ process all ECG leads uniformly, leading to increased computational overhead and redundant feature extraction. Furthermore, interpretability is commonly addressed only as a post-hoc analysis rather than being incorporated into the framework design itself.

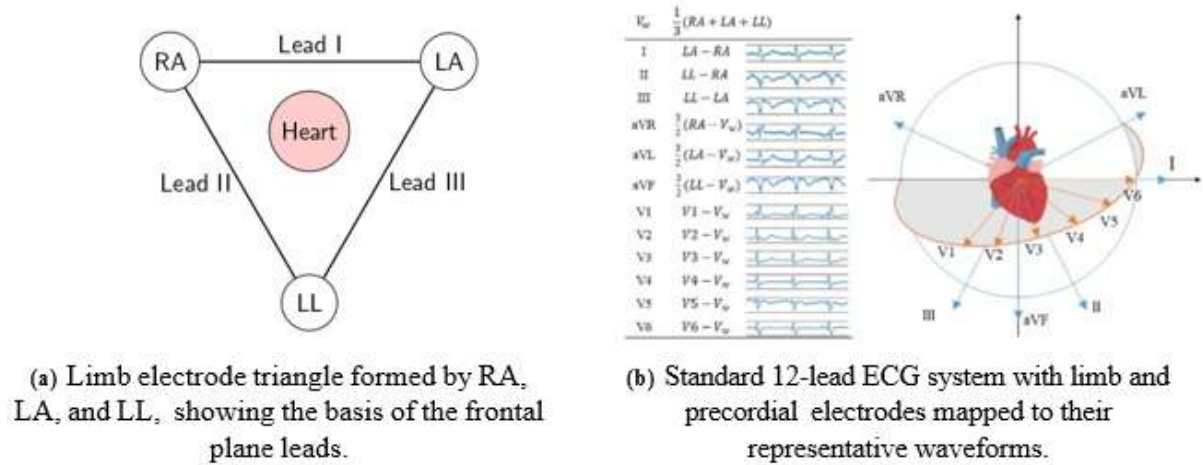


Figure 1. Comprehensive view of the 12-lead ECG system: (a) limb electrode triangle (RA, LA, LL) forming the foundation of bipolar limb leads; (b) full 12-lead configuration combining six limb leads and six precordial leads to provide complementary frontal and horizontal plane perspectives of cardiac electrical activity.

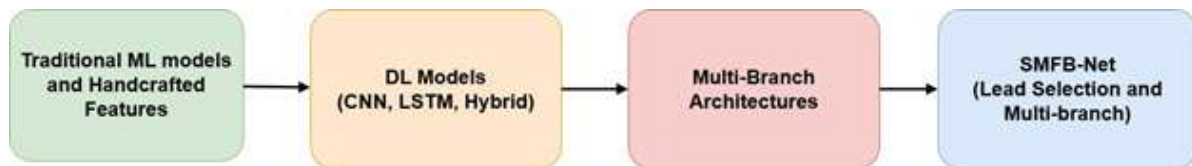


Figure 2. Motivation and contributions of the proposed framework: from cardiovascular disease burden to ECG-based arrhythmia detection, limitations of traditional and deep learning approaches, and the proposed SMFB-Net for efficient and interpretable analysis.

The overall progression of arrhythmia detection approaches and the motivation behind the proposed framework are illustrated in Fig. 2. Arrhythmia detection methods have advanced from classical machine learning methods with handcrafted features toward deep learning-based multi-branch architectures. In view of the shortcomings of existing methods, the proposed SMFB-Net combines adaptive lead selection and multi-branch feature learning, further improving computational scalability and interpretability essential to facilitate multi-lead ECG analysis. To address these challenges, SMFB-Net is proposed for interpretable multi-lead arrhythmia detection:

A selective multi-feature branch architecture of multi-lead arrhythmia detection is proposed in this paper, based on convolutional feature extraction, Inception-based multi-scale learning, and Bi-GRU temporal modeling.

A promising method for the detection of clinically informative ECG leads is proposed based on adaptive lead selection with the sole effect of avoiding redundant feature processing during inference.

A lightweight multi-branch fusion approach provides effective handling of spatial ECG data and the temporal ECG characteristics while preserving computational scalability.

A SHAP-based interpretability framework is incorporated, which lays the foundation for lead-level explainability so that prediction of the model can be considered in clinical context.

The proposed framework offers a suitable-scale, accurate multi-lead arrhythmia classification with extensive experimental results on the CPSC 2018 dataset.

2. RELATED WORK

Automated arrhythmia detection (AAD) has evolved significantly over the decades, progressing from feature engineering based methods to sophisticated deep learning (DL)-based frameworks. Existing studies have explored convolutional, recurrent, hybrid, attention-based, and multi-branch architectures for

improving ECG classification performance. Representative methodologies, findings, and distinguishing characteristics of existing DL-based arrhythmia detection models are summarized in Table 1.

Table 1. Methodological features and key findings of DL models for AAD task.

Reference	Model	Key Findings	Unique Features
Rajpurkar et al. ⁶	1D CNN	Achieved cardiologist-level performance on single-lead arrhythmia detection on large ambulatory datasets.	Large-scale supervised training on long ECG records; end-to-end learning from raw signal.
Hannun et al. ⁷	Deep residual CNN	Demonstrated cardiologist-level arrhythmia classification in ambulatory ECG monitoring.	Extensive annotated ambulatory dataset; clinical evaluation and reader study.
Yildirim et al. ⁸	CNN on ECG spectrograms	Improved robustness using spectrogram-derived features and data augmentation.	Spectrogram-based input, novel augmentation pipeline for ECG variability handling.
Acharya et al. ⁹	CNN on varied tachycardia segments	Accurate arrhythmia detection across segments of different lengths.	Segment-length-aware training; emphasis on tachycardia interval selection.
Kiranyaz et al. ¹⁰	1D patient-specific CNN	Real-time, patient-specific classification with low-latency inference.	Lightweight 1D CNN tailored for patient-specific adaptation and real-time deployment.
Alamatsaz et al. ¹¹	Lightweight CNN-LSTM hybrid	8 classes + normal rhythm classification; ~98.2% accuracy.	No handcrafted feature extraction; small-segment CNN-LSTM design.
Yao et al. ¹²	ATICNN	Improved detection of paroxysmal arrhythmias by handling variable-length ECGs.	Attention-based time-incremental CNN, adaptive processing.
Liu et al. ¹³	CNN + expert features	Combines expert ECG features with CNN for robust arrhythmia detection.	Hybrid architecture that fuses handcrafted features with learned CNN representations.
Yao et al. ¹⁴	Time-incremental CNN + Attention	Handles varied-length 12-lead ECGs with attention for better temporal focus.	Time-incremental processing; attention mechanism to prioritize informative segments.
Wang et al. ¹⁵	Multi-scale fusion CNN	High multi-class F1 via multi-scale temporal feature fusion.	Multi-scale convolutional branches and feature fusion for varied morphology capture.
Zhang et al. ¹⁶	CNN + Attention + Bi-GRU	Strong multi-class performance using spatio-temporal features from 12-lead ECGs.	Spatio-temporal attention to focus on lead-wise information; Bi-GRU for temporal context.
Oh et al. ¹⁷	CNN + Bi-GRU + Multi-Head Attention	Enhanced focus on salient features; strong classification across arrhythmia classes.	Attention module and bidirectional GRU; stronger than simple CNN or RNN.
Zhang et al. ¹⁹	Multi-Lead-Branch Fusion Network (MLBF-Net)	Improved multi-class arrhythmia classification using lead-branch specialization.	Dedicated lead-branch feature extractors and cross-lead fusion strategy.

2.1 Summary of Prior Work

CNN-based approaches: Early Deep Learning studies demonstrated the effectiveness of convolutional neural networks (CNNs) for end-to-end ECG feature learning. Rajpurkar et al.⁶ proposed a deep 1-D CNN architecture for single-lead ECG analysis and achieved cardiologist-level arrhythmia classification performance. Similarly, Hannun et al.⁷ applied residual CNN architectures and validated their impact via

clinical reader studies. These techniques have established CNNs as robust tools for automatic feature extraction of ECG signals and arrhythmia detection.

Spectrogram and segment-based methods: to improve robustness to variable-length ECG signals, Yildirim et al.⁸ proposed spectrogram-based CNN models, whereas Acharya et al.⁹ generated segment-oriented CNN models focusing on tachycardia intervals. These methods showed the importance of transformed and localized ECG representations for improving classification performance when morphologically different ECG signals were given as input.

Patient-specific and lightweight models: Many works focused on efficiency in computation and the ability to make use of patient-specific ECG analysis. Kiranyaz et al.¹⁰ applied a small-scaled patient-specific CNN architecture to real-time ECG classification, while Alamatsaz et al.¹¹ proposed a lightweight CNN-LSTM hybrid for multi-class arrhythmia detection.

However, although the computational complexity of these approaches was reduced, there was a scarcity in their ability to generalize to various subgroups of the population.

Multi-lead and attention-based frameworks: As the availability of large-scale 12-lead ECG datasets rapidly increased, research gradually shifted to investigating inter-lead spatial and temporal relationships. Yao et al.¹⁴ introduced a time-incremental CNN with attention method based on variable-length ECG measurements, and Zhang et al.¹⁹ proposed MLBF-Net using lead-specific branch learning and cross-lead feature fusion. Similarly, Wang et al.¹⁵ has used multi-scale convolutional branches to enhance morphological feature extraction. Despite the fact that these techniques improved classification performance, the majority of the methods were homogeneous in terms of ECG leads and could cause computational overheads, especially redundant feature extraction.

Hybrid and interpretable DL models: Emerging research has introduced recurrent learning, attention mechanisms and explainable AI techniques to improve temporal modeling and interpretability²⁰. Liu et al.¹³ integrated CNN-based learning and expert-driven features for reliable ECG analysis, while Oh et al.¹⁷ incorporated CNNs, Bi-GRUs and attention modules for better temporal feature representation. In addition, Ribeiro et al.²¹ explored explainable AI methods to promote transparency in ECG classification systems. Despite these optimizations, interpretability is mostly still a post-hoc analysis, not an in-line method, and is rarely built into AAD models.

2.2 Research Gap

Major breakthroughs were made in DL-based arrhythmia detection but still present many tough challenges that still persist today. Current multi-lead approaches typically process all ECG leads simultaneously (i.e., processing multiple ECGs at once), leading to higher computational overhead and redundant feature extraction. Moreover, other strategies are mainly based on classification and have a weak focus on adaptive lead exploitation. Moreover, interpretable features are often added after the model training rather than embedded into the framework design itself. These constraints inspire the development of lightweight and interpretable multi-lead architectures that can balance classification performance, computation efficiency, and clinical explainability.

3. Methods

SMFB-Net is capable of multi-class arrhythmia detection by processing only the diagnostically informative leads detected by the Lead Selection Module (LSM). As illustrated in Fig. 3 the proposed architecture has parallel BranchNet modules with convolutional and Inception blocks processing spatial features and Bi-GRUs modeling the temporal dependencies. The results from all of the branches are fused via a fusion layer to obtain the merged representation prior to classification. The general design is made up of three main parts, the content of which is given elaborately in the subsections.

3.1 Lead Selection Module (LSM)

Although 12-lead ECGs offer insightful diagnostic information, not all leads contribute equally to arrhythmia classification. The LSM identifies and dynamically selects the most informative leads during training to optimize both accuracy and computational load. Leads with scores above or equal to 0.5 are retained for further processing.

3.1.1 Mathematical Formulation

Lead scores are computed using a learnable lead-attention mechanism, followed by a modified Step ReLU operation. Let $X \in \mathbb{R}^{12 \times T}$ denote the input 12-lead ECG signal containing T time samples. The lead-wise transformed representation is first obtained as

$$\mathbf{z} = \mathbf{W}\mathbf{X} + \mathbf{b}, \quad (1)$$

$$\mathbf{z}^- = \frac{1}{T} \sum_{t=1}^T \mathbf{z}(t) \quad (2)$$

$$\mathbf{a} = \sigma(\mathbf{w}_1 \odot \mathbf{z}^-) \quad (3)$$

$$\mathbf{X}_{\text{score}} = \text{StepReLU}(\mathbf{a}). \quad (4)$$

Here, $\mathbf{W} \in \mathbb{R}^{12 \times 12}$ is a learnable weight matrix, $\mathbf{b} \in \mathbb{R}^{12 \times 1}$ is the learnable bias vector, and $\mathbf{z} \in \mathbb{R}^{12 \times T}$ represents the transformed ECG features. Temporal averaging of $\mathbf{z}$ produces $\mathbf{z}^- \in \mathbb{R}^{12}$, where each element represents the temporally aggregated information corresponding to one ECG lead. The learnable vector $\mathbf{w}_1 \in \mathbb{R}^{12}$ assigns lead-specific importance, $\odot$ denotes element-wise multiplication, and $\sigma(\cdot)$ denotes the sigmoid activation function. Consequently, $\mathbf{a} = [a_1, a_2, \dots, a_{12}]^T \in \mathbb{R}^{12}$ contains one relevance score for each ECG lead.

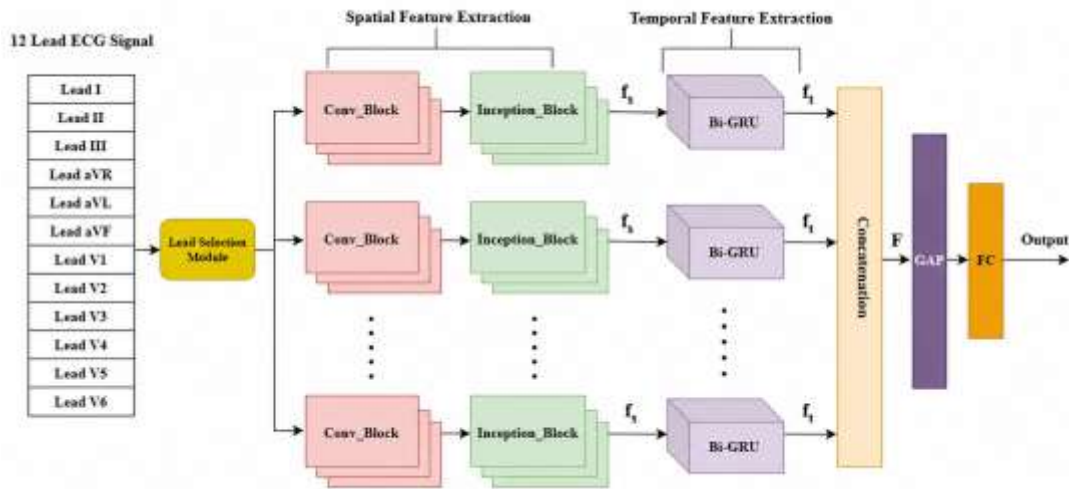


Figure 3. Architecture of SMFB-Net: adding Lead Selection Module (LSM), parallel Branch-Nets for spatial and temporal feature extraction, and a Feature Fusion unit for multi-class arrhythmia classification

The modified Step ReLU operation is defined element-wise as

$$\text{StepReLU}(a_i) = \begin{cases} a_i, & \text{if } a_i \geq \tau, \\ 0, & \text{if } a_i < \tau, \end{cases} \quad \tau = 0.5, \quad (5)$$

where τ denotes the lead-selection threshold. Thus, the filtered lead-score vector is

$$\mathbf{X}_{\text{score}} = [\text{StepReLU}(a_1), \dots, \text{StepReLU}(a_{12})]^T \in \mathbb{R}^{12}. \quad (6)$$

The final set of selected informative leads is defined as

$$\mathbf{S} = \{i \in \{1, \dots, 12\} \mid a_i \geq \tau\}, \quad \tau = 0.5. \quad (7)$$

Therefore, only the ECG leads whose learned relevance scores are greater than or equal to the threshold are forwarded to the subsequent Branch-Net modules.

3.1.2 Algorithmic Implementation

For each ECG recording, the LSM first transforms the input signals and aggregates the resulting features over the temporal dimension. A learnable lead-attention vector is then applied element-wise to the temporally averaged representation to generate one relevance score for each of the 12 ECG leads. The modified Step ReLU operation suppresses leads with relevance scores below the predefined threshold $\tau = 0.5$. Only the retained informative leads are forwarded to the subsequent feature-extraction stages. Algorithm 1 summarizes the complete lead-selection procedure using Eqs. (1)–(7). The symbols used in the algorithm are summarized in Table 2.

Algorithm 1 Lead Selection Module (LSM)**Require:** Input ECG matrix $\mathbf{X} \in \mathbb{R}^{12 \times T}$, weight matrix $\mathbf{W}$, bias vector $\mathbf{b}$, lead-attention vector $\mathbf{w}_1$, threshold $\tau = 0.5$ **Ensure:** Selected lead set $\mathbf{S}$

- 1: Compute transformed lead features: $\mathbf{z} \leftarrow \mathbf{WX} + \mathbf{b}$
- 2: Compute temporal average: $\bar{\mathbf{z}} \leftarrow \frac{1}{T} \sum_{t=1}^T \mathbf{z}_{:,t}$
- 3: Compute lead-wise relevance scores: $\mathbf{a} \leftarrow \sigma(\mathbf{w}_1 \odot \bar{\mathbf{z}})$
- 4: Apply Step ReLU: $\mathbf{X}_{\text{score}} \leftarrow \text{StepReLU}(\mathbf{a})$
- 5: Select informative leads: $\mathbf{S} \leftarrow \{i \in \{1, \dots, 12\} \mid a_i \geq \tau\}$
- 6: **return** $\mathbf{S}$

Table 2. Description of Symbols Used in the Lead Selection Module (LSM)

Symbol	Description
$\mathbf{X} \in \mathbb{R}^{12 \times T}$	Input ECG matrix containing 12 leads and T time samples
$\mathbf{W} \in \mathbb{R}^{12 \times 12}$	Learnable weight matrix for lead transformation
$\mathbf{b} \in \mathbb{R}^{12 \times 1}$	Learnable bias vector
$\mathbf{z} \in \mathbb{R}^{12 \times T}$	Transformed ECG feature matrix
$\bar{\mathbf{z}} \in \mathbb{R}^{12}$	Temporally averaged lead-feature vector
$\mathbf{w}_1 \in \mathbb{R}^{12}$	Learnable lead-attention vector
$\mathbf{a} \in \mathbb{R}^{12}$	Lead-wise relevance scores after sigmoid activation
$\mathbf{X}_{\text{score}} \in \mathbb{R}^{12}$	Lead scores after Step ReLU thresholding
T	Lead-selection threshold, set to 0.5
$\mathbf{S}$	Set of selected informative lead indices
$\sigma(\cdot)$	Sigmoid activation function
$\odot$	Element-wise multiplication
$\text{StepReLU}(\cdot)$	Threshold operation retaining scores $a_i \geq T$

3.2 Branch-Net Architecture

After choosing a lead, any retained ECG lead is processed individually in a Branch-Net module. Each Branch-Net consists of convolutional, Inception, and recurrent units in order to be able to extract spatial and temporal information over a large series of iterations. This modularity allows SMFB-Net to take advantage of the unique morphological and dynamic signatures, which are present in particular ECG leads, while maintaining flexibility and scalability over variable length signals.

3.2.1 Convolutional Block

Each Branch-Net employs three successive one-dimensional convolutional (Conv1D) layers with 32, 64, and 128 filters and kernel sizes of 7, 5, and 3, respectively. The corresponding strides are 2, 2, and 1. The progressively decreasing kernel sizes enable the network to capture ECG characteristics at different temporal resolutions, ranging from broader waveform patterns to finer morphological variations. Each convolutional layer is followed by Leaky ReLU activation and batch normalization (BN)²⁴ to introduce non-linearity and stabilize feature learning. Max-pooling operations are additionally employed for temporal downsampling and reduction of computational complexity. Following the convolutional and subsequent Inception-based feature extraction stages, the resulting spatial features are combined with the temporal features learned by the Bi-GRU. Global Average Pooling (GAP)²⁸ is applied only after spatial-temporal feature fusion to obtain a fixed-dimensional representation for final classification.

3.2.2 Inception Block

Inception Block handles input feature maps in parallel over multiple branches and simulates multi-scale feature representations. All paths use convolutional filters on different receptive fields (1×1 , 3×3 , 5×5), or an additional max-pool operation of coarser and finer feature extraction. In these pipelines, a convolution over 1×1 , inspired by the GoogLeNet architecture²⁵, is computed to reduce dimensionality before larger kernels without the loss of representational richness and computational convenience. This method is especially well-suited to the field of electrocardiogram (ECG) signal processing, where clinical interest features present temporal layers with different scales.

3.2.3 Bidirectional GRU Layer

Temporal features in ECG signals are bidirectional, which are largely influenced by previous and subsequent time periods. Bidirectional Gated Recurrent Unit (Bi-GRU) model architecture is used for this task²⁶. The Bi-GRU analyzes input sequences in both forward and reverse directions as mentioned in following Eqs. (8)–(10):

$$\mathbf{h}_t^f = \text{GRU}_f(\mathbf{x}_t, \mathbf{h}_{t-1}^f), \quad (8)$$

$$\mathbf{h}_t^b = \text{GRU}_b(\mathbf{x}_t, \mathbf{h}_{t+1}^b), \quad (9)$$

$$\mathbf{h}_t = [\mathbf{h}_t^f, \mathbf{h}_t^b], \quad (10)$$

where $\mathbf{h}_t$ denotes the concatenated hidden state at time t . To enhance generalization, the Bi-GRU outputs are passed through a Leaky ReLU activation, followed by batch normalization and a dropout layer having 0.3 as dropout rate. This combination ensures stable temporal feature learning while mitigating overfitting risks.

Table 3. Simplified architecture of SMFB-Net with layer-wise configuration.

Stage	Layer / Block	Configuration / Output
Input	ECG Input	12-lead segment, size $L \times 12L \times 12$
	Lead Selection (LSM)	Sigmoid attention + Step ReLU; retain leads with score ≥ 0.5
Branch-Net (per selected lead)	Conv1D Layers	Filters: 32, 64, 128; Kernel sizes: 7, 5, 3; Strides: 2, 2, 1
	Inception Block	Multi-scale convolutions with kernel sizes 1, 3, and 5
	Bi-GRU	64 hidden units in each direction
Fusion and Classification	Feature Fusion	Concatenation of spatial and temporal features
	Global Avg. Pooling	Sequence-level pooling to obtain a fixed-length representation
	Fully Connected Layer	Dense + Softmax activation; 9-class prediction

3.3 Justification of Design Choices

The architecture in Table 3 balanced representational power and computational efficiency. The Lead Selection Module (LSM) only processes channels with diagnostically informative information to decrease redundancy. To improve performance, Conv1D layers employ progressively larger filter banks (32–128) and decreasing kernel sizes (7–3), minimizing temporal resolution and capturing low- to high-level ECG morphology in a low-throughput approach. Multi-scale kernels in the Inception blocks facilitate detection of localized as well as long-duration arrhythmia. To capture temporal dependence based on the onset and offset of arrhythmia without promoting excessive parameter expansion, the Bi-GRU consists of 64 hidden units in each direction. Through feature fusion, a 512-dimensional spatial-temporal combination is created that is then pooled into a fixed-length vector to support variable-length ECGs. Finally, one fully connected Softmax layer generates predictions across nine clinically relevant arrhythmia classifications. These options are robust, expressive, and computationally manageable enough for real-time deployment.

3.4 Feature Extraction and Fusion

One Branch-Net returns spatial features $\mathbf{f}_s$ obtained from the Inception Block and a temporal feature $\mathbf{f}_t$ from the Bi-GRU layer. To summarize results across all leads, the spatial and temporal features are aggregated as described in (11).

$\mathbf{F} = \text{concat}(\mathbf{f}_s, \mathbf{f}_t)$ for all selected branches (11)

The fused feature representation $\mathbf{F}$ is also processed through Global Average Pooling (GAP) layer which compresses the few dimensions but retains the most salient data.

3.5 Computational Complexity Analysis

The SMFB-Net is built with the goal of enabling dynamic lead selection mechanisms to improve the computational scale during training and inference. Let $O(n \times c)$ be the input complexity, n the number of samples, and c the ECG lead number. By default, the lead selection module adaptively reduces the effective complexity to $O(n \times c')$ and the $c' < c$ decreases with an increasing number of informative leads that are used in different instances. The use of lightweight branches and adaptable lead utilization of SMFB-Net increases the computational efficiency of SMFB-Net for resource-constrained and real-time ECGs analysis applications while maintaining classification performance.

4. EXPERIMENTATION

4.1 Dataset

To evaluate and validate performance of our model CPSC 2018 dataset²⁹ is being used. It comprises 6,877, 12-lead ECG recordings from different individuals with time durations from 6 to 60 seconds, sampled at 500 Hz. It is annotated for 8 different cardiac conditions and a sinus rhythm. Detailed class distributions are shown in Table 4.

Table 4. Class Distribution of the CPSC 2018 Dataset.

Class	Number of Records
SNR	918
AF	1098
I-AVB	704
LBBB	207
RBBB	1695
PAC	556
PVC	672
STD	825
STE	202
Total	6877

4.2 Evaluation Metrics

We evaluate the effectiveness of our method in terms of precision, recall and F1 score on each class. They are calculated as per Eqs. (12)–(14). Here TP, FP, and FN denote true positives, false positives, and false negatives, respectively.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (12)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (13)$$

$$\text{F1 Score} = \frac{2 \times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (14)$$

4.3 Preprocessing

The original 12-lead ECG measurements obtained with a sample frequency of 500 Hz are downsampled to 250 Hz to achieve an accurate signal frequency to ease computation and maintain diagnostic frequency and for analysis of arrhythmias. The vast majority of clinically relevant ECG features are measured at low frequencies leading to lower computation load without losing considerable data, represented as information loss. In order to ensure a uniform input type size for batch-wise model training, we truncate the size of recordings larger than 60 seconds and zero-padded all shorter recordings to defined sequence length. This preprocessing enabled stable batch-wise training and consistent feature scaling across recordings. In addition, the signal normalization is performed separately for each ECG lead to reduce the amplitude differences between the recordings and increase numerical stability in the optimization step. The normalized ECG signals facilitate more stable and consistent feature learning, and make the proposed SMFB-Net framework more robust across different arrhythmia classes.

4.4 Training Strategy

The overall training methodology of SMFB-Net consists of model initialization, optimization setup, and iterative feature learning using the proposed Branch-Net framework.

4.4.1 Training Setup

The training configuration of SMFB-Net is summarized in Table 5. The weight matrix W is initialized using Xavier uniform initialization³⁰ to maintain stable gradient propagation during optimization, while w_1 is initialized from a normal distribution and biases are initialized to zero. The model is optimized using the Adam optimizer³¹ with an initial learning rate of 1×10^{-3} . To reduce overfitting and improve generalization, dropout regularization²⁷ is applied after the Bi-GRU layers and fully connected layer.

Table 5. Training configuration for SMFB-Net.

Parameter	Value
Dataset split	Stratified 80/20 (train/test)

Batch size	64
Epochs	100
Optimizer	Adam
Learning rate	1×10^{-3} (decayed by factor 0.1 every 30 epochs)
Loss function	Focal Loss ($\gamma = 2$)
Dropout rate	0.5 (fully connected layer)
Weight initialization	Xavier initialization
Hardware	NVIDIA Tesla V100 GPU, 32 GB RAM

Algorithm 2 Training Process of SMFB-Net for Multi-class Arrhythmia Classification

Require: Dataset $D = (L, y)$, epochs E , batch size B , lead-selection threshold $\tau = 0.5$

Ensure: Trained SMFB-Net model

```

1: Split  $D$  into training set  $D_{Tr}$  (80%) and testing set  $D_{Te}$  (20%) using stratified sampling
2: for epoch = 1 to  $E$  do
3:   for each batch  $B \subseteq D_{Tr}$ ,  $|B| = B$  do
4:     for each ECG sample  $L_i \in B$  do
5:       Compute transformed lead features:  $z_i \leftarrow WL_i + b$ 
6:       Compute temporal average:  $\bar{z}_i \leftarrow \frac{1}{T} \sum_{t=1}^T z_{i,t}$ 
7:       Compute lead-wise relevance scores:  $a_i \leftarrow \sigma(w_1 \odot \bar{z}_i)$ 
8:       Apply Step ReLU:  $X_{score,i} \leftarrow \text{StepReLU}(a_i)$ 
9:       Select informative leads:  $S_i \leftarrow \{j \in \{1, \dots, 12\} \mid a_{i,j} \geq \tau\}$ 
10:      Extract spatial features from selected leads  $S_i$  using stacked Conv1D and Inception1D layers:  $f_{s,i} \leftarrow \text{BranchNet}(L_i[S_i])$ 
11:      Compute temporal features using Bi-GRU as defined in Eqs. (8)–(10):  $f_{t,i} \leftarrow \text{Bi-GRU}(f_{s,i})$ 
12:      Fuse spatial and temporal features:  $F_i \leftarrow \text{Concat}(f_{s,i}, f_{t,i})$  using Eq. (11)
13:      Compute pooled representation:  $g_i \leftarrow \text{GAP}(F_i)$ 
14:      Predict class probabilities:  $\hat{y}_i \leftarrow \text{Softmax}(\text{FC}(g_i))$ 
15:    end for
16:    Compute batch focal loss:  $L \leftarrow \frac{1}{B} \sum_{i=1}^B \text{FocalLoss}(y_i, \hat{y}_i)$ 
17:    Update all learnable parameters using the Adam optimizer
18:  end for
19: end for
20: return Trained SMFB-Net model

```

4.4.2 Training Procedure

The complete training procedure of SMFB-Net is presented in Algorithm 2, while the corresponding hyperparameters and computational settings are summarized in Table 5. To maintain class balance, the data is divided into training and testing subsets using a stratified 80/20 ratio. For each epoch, the Lead Selection Module (LSM) computes the attention-based weights across all 12 leads and suppresses leads with low diagnostic relevance. The retained leads are processed through individual Branch-Net modules, in which Conv1D and Inception layers are used to extract the local morphological features whilst Bi-GRUs represent long-range temporal dependencies. The spatial and temporal representations are subsequently fused, combined into a fixed-length vector and a fully connected layer to map the final nine arrhythmia classes. To resolve the inherent class imbalance in arrhythmia datasets, the focal loss³² is applied during optimization, replacing standard cross-entropy by concentrating on minority classes to emphasize hard-to-classify and minority-class samples during optimization. Model parameters are updated using the Adam optimizer using planned learning rate decay and the training runs for 100 epochs at batch size of 64 until convergence.

5. Results

This section demonstrates the performance of the proposed model on CPSC 2018 dataset²⁹, especially on the lead selection and classification performance, and F1 score distribution in nine arrhythmia classes. The results are validated using SHAP-based interpretability at both patient and population levels.

5.1 Lead Score Analysis

The lead selection module (LSM) evaluates the diagnostic significance of each ECG lead using Eqs. (1)–(5), where leads with scores greater than or equal to 0.5 are retained for further processing. Figure 4 illustrates the normalized lead scores at both patient-specific and population levels, while Table 6 summarizes the percentage frequency of lead selection across the training dataset.

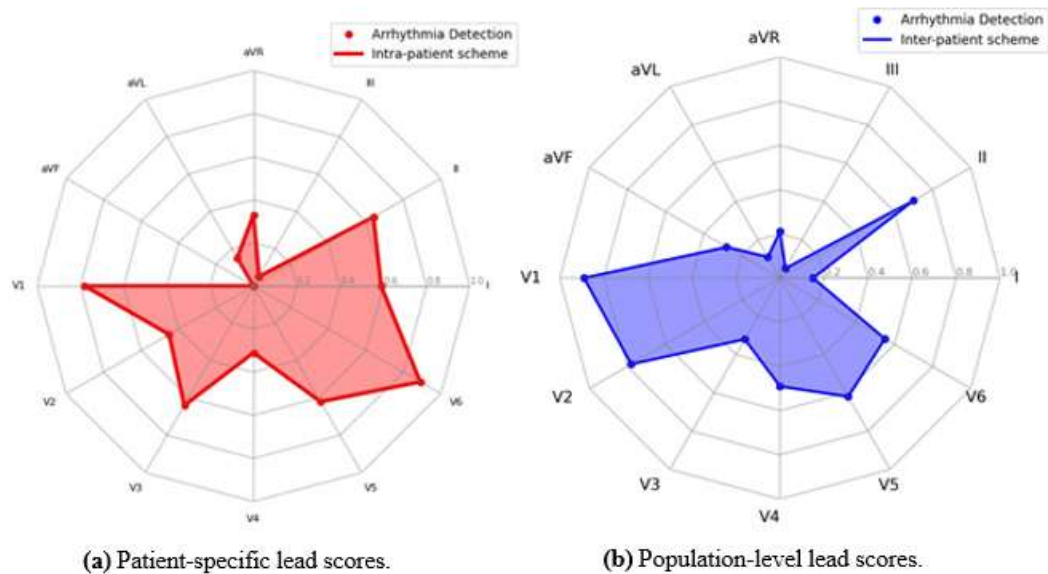


Figure 4. Lead score analysis for SMFB-Net

At the patient level, Fig. 4a shows how the proposed LSM adaptively prioritizes diagnostically relevant leads according to morphological features of specific ECG recordings. Lead scores for leads V1, V2, II, and V5 are higher, suggesting that they contribute strongly towards the characterization of arrhythmia, while leads III and aVR show a relatively lesser importance. Likewise, the population-level visualization in Fig. 4b shows that leads V1, V2, V5, V6, and II overall exhibit higher average lead scores for the series, thereby indicating their superior representational capabilities towards describing the clinically relevant cardiac patterns.

The quantitative data presented in Table 6 further supports the radar plots findings. Leads V1 (95.6%), V2 (93.2%), V5 (92.3%), II (91.2%), and V4 (90.7%) are selected most frequently during training, indicating their consistent clinical relevance in various arrhythmia taxonomies. While V4 demonstrates average attention during radar plot visualization, giving its high selection rate suggests the likelihood that it contributes evenly in a wide variety of the records. Leads aVR (57.4%) and III (62.8%) show only relatively small selection frequencies, indicating that these leads do not make significant contributions to classification. Overall analysis of the visual attention and lead selection frequencies demonstrates that the proposed LSM effectively values clinically informative leads while suppressing less contributive leads, allowing the model to concentrate on clinically informative leads, while diluting the impact of less contributing leads.

5.2 Classification Performance

The classification performance of SMFB-Net was assessed for the nine arrhythmia classes by using binary confusion matrices, as shown in Figure 5.

Table 7 reports the class-wise classification performance of the proposed SMFB-Net model. The model achieved superior recall for AF (98%) and RBBB (97%), suggesting that there is extensive identification of specific ECG-related patterns (e.g., irregular RR intervals for AF, wide QRS for RBBB). However, PAC and STD experienced much larger rates of misclassification (31% and 18%), respectively, during classification, this suggests some clinical characteristics where similarities arise (e.g., slight P-wave changes for PAC and altered ST-segment morphology for STD). These characteristics indicate very good classification ability for the majority of classes, but struggle with subtle arrhythmia detection, which is indicative of the class imbalance of the training data (e.g. PAC: 556 records, STE: 202 records).

Table 6. Lead Selection Frequency Across Training Dataset

Lead	Selection Frequency (%)
I	89.3
II	91.2
III	62.8
aVR	57.4
aVL	60.2
aVF	70.8

V1	95.6
V2	93.2
V3	87.1
V4	90.7
V5	92.3
V6	88.6

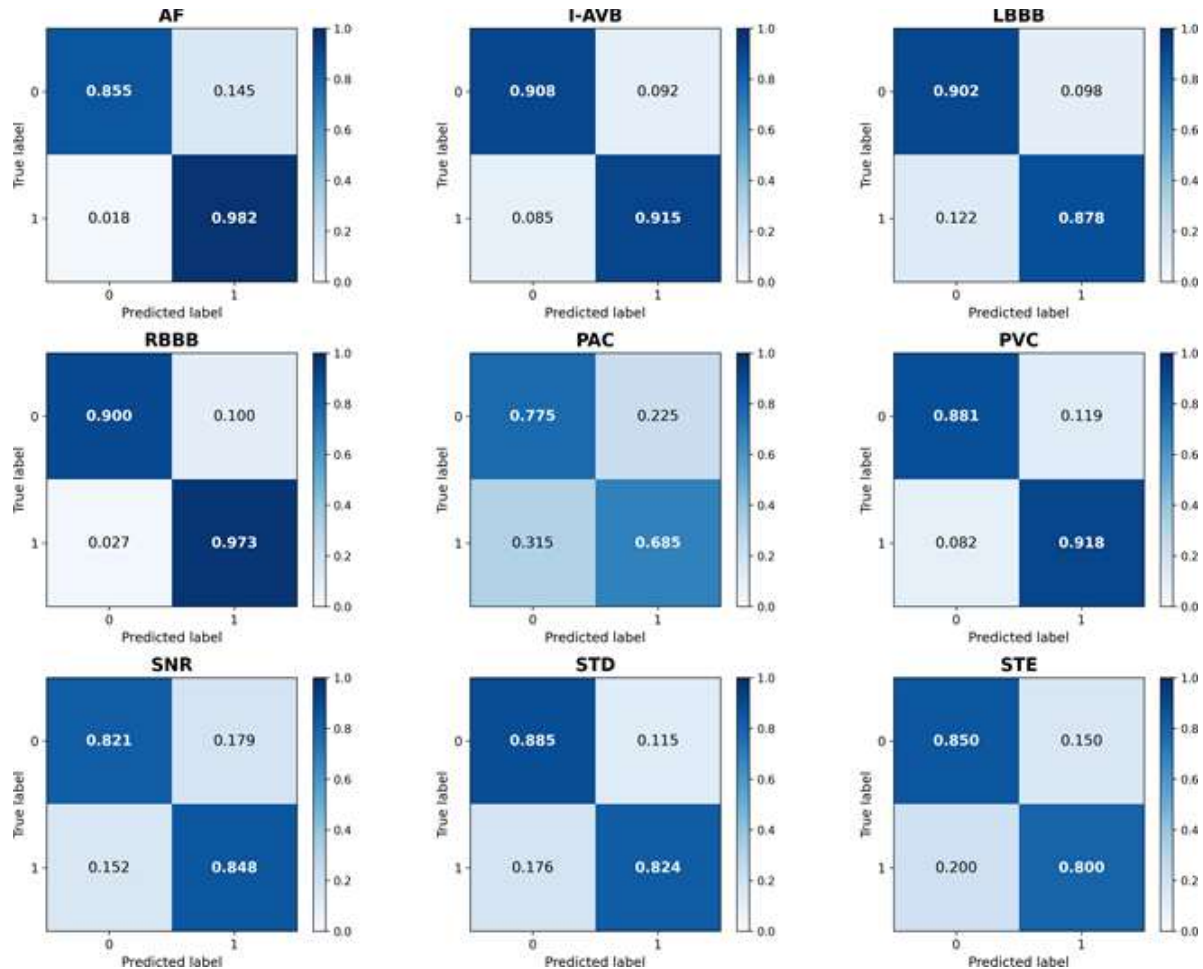


Figure 5. Binary normalized one-vs-rest confusion matrices for nine arrhythmia classes, comparing true labels (rows) to predicted labels (columns).

Table 7. Class-wise classification performance of SMFB-Net on the CPSC 2018 test set.

Class	Precision	Recall	F1-score	Support
AF	0.87	0.98	0.922	220
I-AVB	0.91	0.91	0.912	141
LBBB	0.90	0.88	0.884	41
RBBB	0.91	0.97	0.938	339
PAC	0.75	0.69	0.716	111
PVC	0.88	0.92	0.900	134
SNR	0.83	0.85	0.838	184
STD	0.85	0.82	0.835	165
STE	0.84	0.80	0.816	40
Macro Avg.	0.860	0.869	0.864	1375

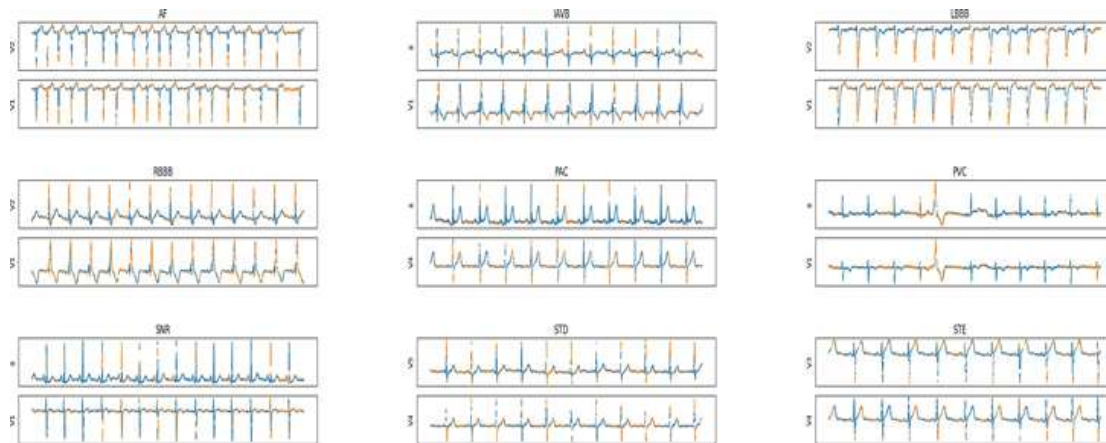


Figure 6. SHAP feature importance visualizations for nine arrhythmias, where each panel highlights the most influential lead for particular arrhythmias.

5.3 Patient-Level Interpretability

The effectiveness of the proposed lead selection module (LSM) was further validated in patient-level SHAP analysis³³ for all nine arrhythmia classes. SHAP provides representative ECG samples alongside the orange areas, yielding the waveform segments with the highest prediction effectiveness for SMFB-Net (Figure 6). These prominent areas are well-aligned with the clinically relevant ECG characteristics, and the quantitative performance metrics are shown in Table 7.

In AF cases, the irregular QRS intervals and lack of clear P waves are observed in leads V1 and V2, both of which are clinically related to atrial fibrillation. This explains the better AF performance of SMFB-Net (F1-score: 0.922). Furthermore, in I-AVB, long PR intervals in leads II and V1 are a sign of delayed atrioventricular conduction, which supports the high F1-score (0.912). For LBBB and RBBB, SHAP focuses mainly on widened QRS regions, as well as abnormal ventricular depolarization trends in leads V1 and V2, which represent clinically identified bundle branch blocks. These relevant regions have high classification accuracy, as shown in F1-score of 0.938 of RBBB. In PAC the previously highlighted areas are features of premature atrial activity and early P-wave morphology. Because PAC waveforms often vary in appearance from the normal sinus rhythms, the model performed relatively worse (F1-score: 0.716). On the other hand, the PVC sample shows marked widening of ventricular complexes as well as compensatory pauses, hence, it has a good PVC classification (F1-score: 0.900). In the case of SNR, the regions highlighted mainly refer to stable sinus rhythm morphology: regular P waves, narrow QRS complexes, which are indicators of the model's ability to distinguish between normal cardiac activity and some pathological rhythms. For the STD and STE, SHAP is focused on clinically relevant ST-segment depression and elevation regions, both of which relate to myocardial ischemia and abnormal ventricular repolarization. The results further clarify that the STD (F1-score: 0.835) and STE (F1-score: 0.816) have better classification results. Despite the above, a strong clinical correspondence is detected for most arrhythmia classes in SHAP analysis even then, however, certain limitations should be considered. Some of the highlighted regions cover wider waveform regions than sharply spaced intervals, as is often the case for deep learning explainability techniques. There are also such differences in the waveform and morphological characteristics (such as PAC) due to the low interpretability consistency. This patient-level SHAP analysis demonstrated that SMFB-Net specifically attends to the clinically appropriate ECG-waveform regions, which reinforces that both the proposed lead selection method and the general interpretability of the framework is effective.

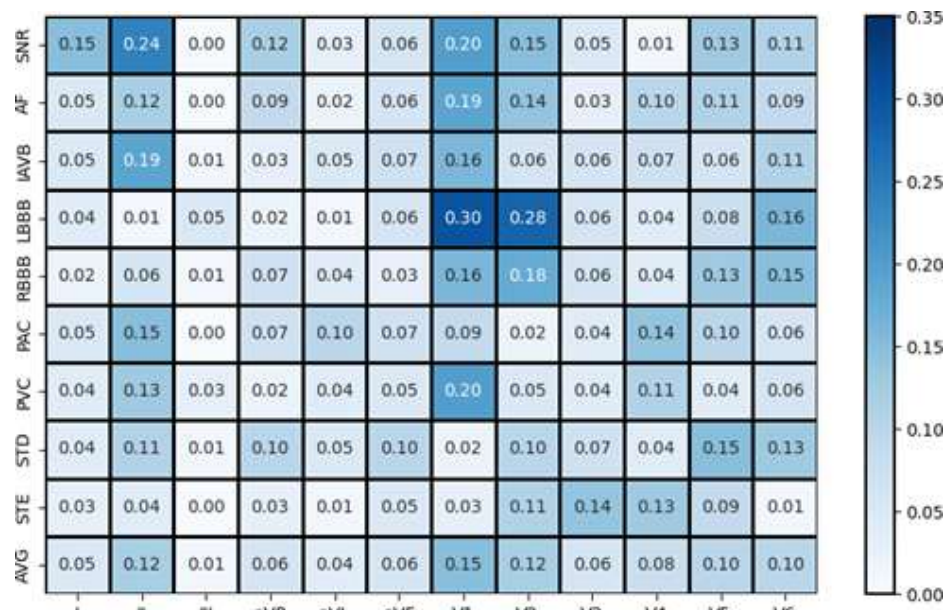


Figure 7. Normalized SHAP heatmap showing lead contributions to nine arrhythmia classes

5.4 Population-Level Interpretability

The population-level SHAP heatmap showing lead-wise contributions across all nine arrhythmia classes is shown in Figure 7. The role of Leads V1 and V2 was found consistently to be significant for AF, LBBB, RBBB, and PVC thus confirming their ability in capturing ventricular depolarization deviations as well as atrial activity deviations. In response to this, Lead II was highly significant for I-AVB and PAC detection, with clinical relevance in both PR-interval and P-wave morphology investigation. In comparison, leads V3 and V4 indicated higher interest for STE detection due to anterior ST-segment elevation patterns while V5 and V6 contributed more prominently in STD analysis. In conclusion, the heatmap shows that precordial leads V1, V2, V5, and V6, leading with lead II, were very informative leads across all the arrhythmia classes, while leads III and aVL, represented smaller contributions. Similar to the patient-level SHAP visualizations in Figure 6, confirming the efficacy of proposed lead selection module on clinically important ECG region identification. The strong lead relevance detected for AF, RBBB, and I-AVB was also associated with the better classification data found in Table 7.

5.5 Comparative Performance Review

Table 8 shows the comparative F1-score of representative deep learning approaches for multi-class arrhythmia classification on the CPSC 2018 dataset. The traditional CNN-based architectures and LSTM-based architectures have presented promising performance for classification of arrhythmia. The CNN+Expert Features model¹³ had an average F1-score of 0.810, although with relatively poor STE and PAC detection performance. For a similar purpose, the CNN+LSTM architecture¹² also improved temporal feature learning and achieved very good RBBB classification performance, although the overall average F1-score remained comparatively lower at 0.772. The CNN+LSTM+Attention model¹⁴ further enhanced classification accuracy by combining attention mechanisms with sequential modeling, obtaining an average F1-score of 0.812. Furthermore, multi-scale CNN architectures improved arrhythmia detection by capturing ECG waveform features at multiple temporal resolutions. Attention-based and recurrent hybrid architectures further improved classification performance through temporal dependency modeling. The CNN+Attention+Bi-GRU and CNN+BiGRU+Attention frameworks^{16,18} used bidirectional recurrent learning for capturing long-range temporal relationships within ECG signals, gaining average F1-scores of 0.835 and 0.837, respectively. In the same manner, the Multi-Lead Branch CNN+BiGRU+Attention architecture¹⁹ employed 12 parallel lead branches for multi-lead feature extraction and obtained an average F1-score of 0.843. These results revealed significant improvements when using attention-guided multi-lead learning as the method for arrhythmia classification compared with traditional convolutional architectures.

Among all compared approaches, SMFB-Net achieved the highest average F1-score of 0.864 and showed the best results on STE and I-AVB classification. This enhanced STE classification performance can be ascribed to the more adaptive lead selection mechanism introduced to the ST-segment based in the prior classes that places more emphasis on clinically relevant precordial leads like V3 and V4, as they are where ST-segment elevation features are prevalent. Moreover, the proposed multi-scale feature extraction strategy coupled with Bi-GRU-based attention modeling will increase the efficiency of capturing temporal

variation and inter-lead interaction patterns dependent on region and time series. The results show that SMFB-Net is successful for robust multi-class arrhythmia classification.

Table 8. Comparison of SMFB-Net with representative deep learning-based methods on multi-class arrhythmia classification on the CPSC 2018 dataset.

Framework	F1 Score for All Classes									Average
	RBBB	AF	I-AVB	PVC	LBBB	STD	SNR	STE	PAC	
CNN+LSTM ¹²	0.922	0.900	0.809	0.832	0.874	0.762	0.753	0.462	0.638	0.772
CNN+Expert Features ¹³	0.910	0.910	0.870	0.820	0.870	0.810	0.820	0.600	0.630	0.810
CNN+LSTM+Attention ¹⁴	0.930	0.920	0.850	0.861	0.872	0.789	0.789	0.556	0.736	0.812
Multi-Scaled CNN+Attention ¹⁵	0.930	0.900	0.860	0.880	0.870	0.800	0.820	0.620	0.780	0.828
CNN+Attention+Bi-GRU ¹⁶	0.926	0.940	0.866	0.865	0.862	0.812	0.819	0.640	0.789	0.835
CNN+Bi-GRU+Attention ¹⁸	0.910	0.930	0.875	0.869	0.880	0.811	0.801	0.624	0.826	0.837
Multi-Lead Branch CNN+Bi-GRU+Attention ¹⁹	0.931	0.930	0.885	0.859	0.862	0.820	0.850	0.646	0.801	0.843
SMFB-Net	0.938	0.922	0.912	0.900	0.884	0.835	0.838	0.816	0.716	0.864

5.6 Evaluation of Computational Efficiency Against Representative Deep Learning Backbones

Importantly, parameter count and FLOPs also serve as the parameter complexity criteria. Even though SMFB-Net keeps parameter sizes relatively small via lightweight convolutional kernels and common branch structures, it applies many-branch temporal operations and Bi-GRU-based sequential feature modeling for processing large multimodal sequence of multi-lead ECG during long duration processing. The overall FLOPs remain high, even though the number of parameters is limited. All computational comparisons were made using similar 12-lead ECG input dimensions and inference settings to have fair performance across architectures. Table 9 compares the architecture of deep learning backbone architectures for detection of an arrhythmia based on the most utilized methods. Comparing computation with the performance of conventional convolutional, and recurrent architectures, the comparison presents the approximate FLOPs, parameter count, inference time, and scalability properties to assess the SMFB-Net's efficacy compared with standard convolutional and recurrent models. Traditional deep convolutional design models such as VGG1634, Inception25, and DenseNet20136 can extract significant features using deeper network architectures and larger parameter spaces. However, these designs involve relatively high computational resources, memory consumption, and inference latency, that might limit their applicability in both real-time and resource-constrained ECG monitoring scenarios. VGG16 especially has the largest computation load (more than 10G FLOPs) and 138M parameters, while DenseNet201 and Inception have larger footprints, 20–25M parameters, inference times are generally between 7–12ms. Hybrid recurrent frameworks like CNN+LSTM12 and VFPred37 enhance temporal sequence modeling with convolutional and recurrent learning modules. However, recurrent processing adds further sequential overhead which incurs high inference time of 6–18ms (though Temporal Dependency Learning is a much preferred architecture), and its scalability may be limited while performing long-duration multi-lead ECG analysis and wearable deployment. When compared to these architectures, MobileNetV235 exhibits enhanced computational efficiency by using lightweight depth-wise separable convolutions over nearly 3.5M parameters, of approximately 1.5G FLOPs, with inference times ranging between about 4–6ms. Nevertheless, those lightweight backbone architectures are mainly optimized for processing image-based features and fail to effectively capture the adaptive inter-lead spatial and temporal relationships of multi-lead ECG signals.

Although MobileNetV2 achieves strong efficiency through depth-wise separable convolutions, the proposed SMFB-Net + LSM features ECG-specific optimizations that reduce inference latency. For a single point of consideration, the lead selection module has the capability to dynamically switch ON only clinically informative leads, thereby avoiding costly branch processing such that it would be costly in duplicate of full 12-lead processing. Second, the architecture leverages lightweight one-dimensional convolution that is designed explicitly for analyzing temporal ECG signals rather than intensive two-dimensional image feature extraction. Third, the multi-branch design relies on shallow branches with compact feature structures and parameters in a minimal space, whereas the Bi-GRU module works on fewer feature representations formed after adaptive lead filtering. All these points reduce the computational burden and make inference latency cheaper albeit with multi-branch attention based processing.

SMFB-Net addresses these shortcomings by light multi-branch feature extraction that combines adaptive lead usage and Bi-GRU-based attention modeling with lightweight filtering. Although we employed multiple-lead ECG processing, our proposed framework remains only over 522K parameters in the total 12-lead configuration compared to classical convolutional and recurrent backbone architectures. Moreover, the lead selection module (LSM) dynamically activates clinically relevant lead groups in real-time. This reduces the computational cost from 2.849G FLOPs to approximately 1.968G FLOPs. The

inference latency is also reduced from 5–7ms to nearly 3–5ms. The relatively higher FLOP count is mainly associated with temporal sequence processing rather than parameter-intensive fully connected operations. The reduced inference latency of SMFB-Net + LSM is due mostly to adaptive lead activation, lightweight 1D branch operations targeted to ECG signal processing. Therefore this optimized lead utilization strategy eliminates redundant branch computation thereby improving our computational scalability, making the framework better suited to use for real-time and resource-bound ECG analysis. The reported computational values are intended to provide relative architectural comparisons under comparable ECG input settings rather than exact implementation-level benchmarking across all studies.

Table 9. Approximate computational characteristics of representative deep learning backbone architectures under comparable ECG input settings

Architecture	Representative FLOPs	Approx. Parameters	Inference Time	Scalability
CNN+LSTM12	~4.5G	~2–5M	12–18ms	Moderate
VGG1634	>10G	~138M	15–25ms	Low
MobileNetV235	~1.5G	~3.5M	4–6ms	High
Inception25	~3.8G	~20–25M	7–10ms	Moderate
DenseNet20136	~5.2G	~20M	8–12ms	Moderate
VFPred37	~2.8G	~1–5M	6–9ms	Moderate
SMFB-Net (Full 12-lead)	2.849G	522K	5–7ms	High
SMFB-Net + LSM	1.968G	510K	3–5ms	Very High

The percentage reduction in FLOPs was computed using Equation 15.

$$\text{FLOP Reduction} = \frac{F_{\text{full}} - F_{\text{LSM}}}{F_{\text{full}}} \times 100 \quad (15)$$

where F_{full} and F_{LSM} denote the FLOPs of the full multi-branch configuration and the adaptive lead selection configuration, respectively. Using $F_{\text{full}} = 2.849\text{G}$ and $F_{\text{LSM}} = 1.968\text{G}$, the resulting computational reduction becomes:

$$\frac{2.849 - 1.968}{2.849} \times 100 = 30.92\%$$

Overall, the proposed SMFB-Net achieves a favorable balance between classification performance, inference speed, scalability, and model complexity. The combination of lightweight branch fusion and adaptive lead utilization makes the framework more suitable for real-time, wearable, edge-based, and resource-constrained arrhythmia monitoring applications compared with conventional deep convolutional and recurrent backbone architectures.

6. Conclusion

In this paper, SMFB-Net, a lightweight multi-branch deep-learning framework was proposed for automatic arrhythmia detection from the CPSC 2018 dataset that uses 12-lead ECG signals. The architecture, under consideration, combines adaptive lead selection, multi-scale spatial and temporal feature extraction. The lead selection module (LSM) we reported above was capable to emphasize those clinically relevant ECG leads such as V1, V2, II, V5, and V6, given the high weight of these in patient-level and population-based study analysis. SHAP-based interpretability further confirmed that our model is based on clinically informative waveform regions among various arrhythmia classes, making our overall framework more interpretable. Successive experiments showed a robust classification of nine arrhythmia classes. SMFB-Net obtained an average F1 score of 0.864 among representative Deep Learning approaches. In particular, the model provided significant performance in dealing with complex arrhythmia classes including I-AVB (0.912), STD (0.835) and STE (0.816), indicating the potential of adaptive lead usage and multi-scale temporal feature learning on subtle ECG abnormality detection. In addition to performing a well-inclined classification, but also efficient LSM computation. The lead selection algorithm resulted in approximately 30.92% lower computational complexity, with lower inference latency, which is well suited for both resource-limited ECG monitoring as well real-time settings. To sum up, SMFB-Net proposed presents an interpretable and computationally efficient framework for automated arrhythmia classification. Future work will include patient-independent validation and evaluation on additional external clinical datasets and real-time deployment in continuous cardiac monitoring applications.

Acknowledgements

The authors would like to thank the organizers of the China Physiological Signal Challenge (CPSC) 2018 for providing the publicly available 12-lead ECG dataset used in this study.

Ethics approval

This study uses publicly available data and does not involve any direct experiments on human or animal subjects; therefore, ethics approval is not applicable.

Consent to participate

The dataset used in this study is publicly available and anonymized; hence, informed consent is not required.

Author Contributions

A.S. conceived the study, designed and implemented the model, conducted the experimental work, analyzed the results, and wrote the manuscript. M.I. contributed to the conceptualization, methodology, and supervision and assisted in experimental design and validation, as well as critically revising the manuscript. All authors reviewed and approved the final manuscript.

Data availability

The CPSC 2018 12-lead ECG dataset used in this study is publicly available through the China Physiological Signal Challenge (CPSC) 2018 repository. Additional processed data supporting the findings of this study are available from the corresponding author upon reasonable request.

Code availability

The code used to develop and evaluate the proposed model is available from the corresponding author upon reasonable request.

Competing interests

The authors declare no competing interests.

References

1. World Health Organization, "Cardiovascular diseases (CVDs)," WHO, 2019. [Online]. Available: [https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-\(cvds\)](https://www.who.int/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds))
2. R. Mehra, "Global public health problem of sudden cardiac death," *Journal of Electrocardiology*, vol. 40, no. 6, Suppl., pp. S118–S122, 2007. doi: 10.1016/j.jelectrocard.2007.06.023.
3. S. Karpagachelvi, M. Arthanari, and M. Sivakumar, "ECG Feature Extraction Techniques – A Survey Approach," arXiv preprint arXiv:1005.0957, May 2010.
4. P. S. Addison, "Wavelet transforms and the ECG: a review," *Physiological Measurement*, vol. 26, no. 5, pp. R155–R199, 2005.
5. Stanislaw Osowski, Linh Tran Hoai, and Tomasz Markiewicz, "Support Vector Machine-Based Expert System for Reliable Heartbeat Recognition," **IEEE Transactions on Biomedical Engineering**, vol. 51, no. 4, April 2004, pp. 582–589. doi:10.1109/TBME.2004.824138
6. P. Rajpurkar, A. Y. Hannun, M. Haghpahani, C. Bourn, and A. Y. Ng, "Cardiologist-Level Arrhythmia Detection with Convolutional Neural Networks," arXiv preprint arXiv:1707.01836, 2017.
7. A. Y. Hannun, P. Rajpurkar, M. Haghpahani, G. H. Tison, C. Bourn, M. P. Turakhia, and A. Y. Ng, "Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network," *Nature Medicine*, vol. 25, no. 1, pp. 65–69, 2019. doi: 10.1038/s41591-018-0268-3.
8. O. Yildirim, "A novel wavelet sequence based on deep bidirectional LSTM network model for ECG signal classification," *Computers in Biology and Medicine*, vol. 96, pp. 189–202, 2018.
9. U. R. Acharya, H. Fujita, O. S. Lih, M. Adam, J. H. Tan, and C. K. Chua, "Automated detection of arrhythmias using different intervals of tachycardia ECG segments with convolutional neural network," *Information Sciences*, vol. 405, pp. 81–90, 2017.
10. S. Kiranyaz, T. Ince, and M. Gabbouj, "Real-time patient-specific ECG classification by 1-D convolutional neural networks," *IEEE Transactions on Biomedical Engineering*, vol. 63, no. 3, pp. 664–675, 2016.
11. N. Alamatsaz, L. S. Tabatabaei, M. Yazdchi, H. Payan, N. Alamatsaz, and F. Nasimi, "A lightweight hybrid CNN-LSTM explainable model for ECG-based arrhythmia detection," *Biomedical Signal Processing and Control*, vol. 90, Art. no. 105884, 2024. doi: 10.1016/j.bspc.2023.105884.

11. Q. Yao, J. Liu, and Y. Li, "Time-incremental convolutional neural network for arrhythmia detection using variable-length ECGs," *Computers in Biology and Medicine*, vol. 102, pp. 122–132, 2018
12. F. Liu, Z. Zhou, J. Shang, and J. Cao, "Multiple-feature-branch convolutional neural network for myocardial infarction diagnosis using electrocardiogram," *Biomedical Signal Processing and Control*, vol. 45, pp. 22–32, 2018. doi: 10.1016/j.bspc.2018.05.018.
13. Q. Yao, R. Wang, X. Fan, J. Liu, and Y. Li, "Multi-class arrhythmia detection from 12-lead varied-length ECG using attention-based time-incremental convolutional neural network," *Information Fusion*, vol. 53, pp. 174–182, 2020.
14. R. Wang, J. Fan, and Y. Li, "Deep multi-scale fusion neural network for multi-class arrhythmia detection," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 9, pp. 2461–2472, 2020. doi: 10.1109/JBHI.2020.2975368.
15. J. Zhang, A. Liu, M. Gao, X. Chen, X. Zhang, and X. Chen, "ECG-based multi-class arrhythmia detection using spatio-temporal attention-based convolutional recurrent neural network," *Artificial Intelligence in Medicine*, vol. 106, Art. no. 101856, 2020. doi: 10.1016/j.artmed.2020.101856.
16. S. L. Oh, E. Y. K. Ng, R. San Tan, and U. R. Acharya, "Automated diagnosis of arrhythmia using combination of CNN and LSTM techniques with variable length heart beats," *Computers in Biology and Medicine*, vol. 102, pp. 278–287, 2018. doi: 10.1016/j.compbiomed.2018.06.002.
17. T.-M. Chen, C.-H. Huang, E. S. C. Shih, Y.-F. Hu, and M.-J. Hwang, "Detection and classification of cardiac arrhythmias by a challenge-best deep learning neural network model," *iScience*, vol. 23, no. 3, Art. no. 100886, Mar. 2020, doi: 10.1016/j.isci.2020.100886.
18. J. Zhang, D. Liang, A. Liu, M. Gao, X. Chen, X. Zhang, and X. Chen, "MLBF-Net: A multi-lead-branch fusion network for multi-class arrhythmia classification using 12-lead ECG," *IEEE Journal of Translational Engineering in Health and Medicine*, vol. 9, pp. 1–11, 2021. doi: 10.1109/JTEHM.2021.3064675.
19. Ansari Y, Mourad O, Qaraqe K, Serpedin E. Deep learning for ECG Arrhythmia detection and classification: an overview of progress for period 2017-2023. *Front Physiol.* 2023 Sep 15;14:1246746. doi: 10.3389/fphys.2023.1246746. PMID: 37791347; PMCID: PMC10542398.
- A. H. Ribeiro et al., "Automatic diagnosis of the 12-lead ECG using a deep neural network," *Nature Communications*, vol. 11, no. 1, p. 1760, 2020. doi: 10.1038/s41467-020-15432-4.
20. Vaswani, N. Shazeer, N. Parmar, et al., "Attention is all you need," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2017, pp. 5998–6008.
21. V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Machine Learning (ICML)*, Haifa, Israel, Jun. 2010, pp. 807–814.
22. S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Machine Learning (ICML)*, 2015, pp. 448–456.
23. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 1–9, 2015. doi: 10.1109/CVPR.2015.7298594.
24. K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder–decoder for statistical machine translation," in *Proc. Conf. Empirical Methods in Natural Language Processing (EMNLP)*, Doha, Qatar, Oct. 2014, pp. 1724–1734.
25. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.*, vol. 15, no. 1, pp. 1929–1958, Jun. 2014.
26. M. Lin, Q. Chen, and S. Yan, "Network in network," in *Proc. Int. Conf. Learning Representations (ICLR)*, 2014. arXiv:1312.4400.
27. F. Liu, C. Liu, L. Zhao, X. Zhang, X. Wu, X. Xu, Y. Liu, J. Ma, S. Wei, Z. He, and E. N. Yin Kwee, "An open access database for evaluating the algorithms of electrocardiogram rhythm and morphology abnormality detection," *Journal of Medical Imaging and Health Informatics*, vol. 8, no. 7, pp. 1368–1373, 2018.
28. X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in *Proc. 13th Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, Sardinia, Italy, May 2010, vol. 9, pp. 249–256.
29. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," in *Proc. Int. Conf. on Learning Representations (ICLR)*, 2015. arXiv:1412.6980.
30. T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, 2017, pp. 2980–2988. doi: 10.1109/ICCV.2017.324.

31. S. Lundberg and S.-I. Lee, "A unified approach to interpreting model predictions," in *Advances in Neural Information Processing Systems (NeurIPS)*, pp. 4765–4774, 2017.
32. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, San Diego, CA, USA, May 2015. [Online]. Available: [arXiv:1409.1556](https://arxiv.org/abs/1409.1556)
33. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 4510–4520.
34. G. Huang, Z. Liu, L. van der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Honolulu, HI, USA, Jul. 2017, pp. 2261–2269.
35. N. Ibtehaz, M. S. Rahman, and M. S. Rahman, "VFPred: A fusion of signal processing and machine learning techniques in detecting ventricular fibrillation from ECG signals," *Biomed. Signal Process. Control*, vol. 49, pp. 349–359, Mar. 2019.
36. Geetanjali, & Singh, A. R. (2025). Precision and smart farming: An approach to efficient resource management in agriculture. *Journal of Artificial Intelligence in Agriculture and Applied Research*, 1(1), 25–33. <https://agrpjournals.com/journals/journal/index.php/jaiaar/article/view/6>