Classification of ECG Anomaly with Dynamically-biased LSTM for Continuous Cardiac Monitoring

Jinhai Hu[1,2], Wang Ling Goh[1] and Yuan Gao[2]

[1]School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

[2]Institute of Microelectronics, Agency for Science, Technology and Research (A*STAR), Singapore

Email: jinhai001@e.ntu.edu.sg, ewlgoh@ntu.edu.sg, gaoy@ime.a-star.edu.sg

Abstract— **This paper presents an electrocardiogram (ECG) signal classification model based on dynamically-biased Long Short-Term Memory (DB-LSTM) network. Compared to conventional LSTM networks, DB-LSTM introduces a set of parameters *C* which save the previous time-step cell gate states of the unit cell. Hence, more feature information is preserved and a smaller size network is required for the classification task. Comprehensive simulations using MIT-BIH ECG datasets show that this model can perform ECG feature classification with shorter time window, faster training convergence while achieving comparable training and classification accuracy with much lower weigh resolution. Compared to the other state-of-art ECG analysis algorithms, this model only requires 4 layers, and it achieved 96.74% accuracy when weights are truncated from FP32 to INT4 with only 2.4% accuracy degradation. Implemented on Xilinx Artix-7 FPGA, the proposed design is estimated to consume only 40μW dynamic power, which is a promising candidate for resource constrained edge devices.**

Keywords— Long Short-Term Memory (LSTM), DB-LSTM, ECG Classification

I. Introduction

Cardiovascular diseases (CVDs) such as heart failure and atrial fibrillation are the leading causes of death globally [1]. Electrocardiogram (ECG) is the golden standard for cardiac monitoring by measuring the strength and timing of the electrical activity in the heart [2]. Some of the cardiovascular events such as heart rate variability (HRV) and atrial fibrillation are rare and not predictable. Therefore, long-term continuous ECG monitoring is required to capture such events for cardiac risk evaluation. Current medical practices require patient to wear a portable recording device for 24 hours or longer period. The recorded data are either stored in the device or uploaded to the cloud for clinician to conduct off-line analysis. However, there are rising needs for online real-time ECG analysis and classification. The purpose is to detect the onset of symptoms such as arrythmia or cardiac arrest timely, so that early alarm and necessary medical intervention can be triggered in time to improve the therapeutic outcome.

ECG signal is time-variant in nature. To classify a time-variant sequence, the algorithm needs to identify the key features from the data series and map them to the respective pattern categories. There are a few key ECG anomaly patterns illustrated in Fig. 1, including Supraventricular ectopic beat (S), Ventricular ectopic beat (V), Fusion beat (F) and Unknown beat (Q). Together with the normal beat (N), five categories suggested by AAMI standard [3] are required for meaningful ECG classification.

This work was supported by the Agency for Science, Technology and Research (A*STAR), Singapore under the Cyber-Physiochemical Interface programme, grant No. A18A1b0045.

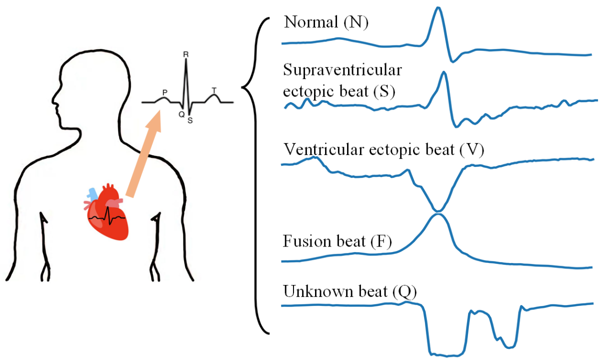


Fig. 1. Illustration of ECG signal with five key types of waveform pattern.

There are many prior works on deep-learning based ECG analysis algorithms [4-9]. Convolutional neural network (CNN), both in one dimensional (1D) [6-8] and two-dimensional (2D) [9] formats have been applied in ECG analysis. In 2D-CNN, ECG signal is converted to image and then perform image processing [9, 10]. To achieve satisfactory results, large-scale deep CNN is required. For example, the CNN model in [9] contains more than 10 million weight parameters and requires huge computation power, therefore it is very challenging to implement such kind of algorithm in resource constraint portable devices. On the other hand, Long Short-Term Memory (LSTM) [11], which is a kind of recurrent neural network, has demonstrated superior classification accuracy in time series analysis. Although LSTM can achieve similar performance with smaller network size as compared to CNN, deeper networks such as bidirectional LSTM are required for practical ECG applications [12], which is still a huge computation burden for edge devices. In addition, the conventional LSTM and CNN performance will degrade significantly when the weights are truncated to lower resolution for hardware implementation. Longer or even nonconvergence of training processes may occur, resulting in stability issue when on-line training is necessary to build personalized ECG classification model.

In this paper, a new ECG signal classification model based on the dynamically-biased LSTM (DB-LSTM) [13] is proposed to improve classification accuracy with shorter window size and less training parameters. Results show that the proposed model can achieve maximum 97% accuracy for 5-category classification with MIT-BIH ECG datasets [14]. With the same number of neurons, the conventional LSTM can achieve 89.50% accuracy. On the other hand, deep CNN costs 77% more training weight parameters to reach similar accuracy. This paper is organized as follows: Section II introduces the proposed model and training procedure. The results for ECG classification, fixed-point weight truncation, and FPGA implementation are discussed in Section III, followed by a summary in Section IV.

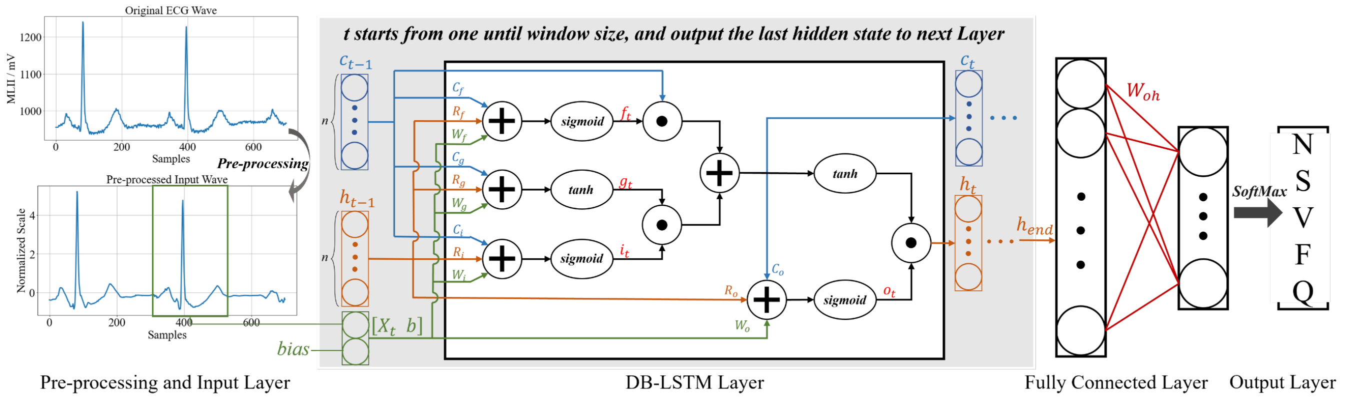


Fig. 2. Block diagram of the proposed DB-LSTM algorithm model.

TABLE I DB-LSTM CELL STRUCTURE PARAMETERS

X_t	Current input vector
b	Bias
c_t	Current cell state
h_t	Current cell output (hidden state)
c_{t-1}	Previous cell state
h_{t-1}	Previous cell output (hidden state)
h_{end}	Cell ouput (hidden state) at last time step
f_t, g_t, i_t, o_t	Current Forget, Generate, Input, Output statues
$W_{f,g,i,o}$	Input weights for *f, g, i, o*
$R_{f,g,i,o}$	Recurrent weights for *f, g, i, o*
$C_{f,g,i,o}$	Cell weights for *f, g, i, o*
W_{oh}	Weights for fully connected layer

II. PROPOSED MODEL

A. Model Architecture

The proposed DB-LSTM ECG classification algorithm block diagram is illustrated in Fig. 2, the corresponding network parameters are listed in Table I. As shown in Fig. 2, the network consists of four layers, namely input layer, DB-LSTM layer, fully connected layer and output layer. The input data first go through three steps of pre-processing, including denoising, z-score normalization and equalization, to enhance the ECG signal features. Since most types of ECG signal consist of a distinct R peak, z-score normalization can maintain useful information about signal envelope and make the algorithm less sensitive to amplitude variation, compared to min-max scaling. Lastly, it is important to include all types of target ECG features into the training process with equal percentage of occurrence, so that the trained model has a balanced feature weighting.

As illustrated in Fig. 2, the input layer receives the pre-processed ECG signal with m features and window size of k. A shareable random number is embedded as dynamic bias, which is split into four different biases due to the $W_{f,g,i,o}$ adjustment during the training. In this way, the bias can be fine-tuned to provide two-fold benefits, for the gate output to be more accurate, and to smoothen the back-propagation loss gradient. Thereafter, the processed input vector is sent to the DB-LSTM layer that is made up of DB-LSTM cell consisting of four statuses which discard non useful message, update memory, and control the output, respectively. At time step t, the input vector is $[X_t\ \ b]$ with dimension $1 \times (m+1)$, and its corresponding $W_{f,g,i,o}$ has the size $n \times (m+1)$. Since the window size is set to k, which is same as sequence length, the total training steps within one epoch is k. When the current time step t is less than k, the output of the DB-LSTM cell, c_t and h_t with dimension $n \times 1$, will be redirected to the recurrent input in the next time step, as shown in the shaded block in Fig. 2. Otherwise, the final hidden state h_{end} is sent to the fully connected layer with n neurons to centralize these n hidden features into five output neurons, whose weight W_{oh} has size $5 \times n$. The output layer performs SoftMax function to the five output neurons and generate five categories.

Different from conventional LSTM structure, *f*, *g*, *i*, *o* statuses in DB-LSTM have one additional parameter, the cell state c_t. *f*, *g* and *i* gates consider both h_{t-1} and c_{t-1}, while *o* gate counts in the current cell state c_t instead of c_{t-1} [13]. As a result, *f* and *i* gates will analyse current input, previous cell state and the output to yield a combined output which includes information of both previous and current states. Furthermore, the cell state parameter in *g* allows DB-LSTM to hold a longer memory and reduce the effect of single direction LSTM's deficiency. Similarly, the output gate *o* monitors the current cell state to determine the hidden output. The difference between the selection of c_t and c_{t-1} is that *o* is the gating machine working on the process from the current cell state to the hidden state, while *f* and *i* are deciding on how the old data can affect the current *c*.

The optimized network size is single neuron with two inputs and 32 hidden outputs in DB-LSTM layer, followed by 32 hidden neurons in fully connected layer, and finally connected by five output neurons. Thus, the number of weight parameters are 8448 and 160 in DB-LSTM and fully connected layers, respectively.

B. Training with Fixed-point Weights Specifications

A personalized online training process is preferred since individual user has his or her unique ECG patterns. For wearable devices, shorter training period allows the model to be updated to new users' condition quickly. Back propagation (BP) is used to adjust the three groups of weight parameters for rapid training convergence. The input for BP is the output residual error which is expressed as the cross-entropy loss function shown in (1).

$$E = -\sum_{i=1}^{n} y_i \log(p_i) \tag{1}$$

where n stands for the number of classes, y_i is the truth label and p_i is the SoftMax probability for the i^{th} class. After the model is fully trained, the DB-LSTM can classify full heartbeat ECG signal efficiently.

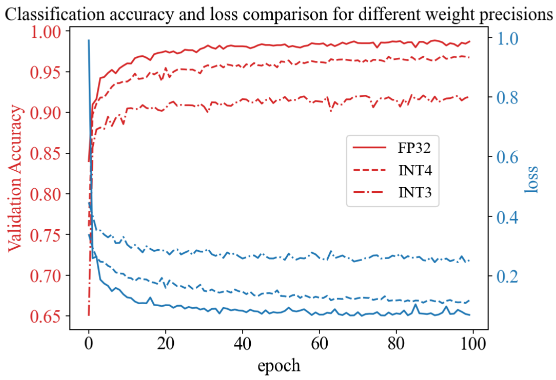


Fig. 3. Comparison of training process with different weight resolution.

TABLE II MULTI MODEL PERFORMANCE COMPARASION

Topology	Number of NN Layer	Window Size	Number of Weight Parameters	Validation Accuracy	Weight Quantization
Deep CNN	9	360	22,181	99.82%	FP32
Deep CNN	9	180	14,981	97.58%/ 99.46%	INT4/FP32
LSTM	4	360	4,517	89.50%	FP32
LSTM	4	180	4,517	80.40%/ 88.62%	INT4/FP32
BiLSTM	4	180	9,029	98.92%	FP32
DB-LSTM	4	180	8,608	96.74%/ 99.10%	INT4/FP32

For hardware implementation, the weights are mapped to fixed-point integer numbers, which are quantized into signed 2×2^n states (s) including negative values, where n is the bit length. The quantization procedure is to firstly find the maximum and minimum numbers in that weight matrix in absolute value, denoted as w_{max} and w_{min}, respectively. They set the boundary of states, and the resolution between these states is given by

$$\Delta w = \frac{w_{max} - w_{min}}{2^n - 1} \tag{2}$$

For each weight entry w_i within that weight matrix, the mapping process follows the following equation,

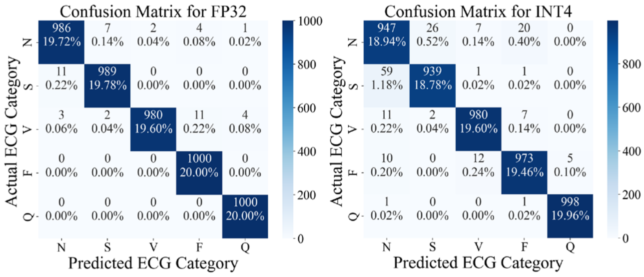
$$w_i = s_{j+1}, \qquad if\ |w_i - (s_1 + j \times \Delta w)| \le \frac{\Delta w}{2} \tag{3}$$

where $j = 0, 1, 2, \dots 2^n - 1$. Fig. 3 shows the training accuracy and loss with different weight precisions from floating-point FP32 to fixed-point integer of 4-bit (INT4) and 3-bit (INT3). It can be observed that FP32 weights achieved the fastest convergence and highest accuracy. However, INT4 also achieved comparable accuracy and loss. With lower fixed-point weight precision INT3, the convergence speed became slower, and the degradation of accuracy is around 9%. Thus, INT4 will be the selected weight precision for future hardware implementation.

III. SIMULATION AND IMPLEMENTATION RESULTS

A. Simulation Results and Benchmarking

The performance of the proposed algorithm is evaluated with MIT-BIH ECG database which includes 48 sets of 30-minute-long ECG data. Each dataset contains both normal and abnormal ECG features (N, S, V, F, Q) shown in Fig. 1 recorded with a fixed sampling frequency of 360 Hz and 11-bit data resolution.



(a) (b)

Fig. 4. Classification confusion matrices for different weight precisions, (a) FP32, and (b) INT4.

After pre-processing, the training dataset contains 20,000 samples with 4000 samples in each category and the validation dataset has 5000 samples with 1000 samples in each category. The output size of DB-LSTM cell is set as 32, thus the total number of training parameters is 8608. With 0.01 learning rate, 0.05 weight gradient, 0.01 weight penalty and cross-entropy loss function, the validation accuracy for five-category classification with FP32 weight precision is 99.1%, the corresponding confusion matrix is shown in Fig. 4 (a). Furthermore, after weight being quantized into INT4, the validation accuracy dropped to 96.7%, as shown in Fig. 4 (b). For each confusion matrix, the horizontal axis stands for the predicted category and the vertical axis gives the actual label. For each entry in the matrix, the upper data is the number of predicted samples, while the lower one is the percentage out of the total validation dataset. The summation of diagonal entries is the total validation accuracy.

Various Deep CNN, conventional LSTM and bidirectional LSTM models are built to compare the performance, the results are tabulated in Table II. Different window sizes are used to characterize the influence on accuracy and training parameters. With the same deep CNN structure, double window size will result in almost double in total training parameters, but only 2% improvement of accuracy. On the other hand, the training parameters of LSTM will not be affected by the variance of window size, which only depends on the input and output size. Comparing to conventional LSTM, bidirectional LSTM doubles the number of parameters and enhance the accuracy by 9.1%. The drawback of it is that bidirectional cannot perform continuous time monitoring, which yields a large latency in real-time classification. DB-LSTM yields short window size, continuous time processing, less training parameters and higher classification accuracy than other models. It reached similar classification result as Deep CNN with same window size, but a smaller number of neural network layers and 43.6% less of training parameters.

Supraventricular ectopic beat (SVEB) and ventricular ectopic beat (VEB) are the two most common arrhythmias suggested by AAMI standard [3], whose classification results involved by *Acc, Spe, Sen, Ppr* and *F1* [19], are the criteria to evaluate the performance of the models. The classification accuracy of VEB, SVEB, and the overall classification accuracy can be calculated, as shown in Table III. The overall accuracy is 99.1% and 96.74% for FP32 and INT4, respectively. The performance of the proposed model is summarized and compared with other state-of-arts in Table IV. Compared to those designs implemented with FP32 weight resolution, the proposed algorithm used one of the smallest network sizes and achieved comparable accuracy. When the weight is quantized to INT4, the accuracy only has minor degradation of 2.4%.

TABLE III CLASSIFICATION ACCURACY

	Overall	VEB					SVEB				
	Accuracy	*Acc*	*Sen*	*Spe*	*Ppr*	*F1*	*Acc*	*Sen*	*Spe*	*Ppr*	*F1*
MeMeA 2018 [15]	97.6	NA	97.0	NA	97.0	97.0	NA	96.0	NA	75.0	84.0
TBioCAS 2019 [16]	97.3	97.9	80.2	99.8	97.3	88.0	NA	NA	NA	NA	NA
TBioCAS 2019 [17]	98.4	99.1	91.8	99.6	95.3	93.5	99.4	79.5	99.9	96.3	87.2
JBHI 2020 [18]	97.6	98.9	93.1	99.5	95.6	94.3	96.7	78.3	99.7	92.5	84.8
JBHI 2021 [19]	99.1	99.8	97.6	99.9	98.8	98.2	99.7	89.5	99.9	97.2	93.2
This Work (FP32)	**99.10**	**99.56**	**98.00**	**99.38**	**99.80**	**98.89**	**99.60**	**98.90**	**99.15**	**99.10**	**99.00**
This Work (INT4)	**96.74**	**99.20**	**98.00**	**96.43**	**98.00**	**98.00**	**98.22**	**93.90**	**97.45**	**97.10**	**95.47**

TABLE IV BENCHMARK TABLE

	CISP-BMEI 2017 [6]	TBioCAS 2019 [17]	ICECS 2021 [8]	TSP 2021 [9]	Sci Rep 2021 [20]	ISCAS 2021 [21]	JBHI 2022 [19]	**This Work**
Dataset	MIT-BIH	MIT-BIH	MIT-BIH	Fantasia	MIT-BIH	PTB Diagnostic	MIT-BIH	**MIT-BIH**
Topology	1-D CNN	Deep CNN	1-D CNN	Deep CNN	SVM+MLP	LSTM	Deep CNN	**DB-LSTM**
Number of Layers	10	10	10	42	4	3	12	**4**
No. of Weight Parameters	19,750	198,037	49,216*	>10M	NA (Input:32)	NA	8,157	**8,608**
Window Size	260	400	260	160	NA	NA	400	**180**
No. of Classes	5	5	5	40	5	2	5	**5**
Validation Accuracy	93.47%	98.40%	98.12%	99.5%	79.33%	84.63%	99.10%	**96.7% / 99.1%**
Weight Resolution	FP32	FP32	FP32	FP32	FP32	INT1	FP32	**INT4 / FP32**

* Calculated based on information provided in paper.

B. Hardware Implementation

The block diagram of the overall DB-LSTM model implemented in FPGA is shown in Fig. 5 (a), which includes a feedback loop of DB-LSTM as well as a fully connected classifier. Both cell state and hidden state of each time step will be saved in the memory for future usage. The piecewise sigmoid and tanh activation function are also illustrated in Fig. 5 (b) and (c).

Since energy and resource efficiencies are the critical evaluation aspects for edge devices, hardware resources are examined by using Vivado shown in Table V. The total number of clock cycles to classify a full heartbeat is 721 under 40KHz clock frequency. With bit length of INT4 for weight and gate I/O, as well as piece wised activation function, it does not require any digital signal processing (DSP) blocks while using relatively smaller LUT and registers resources, whose utilization is comparable with other works benchmarked in Table V. Dynamic power is the power consumed during the computation. This ECG classification model consumed $40\mu W$ dynamic power when running at 40kHz clock frequency. The overall static power due to FPGA board peripheral circuitry is significant, the total power consumption is 0.091W on FPGA platform Artix-7. Nevertheless, the actual circuit power consumption can be reduced significantly if implemented in application specific integrated circuit (ASIC).

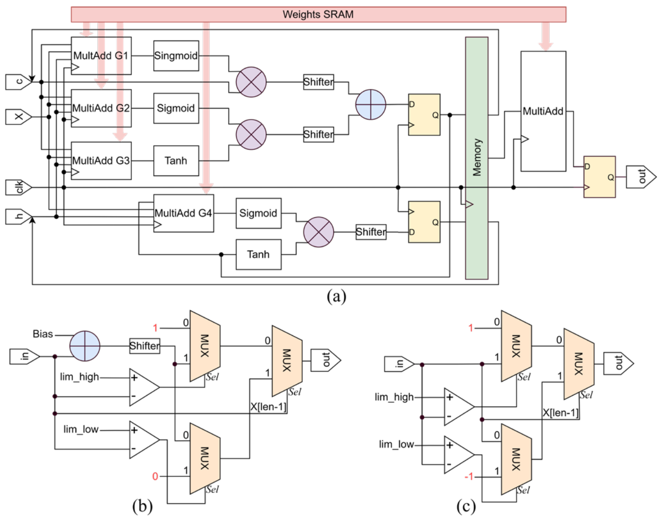


Fig. 5. Schematic diagram for (a) DB-LSTM model, (b) sigmoid activation function, and (c) tanh activation function.

TABLE V PERFORMANCE AND UTILIZATION

	LUT Util.	REG Util.	DSP Blocks	Clock Freq.	Dynamic Power	Total Power	Platform
ISCAS 2017 [22]	1,932	1,926	32	100MHz	NA	0.124W	Zynq XC7Z020
NCA 2020 [23]	7,298	2,474	214	98.2MHz	NA	NA	Artix-7
TBioCAS 2022 [24]	4,223	2,397	8	100kHz	$55\mu W$	$432\mu W$	iCE40UP5k
This Work	**3,980**	**2,376**	**0**	**40kHz**	**$40\mu W$**	**0.091W**	**Artix-7**

IV. CONCLUSION

In this paper, an ECG classification model based on DB-LSTM is proposed for resource constraint edge devices. DB-LSTM can effectively extract the classification information with reduced overall network computation complexity. The proposed method achieved 99.1% accuracy for five categories ECG classification with MIT-BIH database. Compared with conventional LSTM model, the DB-LSTM yields shorter window size and higher performance. The next step is to implement this real-time classification algorithm in CMOS ASIC format to provide ECG anomaly classification for continuous cardiac monitoring.

References

[1] S. Mendis, P. Puska, and B. Norrving, Global atlas on cardiovascular disease prevention and control. *World Health Organization*, 2011.

[2] L. S. Lilly, Pathophysiology of heart disease: A collaborative project of medical students and faculty, 6th ed. *Wolters Kluwer*, 2016.

[3] ANSI/AAMI "Testing and reporting performance results of cardiac rhythm and ST segment measurement algorithms," *Assoc. Adv. Med. Instrument*, no. EC57, 1998.

[4] N. A. Trayanova, D. M. Popescu, en J. K. Shade, "Machine learning in arrhythmia and electrophysiology", *Circulation Research*, vol 128, no 4, pp. 544–566, 2021.

[5] S. Raghunath, *et al*., "Deep neural networks can predict new-onset atrial fibrillation from the 12-lead ECG and help identify those at risk of atrial fibrillation–related stroke", *Circulation*, vol 143, no 13, pp. 1287–1298, 2021.

[6] U. R.. Acharya, *et al*., "A deep convolutional neural network model to classify heartbeats," *Comput. Biol. Med.*, vol. 89, pp. 389-396, 2017.

[7] S. Kiranyaz, T. Ince and M. Gabbouj, "Real-Time Patient-Specific ECG Classification by 1-D Convolutional Neural Networks," *IEEE Trans. on Biomed. Eng.*, vol. 63, no. 3, pp. 664-675, Mar. 2016.

[8] L. Xiaolin, *et al*, "A 1D convolutional neural network for heartbeat classification from Single Lead ECG," *IEEE Int. Conf. on Electron., Circuit. and Syst.*, pp. 1-2, 2020.

[9] C. Cordoș, L. Mihailă, P. Faragó and S. Hintea, "ECG signal classification using convolutional neural networks for biometric identification," *IEEE Int. Conf. on Telecom. and Signal Process. (TSP)*, pp. 167-170, 2021.

[10] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision", *IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, pp. 2818-2826, 2016.

[11] Hochreiter, S., and J. Schmidhuber. "Long short-term memory," *Neural Comput. Appl.*, vol. 9, no. 8, pp.1735-1780, 1997.

[12] M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. on Signal Process.*, vol. 45, no. 11, pp. 2673-2681, Nov. 1997.

[13] J. Hu, *et al*, "Dynamically-biased fixed-point LSTM for time series processing in AIoT edge device," *IEEE Int. Conf. on Artif. Intell. Circuits and Syst. (AICAS)*, pp. 1-4, 2021.

[14] Moody GB, Mark RG., "The impact of the MIT-BIH arrhythmia database," *IEEE Eng. in Med. and Biol.*, 20(3):45-50, 2001.

[15] B. Murugesan, *et al*., "ECGNet: Deep network for arrhythmia classification," *IEEE Int. Symp. Med. Meas. Appl.*, pp. 1-6, Jun. 2018.

[16] A. Amirshahi and M. Hashemi, "ECG classification algorithm based on STDP and R-STDP neural networks for real-timemonitoring on ultra low power personal wearable devices," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 6, pp. 1483–1493, Dec. 2019.

[17] N. Wang, J. Zhou, G. Dai, J. Huang, and Y. Xie, "Energy-efficient intelligent ECG monitoring for wearable devices," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 5, pp. 1112–1121, Oct. 2019.

[18] S. S. Xu, M. -W. Mak, and C. -C. Cheung, "I-vector-based patient adaptation of deep neural networks for automatic heartbeat classification," *IEEE J. Biomed. Health Inform.*, vol. 24, no. 3, pp. 717–727, Mar. 2020.

[19] J. Xiao, *et al*., "ULECGNet: An ultra-lightweight end-to-end ECG classification neural network," *IEEE J. Biomed. Health Inform.*, vol. 26, no. 1, pp. 206-217, Jan. 2022.

[20] A. Saira, *et al*, "ECG-based machine-learning algorithms for heartbeat classification," *Sci. Rep.*, 2021.

[21] S. A. Mirsalari, *et al*, "ELC-ECG: Efficient LSTM cell for ECG classification based on quantized architecture," *IEEE Int. Symp. Circuits Syst (ISCAS)*, pp. 1-5, 2021.

[22] M. Wess, P. D. S. Manoj, and A. Jantsch, "Neural network based ECG anomaly detection on FPGA and trade-off analysis," *IEEE Int. Symp. Circuits Syst. (ISCAS)*, pp. 1–5, 2017.

[23] H. Zairi, *et al*, "FPGA-based system for artificial neural network arrhythmia classification," *Neural Comput. Appl.*, vol. 32, no. 8, pp. 4105–4120, 2020.

[24] D. L. T. Wong, *et al*, "Low complexity binarized 2D-CNN Classifier for Wearable Edge AI Devices," *IEEE Trans. Biomed. Circuits Syst*., early access, 2022.