IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 26, 2021, accepted April 30, 2021, date of publication May 11, 2021, date of current version May 26, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3079155

Hardware Acceleration of High Sensitivity Power-Aware Epileptic Seizure Detection System Using Dynamic Partial Reconfiguration

HEBA ELHOSARY[1], MICHAEL H. ZAKHARI[2], MOHAMED A. ELGAMMAL[2], KHALED A. HELAL KELANY[2], MOHAMED A. ABD EL GHANY[1,5], KHALED N. SALAMA[3], (Senior Member, IEEE), AND HASSAN MOSTAFA[2,4], (Senior Member, IEEE)

[1]Department of Electronics, German University in Cairo (GUC), New Cairo 11835, Egypt
[2]Department of Electronics and Communications Engineering, Cairo University, Giza 12613, Egypt
[3]Computer, Electrical and Mathematical Science and Engineering Division, King Abdullah University of Science and Technology (KAUST), Thuwal 23955, Saudi Arabia
[4]University of Science and Technology, Nanotechnology and Nanoelectronics Program, Zewail City of Science and Technology, Giza 12578, Egypt
[5]Integrated Electronic Systems Laboratory, TU Darmstadt, 64289 Darmstadt, Germany

Corresponding author: Hassan Mostafa (hmostafa@uwaterloo.ca)

This work was supported in part by the Opto-Nano-Electronics (ONE) Lab at Zewail City of Science and Technology and at Cairo University, in part by the Information Technology Industry Development Agency (ITIDA), and in part by the Academy of Scientific Research and Technology (ASRT).

ABSTRACT In this paper, a high-sensitivity low-cost power-aware Support Vector Machine (SVM) training and classification based system, is hardware implemented for a neural seizure detection application. The training accelerator algorithm, adopted in this work, is the sequential minimal optimization (SMO). System blocks are implemented to achieve the best trade-off between sensitivity and the consumption of area and power. The proposed seizure detection system achieves 98.38% sensitivity when tested with the implemented linear kernel classifier. The system is implemented on different platforms: such as Field Programmable Gate Array (FPGA) Xilinx Virtex-7 board and Application Specific Integrated Circuit (ASIC) using hardware-calibrated UMC 65nm CMOS technology. A power consumption evaluation is performed on both the ASIC and FPGA platforms showing that the ASIC power consumption is lower by at least 65% when compared with the FPGA counterpart. A power-aware system is implemented with FPGAs by the adoption of the Dynamic Partial Reconfiguration (DPR) technique that allows the dynamic operation of the system based on power level available to the system at the expense of degradation of the system accuracy. The proposed system exploits the advantages of DPR technology in FPGAs to switch between two proposed designs providing a decrease of 64% in power consumption.

INDEX TERMS Low power, support vector machine (SVM), sequential minimal optimization (SMO), accelerator IP, feature extraction, classification, FPGA, dynamic partial reconfiguration (DPR), ASIC.

I. INTRODUCTION

Epilepsy is a brain disorder that is accompanied by uncontrolled shaking movements of different body parts with a chance of losing consciousness. These shaking movements usually happen as a result of abnormal electrical discharges in the brain neurons. Epilepsy affects 1% of the population worldwide [1]. Many epileptic patients are treated with a daily medication. Regardless of the intensive efforts to develop new pharmacotherapies antiepileptic drugs fail to adequately treat approximately one-third of the patients with epilepsy [2], and even responsive subjects often suffer from side effects as medication are experimental and their concentrations are adopted for each patient individually [3]. Surgical removal of epileptic focus, which is the part of the brain where the seizure is originated, is an option for some patients with medically-resistant epilepsy, but carries the risk of irreversible functional impairment. Thus, new therapeutic approaches are needed to overcome the shortages in other mentioned techniques.

A promising alternative has become booming in the last couple of years which is the intracranial electrical stimulation [4] after detecting the seizure onset. However, the detection of epileptic seizures is usually done by visual observation

The associate editor coordinating the review of this manuscript and approving it for publication was Yasar Amin.

of EEG signals by a trained professional, a process that has several deficiencies. It is a time-consuming procedure, sensitive to bias and can affect the accuracy of the result, consequently, automatic seizure detection algorithms have evolved [5]. The need for these automatic detection systems that would alert the patient to take any needed precautions is now of great importance.

An implantable device that can be inserted in the patient's scalp providing an electrical stimulation as soon as a seizure occurs can be useful for the patient. Recently, machine learning techniques are exploited in automatic seizure detection algorithms as reported in [6]–[8]. Machine learning is the science of teaching computers how to deal with different situations and to perform some complicated tasks without being programmed. Machine learning techniques vary in their complexity. Multiple optimization techniques have evolved to deal with the machine learning techniques increasing complexity as proposed in [9] and [10]. One of the exploited machine learning techniques in seizure detection is the supervised machine learning algorithm. A supervised machine learning algorithm SVM (support vector machine) that was first introduced by Vladimir N. Vapnik *et al.* in 1963 [11] is used in the implemented design. SVM is widely used in statistical classification and regression analysis generally and as it has produced very promising results in detecting and predicting seizures onset [7], [12], [13].

Learning in SVM is a process in which a hyperplane that separates two labeled sets of training examples is determined. SVM searches for the hyperplane that gives the largest margin between the two sets. In this paper, a hardware implemented automatic seizure detection system using supervised machine learning that utilizes EEG signals is proposed. The proposed system is consolidated as follows: First, Features Extraction for training using Sequential Minimal Optimization (SMO) training accelerators which is used in [13]. Then, Feature extraction for classification through linear Support Vector Machine (SVM) classifier, after that a phase of validation is executed to verify the quality of the implemented system. The complete flow of training and validating the system of supervised machine is shown in Figure 1.

The process begins with the training phase in which the data is inserted through a feature extractor module which extracts important information from the input signal. These features in addition to the data and corresponding labels are inserted to the training algorithm. Following that, a classifier is used with the extracted features from the unlabeled testing examples to detect seizures and to classify these inputs into one of the two classes whether it's a seizure or not [14]. A power-aware system is implemented in order to maintain the longevity of the battery life. This power-aware system is achieved using the new capabilities of Field Programmable Gate Array (FPGA).

The rest of the paper is organized as follows. Section II articulates the materials and methods used. Section III provides the implementation of the classifier. Section IV shows the results and the comparison with prior work. Section V discusses the implementation of a power aware system using FPGA. Finally, a comparison with prior work is depicted in section VI, and the conclusion is drawn in Section VII.

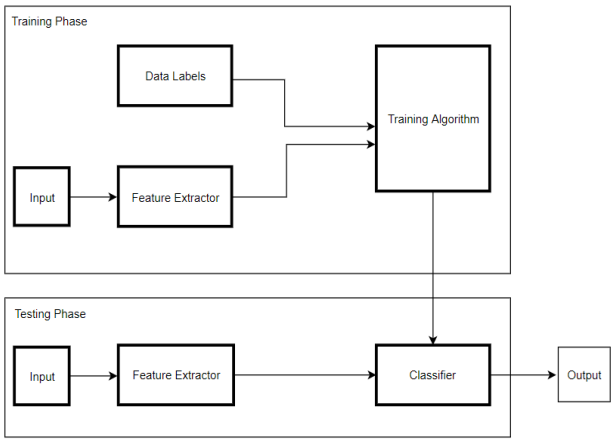


FIGURE 1. Supervised learning training and testing structure.

II. MATERIALS AND METHODS

CHB (Children Hospital Boston) data set is adopted for testing [15]. The dataset is collected from 22 patients with intractable seizures (5 males, and 17 females). From each patient, 23 channels from different electrodes are recorded. The data is sampled at a frequency of 256 Hz with a 16-bit resolution. All signal processing and manipulation are carried out on MATLAB 2017a before the system hardware implementation.

The feature extractor is composed of three feature extraction techniques, namely, fractal dimension, Hurst exponent, and Coastline. The features selection is based on the best performing features obtained in [16]. The three features along with linear Support Vector Machine (SVM) are tested with sequential minimal optimization (SMO) training that is carried out offline.

In this section the features equations are first illustrated, then two feature extractors are proposed.

One performs the exact techniques with minor approximations that does not affect the classification result, while the other computes approximated features which affects the sensitivity obtained in the favor of the power consumption and area utilization of the design.

A. FEATURES' EQUATIONS

1. Fractal Dimension: It is a measurement for the complexity of the input EEG signal over multiple scales. In other words, it is a measure of how many times a pattern can be found in a signal. Higuchi's algorithm with k = 5 is used to calculate the fractal dimension.

$$L_m(k) = \frac{\sum_{i=1}^{\frac{N-m}{k}} |x(m+ik) - x(m+(i-1)*k)|}{N} * \frac{1}{mk} \quad (1)$$

where x is a time series that consists of N data points, m is a constant that varies from 1 to k, and $M = \frac{n-m}{k}$.

$$FD = \sum_{m=1}^{k} \frac{\ln(L_m(k))}{\ln(\frac{1}{k})} \tag{2}$$

2. Hurst Exponent. It is a technique that quantifies the meaningfulness of the input signal. If the output value is in the range from 0.5 to 1 then the input EEG signal contains meaningful patterns, on the other hand if the output value equals 0.5 it is considered noise.

$$MAV = \frac{1}{N} \sum_{i=1}^{N} |x_i| \tag{3}$$

$$R = ||\max(x - MAV)| - |\min(x - MAV)|| \tag{4}$$

$$S = \sqrt{\frac{\sum_{i=1}^{N} (x_i - mean(x))^2}{N-1}} \tag{5}$$

$$H = \frac{\log(\frac{R}{S})}{\log(T)} \tag{6}$$

where x is a time series that consists of N data points, MAV is the mean absolute value, R is the range of the cumulative deviation from the mean, S is the standard deviation, and T is the sampling period.

3. Coastline: It quantifies the amount of fluctuations in the given epoch. Seizures are identified by recurrent discharges in brain neurons, which means higher frequency of fluctuations than usual.

$$Coastline = \sum_{i=1}^{N} |x_{i+1} - x_i| \tag{7}$$

where x is a time series that consists of N data points

B. OPTIMIZED FEATURE EXTRACTOR

As observed in the features' equations, many stages have a division by a certain constant. For instance, the division by the constant T in the final stage of calculating the Hurst exponent depicted in equation (7), or the division by the constant MK in each of the k loops of the fractal dimension shown in equation (1), and also the division by $\ln(\frac{1}{k})$ in the final stage of calculating the fractal dimension shown in equation (2). In those cases, removing the division does not affect the classification result, because it is just scaling all the points by the same amount linearly. Same concept applies to multiplication by a certain constant; all points are scaled by the same value. In both division and multiplication by constant removal cases, the classifier shifts the hyper-plane it draws by the scaling amount correspondingly. Figure 2 shows the effect of linearly scaling data points by removing multiplication or division by a constant.

In other stages, the division by the constant N (total number of data points in the designated window) as in the mean absolute value (MAV) depicted in equation (3) and the mean in standard deviation (STD) depicted in equation (6) contributes in resulting a shorter vector. Since the output of both MAV and the mean are fed to other following stages, a shorter output vector creates smaller and faster design. If the division can be done through shifting, the complexities of division

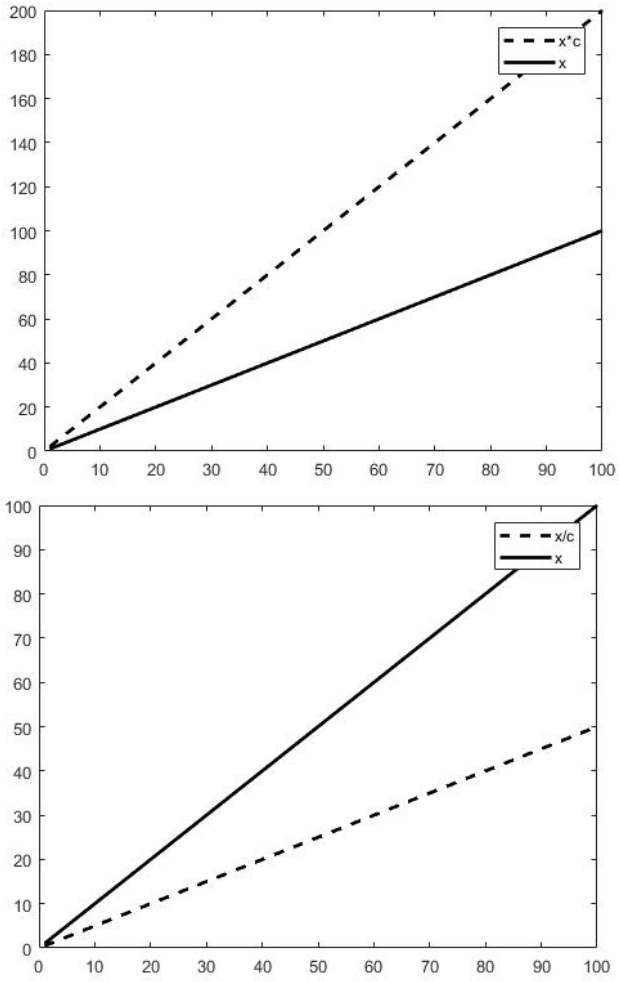


FIGURE 2. Linear scaling of data points through multiplication and division by constants.

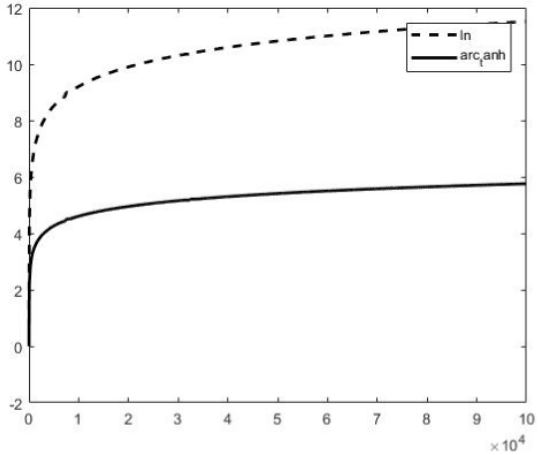


FIGURE 3. Effect of removing the multiplication by the constant 2 on the curve of ln(x).

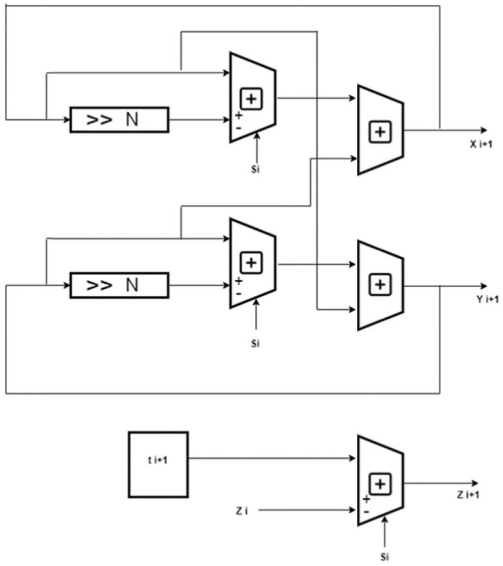


FIGURE 4. Basic building unit of the hyperbolic CORDIC architecture.

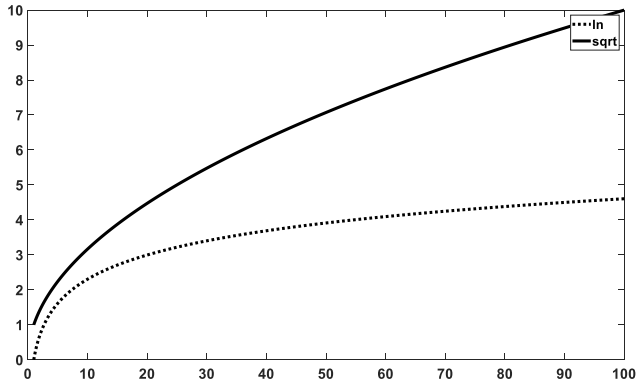


FIGURE 5. The analogy between Ln function and SQRT function.

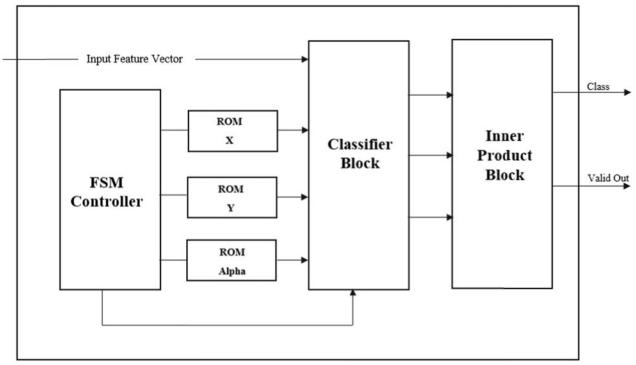


FIGURE 6. Block diagram of the SVM classifier.

can be avoided. For this matter, the window size (number of samples in one window) is chosen such that it is a power of 2.

The only division that cannot be removed, without affecting the sensitivity obtained, is the division by the standard deviation in the Hurst exponent shown in equation (6). The standard deviation is not a constant, but rather varying according to the values of the data samples in each window. However, in order to avoid the complexities of performing a fractional division (as the length of the vector R is shorter than the length of the vector S), the value of R is multiplied by 2^{16}. Linearly scaling the value of R by the constant 2^{16} does not affect the classification results as previously stated in this subsection.

The natural logarithm in the last stages of both fractal dimension and Hurst exponent is realized by calculating the inverse hyperbolic tan of the specified value. The adopted identity for calculating the natural logarithm is $\ln(x) = 2 * tanh^{-1}\left(\frac{x-1}{x+1}\right)$. The multiplication by 2 is, however, not necessarily to be performed as it is only doubling all values and does not affect the classifier decision. The effect of removing the multiplication by the constant 2 is depicted in Figure 3.

C. APPROXIMATE FEATURE EXTRACTOR

The approximate feature extractor considers the power consumption and area utilization more than keeping the sensitivity as originally obtained. A performance degradation of 1-2% in the favor of saving power and area is the objective of the approximate design. Therefore, all the scaling techniques performed in the optimized feature extractor are preserved, and in this section further approximations are adopted.

The inverse hyperbolic tan is a CORDIC design with range expansion. The CORDIC approach is an iterative technique that is utilized for evaluating several elementary mathematical functions that is inverse hyperbolic tan in the proposed design. The technique presents a relatively low-cost solution for evaluating the function in question as it consists of only LUTs, additions, and shifts. The basic building unit of the hyperbolic CORDIC design is depicted in Figure 4.

The CORDIC design of the $tanh^{-1}$ does not only suffer relatively high latency compared to all other utilized modules, but also has the highest dynamic power consumption in the design. Taking the natural logarithm of certain value can be considered as non-linearly scaling this value through a specific manner which is, in this case, $tanh^{-1}$. If there exists a mathematical function that can perform an approximate scaling to the $tanh^{-1}$, then the $tanh^{-1}$ can be approximated using this mathematical function. The closest curve to the $tanh^{-1}$ is the square root curve. Figure 5 shows that the square root function presents itself as an approximately scaled version of the hyperbolic tan function.

A comparison between both inverse hyperbolic tan and square root functions in terms of power consumption, area utilization, and latency when implemented on Virtex-7 FPGA for 24 bit inputs is depicted in Table 1.

The divider exploited in the Hurst exponent to divide the range of the cumulative deviation from the mean by the standard deviation is the second highest power consumer in the proposed feature extractor. Removing this division enhances both occupied area, and power consumption at the expense of accuracy degradation.

TABLE 1. FPGA implementation for 24-bit hyperbolic TAN and SQRT.

	$tanh^{-1}$	SQRT
Power consumption	27.1 mW	3.7 mW
Area utilization	1314 LUT	70 LUT
	451 Register	91 Register
Latency	30 cycles	14 cycles

TABLE 2. Pseudo code of SMO algorithm.

INITIAL	W = 0, A = 0, B = 0, E = 0
	Select i and j
	Read x_i and x_j from memory
	Calculate kernel functions k_{ii}, k_{ij}, k_{jj}
	Calculate the Errors E_i and E_j
Iterate till convergance	Calculate the limits L and H
	$\eta = 2\ k_{ij} - k_{ii} - k_{jj}$
	$\alpha_j^{new} = \alpha_j^{old} + \frac{y_j\,(E_j^{old} - E_j^{new})}{\eta}$
	$\alpha_j^{new} = \alpha_j^{old} + y_i y_j\,(\alpha_j^{old} - \alpha_j^{old,\ clipped})$
	Calculate the bias b
	Check for convergence

III. CLASSIFIER

After the execution of the training phase using the feature extractor discussed earlier, the training can be executed offline afterwards using software methods or by hardware accelerated methods as proposed in [14]. Three important outputs are extracted from the training phase namely: the support vector points (x_j)s, their corresponding labels (y_j)s and the training α's calculated from the SMO algorithm depicted in Table 2.

The three outputs are used in the classification process and the classifier decision for any input vector x_{test} is obtained as shown in equation (8).

$$y_{test} = \sum \alpha_j y_j x_{test} x_j + b \tag{8}$$

Figure 6 shows the architecture of the top-level design of the SVM classifier, which consists of the following blocks: three ROM blocks, classifier block, inner product block and the finite state machine (FSM) controller. The FSM controller is responsible for generating the addresses of the three ROMs and the enable signal of the classifier block. The ROM Support vectors block is used to save the input vectors of the support vector points x_j. The width of this ROM is the same as the data width, while the depth equals to the number of support vectors. The ROM Labels (y_j)s block is used to save the values of the true labels of the support vector points. The width of this ROM is one bit, while the depth is the number of support vectors. The ROM Alpha block is used to save the values of (α)s. The width of this ROM is the same as the data width, while the depth equals the number of support vectors. The classifier block in which each α is multiplied by its corresponding label y and then fed to the inner product calculator. The final block of the SVM classifier is the inner product block which is mainly a MAC (multiply and accumulate) unit is used to calculate the class of size equal to the number of dimensions. The inner product block is used to multiply the input feature vector x_{test} by the corresponding support vector point producing classifier decision y_{test}

The classifier is designed to achieve the lowest area and power consumption while achieving an acceptable performance. The classifier is implemented to operate with the same frequency as the feature extractor module which is 100 MHz and its latency is 16 clock cycles.

Many techniques are used in the implementation of the classifier in order to achieve the mentioned specification. To achieve lower power consumption fixed-point representation is used instead of power hungry floating point based techniques. MATLAB simulations on different word lengths was carried out to decide the optimum length, it is found that a 16-bit word length is enough for achieving the same performance obtained from floating point based algorithm. To achieve higher frequency multipliers are replaced by XOR gates in the multiplication process as the numbering representation exploited is sign-magnitude representation.

IV. RESULTS AND COMPARISON

The results are divided into two sections; Optimized feature extractor results, and approximate feature extractor results. Each section elucidates Software, and Hardware results for each design. For the software results, three performance metrics, namely, sensitivity, specificity, and accuracy are measured for each feature extractor along with Sequential Minimal Optimization (SMO) training accelerator. While the Hardware results discuss the Area utilization, and power consumption of each design on both FPGA and ASIC platforms.

TABLE 3. The effect of changing the window size on the performance metrics for the optimized feature extractor.

WINDOW SIZE	SENSITIVITY	SPECIFICITY	ACCURACY
1 second (256 sample)	96.20%	94.04%	94.05%
2 seconds (512 sample)	97.52%	93.66%	93.67%
4 seconds (1024 sample)	98.39%	92.37%	92.39%

A. OPTIMIZED FEATURE EXTRACTOR

The sensitivity, specificity, and accuracy originally obtained when the exact features using floating point arithmetic are exploited along with SMO training and linear SVM classification are 98.39%, 92.60%, and 92.61% respectively.

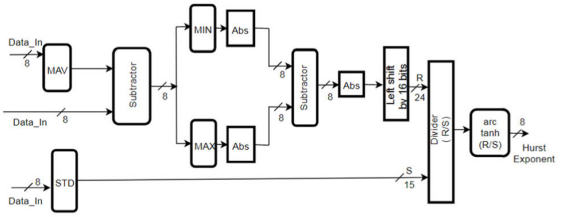
When the discussed divisions and multiplications by constants are removed, and fixed point representation is exploited, the sensitivity, specificity, and accuracy are 98.39%, 92.37%, 92.39%, respectively and correspondingly the metrics degradations due to these approximations are insignificant.

The performance metrics are not only affected by the feature extraction strategies, training approaches and classification techniques, but also affected by changing the window size. Window size is the number of samples per single entry; each EEG signal is divided into epochs of either 1024 sample, 512 sample, or 256 sample which corresponds to 4-second windows, 2-second windows, 1-second windows respectively.

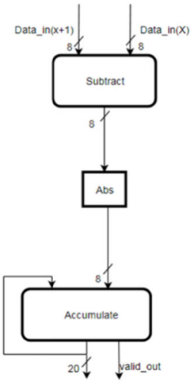
Table 3 lists the performance metrics obtained in response to changing the window size. With 4-second window the specificity significantly increases at the expense of a slight decrease in the sensitivity.

The proposed feature extractor is generic; by changing the window size parameter, the whole design is adjusted to be working with the new window size. The optimized feature extractor is implemented on Vertix-7 FPGA with 100 MHz clock. The block diagram of the proposed design is shown in Figure 7. Figure 7.d depict the standard deviation (STD) block diagram utilized in the Hurst Exponent (HE) block diagram shown in Figure 7.a. Fractal dimension (FD) feature is depicted in Figure 7.c; The inverse hyperbolic tan is applied to the output of the five accumulate windows sequentially, the five resulting values are then added up producing the output. Figure 7.b depicts the coastline (CL) feature.

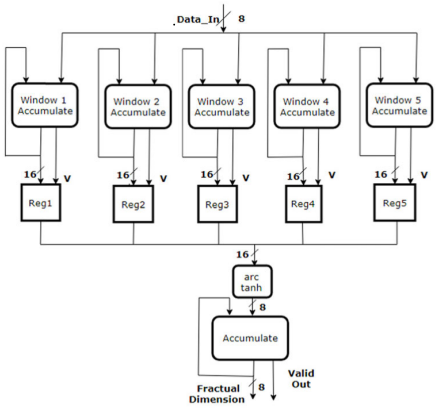
The resource utilization, power consumption, and latency are measured for the proposed design with different window sizes. Table 4 shows the obtained results. It is clearly shown that stretching the window size increases the utilization and consequently the power consumption and latency. However, the more stretched the window size, the higher the sensitivity.



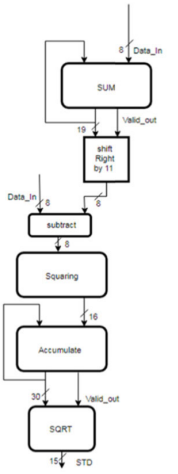
a. Hurst exponent feature



b coastline feature



c. feactal dimension feature



d. standard deviation exploited in Hurst exponent feature

FIGURE 7. Block diagram of the optimized feature extractor.

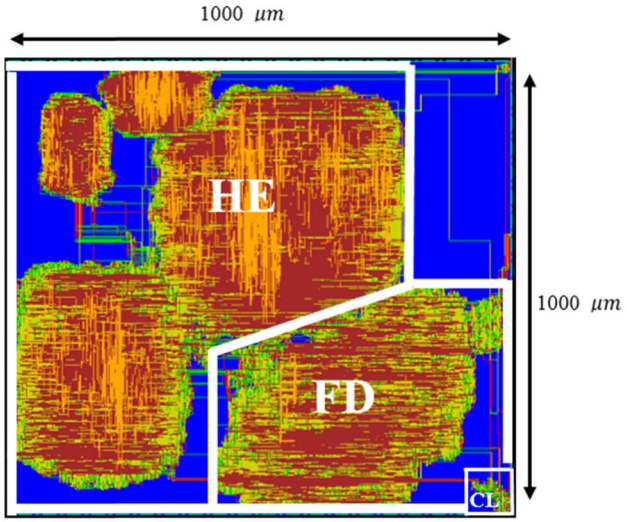


FIGURE 8. ASIC implementation for the optimized feature extractor.

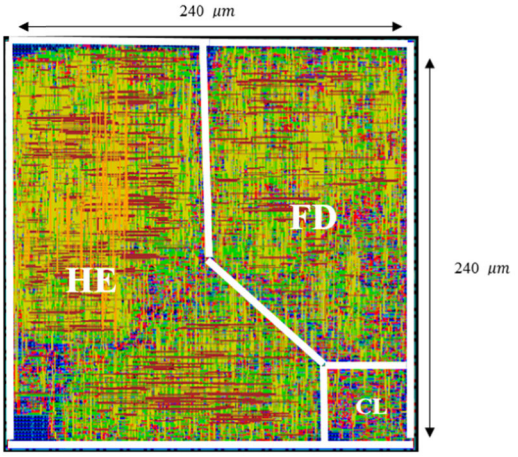


FIGURE 9. ASIC implementation of the approximate feature extractor.

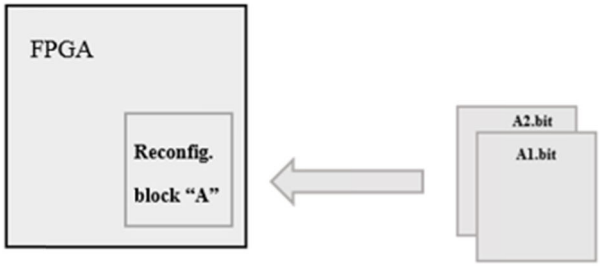
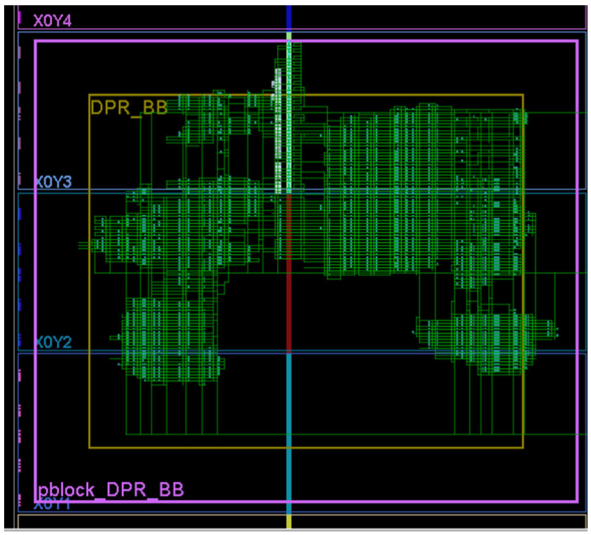
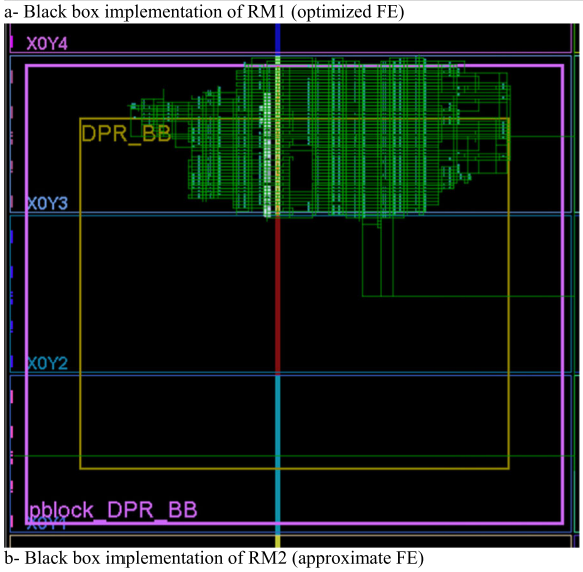


FIGURE 10. Dynamic reconfiguration between optimized and approximate feature extractor [17].

The ASIC implementation of the optimized feature extractor is shown in Figure 8. The design is implemented using UMC 65 nm CMOS technology. It utilizes an area of 0.9 mm^2. The operating frequency of the design is 91 MHz, and the power supply is 0.9 V. The total power consumption at the specified operating frequency and supply is 22.41 *mw*. Extra silicon area is added to solve connectivity errors. Detailed ASIC implementation results are depicted in Table 5.



a- Black box implementation of RM1 (optimized FE)



b- Black box implementation of RM2 (approximate FE)

FIGURE 11. Dynamic reconfiguration of optimized and approximate feature extractors.

B. APPROXIMATE FEATURE EXTRACTOR

After removing the division by the standard deviation replacing every natural logarithm by square root, the sensitivity correspondingly drops to 96.77% (a drop by 1.6%).The sensitivity, specificity, and accuracy obtained when the approximate feature extractor along with linear SVM classifier and SMO training accelerator are utilized for 4-second windows are 96.77%, 90.43%, and 90.42% respectively.

The effect of changing the window size on the performance metrics is studied by testing the design with input windows that varies in size from 1 second to 4 seconds. The performance metrics obtained in response to changing the window size are listed in Table 6. With 4-second window the specificity significantly increases at the expense of a slight decrease in the sensitivity.

Similar to the proposed optimized feature extractor, the proposed approximate feature extractor is generic; by changing the *window-Size* parameter, the whole design is

TABLE 4. FPGA implementation results for different widnow sizes.

	1-SECOND WINDOW (256 SAMPLE)	2- SECOND WINDOW (512 SAMPLE)	4-SECOND WINDOW (1024 SAMPLE)
Resource Utilization	5026 LUTs 2319 Slice Registers	5890 LUTs 2320 Slice Registers	7611 LUTS 2342 Slice Registers
Power Consumption	43.67 mW	47.78 mW	95.822 mW
Latency	8280 ns	15960 ns	31320 ns

TABLE 5. ASIC implementation results of the optimized feature extractor.

BLOCK	NUMBER OF CELLS	AREA (mm^2)	POWER (mW)
Optimized Feature Extractor	99051	0.9	22.41

TABLE 6. Performance metrics with different input window sizes for the approximate feature extractor.

WINDOW SIZE	SENSITIVITY	SPECIFICITY	ACCURACY
1 second (256 sample)	98.31%	90.11%	90.13%
2 seconds (512 sample)	97.52%	90.97%	90.99%
4 seconds (1024 sample)	96.77%	90.41%	90.43%

adapted to be working with any window size. The proposed design is implemented on Vertix-7 FPGA with a clock frequency of 100 MHz. The resource utilization, power consumption, and latency associated with each window size are shown in Table 7. It is clearly shown that increasing the window size increases the utilization and consequently the power consumption and latency. As opposed to the optimized feature extractor, the sensitivity increases with smaller

TABLE 7. FPGA implementation results for the approximate feature extractor.

	1-SECOND WINDOW (256 SAMPLE)	2- SECOND WINDOW (512 SAMPLE)	4-SECOND WINDOW (1024 SAMPLE)
Resource Utilization	1859 LUTs 738 Slice Registers	2119 LUTs 576 Slice Registers	3319 LUTs 576 Slice Registers
Power Consumption	8.1 mW	9.1 mW	12.3 mW
Latency	7820 ns	15500 ns	30860 ns

TABLE 8. ASIC implementation results of the approximate feature extractor.

BLOCK	NUMBER OF CELLS	AREA (mm^2)	POWER (mW)
Approximate Feature Extractor	11171	0.057	2.55

TABLE 9. Dynamic reconfiguration results.

RECONFIGURABLE MODULE	UTILIZATION	DYNAMIC POWER
RM1 (Optimized FE)	8236 LUTs 2466 slice Registers	101 mW
RM2 (Approximate FE)	4132 LUTs 798 slice Registers	36 mW

window sizes in the approximate feature extractor. However, the highest specificity is achieved with the widest window size.

The ASIC implementation of the approximate feature extractor is shown in Figure 9. The operating frequency of the design is 91 MHz and the power supply is 0.9V. The total power consumption at the specified operating frequency is $2.5mW$. The design area is 0.057 mm^2, and detailed resource utilization and power consumption results are listed in Table 8.

TABLE 10. Performance comparision with prior work on different FPGA platforms.

	System blocks	FPGA	Latency #cycles	Area			Performance	
				LUTs	Registers	DSPs	Sensitivity	Specificity
[19]	Feature Extraction + Classification	Virtex - 5	242	2583 0.3KB memory		-	95.24%	N/R
[20]	Feature Extraction	PYNQ-Z1	N/R	17492	16685	46	92.9%	96.3%
[14]	Feature Extraction + Classification	Zed-board	20156	8001	8622	31	98.4%	N/R
[21]	Feature Extraction + classification	Microsemi Igloo	2563	1237 5.56 KB memory		-	92.51%	80.1%
Proposed Optimized system	Feature Extraction + Classification	Virtex - 7	3148	7849	7479	-	98.39%	92.37%
Proposed Approximate system	Feature Extraction + Classification	Virtex - 7	798	2133	875	-	98.31%	90.11%

V. DYNAMIC RECONFIGURATION

The tradeoff between the two proposed designs is sensitivity and specificity achieved versus area utilization and power consumption of the design. While the approximate design has significantly less area utilization and power consumption, its specificity suffers a drop by around 2% in comparison with the optimized design.

For maximum advantage the two designs can be exploited, however, When both designs are placed together the dynamic power consumption is 95 mW and the junction temperature is 125 (junction temperature exceeded).

The proposed power-aware system is achieved through Dynamic Partial Reconfiguration (DPR). Dynamic partial reconfiguration is a technique that allows the exploitation of certain FPGA resources for more than one design at the same time. The two designs can be configured to replace each other depending on the available power and the criticality of the situation. If the power level is high or the patient is doing a very critical activity (like driving), the optimized feature extractor is in control, and when the available power decrease below certain threshold or the patient is doing a less critical activity, the approximate feature extractor becomes responsible for the operation.

Virtex-7 board is used for testing the proposed prototype; it was found that the most suitable controller for the implemented design using this board is partial reconfiguration controller (PRC) based on the comparative study done in [15].

Figure 11 shows the implemented dynamic part of the design. The reconfigurable partition can be reconfigured during run time with the two different reconfigurable modules, the optimized feature extractor and the approximate feature extractor. The reconfigurable partition size is chosen to accommodate the optimized feature extractor, or the approximate feature extractor. Resources utilization and power consumption for each of the two reconfigurable modules are shown in Table 9. It is shown that the power consumption drops by around 64%, moreover, the junction temperature remains within the normal margins.

VI. COMPARISON WITH PRIOR WORK

The two proposed designs achieve a higher sensitivity and specificity then many work in the literature like [4, 7, 8]. Moreover, an FPGA-system level comparison between the proposed systems and other seizure detection systems is depicted in Table 10.

The highest achieved sensitivity by the proposed designs is higher than that in [19], [20], [21], [22], and [23] and approximately equals to the sensitivity achieved in [20]. No DSPs are utilized in the proposed design in opposition to [20], and [21]. Moreover, the utilization of the approximate feature extractor is less than the mentioned previous work [14], [19]–[21], while the utilization of the optimized feature extractor has less utilization than these systems in [14], [20], [21]. Besides the sensitivity and utilization, the proposed system maximum latency is significantly less than that in [24] and [25]. The comparison does not include the power consumption because every work has its own operating frequency, thus the utilization only is compared which gives an insight that the power consumed by the approximate feature extractor is significantly less than other designs. No work in the literature introduced power-aware systems for seizure detection where DPR is utilized.

VII. CONCLUSION

A high sensitivity low cost power-aware seizure detection system that can be exploited in a neural seizure detection applications is implemented on both FPGA and ASIC platforms. Two designs for the purpose of seizure detection are presented. Optimized design that has significantly high performance on the expense of area and power consumption, and approximate design that saves more than 50% the power and area reserved for the optimized design, but suffers 2% decrease in the specificity achieved.

A power-aware system is achieved through dynamic partial reconfiguration between the two designs saving around 64% of the power consumption to ensure long battery life time of the implemented chip while keeping the highest possible performance. The implemented system exploits fractal dimension, Hurst exponent, and coastline for feature extraction, in addition to support vector machine training and classification with linear kernel. The system is tested with SMO training accelerator through a sliding windows varying from 1 to 4 seconds. The maximum achieved specificity of the proposed system is 92.14%, while the highest achieved sensitivity obtained is 98.39% at 100 MHz operating frequency.

REFERENCES

[1] K. Devarajan, S. Bagyaraj, V. Balasampath, J. E., and J. K., "EEG-based epilepsy detection and prediction," *Int. J. Eng. Technol.*, vol. 6, no. 3, pp. 212–216, 2014.

[2] D. Schmidt, "Schmidt DTwo antiepileptic drugs for intractable epilepsy with complex-partial seizures," *J. Neurol., Neurosurg. Psychiatry*, vol. 45, pp. 1119–1124, Oct. 1982.

[3] A. Varsavsky, I. Mareels, and M. Cook, *Epileptic Seizures and the EEG: Measurement, Models, Detection and Prediction*. Boca Raton, FL, USA: CRC Press, 2016.

[4] J. Gotman, "Automatic seizure detection: Improvements and evaluation," *Electroencephalogr. Clin. Neurophysiol.*, vol. 76, no. 4, pp. 317–324, 1990.

[5] J. F. Kerrigan, B. Litt, R. S. Fisher, S. Cranstoun, J. A. French, D. E. Blum, M. Dichter, A. Shetter, G. Baltuch, J. Jaggi, S. Krone, M. Brodie, M. Rise, and N. Graves, "Electrical stimulation of the anterior nucleus of the thalamus for the treatment of intractable epilepsy," *Epilepsia*, vol. 45, no. 4, pp. 346–354, Apr. 2004.

[6] L. Feng, Z. Li, and Y. Wang, "VLSI design of SVM-based seizure detection system with on-chip learning capability," *IEEE Trans. Biomed. Circuits Syst.*, vol. 12, no. 1, pp. 171–181, Dec. 2017.

[7] Q. Yuan, W. Zhou, S. Li, and D. Cai, "Epileptic EEG classification based on extreme learning machine and nonlinear features," *Epilepsy Res.*, vol. 96, nos. 1–2, pp. 29–38, Sep. 2011.

[8] S. Li, W. Zhou, Q. Yuan, S. Geng, and D. Cai, "Feature extraction and recognition of ictal EEG using EMD and SVM," *Comput. Biol. Med.*, vol. 43, no. 7, pp. 807–816, Aug. 2013.

[9] Gammerman and V. Vovk, "Alexey Chervonenkis's bibliography: Introductory comments," *J. Mach. Learn. Res.*, vol. 16, pp. 2051–2066, Oct. 2015.

[10] S. Hassan, S. Attia, K. N. Salama, and H. Mostafa, "EANN: Energy adaptive neural networks," *Electronics*, vol. 9, no. 5, p. 746, May 2020.

[11] H. E. Yantár, A. M. Eltawil, and K. N. Salama, "Efficient acceleration of stencil applications through in-memory computing," *Micromachines*, vol. 11, no. 6, p. 622, Jun. 2020.

[12] C. Zhang, M. B. Altaf, and J. Yoo, "Design and implementation of an on-chip patient-specific closed-loop seizure onset and termination detection system," *IEEE J. Biomed. Health Inform.*, vol. 20, no. 4, pp. 996–1007, Sep. 2016.

[13] H. Elhosary, M. H. Zakhari, M. A. Elgammal, M. A. Abd El Ghany, K. N. Salama, and H. Mostafa, "Low-power hardware implementation of a support vector machine training and classification for neural seizure detection," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 6, pp. 1324–1337, Dec. 2019.

[14] H. Wang, W. Shi, and C.-S. Choy, "Hardware design of real time epileptic seizure detection based on STFT and SVM," *IEEE Access*, vol. 6, pp. 67277–67290, 2018.

[15] H. Shoeb, "Application of machine learning to epileptic seizure onset detection and treatment," Ph.D. dissertation, Harvard Univ.-MIT Division Health Sci. Technol., Massachusetts Inst. Technol., Cambridge, MA, USA, 2009.

[16] M. Elgammal, O. Elkhouly, H. Elhosary, M. Elsayed, A. Mohieldin, K. N. Salama, and H. Mostafa, "Linear and nonlinear feature extraction for neural seizure detection," in *Proc. IEEE 61st Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Oct. 2018, pp. 795–798.

[17] A. Kamaleldin, A. Mohamed, Y. Nagy, A. Gamal, H. Shalash, and J. Mostafa, "Design guidelines for the high-speed dynamic partial reconfiguration based software defined radio implementations on xilinx Zynq FPGA," in *Proc. IEEE Int. Symp. Circuits Syst. (ISCAS)*, May 2017, pp. 1–4.

[18] S. Soliman, M. A. Jaela, A. M. Abotaleb, Y. Hassan, M. A. Abdelghany, A. T. Abdel-Hamid, K. N. Salama, and H. Mostafa, "FPGA implementation of dynamically reconfigurable IoT security module using algorithm hopping," *Elsevier Integr. VLSI J.*, vol. 68, pp. 108–121, Mar. 2019.

[19] A. Page, C. Sagedy, E. Smith, N. Attaran, T. Oates, and T. Mohsenin, "A flexible multichannel EEG feature extractor and classifier for seizure detection," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 62, no. 2, pp. 109–113, Feb. 2015.

[20] D. Serasinghe, D. Piyasena, and A. Pasqual, "A novel low-complexity VLSI architecture for an EEG feature extraction platform," in *Proc. 21st Euromicro Conf. Digit. Syst. Design (DSD)*, Prague, Czech Republic, Aug. 2018, pp. 472–478.

[21] T. Zhan, S. Z. Fatmi, S. Guraya, and H. Kassiri, "A resource-optimized VLSI implementation of a patient-specific seizure detection algorithm on a custom-made 2.2 cm^2 wireless device for ambulatory epilepsy diagnostics," *IEEE Trans. Biomed. Circuits Syst.*, vol. 13, no. 6, pp. 1175–1185, Dec. 2019.

[22] Y. Elhazek, A. Ibrahim, M. Amer, A. Abubakr, and H. Mostafa, "Hardware accelerated epileptic seizure detection system using support vector machine," in *Proc. 8th Int. Conf. Modern Circuits Syst. Technol. (MOCAST)*, Thessaloniki, Greece, May 2019, pp. 1–4.

[23] M. A. Elgammal, H. Mostafa, K. N. Salama, and A. Nader Mohieldin, "A comparison of artificial neural Network(ANN) and support vector Machine(SVM) classifiers for neural seizure detection," in *Proc. IEEE 62nd Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Dallas, TX, USA, Aug. 2019, pp. 646–649.

[24] Y. S. Park, G. R. Cosgrove, J. R. Madsen, E. N. Eskandar, L. R. Hochberg, S. S. Cash, and W. Truccolo, "Early detection of human epileptic seizures based on intracortical microelectrode array signals," *IEEE Trans. Biomed. Eng.*, vol. 67, no. 3, pp. 817–831, Mar. 2020.
[25] P. Boonyakitanont, A. Lek-uthai, and J. Songsiri, "Automatic epileptic seizure onset-offset detection based on CNN in scalp EEG," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Barcelona, Spain, May 2020, pp. 1225–1229.

HEBA ELHOSARY received the B.Sc. degree in electronic engineering from the Faculty of Information Engineering and Technology, German University in Cairo, New Cairo, Egypt, in 2018. She did her bachelor project/thesis at Ruhr-University Bochum, Bochum, Germany, in 2017. She is currently a TA with the Department of Electronics, Faculty of Information Engineering and Technology, German University in Cairo. Her research interests include digital design, computer arithmetic, computer architecture, FPGAs, VLSI, and ASIC.

MICHAEL H. ZAKHARI received the B.Sc. degree (Hons.) in electronic and communication engineering from the Faculty of Engineering, Ain Shams University, Cairo, Egypt, in 2018. He is currently pursuing the M.Sc. degree in electronic and communication engineering with the Faculty of Engineering, Cairo University. His research interests include digital electronics, VLSI, FPGA, and ASIC platforms.

MOHAMED A. ELGAMMAL received the B.Sc. and M.A.Sc degrees (Hons.) in electronics engineering from Cairo University, Cairo, Egypt, in 2016 and 2018, respectively. He is currently pursuing the Ph.D. degree with the ECE Department, University of Toronto, Toronto, ON, Canada. He was a Teaching and Research Assistant with Cairo University. He was a QA FPGA Prototyping Engineer with Mentor Graphics, Egypt. His research interests include FPGA, digital electronics, machine learning, and hardware implementation.

KHALED A. HELAL KELANY received the B.Eng. degree in electronics and electrical communications and the M.Sc. degree from Cairo University, in 2013 and 2016, respectively. He is currently pursuing the Ph.D. degree with the Department of Electrical and Computer Engineering, University of Victoria. He joined the Department of Electrical and Computer Engineering, University of Victoria, in 2017. His research interests include quantum computing, artificial neural networks, parallel computing, and computer architecture.

MOHAMED A. ABD EL GHANY received the B.S. degree (Hons.) in electronics and communications engineering and the master's degree in electronics engineering from Cairo University, Cairo, Egypt, in 2000 and 2006, respectively, and the Ph.D. degree in high-performance VLSI/IC design from German University in Cairo (GUC), New Cairo, Egypt, in 2010. From 2003 to 2006, he was with the National Space Agency of Ukraine's EGYPTSAT-1 Project. From 2008 to 2010, he was an International Scholar with the Electrical Engineering Department, Ohio State University, Columbus, OH, USA. From 2012 to 2014, he was awarded the Alexander von Humboldt Foundation Postdoctoral Fellowship at Technische Universität (TU) Darmstadt, Darmstadt, Germany. He is currently an Associate Professor with GUC. He is a Project Manager of four international projects between TU Darmstadt, Ruhr-University Bochum, and GUC. He is the author of about 55 articles, two book chapters, and two books in the fields of high throughput and low-power VLSI/IC design and NoC/SoC. His research interests include low-power design, embedded system design, network on chip design and related circuit level issues in high-performance VLSI circuits, clock distribution network design, digital ASIC design, and SoC/NoC design and verification. He is a reviewer and a program committee member of many IEEE international journals and conferences.

KHALED N. SALAMA (Senior Member, IEEE) received the B.S. degree from the Department of Electronics and Communications, Cairo University, Cairo, Egypt, in 1997, and the M.S. and Ph.D. degrees from the Department of Electrical Engineering, Stanford University, Stanford, CA, USA, in 2000 and 2005, respectively.

He was an Assistant Professor with the Rensselaer Polytechnic Institute, Troy, NY, USA, from 2005 to 2009. In 2009, he joined the King Abdullah University of Science and Technology (KAUST), Thuwal, Saudi Arabia, where he is currently a Professor, and was the Founding Program Chair, until 2011. He is the Director of the sensors initiative, a consortium of nine universities, including KAUST, MIT, UCLA, GATECH, MIT, UCLA, Brown University, Georgia Tech, TU Delft, Swansea University, the University of Regensburg, and the Australian Institute of Marine Science. He has authored 250 articles and 20 issued U.S. patents on low-power mixed-signal circuits for intelligent fully integrated sensors and neuromorphic circuits using memristor devices. His work on CMOS sensors for molecular detection has been funded by the National Institutes of Health and the Defense Advanced Research Projects Agency, awarded the Stanford-Berkeley Innovators Challenge Award in biological sciences and was acquired by Lumina Inc.

HASSAN MOSTAFA (Senior Member, IEEE) received the B.Sc. and M.Sc. degrees (Hons.) in electronics engineering from Cairo University, Cairo, Egypt, in 2001 and 2005, respectively, and the Ph.D. degree in electrical and computer engineering from the Department of Electrical and Computer Engineering, University of Waterloo, Waterloo, ON, Canada, in 2011. He was an NSERC Postdoctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. He is currently an Associate Professor with the Nanotechnology and Nanoelectronics Program, Zewail City of Science and Technology, Giza, Egypt—on leave from the Department of Electronics and Electrical Communications, Cairo University. His postdoctoral work includes the design of the next generation FPGA in collaboration with the Fujitsu Research Labs in Japan/USA. He has authored/coauthored more than 170 papers in international journals and conferences and is the author/coauthor of five published books. His research interests include neuromorphic computing, the IoT hardware security, software-defined radio, reconfigurable low power systems, analog-to-digital converters, low-power circuits, subthreshold logic, variation-tolerant design, soft error tolerant design, statistical design methodologies, next generation FPGA, spintronics, memristors, energy harvesting, MEMS/NEMS, power management, and optoelectronics. He has been a member of the IEEE Technical Committee of VLSI Systems and Applications, since 2017. He was a recipient of and the University of Waterloo Sandford Fleming TA Excellence Award, in 2008, the Ontario Graduate Scholarship, in 2009, the Waterloo Institute of Nano-Technology Nano Fellowship Research Excellence Award, in 2010, the Natural Sciences and Engineering Research Council of Canada Prestigious Postdoctoral Fellowship, in 2011, and the University of Toronto Research Associate Scholarship, in 2012.