mathematics

MDPI

Article

Efficient and Interpretable ECG Abnormality Detection via a Lightweight DSCR-BiGRU-Attention Network with Demographic Fusion

Kan Luo [1,2,*], Longying Huang [1,2], Haixin He [2], Yu Chen [1,2], Lu You [1,2], Siluo Chen [1,2], Jian Chen [1] and Chengyu Liu [3,*]

[1] School of Electronic, Electrical Engineering and Physics, Fujian University of Technology, Fuzhou 350118, China; 2231908046@smail.fjut.edu.cn (L.H.); 2220104050@smail.fjut.edu.cn (Y.C.); 2231908007@smail.fjut.edu.cn (L.Y.); 2231908009@smail.fjut.edu.cn (S.C.); jianchen@fjut.edu.cn (J.C.)
[2] Fuzhou Industrial Integration Automation Technology Innovation Center, Fuzhou 350118, China; 2221308017@smail.fjut.edu.cn
[3] State Key Laboratory of Digital Medical Engineering, School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China
* Correspondence: luokan@fjut.edu.cn (K.L.); chengyu@seu.edu.cn (C.L.)

Abstract

Deep learning has advanced automated electrocardiogram (ECG) interpretation, yet many models are computationally expensive, opaque, and overlook demographic factors. We propose DBA-ASFNet, a lightweight network that combines depthwise-separable convolutional residual blocks with a BiGRU and an attention mechanism to extract rich spatiotemporal features from 12-lead ECGs while maintaining low computational requirements. The Age-and-Sex Fusion (ASF) module integrates demographic information without enlarging the model, enabling personalized predictions. On the PTB-XL and CPSC2018 datasets, DBA-ASFNet achieves competitive multi-label performance with only ~0.03 million parameters and ~6.43 MFLOPs per inference. Real-time testing on a Raspberry Pi 5 achieved an average inference latency of ~2 ms, supporting deployment on resource-limited devices. Shapley additive explanations (SHAP) analysis shows that the model focuses on clinically meaningful ECG patterns and appropriately incorporates demographic factors, enhancing transparency. These results suggest that DBA-ASFNet is suited for accurate, efficient, and interpretable ECG analysis.

Keywords: electrocardiogram; lightweight model; DSCR; BiGRU; attention mechanism; demographic features; SHAP; interpretability

MSC: 68T45

check for updates

Academic Editor: Vince Grolmusz

Received: 3 October 2025
Revised: 19 November 2025
Accepted: 2 December 2025
Published: 3 December 2025

Citation: Luo, K.; Huang, L.; He, H.; Chen, Y.; You, L.; Chen, S.; Chen, J.; Liu, C. Efficient and Interpretable ECG Abnormality Detection via a Lightweight DSCR-BiGRU-Attention Network with Demographic Fusion. *Mathematics* **2025**, *13*, 3882. https://doi.org/10.3390/math13233882

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Cardiovascular disease (CVD) remains the leading cause of death worldwide [1]. In 2022, it accounted for an estimated 19.8 million deaths [2]. The electrocardiogram (ECG) is a fundamental, noninvasive, and cost-effective diagnostic tool that is widely used in clinical practice and for daily heart monitoring [3], generating over 300 million recordings globally each year [4]. However, the enormous volume of data generated far exceeds clinicians' capacity for manual interpretation, creating an urgent need for automated, efficient methods to detect cardiac anomalies.

Deep learning (DL) has demonstrated its powerful capabilities in ECG analysis, excelling in tasks such as arrhythmia detection, myocardial ischemia identification, and heart failure classification [5]. DL models can efficiently extract intricate temporal and morphological patterns directly from raw ECG signals, often outperforming traditional methods that rely on handcrafted features [6]. Due to these advantages, DL techniques have become the mainstream technology in ECG interpretation. However, the deployment of current DL-based ECG models in real-world clinical and ambulatory settings faces several practical challenges. Many existing models achieve high performance by using very deep architectures with large parameter counts, which results in high computational costs that limit their use in environments with limited resources, such as wearable devices or mobile monitors [7]. Additionally, most models operate as "black boxes" with limited transparency, which makes it difficult for clinicians to understand or trust the diagnostic process [8]. Furthermore, existing models typically neglect patient-specific factors such as age and sex, which are known to influence ECG characteristics and disease risk [9]. These limitations highlight the need for ECG analysis methods that not only maintain high diagnostic accuracy but are also lightweight enough for deployment on resource-constrained platforms, while being more personalized and interpretable to support clinical decision-making.

To address these challenges, we propose DBA-ASFNet, a lightweight model for ECG anomaly detection that integrates demographic features and offers clear interpretability. DBA-ASFNet combines depthwise separable convolutional residual (DSCR) blocks, bidirectional gated recurrent units (BiGRUs), and attention mechanisms to efficiently extract spatiotemporal features from multi-lead ECG signals with low computational cost. The model introduces an Age-and-Sex Fusion (ASF) module to seamlessly integrate patient demographic information, thereby improving generalization and personalized anomaly detection. Extensive evaluation on the PTB-XL and CPSC2018 datasets shows that DBA-ASFNet delivers competitive multi-label classification performance and can be deployed on resource-constrained devices, such as the Raspberry Pi 5, for real-time ECG analysis. Using the SHAP framework, we can examine the rationale behind the model's decisions and quantify the contributions of both ECG and demographic features to each prediction.

The main contributions of this work are as follows:

1. We designed a lightweight network architecture that combines DSCR, BiGRU, and attention modules, significantly reducing parameter count and computational cost while maintaining high classification accuracy.
2. We introduced the ASF module, which integrates age and sex information to enhance the recognition of ECG abnormalities and improve personalized diagnostic performance.
3. We validated DBA-ASFNet on the PTB-XL and CPSC2018 datasets, achieving competitive multi-label classification performance. Its real-time inference capability was further confirmed on a Raspberry Pi 5.
4. We used the SHAP framework to interpret model predictions at both the patient and cross-patient levels, revealing the impact of demographic features and further enhancing model transparency and clinical applicability.

2. Related Works

Recent studies emphasize the progress of ECG modeling in terms of lightweighting, demographic fusion, and interpretability. Table 1 summarizes representative works from these areas.

Table 1. Overview of representative ECG deep-learning studies.

Hotspots	Model	Data Set	Method	Metrics & Results
Lightweighting	Lightx3ECG [10]	CPSC2018 Chapman	3-branch 1D-CNN attention fusion pruning	9 classes: Pre 82.09, Rec 78.62, F1 80.04, Acc 96.28 4 classes: Pre 97.36, Rec 97.03, F1 97.18, Acc 98.73
	CNN-LSTM [11]	MIT-BIH & LTAF	Parallel shallow 1D-CNN 1-layer LSTM	9 classes: Acc 98.24, Sen 86.1, Spe 97.5
	CNN-SE-LSTM [12]	Chapman	Knowledge distillation	4 classes: Pre 82.09, Rec 78.62, F1 80.04, Acc: 96.28
	MTECG [13]	Private dataset	MAE-style pretraining	28 classes: F1 76.5
Demographic fusion	TransECG [14]	MIT-BIH	age, sex	2 classes: Acc 89.9, Pre 90.0, F1 89.9 5 classes: Acc 89.9, Pre 90.1, F1 89.9
	CNN [15]	Private dataset	age, sex	-
	AlexNet [16]	PTB-XL	HRV, age, sex	2 classes: Sen 92.25, Auc 96.29, Spe 92.03
Interpretability	CNN-GRU [17]	MIT-BIH	B-LIME	-
	HAN-ECG [18]	MIT-BIH AFIB	Attention mechanism	-
	xECGArch [19]	PTB-XL CPSC 2018 Chapman	SHAP	-

2.1. Lightweighting

Researchers have explored various strategies to develop lightweight ECG deep learning models without compromising performance. For example, LightX3ECG employs three leads and a pruned lightweight 1D-CNN with attention, while maintaining accuracy comparable to larger 12-lead models [10]. Other efforts include using hybrid CNN–LSTM architectures for joint spatial–temporal modeling [11], using knowledge distillation to compress large models into lightweight versions [12], and using transformer-based architectures for enhancing long-range feature learning [13]. These studies collectively emphasize the importance of designing models that maintain diagnostic performance while remaining computationally efficient for real-time deployment.

2.2. Demographic Fusion

Beyond model architecture, integrating patient demographic data has emerged as a promising approach to personalized ECG analysis. Clinical evidence confirms that demographic factors such as age and sex impact cardiac electrical activity and disease manifestation. Recent models have accounted for these factors. For example, transformer-based architectures have successfully incorporated age and sex features into ECG modeling [14]. Similarly, deep learning-based "physiological age" estimations from ECGs have shown

clinical utility, highlighting the predictive relevance of demographic factors [15]. Furthermore, combining patient metadata with ECG data enhances diagnostics [16]. Developing lightweight models that effectively fuse demographic information with ECG features is therefore of practical significance.

2.3. Interpretability

Interpretability is also increasingly recognized as essential for clinical acceptance. Clinicians are more likely to trust and adopt AI systems if they can understand the rationale behind each prediction [20]. Although methods such as B-LIME [17] and attention mechanism [18] offer some explanatory power, they typically provide only local or coarse-grained insight, and fall short of delivering a comprehensive understanding of model decision-making. Recently, the Shapley additive explanations (SHAP) framework has emerged as a widely recognized and robust approach to model interpretability. It provides consistent and fair attributions of model outputs to input features, offering both individualized and global explanations [21]. SHAP has been successfully applied to highlight clinically meaningful ECG features and align model decisions with medical knowledge [19]. Together, these works highlight the importance of developing interpretable ECG models that can provide transparent and clinically aligned explanations.

3. Materials and Methods

3.1. Datasets

3.1.1. PTB-XL Dataset

PTB-XL is one of the largest and most richly annotated ECG datasets [22]. It contains 21,837 12-lead ECG records from 18,885 patients ranging in age from 0 to 95 years old. The sex distribution is balanced (52% male and 48% female). Each record is 10 s long, and has a sampling rate of 100 and 500 Hz. The labels follow the standard communication protocol ECG standard [23], and include 71 categories, such as 44 diagnostics, 19 form, and 12 rhythm labels. The diagnostic labels are divided into 5 superclasses and 23 subclasses, which facilitates multi-level anomaly modeling.

3.1.2. CPSC2018 Dataset

The China physiological signal challenge (CPSC) released the CPSC2018 dataset, which contains 6877 12-lead ECG records collected from patients of different regions and age groups [24]. The dataset includes expert annotations for nine common arrhythmias. The signals are sampled at 500 Hz, and the records range from six to 60 s in duration. While most records have a single label, some include two or three labels, reflecting the complexity of coexisting anomalies encountered in real clinical settings. This dataset is widely used for multi-label ECG classification and algorithm evaluation [25].

3.2. Data Preprocessing

To enhance stability and efficiency, we applied a fifth-order Butterworth high-pass filter at 0.5 Hz to remove baseline wander [26], followed by z-score normalization [27] across all 12 leads. To standardize sampling and reduce computation, we used PTB-XL at 100 Hz and downsampled CPSC2018 from 500 Hz to 100 Hz. All ECG records were truncated or zero-padded to 10 s [28]. Patient demographic information of age and sex was structurally encoded to improve sensitivity to inter-individual differences. As shown in Figure 1, age was linearly scaled by dividing it by 10, rather than using min-max normalization. This simple rescaling brings age into a numerically stable range that is well suited for gradient-based optimization, while avoiding the sensitivity of min-max scaling to dataset-specific extremes and potential distribution shifts between training and

deployment cohorts [29]. Sex was one-hot encoded (male = [1, 0] and female = [0, 1]), and missing demographic entries were handled using binary mask indicators (1 = present, 0 = missing) to ensure robustness.

Age	Sex
26	Male
75	Female
30	-
-	Male
-	Female
-	-

$Mask_{age}$	X_{age}	$Mask_{sex}$	X_{sex}	
1	2.6	1	1	0
1	7.5	1	0	1
1	3.0	0	0	0
0	0	1	1	0
0	0	1	0	1
0	0	0	0	0

Figure 1. Example of structured demographic feature encoding.

3.3. DBA-ASFNet

DBA-ASFNet is an efficient model that combines multi-scale temporal feature extraction, sequence modeling, attention weighting, and a demographic fusion. Figure 2a shows the DBA-ASFNet architecture. It has two cooperative branches. For ECG feature extraction, the DBA backbone begins with a 1D convolution and then a series of depthwise separable residual (DSCR) blocks. These blocks capture the multi-scale waveform morphology. Next, a bidirectional gated recurrent unit (BiGRU) and a lightweight additive attention layer aggregate informative time steps into a fixed-length vector. In parallel, the demographic branch uses a fully connected (FC) block to encode age and sex, forming a compact demographic embedding. Finally, the two resulting vectors are then concatenated and fed into a final FC layer to produce multi-label predictions.

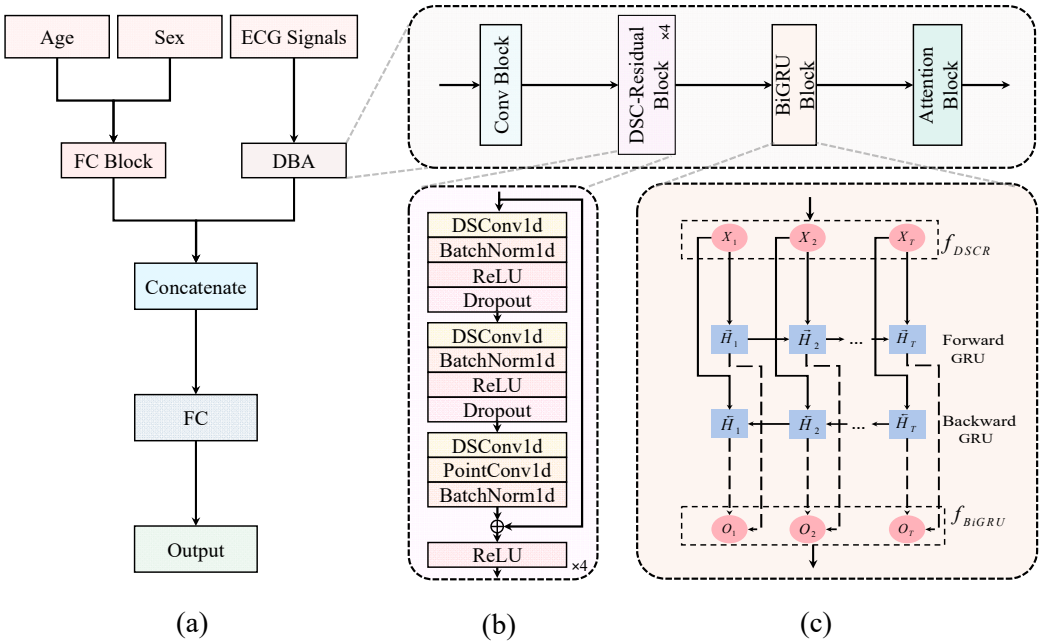

Figure 2. DBA-ASF Network. (**a**) Network architecture; (**b**) DSCR module; (**c**) BiGRU module. The model and code are available at https://github.com/Talitaaa1/DBA-ASFNet (accessed on 23 November 2025).

The core components of the DBA backbone network are the DSCR, BiGRU, and the attention module. As shown in Figure 2b, each DSCR block contains three depthwise-pointwise convolution stacks with varying kernel sizes, which are designed to capture multi-scale ECG morphologies. The first two stacks use BatchNorm, ReLU, and dropout, while the third stack uses only BatchNorm. The output of the third stack is added to the

block input via a residual shortcut to stabilize training, followed by a ReLU activation. Stride is applied in the depthwise convolution when downsampling is required. This multi-scale residual design reduces computation while maintaining stable training and a strong feature representation. The BiGRU (Figure 2c) takes the feature sequence $X = \{x_t\}_{t=1}^{T}$ generated by the preceding DSCR stack as input. Let $GRU(\cdot)$ represent the gated recurrent unit operation [30]. At each time step t, the forward state $\overrightarrow{h}_t$ and backward state $\overleftarrow{h}_t$ are updated, respectively.

$$\overrightarrow{h}_t = GRU(x_t, \overrightarrow{h}_{t-1}) \tag{1}$$

$$\overleftarrow{h}_t = GRU(x_t, \overleftarrow{h}_{t-1}) \tag{2}$$

The BiGRU output at time step t is $h_t = [\overrightarrow{h}_t, \overleftarrow{h}_t]$, and over T time steps, the module produces the context-aware sequence $H = \{h_t\}_{t=1}^{T}$. Then, a lightweight temporal attention layer adapts the contribution of each time step before classification, and attention scores are computed as follows:

$$u_t = \tanh(Wh_t + b) \tag{3}$$

$$a_t = \frac{\exp(u_t^\top u_w)}{\sum_{t=1}^{T} \exp(u_t^\top u_w) + \varepsilon} \tag{4}$$

the parameters W, u_w, and b are all randomly initialized trainable parameters. A small regularization term $\varepsilon = 1 \times 10^{-8}$ is introduced to prevent division by zero and ensure numerical stability. The attention-pooled feature is computed as follows:

$$z = \sum_{t=1}^{T} a_t h_t \tag{5}$$

Let C be the number of final prediction classes and let X, H, z, and g denote the outputs of the DSCR, BiGRU, attention layer, and ASF, respectively. The kernel sizes in the DSCR blocks are set to 7, 5, and 3. More detailed parameters and configurations of DBA-ASFNet are provided in Table 2. The data processing logic of the model is as follows:

(1) DBA branch: 10-s, 12-lead ECGs pass through an initial convolutional block. Then, four stacked DSCR blocks produce a compact temporal feature map X. A BiGRU with 32 hidden units per direction processes sequence X to obtain H. Finally, H is attention-pooled into a sixty-four-dimensional vector $z \in \mathbb{R}^{64\times1}$

(2) Demographic branch: Age is normalized (age/10) and sex is one-hot encoded into a two-dimensional vector. Each is paired with a presence mask, resulting in a combined five-dimensional input vector. A fully connected block maps this vector to an embedding vector $g \in \mathbb{R}^{4\times1}$.

(3) Fusion and classification: The vectors z and g are concatenated $[z; g]$ and fed into a final FC layer with a sigmoid activation function to generate the C multi-label predictions.

3.4. SHAP Interpretability Analysis

To interpret the decision-making of DBA-ASFNet in ECG classification, we used the SHAP method to analyze model predictions at the individual (patient) and global (cross-patient) levels. SHAP is a model-agnostic, post hoc interpretability approach based on game theory that uses Shapley values to quantify the marginal contribution of each feature to the model output [21]. SHAP provides a unified explanation framework for both individual predictions and overall feature importance. Since SHAP does not affect model inference efficiency but increases the computational cost of the interpretation, it is recommended for analyses where real-time constraints are not a factor.

Table 2. Configuration and parameters of the DBA-ASFNet Network.

Stage	DBA Backbone	Output	ASF	Output
Input	12-lead ECG	12 × 1000	Masked [Age, Sex]	5 × 1
Conv	Conv (7, 32), stride 2	32 × 500	FC (5 → 4)	g: 4 × 1
DSCR 1	DSConv (7/5/3, 32), stride 1/1/1	32 × 500		
DSCR 2	3 × {DSConv (7/5/3, 32), stride 1/1/2}	X:32 × 63		
BiGRU	BiGRU (hidden = 32)	H:63 × 64		
Attention	Equation (5)	z:64 × 1		
Concatenate	[z; g]	68 × 1	—	—
Classifier	FC (68 → C), Sigmoid	C × 1	—	—

Figure 3 illustrates the SHAP-based interpretation of the model at the patient and cross-patient levels. In our analysis, we treat ECG signals from all leads, along with age and sex, as explanatory features. By applying SHAP, we can quantify the contributions of different ECG leads and specific temporal segments, as well as age and sex, to predicting different ECG abnormalities. This allows clinicians to align the model's reasoning with their own clinical expertise and determine if the predicted abnormalities are supported by physiologically meaningful evidence. Formally, let F represent the set of all input features. The SHAP value for a specific feature i is defined as follows:

$$\Phi_i = \sum_{S \subseteq F \setminus \{i\}} \frac{|S|!(|F|-|S|-1)!}{|F|!} \left[f_{S\cup\{i\}}\left(x_{S\cup\{i\}}\right) - f_S(x_S) \right] \quad (6)$$

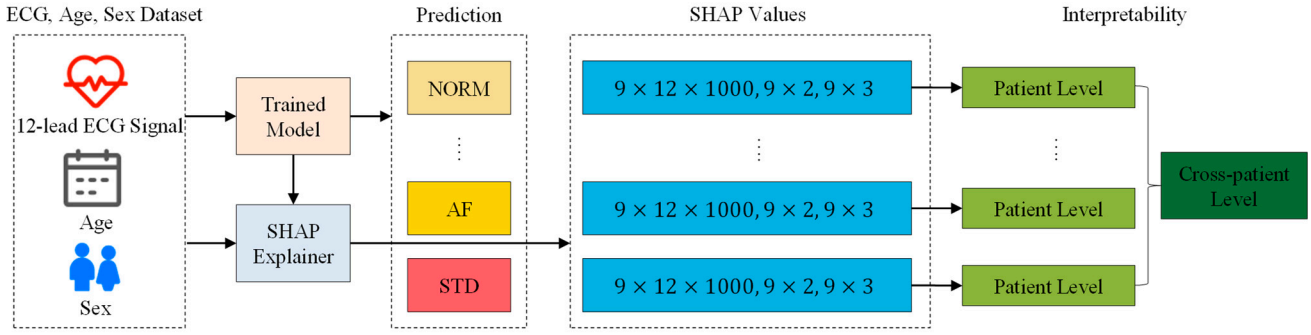

Figure 3. SHAP-based interpretation of the model at both patient and cross-patient levels.

Here, S denotes any subset of features excluding i, $f_S(x_S)$ is the model output considering only features in S, and $f_{S\cup\{i\}}\left(x_{S\cup\{i\}}\right)$ is the output after adding feature i to S. Larger SHAP values indicate greater importance of the feature.

At the individual level, interpretation involves examining how ECG and structured demographic features contribute to each diagnostic category for a given sample. For example, the ECG features are $X_{ecg} \in \mathbb{R}^{12\times1000}$, structured demographic features are $X_{age} \in \mathbb{R}^2$ and $X_{sex} \in \mathbb{R}^3$, and the model output is $y \in \mathbb{R}^{1\times9}$. They serve as inputs fed into the SHAP explainer, and then the explainer produces Shapley value matrices of $\Phi_{ecg} \in \mathbb{R}^{9\times12\times1000}$ $\Phi_{age} \in \mathbb{R}^{9\times2}$ and $\Phi_{sex} \in \mathbb{R}^{9\times3}$ to quantify the contribution of each feature to the predictions for that individual sample.

At the global level, SHAP values are aggregated across all samples to analyze feature contributions. For N samples ECG with SHAP value matrix Φ, the contribution $C_{j,k}$ of lead j to diagnostic class k is calculated as follows:

$$C_{j,k} = \sum_{n=1}^{N} \Phi_{j,k}^{(n)} \tag{7}$$

The normalized contribution rate $r_{j,k}$ of lead j to class k is defined as follows:

$$r_{j,k} = \frac{\left|C_{j,k}\right|}{\sum_{j=1}^{12}\left|C_{j,k}\right|} \tag{8}$$

Similarly, demographic features (age and sex) are analyzed at a global level. Given that age is divided into g groups and sex into m groups (where m = 2 for male and female), the average SHAP contribution of each age or sex group to diagnostic category k is calculated as follows:

$$A_{g,k} = \frac{1}{P}\sum_{n\in P} \Phi_{age,k}^{(n)} \tag{9}$$

$$G_{m,k} = \frac{1}{Q}\sum_{n\in Q} \Phi_{sex,k}^{(n)} \tag{10}$$

For class k, $\Phi_{age,k}$ and $\Phi_{sex,k}$ denote the SHAP matrices of age and sex, respectively. P and Q represent the number of samples in the g-th age group and the m-th sex group, respectively. The global average contribution rates of ECG signals and demographic features reveal the importance of each feature and highlight potential interactions that influence model predictions.

4. Experiments and Results

4.1. Evaluation Metrics and Settings

To comprehensively address label imbalance and robustly evaluate multi-label classification performance, we used the Macro Area Under the Curve (Macro AUC) evaluation metric, as recommended for the PTB-XL dataset [31]. Macro AUC is calculated by first computing the AUC for each label, and then averaging these values across all labels. The Macro AUC is defined as follows:

$$Macro\ AUC = \frac{1}{L}\sum_{i=1}^{L} AUC_i \tag{11}$$

where L represents the total number of labels, and AUC_i is the area under the ROC curve for label i. This metric mitigates biases toward labels with more abundant samples.

We conducted experiments in Python 3.9 using PyTorch 1.7.1 on a workstation with an Intel Xeon Silver 4216 CPU, 128 GB RAM, and an NVIDIA RTX 3090 GPU. For the PTB-XL dataset, we used the official fold partition (Folds 1–10). Specifically, Folds 1–8 were used for training, Fold 9 for validation, and Fold 10 for testing, which closely corresponds to an 8:1:1 split. Although the PTB-XL dataset contains multiple ECG recordings per patient (21,837 ECGs from 18,885 patients), the predefined folds are constructed at the patient level. This ensures that all ECGs from the same patient are assigned to the same fold. Therefore, no patient appears in more than one subset, and our split is fully consistent with the official benchmark and prior studies [32,33]. A similar 8:1:1 grouping strategy was applied to the CPSC2018 dataset, ensuring that no patients overlap between the training, validation, and test sets. We optimized the model with the binary cross-entropy (BCE) [34] loss function and the Adam optimizer, setting the initial learning rate to 0.001 and the batch size to 64. The model was trained for 100 cycles with a batch size of 64.

4.2. Comparative Experiments

To validate the comprehensive advantages of our proposed DBA-ASFNet model in terms of performance and efficiency, we compared it to several state-of-the-art (SOTA) models on two multi-label ECG datasets (PTB-XL and CPSC2018). The selected baseline models included FCN_wang [35], ResNet1d_wang [35], InceptionTime [36], XResNet1d101 [37] and MobileNetV3 [38], ATI-CNN [39], Chen et al. [40], and DCRR-Net [41]. Table 3 summarizes the comparison results across multiple metrics, including Macro AUC, model size, and computational complexity. In addition, we conducted experiments on the CPSC2018 dataset using ECG signals with different sampling rates to evaluate the impact of sampling rate variation on model performance. The optimal and second-best results are indicated by bold and underlined text, respectively.

Table 3. Comparison of Macro AUC (%) performance, number of parameters, and computational complexity of different models on the PTB-XL and CPSC2018 datasets.

Method	PTB-XL						CPSC2018	Param (M)	Flops (M)
	All	Diag.	Sub-Diag.	Super-Diag.	Form	Rhythm			
Fcn_wang [35] *	89.06	91.72	90.95	91.26	79.89	87.46	89.97	0.28	276.33
Resnet1d_wang [35] *	<u>91.15</u>	**92.98**	<u>92.74</u>	91.67	83.33	88.41	93.28	0.29	33.40
InceptionTime [36] *	90.78	91.96	**92.78**	<u>91.88</u>	**85.47**	92.15	93.22	0.47	475.52
Xresnet1d101 [37] *	90.83	90.53	91.27	89.76	79.72	92.96	92.36	1.53	140.64
MobileNetV3 [38] *	89.96	87.58	88.28	90.52	76.63	94.27	93.04	1.48	<u>20.67</u>
ATI-CNN [39]	89.29	91.11	89.89	**92.11**	82.71	**96.84**	94.65	5.00	287.34
Chen et al. [40]	-	85.05	88.02	91.30	-	-	-	3.75	-
DCRR-Net [41]	-	-	-	-	-	93.60	-	<u>0.17</u>	-
Proposed (100 Hz)	**92.48**	<u>92.13</u>	90.32	91.66	<u>83.91</u>	<u>95.88</u>	<u>94.92</u>	**0.03**	**6.43**
Proposed (250 Hz)	**-**	-	-	-	-	-	**95.03**	0.03	16.07
Proposed (500 Hz)	**-**	-	-	-	-	-	94.18	0.03	32.12

* The models are open source; they were evaluated in our local environment.

DBA-ASFNet demonstrated solid robustness and superior efficiency across multiple ECG classification tasks. On the PTB-XL dataset, it achieved the best Macro AUC score of 92.48% in the "all" task, and ranked second in the "diag.", "form", and "rhythm" tasks, with scores of 92.13%, 83.91%, and 95.88%, respectively. Although it did not lead in the "sub-diag." (90.32%) and "super-diag." (91.66%) categories, its performance remained close to the best models, with differences of only 2.46% and 0.45%, respectively. On the CPSC2018 dataset, DBA-ASFNet achieved the highest Macro AUC score of 94.92%. Overall, its performance was comparable to or better than that of current SOTA models. Furthermore, DBA-ASFNet offers a lightweight design. As shown in Table 3, large-scale models such as XResnet1d101 (1.53 M parameters), ATI-CNN (5.00 M), and Chen et al. (3.75 M) require substantial computational resources. Intermediate-scale models, including Resnet1d_wang (0.29 M), FCN_wang (0.28 M), DCRR-Net (0.17 M), and InceptionTime (0.47 M), strike a balance between complexity and performance. In contrast, DBA-ASFNet achieves competitive accuracy with only 0.03 M parameters. It also requires only 6.43 MFLOPs per inference, compared to 475.52, 287.34, and 276.33 MFLOPs/inference for InceptionTime, ATI-CNN, and FCN_wang, respectively. Evaluation across different sampling rates shows that the 100 Hz configuration achieves a Macro-AUC of 94.92% with 0.03 M parameters and only 6.43 MFLOPs, providing the best balance between accuracy and efficiency. Increasing the sampling rate to 250 Hz yields a slightly higher Macro-AUC of 95.03%, but at more than double the computational cost. At 500 Hz, the model achieves a Macro-AUC of 94.18% with 32.12 MFLOPs, offering no performance gain despite the substantially higher complexity. These findings indicate that a 100 Hz sampling rate effectively preserves diagnostic information while significantly reducing computation. Overall, the results show that the proposed DBA-ASFNet has good classification performance, a small parameter size, and a

low computational cost. This makes it highly suitable for deployment in mobile, embedded, and other resource-constrained environments.

4.3. Ablation Experiments

Ablation experiments were conducted on the PTB-XL and CPSC2018 datasets to evaluate the impact of DSCR block depth and GRU/BiGRU selection on Macro AUC and computational efficiency, as well as to evaluate the effectiveness of the ASF module.

Table 4 presents the impact of DSCR block depth and the choice of GRU versus BiGRU units on model performance. The highest Macro AUC on the comprehensive "all" task is achieved with four DSCR blocks paired with BiGRU (92.48%), followed by two blocks (91.83%). A similar trend is observed for the "diag." task and on the CPSC2018 dataset, indicating that four DSCR blocks offer the most balanced trade-off between accuracy and computational cost (~6 MFLOPs). Notably, performance does not increase monotonically with DSCR depth. Very shallow (1–2 blocks) or very deep (5–6 blocks) configurations lead to fluctuations across tasks, suggesting that insufficient depth limits feature extraction, whereas excessive depth may introduce redundancy. Replacing GRU with BiGRU consistently improves performance on major PTB-XL tasks without increasing computational overhead. Comparable gains appear on CPSC2018, where the BiGRU-based model achieves a higher Macro AUC (94.92 vs. 93.66). This improvement is likely due to BiGRU's ability to capture both forward and backward temporal dependencies, enriching contextual representation and enhancing classification accuracy. Overall, a moderate DSCR depth ($\times 4$) combined with BiGRU provides an effective balance between representational capacity and efficiency, forming the backbone of the proposed DBA-ASFNet architecture.

Table 4. Effect of DSCR Block depth and GRU/BiGRU selection on Macro AUC (%) and computational efficiency.

Model	PTB-XL						CPSC2018	Param (M)	Flops (M)
	All	Diag.	Sub-Diag.	Super-Diag.	Form	Rhythm			
DSCR Block $\times 1$ + BiGRU	91.70	91.30	90.99	91.56	**84.89**	93.31	94.38	**0.02**	9.94
DSCR Block $\times 2$ + BiGRU	91.83	91.46	**92.23**	91.16	80.36	94.78	94.35	0.02	7.93
DSCR Block $\times 3$ + BiGRU	91.54	91.16	92.08	91.70	83.35	95.23	94.63	0.03	6.92
DSCR Block $\times 5$ + BiGRU	90.89	91.65	90.89	91.18	83.68	**96.05**	94.33	0.04	6.68
DSCR Block $\times 6$ + BiGRU	91.42	91.92	90.92	91.72	84.10	95.66	94.82	0.04	6.93
DSCR Block $\times 4$ + GRU	91.12	90.90	90.76	**91.73**	80.29	96.00	93.66	0.03	**6.02**
DSCR Block $\times 4$ + BiGRU	**92.48**	**92.13**	90.32	91.66	83.91	95.88	**94.92**	0.03	6.43

Table 5 reports Macro AUC values together with 95% confidence intervals and DeLong p-values. It also shows the number of parameters and megaflops (MFLOPs). For the PTB-XL, adding the ASF module increased the Macro-AUC of the "all" task from 91.54% to 92.48% (95% CI: 90.45–92.58 vs. 91.51–93.30, $p = 0.0058$) and of "diag." task from 91.42% to 92.13% (95% CI: 90.12–92.57 vs. 91.15–93.09, $p = 0.0467$). These results indicate modest yet statistically significant improvements in performance on these two diagnostic tasks. For the remaining PTB-XL task categories ("sub-diag.", "super-diag.", "form", and "rhythm"), the differences between DBA and DBA-ASF were very small ($\leq$0.4%) and not statistically significant ($p > 0.09$). This suggests that the ASF module does not materially alter performance in these settings. On the CPSC2018 dataset, DBA-ASF achieved a slightly higher Macro AUC than DBA (94.92% vs. 94.56%). Although the p-value of 0.081 does not meet the conventional threshold of 0.05 for statistical significance, the results reflect a positive trend in favor of the ASF module. Overall, these results suggest that incorporating structured demographic information through the ASF module yields modest improvements in Macro AUC metrics, particularly for the "all" and "diag." tasks, while maintaining an unchanged model size of 0.03 M parameters and a low computational cost of 6.43 MFLOPs per inference. We therefore consider the ASF to be a positive addition.

However, we acknowledge that its performance benefits are modest and not significant across all task categories.

Table 5. Impact of the ASF module on Macro AUC (%), number of parameters and computational complexity (PTB-XL and CPSC2018 Datasets).

Model	PTB-XL						CPSC2018	Param (M)	Flops (M)
	All	Diag.	Sub-Diag.	Super-Diag.	Form	Rhythm			
DBA	91.54	91.42	90.72	91.32	83.80	95.81	94.56	0.03	6.43
95%CI	90.45–92.58	90.12–92.57	89.18–92.08	90.54–92.12	80.99–86.04	93.63–97.22	92.95–95.74		
DBA + ASF	92.48	92.13	90.32	91.66	83.91	95.88	94.92	0.03	6.43
95%CI	91.51–93.30	91.15–93.09	88.22–92.26	90.82–92.44	81.71–86.42	93.21–97.47	93.69–96.03		
p-values	0.0058	0.0467	0.9886	0.0954	0.3840	0.9546	0.0810	-	-

4.4. Interpretability of Model Output

4.4.1. Patient Individual Level

Figure 4 shows the SHAP-based visualization results for DBA-ASFNet, which illustrate the interpretability of the model for four ECG abnormalities. The highlighted red regions indicate ECG segments with high SHAP values, indicating strong contributions to the model predictions. Specifically, Figure 4a shows premature ventricular contractions (PVCs), which are typically identified by intermittent abnormal waveforms [42]. The SHAP analysis accurately identifies these critical segments, and aligns well with clinical diagnoses. Figure 4b illustrates left bundle branch block (LBBB), which is often recognized clinically by a prominent, deep S-wave in lead V1 [43]. SHAP visualization clearly captures and highlights this feature. Similarly, Figure 4c demonstrates right bundle branch block (RBBB), which is usually identified by an RSR′ QRS complex in lead V1 [44], and precisely identified by SHAP output. Figure 4d represents atrial fibrillation (AF), which is characterized by an absence of P-wave and irregular F-wave oscillations [45]. Again, SHAP effectively highlights these defining patterns.

To quantify the importance of the ECG waveform, age, and sex, we refer to the SHAP value definitions and matrix representations described in Section 3.4. The SHAP matrices Φ_{ecg}, Φ_{age} and Φ_{sex} summed over all entries to obtain the total contribution of each feature type. Then, a natural logarithm transformation was applied. The bars in Figure 4 visualize these log-transformed contributions, and reflect the relative impact of each feature on individual prediction outcomes. The results indicate that ECG features are the dominant influence, followed by age, and sex contributes minimally. These findings are consistent with prior research findings [43]. All results confirm the transparency of the decision-making process and the clear alignment of the proposed DBA-ASFNet with clinical knowledge, reinforcing its diagnostic reliability and applicability.

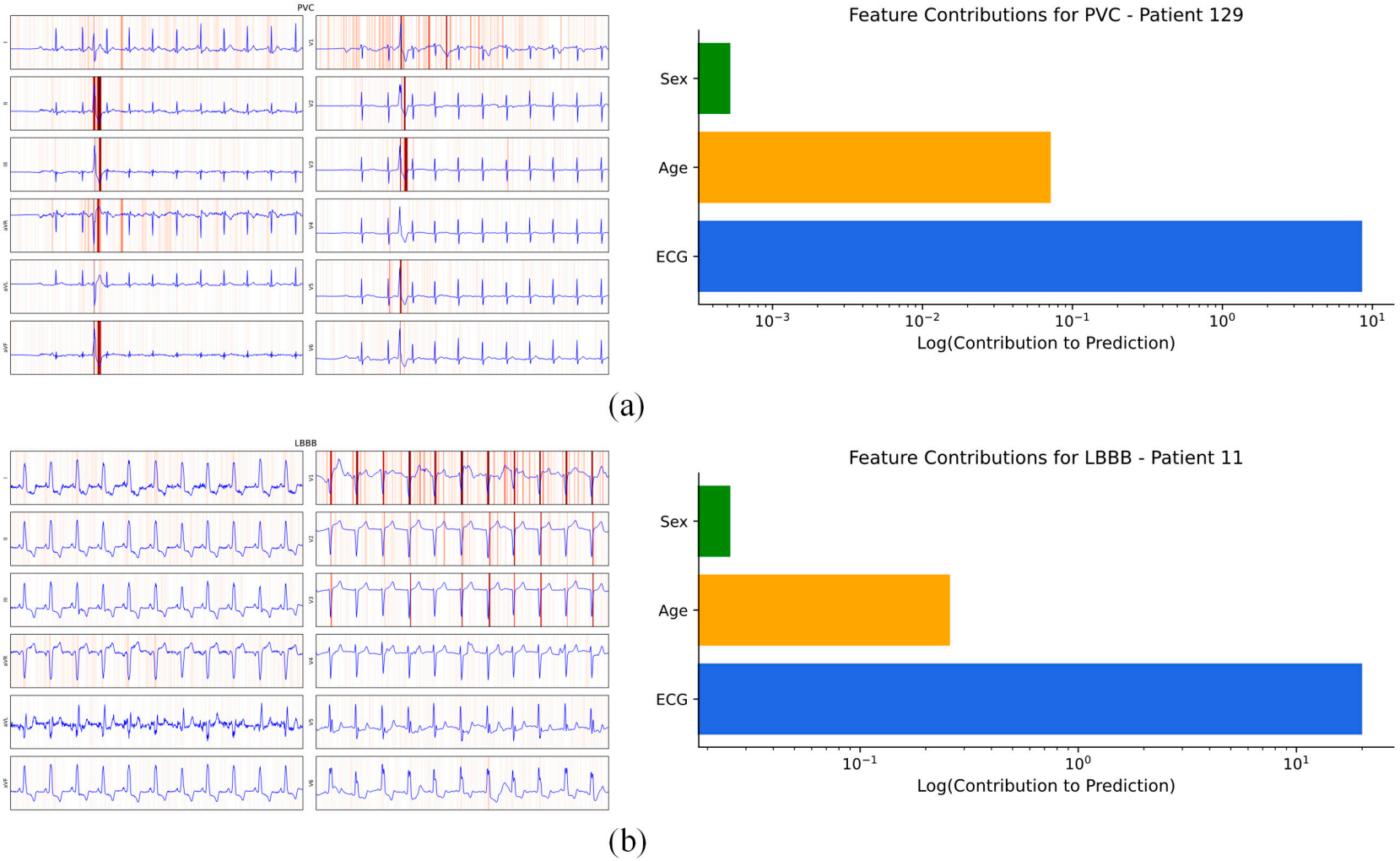

Figure 4. *Cont.*

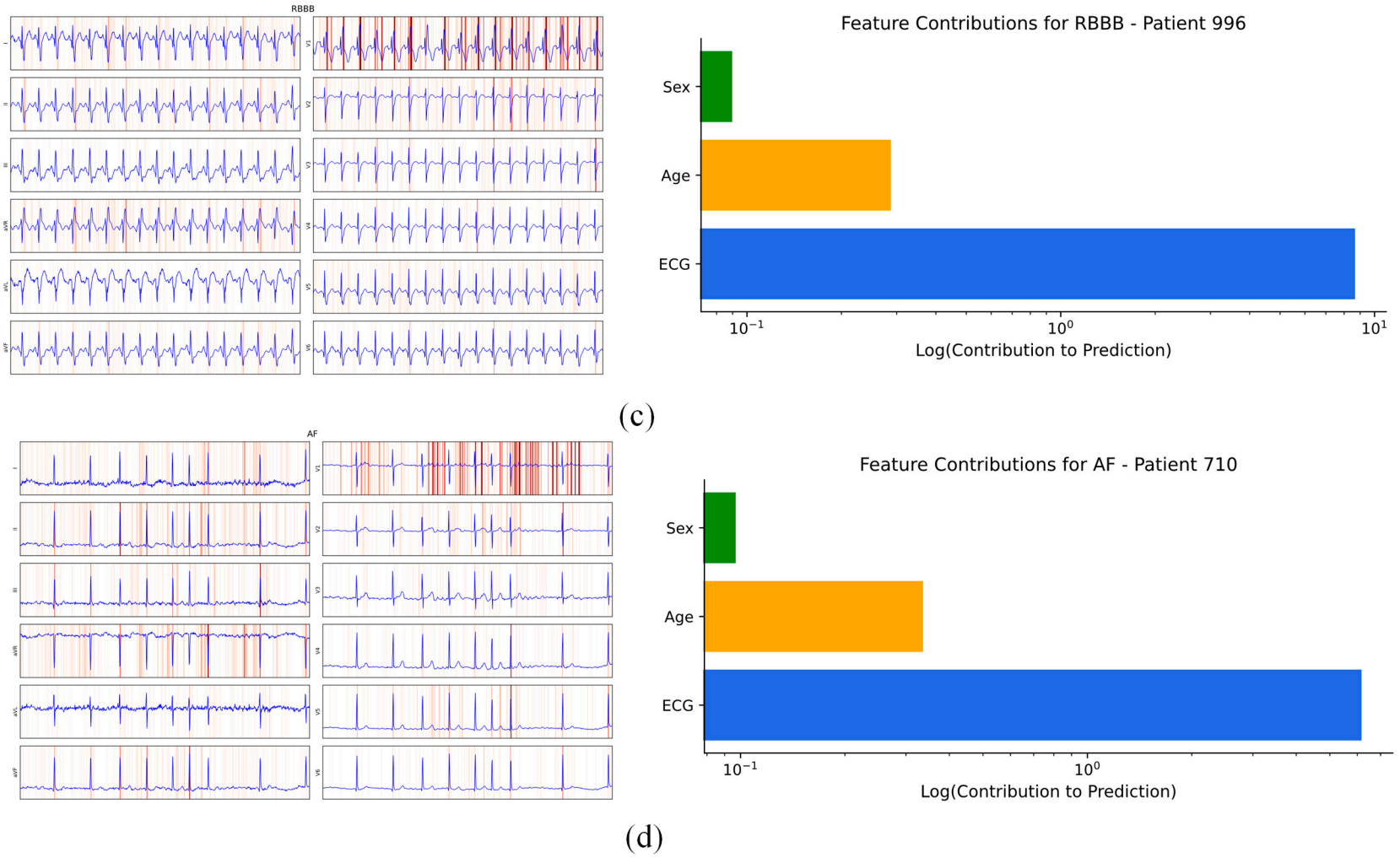

Figure 4. Multi-category recognition ability of DBA-ASFNet on typical ECG abnormalities. (**a**) Early ventricular contraction (PVC), ECG ID 129; (**b**) Left bundle branch block (LBBB), ECG ID 11; (**c**) Right bundle branch block (RBBB), ECG ID 996; (**d**) Atrial fibrillation (AF), ECG ID 710.

4.4.2. Cross-Patient Global Level

Figure 5 illustrates the contribution rates $r_{j,k}$ of the 12-lead ECG to the diagnostic categories. Darker shading indicates greater importance. Lead V1 had the highest mean contribution (0.27), signifying its strong overall diagnostic relevance. Other significant leads included II (0.11), aVR (0.13), I (0.09), V2 (0.10), and V5 (0.08). Specifically, V1 (0.40) and V2 (0.23) were dominant in LBBB diagnosis, aligning with clinical criteria that emphasize QRS morphology. Similarly, for RBBB, leads V1 (0.36) and V2 (0.13) were pivotal, further confirming their clinical relevance [43,44]. For AF, V1 (0.24) and aVR (0.15) contributed significantly, highlighting V1's sensitivity to subtle atrial rhythm disturbances [37]. In addition, V1 showed notable contributions to PAC (0.31) and PVC (0.20), demonstrating its ability to detect abnormal beats [42]. In contrast, leads III (0.04), aVL (0.02), aVF (0.05), and V3/V4/V6 (<0.05) exhibited relatively minor contributions, likely due to redundancy or weaker diagnostic roles [46]. Thus, cross-patient SHAP analysis highlights V1's prominent role in diagnosing multiple abnormalities. This finding aligns well with clinical knowledge and facilitates interpretability and future feature selection.

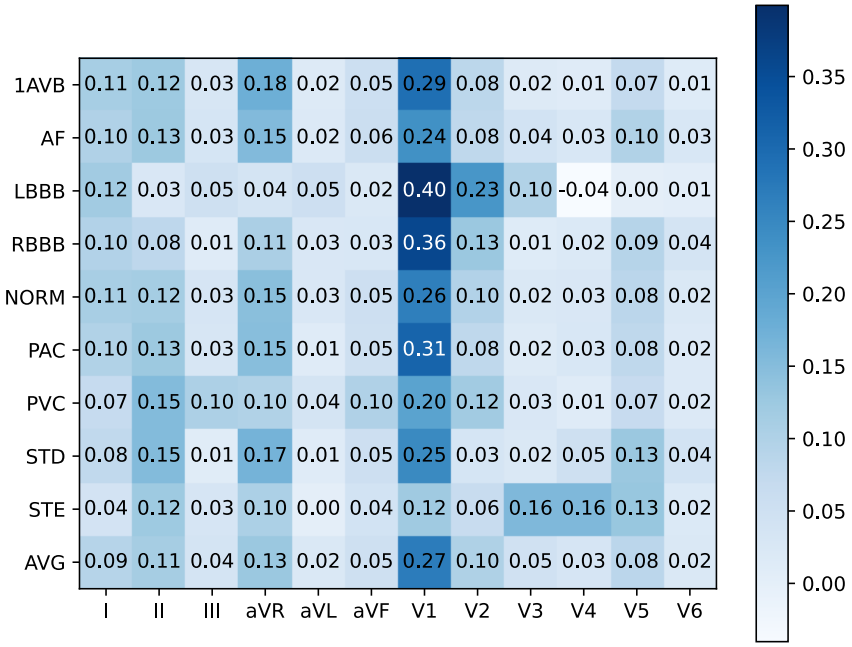

Figure 5. Contribution of 12-lead ECG to diagnostic categories. (1AVB: First-Degree Atrioventricular Block; AF: Atrial Fibrillation; LBBB: Left Bundle Branch Block; RBBB: Right Bundle Branch Block; NORM: Normal ECG; PAC: Premature Atrial Contraction; PVC: Premature Ventricular Contraction; STD: ST-segment Depression; STE: ST-segment Elevation; AVG: Average).

Furthermore, SHAP analysis revealed the influence of structured demographic features. Figure 6 shows the average SHAP contribution values for each age and sex group across diagnostic categories. The blank entry for male category in the STE group represents the absence of corresponding samples. As shown in Figure 6a, sex information contributed positively overall, particularly in the NORM, RBBB, and PAC categories. Clinically, sex differences significantly impact cardiac electrophysiological properties [39]. Studies in the NORM category have shown that females typically have shorter RR intervals, lower heart rate variability, and smoother ECG rhythms. In contrast, males display greater waveform variability [47,48]. Consistent with this, DBA-ASFNet assigned higher SHAP values to male samples in the NORM category (0.221), indicating that sex information improves diagnostic reliability. In RBBB diagnosis, females often present waveform distortions due to the higher placement of leads V1 and V2, both of which are critical for RBBB identifica-

tion [49]. As shown in Figure 6a, the average SHAP value for females in this category was 0.126, indicating that the model relied more heavily on sex information. In contrast, sex contributed minimally to PAC classification, consistent with clinical findings of a low sex correlation [50].

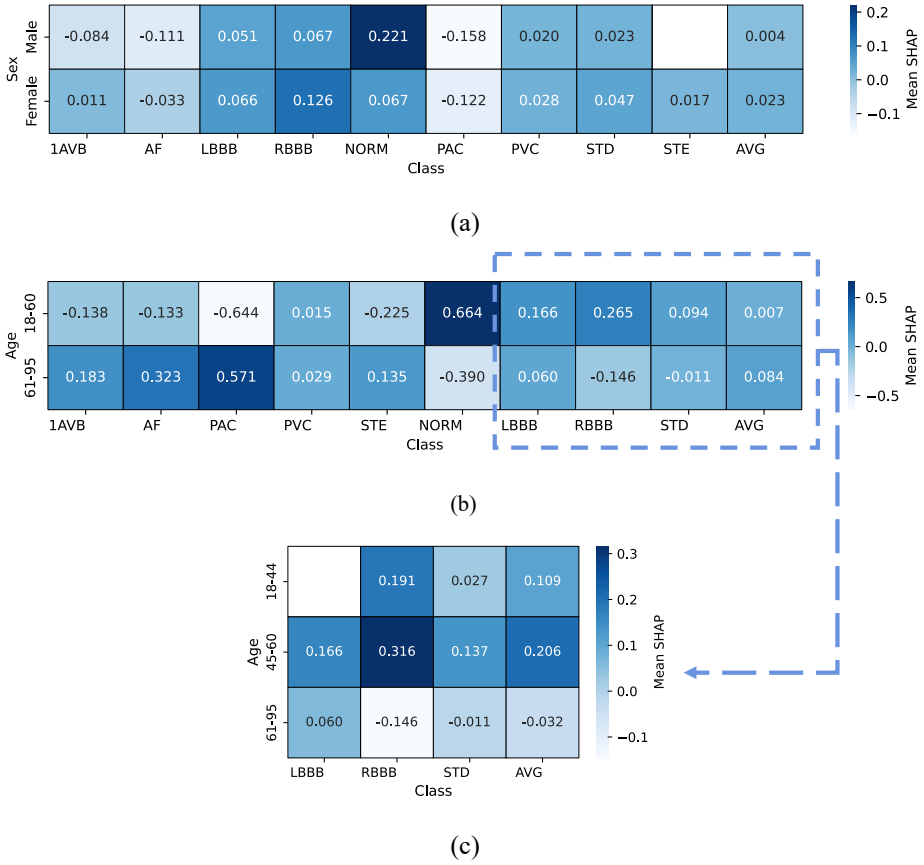

Figure 6. Contribution of sex and age to diagnostic categories. (**a**) Contribution of gender to diagnostic categories; (**b**) Contribution of the two age groups to each diagnostic category; (**c**) Contribution of the three age groups to each diagnostic category.

Age contributions are shown in Figure 6b, excluding the 0–17 group due to insufficient samples. Older patients (ages 61–95) generally had higher age SHAP values across most categories, indicating that the model relies more on age for elderly populations. Older individuals are more susceptible to rhythm-related conditions (e.g., AF, PAC, IAVB) due to conduction system degeneration and electrophysiological disturbances [51–53]. Such abnormalities often exhibit subtle waveform features, so the DBA-ASFNet relies on age information. This is evident by the higher SHAP values observed in the older group. Interestingly, the younger group (ages 18–60) showed higher age SHAP contributions in the LBBB, RBBB, and STD categories, despite these diseases commonly occurring in older patients clinically [51]. Further analysis (Figure 6c) dividing the age into 18–44, 45–60, and 61–95 revealed that the 45–60 subgroup had the highest age contribution. This suggests that the DBA-ASFNet uses age more heavily in middle-aged patients due to atypical, early-stage waveforms. In contrast, elderly waveforms typically exhibit clearer, more diagnostic characteristics, reducing the dependence on age features.

5. Discussion

This study proposes DBA-ASFNet to address three major challenges in automated ECG diagnosis. By combining DSCR blocks for efficient feature extraction with BiGRU for

temporal modeling, DBA-ASFNet reduces the number of parameters to 0.03 M and requires only 6.43 MFLOPs per inference, making it well suited for real-time or continuous monitoring scenarios with constrained hardware resources. Real-time evaluation on a Raspberry Pi 5 further demonstrates an average inference time of 2.27 ± 0.55 ms, confirming its practical feasibility on low-power devices. Given that Raspberry Pi 5 operates at the GFLOPs level, whereas modern wearables (e.g., Apple Watch, Samsung smartwatches) include NPUs operating in the TFLOPs range, the proposed model can be readily deployed on edge or wearable platforms for real-time ECG monitoring (The Raspberry Pi OS image file and testing video are provided at https://github.com/Talitaaa1/DBA-ASFNet) (accessed on 23 November 2025).

Another contribution of this work is the integration of demographic features through the ASF module. By incorporating age and sex, DBA-ASFNet provides more individualized predictions without increasing model size or computational cost. Experimental results indicate modest but statistically significant improvements on the PTB-XL "all" and "diag." tasks, suggesting that even lightweight architectures can benefit from patient-specific information, particularly when ECG morphology alone is ambiguous.

Interpretability is also a key strength of DBA-ASFNet. SHAP analyses show clear alignment between the model's decision patterns and established clinical knowledge. At the individual level, the model highlights clinically relevant ECG segments for each diagnosis, while at the population level it identifies key leads and waveform characteristics consistent with cardiology practice. This transparent decision process enhances the model's clinical trustworthiness and supports its potential for real-world deployment.

Despite the promising results, several limitations remain in our study. First, the model makes errors in challenging cases. As shown in Table 6, misclassifications usually occur in ECGs with overlapping abnormalities, borderline morphology, or mixed rhythms. These patterns are inherently difficult and continue to limit the accuracy of multi-label classification. Second, the model exhibits limited robustness to artifacts and cross-domain shifts. Our preprocessing pipeline relies only on a 0.5 Hz high-pass filter to remove baseline wander, whereas real clinical settings often involve more complex noise and motion artifacts that may degrade performance. Moreover, although DBA-ASFNet is not restricted to the label spaces of PTB-XL and CPSC2018, cross-dataset evaluation revealed a marked decline in performance, with Macro-AUC dropping to 72.25%. These results highlight the importance of improving the model's resilience to noise, device variability, and population differences in future studies. Third, the benefits of the ASF module are inconsistent across all diagnostic tasks. Experiment results show that performance in the "sub-diag." task slightly decreased, and the gains observed in the "super-diag.", "form", and "rhythm" tasks were small and not statistically significant. In addition, the current demographic fusion strategy is intentionally simple, relying only on age and sex. Incorporating richer clinical information and exploring more advanced fusion mechanisms remain important directions for future improvement. Furthermore, while SHAP provides meaningful post hoc explanations by identifying influential features, it primarily addresses contributions to a prediction rather than why certain physiological patterns drive the model's decision. Figure 7 illustrates this distinction by comparing two attribution approaches for an AF example. The gradient $\times$ input map highlights broader temporal regions, particularly around QRS complexes and their surrounding intervals capturing the irregular R-R interval and the P waves changes caused by AF. This produces a smooth, rhythm-level attribution pattern that aligns well with clinical reasoning. In contrast, the SHAP $\times$ input map yields a much finer, point-wise distribution of importance values. Although SHAP also identifies key contributions near QRS complexes, its highly localized attribution makes it more challenging to interpret the underlying rhythm structure. Together, these observations

suggest that while SHAP is effective for feature-level explanations, gradient-based methods provide better insight into how the model perceives temporal patterns within the ECG. Combining SHAP with gradient-based or intermediate-layer visualization techniques [45] may therefore offer a more comprehensive understanding of the model's internal reasoning. Finally, we did not examine how the model would be integrated into an actual clinical workflow. Its real-time deployment in clinical environments will be an important direction for future work.

Table 6. Typical misclassification cases selected from different diagnostic tasks.

Case	Record No.	Labels	Predict
1	PTB-XL (all), ECG ID 9	'NORM', 'SR'	'ABQRS', 'SR'
2	PTB-XL (diag.), ECG ID 299	'ISC_', 'LAO/LAE','LVH'	'IMI','ISC_', 'LVH'
3	PTB-XL (sub-diag.), ECG ID 218	'NORM', '_AVB'	'NORM'
4	PTB-XL (super-diag.), ECG ID 38	'NORM'	'MI'
5	PTB-XL (form), ECG ID 63	'ABQRS'	'ABQRS', 'PVC'
6	PTB-XL (rhythm), ECG ID 347	'AFLT', 'SR'	'SR'
7	CPSC2018, ECG ID 11	'CLBBB'	'AFIB','CLBBB'

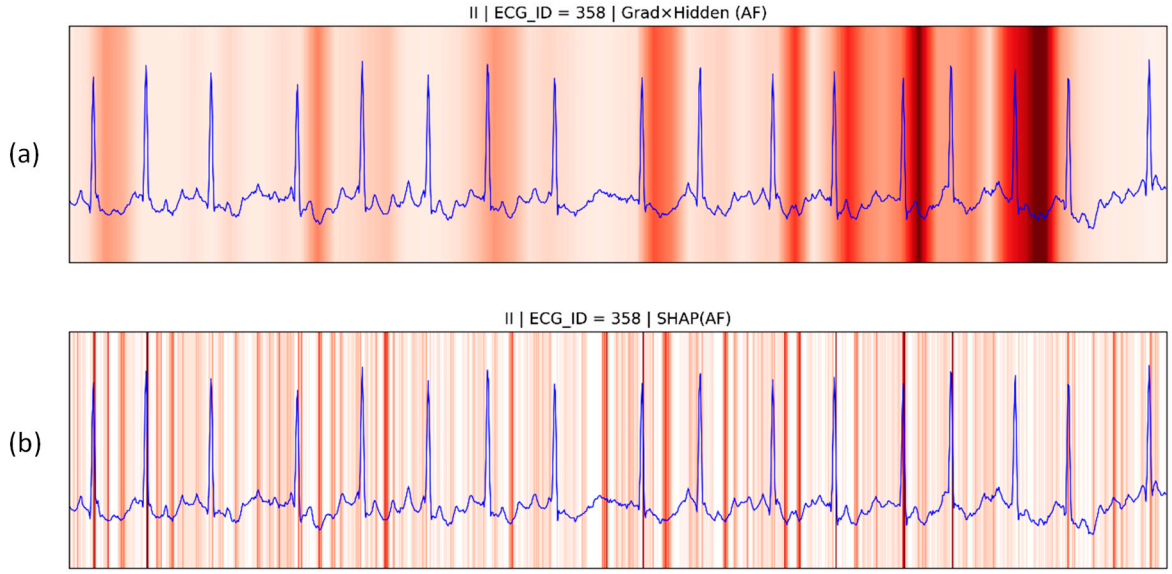

Figure 7. Comparison of (**a**) the gradient × input attribution map of the final BiGRU hidden states and (**b**) the SHAP × input attribution map for an AF example (ECG ID 358, Lead II, PTB-XL).

6. Conclusions

We present DBA-ASFNet, a deep neural network that strikes a strong balance between accuracy, efficiency, and interpretability for multi-lead ECG anomaly detection. Built on DSCR convolutional blocks, a BiGRU, and an attention mechanism, DBA-ASFNet integrates patient age and sex information through the ASF module. Despite having only 0.03 million parameters, the model achieves performance comparable to some SOTA models on the PTB-XL and CPSC2018 datasets while maintaining low computational complexity. Including demographic features provides modest performance gains in certain diagnostic tasks without increasing computational cost. SHAP-based interpretability analysis demonstrates that the decision-making process of DBA-ASFNet aligns with established clinical knowledge, improving transparency and clinician trust. Due to its combination of high accuracy, low resource requirements, and clear interpretability, DBA-ASFNet is a potential candidate for mobile, and wearable ECG monitoring applications.

Author Contributions: This work was conducted in collaboration with all authors. K.L.: Conceptualization, Methodology, Supervision, Writing—original draft, Writing—review and editing, Project administration. L.H.: Conceptualization, Data curation, Methodology, Software, Writing—original draft. H.H.: Methodology, Validation, Writing—original draft, Writing—review and editing. Y.C.: Methodology, Validation, Writing—original draft. L.Y.: Writing—review and editing, Validation. S.C.: Methodology, Validation. J.C.: Writing—review and editing, Validation, Supervision. C.L.: Conceptualization, Methodology, Writing—review and editing, Validation. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by Grant 61601124 from the NSFC, China, Grant 2023C007 from the Ningde Science and Technology Bureau, China.

Data Availability Statement: The model and code are available at https://github.com/Talitaaa1/DBA-ASFNet (accessed on 23 November 2025).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Mensah, G.A.; Fuster, V.; Murray, C.J.L.; Roth, G.A.; Mensah, G.A.; Abate, Y.H.; Abbasian, M.; Abd-Allah, F.; Abdollahi, A.; Abdollahi, M.; et al. Global Burden of Cardiovascular Diseases and Risks, 1990–2022. *J. Am. Coll. Cardiol.* **2023**, *82*, 2350–2473. [CrossRef]
2. Mensah, G.A.; Fuster, V.; Roth, G.A. A Heart-Healthy and Stroke-Free World. *J. Am. Coll. Cardiol.* **2023**, *82*, 2343–2349. [CrossRef] [PubMed]
3. Denysyuk, H.V.; Pinto, R.J.; Silva, P.M.; Duarte, R.P.; Marinho, F.A.; Pimenta, L.; Gouveia, A.J.; Gonçalves, N.J.; Coelho, P.J.; Zdravevski, E.; et al. Algorithms for Automated Diagnosis of Cardiovascular Diseases Based on ECG Data: A Comprehensive Systematic Review. *Heliyon* **2023**, *9*, 13601. [CrossRef]
4. Zhu, H.; Cheng, C.; Yin, H.; Li, X.; Zuo, P.; Ding, J.; Lin, F.; Wang, J.; Zhou, B.; Li, Y.; et al. Automatic Multilabel Electrocardiogram Diagnosis of Heart Rhythm or Conduction Abnormalities with Deep Learning: A Cohort Study. *Lancet Digit. Health* **2020**, *2*, 348–357. [CrossRef]
5. Wu, Z.; Guo, C. Deep Learning and Electrocardiography: Systematic Review of Current Techniques in Cardiovascular Disease Diagnosis and Management. *Biomed. Eng. OnLine* **2025**, *24*, 23. [CrossRef] [PubMed]
6. Liu, X.; Wang, H.; Li, Z.; Qin, L. Deep Learning in ECG Diagnosis: A Review. *Knowl.-Based Syst.* **2021**, *227*, 107187. [CrossRef]
7. Huang, Z.; Herbozo Contreras, L.F.; Leung, W.H.; Yu, L.; Truong, N.D.; Nikpour, A.; Kavehei, O. Efficient Edge-AI Models for Robust ECG Abnormality Detection on Resource-Constrained Hardware. *J Cardiovasc. Transl. Res.* **2024**, *17*, 879–892. [CrossRef]
8. Xu, H.; Shuttleworth, K.M.J. Medical Artificial Intelligence and the Black Box Problem: A View Based on the Ethical Principle of "Do No Harm". *Intell. Med.* **2024**, *4*, 52–57. [CrossRef]
9. Kaur, D.; Hughes, J.W.; Rogers, A.J.; Kang, G.; Narayan, S.M.; Ashley, E.A.; Perez, M.V. Race, Sex, and Age Disparities in the Performance of ECG Deep Learning Models Predicting Heart Failure. *Circ. Heart Fail.* **2024**, *17*, 010879. [CrossRef]
10. Le, K.H.; Pham, H.H.; Nguyen, T.B.T.; Nguyen, T.A.; Thanh, T.N.; Do, C.D. LightX3ECG: A Lightweight and eXplainable Deep Learning System for 3-Lead Electrocardiogram Classification. *Biomed. Signal Process. Control* **2023**, *85*, 104963. [CrossRef]
11. Alamatsaz, N.; Tabatabaei, L.; Yazdchi, M.; Payan, H.; Alamatsaz, N.; Nasimi, F. A Lightweight Hybrid CNN-LSTM Explainable Model for ECG-Based Arrhythmia Detection. *Biomed. Signal Process. Control* **2024**, *90*, 105884. [CrossRef]
12. An, X.; Shi, S.; Wang, Q.; Yu, Y.; Liu, Q. Research on a Lightweight Arrhythmia Classification Model Based on Knowledge Distillation for Wearable Single-Lead ECG Monitoring Systems. *Sensors* **2024**, *24*, 7896. [CrossRef]
13. Zhou, Y.; Diao, X.; Huo, Y.; Liu, Y.; Fan, X.; Zhao, W. Masked Transformer for Electrocardiogram Classification. *arXiv* **2024**, arXiv:2309.07136. [CrossRef]
14. Wang, Z.; Khatibi, E.; Kazemi, K.; Azimi, I.; Mousavi, S.; Malik, S.; Rahmani, A.M. TransECG: Leveraging Transformers for Explainable ECG Re-Identification Risk Analysis. *arXiv* **2025**, arXiv:2503.13495.
15. Chang, C.-H.; Lin, C.-S.; Luo, Y.-S.; Lee, Y.-T.; Lin, C. Electrocardiogram-Based Heart Age Estimation by a Deep Learning Model Provides More Information on the Incidence of Cardiovascular Disorders. *Front. Cardiovasc. Med.* **2022**, *9*, 754909. [CrossRef] [PubMed]
16. Rawshani, A.; Rawshani, A.; Smith, G.; Boren, J.; Bhatt, D.L.; Börjesson, M.; Engdahl, J.; Kelly, P.; Louca, A.; Ramunddal, T.; et al. Integrating Deep Learning with ECG, Heart Rate Variability and Demographic Data for Improved Detection of Atrial Fibrillation. *Open Heart* **2025**, *12*, 003185. [CrossRef] [PubMed]
17. Abdullah, T.A.A.; Zahid, M.S.M.; Ali, W.; Hassan, S.U. B-LIME: An Improvement of LIME for Interpretable Deep Learning Classification of Cardiac Arrhythmia from ECG Signals. *Processes* **2023**, *11*, 595. [CrossRef]

18. Mousavi, S.; Afghah, F.; Acharya, U.R. HAN-ECG: An Interpretable Atrial Fibrillation Detection Model Using Hierarchical Attention Networks. *Comput. Biol. Med.* **2020**, *127*, 104057. [CrossRef]
19. Goettling, M.; Hammer, A.; Malberg, H.; Schmidt, M. xECGArch: A Trustworthy Deep Learning Architecture for Interpretable ECG Analysis Considering Short-Term and Long-Term Features. *Sci. Rep.* **2024**, *14*, 13122. [CrossRef]
20. Abgrall, G.; Holder, A.L.; Chelly Dagdia, Z.; Zeitouni, K.; Monnet, X. Should AI Models Be Explainable to Clinicians? *Crit. Care* **2024**, *28*, 301. [CrossRef]
21. Michiels, J.; Suykens, J.; De Vos, M. Explaining the Model and Feature Dependencies by Decomposition of the Shapley Value. *Decis. Support Syst.* **2024**, *182*, 114234. [CrossRef]
22. Wagner, P.; Strodthoff, N.; Bousseljot, R.-D.; Kreiseler, D.; Lunze, F.I.; Samek, W.; Schaeffter, T. PTB-XL, a Large Publicly Available Electrocardiography Dataset. *Sci. Data* **2020**, *7*, 154. [CrossRef]
23. Willems, J.L.; Zywietz, C.; Rubel, P.; Degani, R.; Macfarlane, P.W.; Van Bemmel, J.H. A Standard Communications Protocol for Computerized Electrocardiography. *J. Electrocardiol.* **1991**, *24*, 173–178. [CrossRef]
24. Liu, F.; Liu, C.; Zhao, L.; Zhang, X.; Wu, X.; Xu, X.; Liu, Y.; Ma, C.; Wei, S.; He, Z.; et al. An Open Access Database for Evaluating the Algorithms of Electrocardiogram Rhythm and Morphology Abnormality Detection. *J. Med. Imaging Health Inform.* **2018**, *8*, 1368–1373. [CrossRef]
25. Huang, W.; Wang, N.; Feng, P.; Wang, H.; Wang, Z.; Zhou, B. A Multi-Resolution Mutual Learning Network for Multi-Label ECG Classification. In Proceedings of the 2024 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Lisbon, Portugal, 3 December 2024; pp. 3303–3306.
26. Van Alsté, J.A.; Van Eck, W.; Herrmann, O.E. ECG Baseline Wander Reduction Using Linear Phase Filters. *Comput. Biomed. Res.* **1986**, *19*, 417–427. [CrossRef]
27. Sharma, V. A Study on Data Scaling Methods for Machine Learning. *Int. J. Glob. Acad. Sci. Res.* **2022**, *1*, 31–42. [CrossRef]
28. Zhou, F.; Chen, L. Leadwise Clustering Multi-Branch Network for Multi-Label ECG Classification. *Med. Eng. Phys.* **2024**, *130*, 104196. [CrossRef]
29. Papageorgiou, G.; Tjortjis, C. Adaptive Sliding Window Normalization. *Inf. Syst.* **2025**, *129*, 102515. [CrossRef]
30. Cho, K.; Van Merrienboer, B.; Gulcehre, C.; Bahdanau, D.; Bougares, F.; Schwenk, H.; Bengio, Y. Learning Phrase Representations Using RNN Encoder–Decoder for Statistical Machine Translation. In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP), Doha, Qatar, 3 June 2014; pp. 1724–1734.
31. Strodthoff, N.; Wagner, P.; Schaeffter, T.; Samek, W. Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL. *IEEE J. Biomed. Health Inform.* **2021**, *25*, 1519–1528. [CrossRef]
32. Jyotishi, D.; Dandapat, S. An Attentive Spatio-Temporal Learning-Based Network for Cardiovascular Disease Diagnosis. *IEEE Trans. Syst. Man Cybern. Syst.* **2023**, *53*, 4661–4671. [CrossRef]
33. Li, Z.; Zhang, H. Fusing Deep Metric Learning with KNN for 12-Lead Multi-Labelled ECG Classification. *Biomed. Signal Process. Control* **2023**, *85*, 104849. [CrossRef]
34. Elharrouss, O.; Mahmood, Y.; Bechqito, Y.; Serhani, M.A.; Badidi, E.; Riffi, J.; Tairi, H. Loss Functions in Deep Learning: A Comprehensive Review. *arXiv* **2025**, arXiv:2504.04242.
35. Wang, Z.; Yan, W.; Oates, T. Time Series Classification from Scratch with Deep Neural Networks: A Strong Baseline. In Proceedings of the 2017 International Joint Conference on Neural Networks (IJCNN), Anchorage, AK, USA, 14–17 May 2017; pp. 1578–1585.
36. Fawaz, H.I.; Lucas, B.; Forestier, G.; Pelletier, C.; Schmidt, D.F.; Weber, J.; Webb, G.I.; Idoumghar, L.; Muller, P.-A.; Petitjean, F. InceptionTime: Finding AlexNet for Time Series Classification. *Data Min. Knowl. Discov.* **2020**, *34*, 1936–1962. [CrossRef]
37. He, T.; Zhang, Z.; Zhang, H.; Zhang, Z.; Xie, J.; Li, M. Bag of Tricks for Image Classification with Convolutional Neural Networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 558–567.
38. Howard, A.; Sandler, M.; Chen, B.; Wang, W.; Chen, L.-C.; Tan, M.; Chu, G.; Vasudevan, V.; Zhu, Y.; Pang, R.; et al. Searching for MobileNetV3. In Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), Seoul, Republic of Korea, 27 October–2 November 2019; pp. 1314–1324.
39. Yao, Q.; Wang, R.; Fan, X.; Liu, J.; Li, Y. Multi-Class Arrhythmia Detection from 12-Lead Varied-Length ECG Using Attention-Based Time-Incremental Convolutional Neural Network. *Inf. Fusion* **2020**, *53*, 174–182. [CrossRef]
40. Chen, Q.; Lian, C.; Xu, B.; Zhou, Q.; Su, Y.; Zeng, Z. Large Language Model-Assisted Multi-Scale Hierarchical Classification of ECG Signals. *Knowl.-Based Syst.* **2025**, *324*, 113807. [CrossRef]
41. Jiang, R.; Fu, B.; Li, R.; Li, R.; Chen, D.Z.; Liu, Y.; Xie, G.; Li, K. A Dual-Branch Convolutional Neural Network with Domain-Informed Attention for Arrhythmia Classification of 12-Lead Electrocardiograms. *Eng. Appl. Artif. Intell.* **2025**, *139*, 109480. [CrossRef]
42. Ahn, M.-S. Current Concepts of Premature Ventricular Contractions. *J. Lifestyle Med.* **2013**, *3*, 26–33.
43. Tan, N.Y.; Witt, C.M.; Oh, J.K.; Cha, Y.-M. Left Bundle Branch Block: Current and Future Perspectives. *Circ. Arrhythm. Electrophysiol.* **2020**, *13*, 008239. [CrossRef]

44. Alventosa-Zaidin, M.; Guix Font, L.; Benitez Camps, M.; Roca Saumell, C.; Pera, G.; Alzamora Sas, M.T.; Forés Raurell, R.; Rebagliato Nadal, O.; Dalfó-Baqué, A.; Brugada Terradellas, J. Right Bundle Branch Block: Prevalence, Incidence, and Cardiovascular Morbidity and Mortality in the General Population. *Eur. J. Gen. Pract.* **2019**, *25*, 109–115. [CrossRef]
45. Ayano, Y.M.; Schwenker, F.; Dufera, B.D.; Debelee, T.G.; Ejegu, Y.G. Interpretable Hybrid Multichannel Deep Learning Model for Heart Disease Classification Using 12-Lead ECG Signal. *IEEE Access* **2024**, *12*, 94055–94080. [CrossRef]
46. Ramirez, E.; Ruiperez-Campillo, S.; Casado-Arroyo, R.; Merino, J.L.; Vogt, J.E.; Castells, F.; Millet, J. The Art of Selecting the ECG Input in Neural Networks to Classify Heart Diseases: A Dual Focus on Maximizing Information and Reducing Redundancy. *Front. Physiol.* **2024**, *15*, 1452829. [CrossRef] [PubMed]
47. De Maria, B.; Parati, M.; Dalla Vecchia, L.A.; La Rovere, M.T. Day and Night Heart Rate Variability Using 24-h ECG Recordings: A Systematic Review with Meta-Analysis Using a Gender Lens. *Clin. Auton. Res.* **2023**, *33*, 821–841. [CrossRef]
48. Prajapati, C.; Koivumäki, J.; Pekkanen-Mattila, M.; Aalto-Setälä, K. Sex Differences in Heart: From Basics to Clinics. *Eur. J. Med. Res.* **2022**, *27*, 241. [CrossRef]
49. Chen, T.-M.; Huang, C.-H.; Shih, E.S.C.; Hu, Y.-F.; Hwang, M.-J. Detection and Classification of Cardiac Arrhythmias by a Challenge-Best Deep Learning Neural Network Model. *iScience* **2020**, *23*, 100886. [CrossRef]
50. García-Isla, G.; Mainardi, L.; Corino, V.D.A. A Detector for Premature Atrial and Ventricular Complexes. *Front. Physiol.* **2021**, *12*, 678558. [CrossRef]
51. Zhang, J.; Liu, J.; Ye, M.; Zhang, M.; Yao, F.; Cheng, Y. Incidence and Risk Factors Associated with Atrioventricular Block in the General Population: The Atherosclerosis Risk in Communities Study and Cardiovascular Health Study. *BMC Cardiovasc. Disord.* **2024**, *24*, 509. [CrossRef]
52. Elliott, A.D.; Middeldorp, M.E.; Van Gelder, I.C.; Albert, C.M.; Sanders, P. Epidemiology and Modifiable Risk Factors for Atrial Fibrillation. *Nat. Rev. Cardiol.* **2023**, *20*, 404–417. [CrossRef]
53. Verardi, R.; Iannopollo, G.; Casolari, G.; Nobile, G.; Capecchi, A.; Bruno, M.; Lanzilotti, V.; Casella, G. Management of Acute Coronary Syndrome in Elderly Patients: A Narrative Review through Decisional Crossroads. *J. Clin. Med.* **2024**, *13*, 6034. [CrossRef]

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.