MDPI

Article

General Methodology for the Design of Bell-Shaped Analog-Hardware Classifiers

Vassilis Alimisis *, Nikolaos P. Eleftheriou, Argyro Kamperi, Georgios Gennis, Christos Dimas and Paul P. Sotiriadis *

Department of Electrical and Computer Engineering, National Technical University of Athens, 15773 Athens, Greece; eleftheriou_nikos@hotmail.com (N.P.E.); argykaberi@gmail.com (A.K.); giorgosyennis@gmail.com (G.G.); chdim@central.ntua.gr (C.D.)
* Correspondence: alimisisv@gmail.com (V.A.); pps@ieee.org (P.P.S.)

Abstract: This study introduces a general methodology for the design of analog integrated bell-shaped classifiers. Each high-level architecture is composed of several Gaussian function circuits in conjunction with a Winner-Take-All circuit. Notably, each implementation is designed with modularity and scalability in mind, effectively accommodating variations in classification parameters. The operating principles of each classifier are illustrated in detail and are used in low-power, low-voltage, and fully tunable implementations targeting biomedical applications. The realization of this design methodology occurred within a 90 nm CMOS process, leveraging the Cadence IC suite for both electrical and layout design aspects. In the verification phase, post-layout simulation outcomes were meticulously compared against software-based implementations of each classifier. Through the simulation results and comparison study, the design methodology is confirmed in terms of accuracy and sensitivity.

Keywords: analog VLSI; general design methodology; biomedical applications; wake-up circuit; analog classifiers

1. Introduction

The rapid growth of the Internet of Things (IoT) is leading to the proliferation of devices and sensors that often operate solely on batteries [1,2]. Numerous consumer and industrial applications rely on IoT devices, some of which lack online recharging capabilities. Consequently, hardware designers are increasingly dependent on power management solutions to effectively handle the required power. The advancement of technology has facilitated the integration of an increasing number of sensors into modern smart sensor systems. This progress has led to promising developments in miniaturization and power efficiency, enabling these systems to sense a wide range of physical variables [3]. Integrated circuit (IC) technologies have played a crucial role in addressing the challenges faced by smart sensor systems. These ICs are complex but designed to be power- and area-efficient.

Given the dependency on batteries and the need for efficient use of the area, there is a growing need for new computing paradigms [1]. Edge computing, which includes analog computing [4], holds great promise for power-hungry systems with high latency. By exploiting the physical laws describing the transistor, such as analog translinear circuits, it becomes possible to approximate various mathematically established models [4]. The utilization of the sub-threshold region, along with these advantages, leads to the development of architectures that are more power efficient [5]. Furthermore, analog ICs can perform high-performance computations based on the physical laws of MOS or BJT transistors [4,6].

Biomedical engineering is a widely researched area within the IoT domain [7]. Smart wearable sensors based on IoT technology offer a cost-effective, reliable, and energy-efficient solution for clinical patient monitoring and disease detection [8]. For instance, thyroid disease prediction or detection has recently emerged as an important task. The

thyroid gland, resembling the shape of a butterfly, is situated in the front of the neck [9,10]. More specifically, it is a small organ located below the Adam's apple surrounding the trachea (windpipe). Functioning as part of the endocrine system, which comprises glands responsible for producing, storing, and releasing hormones into the bloodstream, the thyroid plays a crucial role in regulating metabolism and various metabolic processes [9,10]. Additionally, it contributes to muscle control, brain development, mood regulation, and digestive function.

Thyroid disorders refer to conditions that affect the normal functioning of this gland. Abnormal hormone production by the thyroid can lead to a range of issues [11]. The two most prevalent types of thyroid disease are hyperthyroidism (excessive hormone production) and hypothyroidism (insufficient hormone production). Other abnormal conditions associated with the thyroid include thyroiditis, thyroid nodules, goiter, and thyroid cancer. The specific form of thyroid disease detected determines the available treatment options, which may involve medications, radioactive iodine, and, in some cases, surgery.

Another interesting application is epileptic seizure prediction [12,13]. An epileptic seizure occurs due to a sudden surge of neural activity or an electrical disruption within the brain [14,15]. Individuals diagnosed with epilepsy display symptoms that can vary from hardly noticeable to severe, potentially even leading to fatality. The unpredictable frequency and nature of these seizures significantly impact the overall quality of life for patients; therefore, improving it involves the anticipation and early warning of impending epileptic episodes. Accurately predicting an upcoming seizure could empower individuals to take necessary precautions and prevent engaging in hazardous activities like driving.

The anticipation of epileptic seizures relies on the analysis of patients' well-being through the utilization of bio-signal acquisition techniques. Epileptic seizures consist of four distinct phases: pre-ictal, ictal, post-ictal and inter-ictal [14,15] The first three phases correspond to the periods shortly preceding, during, and immediately following a seizure, respectively, while the fourth phase pertains to the interval between two seizures during which the patient's condition returns to a normal state. According to the findings outlined in [16], the duration of both the pre-ictal and post-ictal phases fluctuates between 30 min and 2 h. The precise and real-time recognition of the pre-ictal phase holds immense significance, as it is equivalent to successfully predicting an imminent seizure.

Motivated by the need for low-power smart biosensors [17,18], we combine sub-threshold-based analog computing techniques with machine learning (ML) ones [5]. To this end, in this work, a general methodology for the design of low-voltage (0.6 V), low-power (less than 44.7 μW) bell-shaped analog classifiers (CLFs) is introduced and tested on real-world biomedical classification problems. It is realized based on a Bayesian mathematical model [19] using two main sub-circuits. Specifically, the employed main building blocks are ultra-low-power Gaussian function circuits [20] and an argmax operator circuit [21]. Post-layout simulation results are conducted on a TSMC 90 nm CMOS process using the Cadence IC suite and compared with a software-based implementation. Furthermore, the architecture's effectiveness is validated through Monte Carlo analysis, affirming its sensitivity and performance.

The remainder of this paper is organized as follows. Section 2 refers to the background of this work. More specifically, the literature review and the mathematical model are analyzed. The proposed design methodology along with the high-level architecture of the analog CLFs are presented in Section 3. In Section 4, the main building blocks of the analog bell-shaped CLFs are presented. The validation of the proposed design methodology is carried out using real-world biomedical datasets in Section 5. This section also includes a comparison between the hardware and software implementations along with sensitivity tests. A comparison study and discussion are provided in Section 6. Finally, Section 7 presents concluding remarks summarizing the findings and implications of this study.

2. Background

2.1. Literature Review

The global landscape is saturated with an abundance of diverse data forms—text, images, videos, and more—showing no signs of slowing down in the near future [22,23]. ML offers the tantalizing prospect of extracting significance from this vast data expanse. ML, an interdisciplinary domain, interweaves with mathematical fields such as statistics, information theory, game theory, and optimization [19,24]. This amalgamation of tools and technology constitutes a means to effectively process this data deluge. Furthermore, automated techniques or algorithms are able to discern meaningful patterns or hypotheses that might evade human observation. While these algorithms are traditionally executed through software, a trend has surfaced where hardware-friendly implementations are pursued to realize these algorithms and models [25,26].

Three distinct hardware design strategies exist, each with its own merits and drawbacks. These strategies encompass analog, digital, and mixed-mode implementations. Digital circuits, commonly employed in ML applications, boast advantages in achieving elevated classification accuracy, adaptability, and programmability. Nevertheless, they exhibit substantial power consumption and spatial requirements due to intensive data transactions and rapid operations. Conversely, dedicated analog-hardware ML architectures facilitate cost-effective parallelism via low-power computation, yet imprecise circuit parameters arising from noise and limited precision undermine accuracy. A number of mixed-mode architectures exploit both analog and digital methodologies to attain reduced power consumption and compact footprints; however, these solutions contend with overhead costs related to domain conversion.

Dedicated analog-hardware architectures geared toward ML algorithms and models utilize circuits based on Gaussian functions. The salient attributes of system-level implementations, coupled with Gaussian function circuitry, are summarized in this subsection. Proposed ML systems encompass radial basis function neural networks (RBF NNs) [27–37] with a general design framework, in addition to other neural networks like multi-layer perceptron (MLP); radial basis function network (RBFN) [32,38]; Gaussian RBF NN (GRBF NN) [39,40]; Gaussian mixture model (GMM) [41]; Bayesian [42]; K-means-based [43]; and voting [44], fuzzy [45], and threshold [46] and centroid [47] classifiers. Support vector machine (SVM) [48–50], support vector regression (SVR) [51], domain description (SVDD) [52] algorithms, pattern-matching classifiers [53,54], vector quantizers [55,56], a deep ML (DML) engine [57], a similarity evaluation circuit [58], a long short-term memory (LSTM) [59–62] and a self-organizing map (SOM) [63] comprise other instances. Gaussian function circuits serve as the bedrock for implementing two pivotal functions beneficial to myriad ML algorithms: (a) kernel density and (b) distance computation. A majority of these applications cater to input dimensions lower than 65 dimensions, with select cases omitting an upper limit specification [28,38,39,55], thus accommodating high-definition image classification.

2.2. Mathematical Model

In this subsection, the mathematical model of the Gaussian Mixture Model (GMM) is presented. This theory is comprehensive and offers a step-by-step approach to modeling this general methodology. The remaining models are founded on modifications to this foundational one, as well as potential simplifications. Hence, it is recommended that the reader refer to the pertinent streamlined theory for the implementation of a particular model. The GMM introduces a novel approach to characterizing the probability density of an N-dimensional random variable. It does so by presenting the variable as a meticulously weighted aggregation of K Gaussian densities. This ingenious formulation not only surpasses the expressiveness of a singular Gaussian (normal distribution) [19,64], but also broadens its range of applications. Each instance of a GMM, denoted as λ_c, is uniquely identified by several key parameters. These encompass the component count K, the weight factors $[w_i^c]_{i=1}^K$, the mean value vectors $[\mathbf{M}_i^c]_{i=1}^K$ (where $\mathbf{M}_i^c \in \mathbb{R}^N$), and the covariance matrices $[\mathbf{\Sigma}_i^c]_{i=1}^K$ (where $\mathbf{\Sigma}_i^c \in \mathbb{R}^{N \times N}$) corresponding to each distinct Gaussian constituent.

For further clarity, consider an input vector $\mathbf{X}$ with dimensions in N denoted as $\mathbf{X} \in \mathbb{R}^N$. The derivation of the probability density function (PDF) of $\mathbf{X}$ via the utilization of λ_c is explicitly elaborated in [64]:

$$p(\mathbf{X}|\lambda_c) = \sum_{i=1}^{K} w_i^c \cdot \mathcal{N}(\mathbf{X}|\mathbf{M}_i^c, \mathbf{\Sigma}_i^c). \tag{1}$$

In this context, it is imperative to observe that the conditions $\sum_{i=1}^{K} w_i^c = 1$ and $0 \leq w_i^c \leq 1$ for $i = 1, 2, \ldots, K$ are upheld. Designating the i-th N-dimensional Gaussian component within λ_c as $\mathcal{N}(\mathbf{X}|\mathbf{M}_i^c, \mathbf{\Sigma}_i^c)$, its specific numerical expression is determined as follows:

$$\mathcal{N}(\mathbf{X}|\mathbf{M}_i^c, \mathbf{\Sigma}_i^c) = \frac{e^{-\frac{1}{2} \cdot (\mathbf{X}-\mathbf{M}_i^c)^T \cdot (\mathbf{\Sigma}_i^c)^{-1} \cdot (\mathbf{X}-\mathbf{M}_i^c)}}{\sqrt{(2\pi)^N |\mathbf{\Sigma}_i^c|}}, \tag{2}$$

where $|.|$ denotes the Euclidean norm. The expression provided above can be streamlined in the case of a diagonal matrix $\mathbf{\Sigma}_i^c$ as follows:

$$\mathcal{N}(\mathbf{X}|\mathbf{M}_i^c, \mathbf{\Sigma}_i^c) = \prod_{n=1}^{N} \mathcal{N}(x_n | \mu_n^c, (\sigma_n^c)^2), \tag{3}$$

In this context, $(\sigma_n^c)^2$ denotes the squared value of x_n and μ_n^c represents a scalar taken as the n-th component of vectors $\mathbf{X}$ and $\mathbf{M}_i^c$. The (n, n)-th entry of the matrix $\mathbf{\Sigma}_i^c$ is referred to as well. The univariate Gaussian distribution for scalar inputs x_n is:

$$\mathcal{N}(x_n | \mu_n^c, (\sigma_n^c)^2) = \frac{1}{\sqrt{(2\pi) \cdot (\sigma_n^c)^2}} e^{-\frac{1}{2} \cdot \frac{(x_n - \mu_n^c)^2}{(\sigma_n^c)^2}}. \tag{4}$$

When GMMs are applied in an unsupervised manner, they are adept at identifying distinctive clusters within the analyzed dataset. This intrinsic capability aligns well with tasks involving clustering. However, in scenarios involving classification, as is the focus of the proposed framework, a plurality of GMMs is harnessed. Within this context, each class is associated with an individual GMM exclusively responsible for clustering data pertaining to that specific class. The selection of the optimal number of components is determined by the intricacies of the distribution inherent in the dataset. For any given input vector $\mathbf{X}$ and a total of C classes, the posterior probabilities $p(\lambda_c|\mathbf{X})$ are computed individually for each GMM $[\lambda_c]_{c=1}^{C}$ using the principles outlined in the Bayes theorem:

$$p(\lambda_c|\mathbf{X}) = \frac{p(\lambda_c)p(\mathbf{X}|\lambda_c)}{p(\mathbf{X})}. \tag{5}$$

The term $p(\lambda_c)$ represents the prior probability while $p(\mathbf{X})$ denotes the probability of evidence. When conducting a comparison of the posterior probabilities associated with two different classes, the evidence probability becomes immaterial. This is due to its inherent independence from the selected class and its function as a mere normalization constant. Hence, the conclusive determination of the winning class is orchestrated by the overarching CLF through the following expression:

$$y = \underset{c \in [1,C]}{\mathrm{argmax}} \{ p(\lambda_c)p(\mathbf{X}|\lambda_c) \}. \tag{6}$$

3. Proposed Design Methodology

In this section, we delve into the overarching design of the bell-shaped CLFs that has been put forth. To shed light on the rationale underpinning this design, we consider a scenario involving the classification of N_{cla} distinct classes (CLF cells) and N_d inputs

in each CLF. In this general methodology, the CLF's high-level configuration involves a hyperparameter denoted as N_{clu}, which signifies the number of CLF sub-cells.

This parameter is determined through an exploratory analysis of the specific design methodology, showcasing the versatility of the proposed design. This adaptability extends to accommodating various input dimensions, classes, or clusters/centroids.

The structure of the suggested block for the analog bell-shaped CLF is depicted in Figure 1. According to the formulation of the classification problem described earlier (scenario), the CLF consists of a single Winner-Take-All (WTA) block with N_{cla} inputs and N_{cla} CLF cells. Each CLF cell comprises N_{clu} sub-cells, which describe either a cluster or a centroid (depending on the type of the implemented CLF). These sub-cells are essentially circuits representing multidimensional Gaussian functions with N_d inputs. Each cell calculates the probability of an input vector **X** belonging to a specific cluster/centroid by employing the Gaussian probability density function (PDF) of the sub-cell, as defined in the mathematical model.
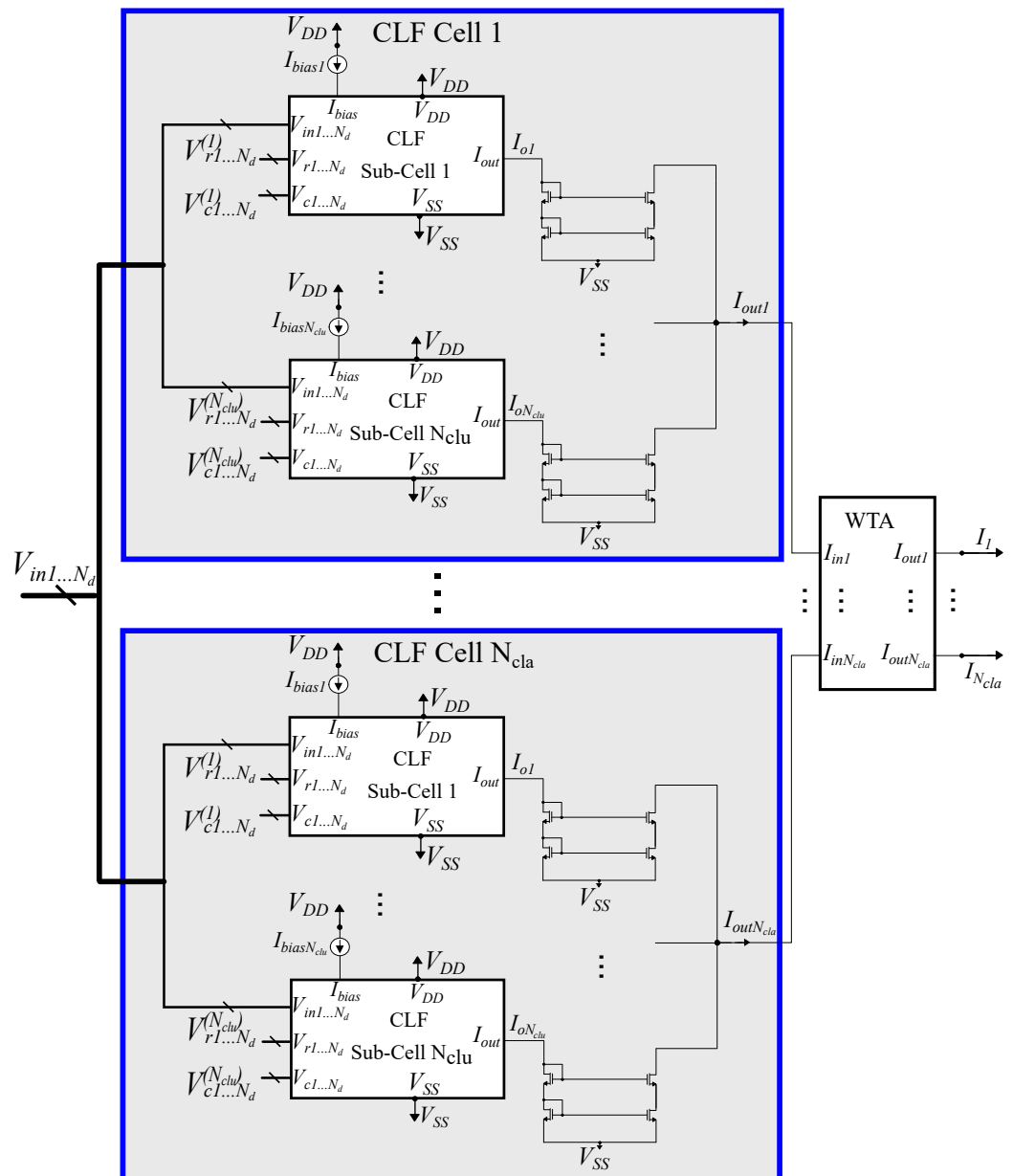

Figure 1. Analog bell-shaped classifier with N_{cla} classes, N_{clu} clusters/centroids, and N_d features. This is a conceptual design describing the general methodology.

In accordance with Equation (1), the probability of **X** being associated with a particular class is obtained by summing the probabilities of the sub-cells that constitute that class. To maintain accuracy and minimize potential distortions, this summation is executed within a CLF cell through the utilization of current mirrors.

Cascode current mirrors are used instead of high-precision ones, since they are compact and provide high accuracy for the application's requirements. The transistors' sizes for every cascode current mirror are $\frac{W}{L} = \frac{3.2\ \mu m}{1.6\ \mu m}$. The WTA block implements the argmax operator. Through Equation (6), it compares the probabilities of different classes and identifies the class with the highest probability (winning class). Furthermore, the employment of a conventional WTA circuit facilitates the identification of the winning class through a digital one-hot vector $[I_1, \ldots, I_{N_{cla}}]$, where the currents $[I_i]_{i=1}^{N_{cla}}$ are represented in binary format [21]. Consequently, the output of the entire CLF is in a digital format.

The employed foundational components introduce certain limitations concerning the upper bounds on the quantity of classes, clusters/centroids, and input dimensions. To elaborate, the range of permissible classes is constrained by the WTA circuit's capacity to effectively compare a substantial number of inputs. Correspondingly, augmenting the number of individual currents amalgamated at a node also augments undesired distortion. Consequently, the utmost number of clusters is confined by the fidelity of this summation process. The volume of input dimensions is contingent on the multidimensional Gaussian function circuit. While there exist various circuits that yield PDFs, extant literature confines their application to dimensions of a modest scale, typically less than 16 [20,65].

4. Circuit Implementation

The main building circuits for the implementation of bell-shaped CLFs are thoroughly analyzed in this section. Based on Section 3, each CLF cell requires two main blocks: a CLF cell and an argmax operator circuit which is called a Winner-Take-All (WTA) circuit. Moreover, each CLF cell requires two main cells: the sub-cells and cascode current mirrors. The sub-call cells are in fact multidimensional Gaussian function circuits with N_d inputs. By arranging a series of $N_d = N$ basic bump circuits in sequence, as shown in Figure 2, the final output of this arrangement provides the behavior of an N-dimensional Gaussian function. The individual parameters (V_r, V_c, I_{bias}) for each bump circuit are adjusted independently. This design methodology of the architecture focuses on the utilization of ultra-low-power circuits as foundational elements for constructing the primary cells. As a result, all transistors within the architecture operate within the sub-threshold range. The classification result is not affected by the current noise of the implemented circuits, since its worst-case value, as extracted after simulations, is less than 20 pA within the frequency operating range (<1 kHz). In order to enhance the CLF's relevance in scenarios where battery dependence is a concern, the power supply voltages are configured to be $V_{DD} = -V_{SS} = 0.3$ V.
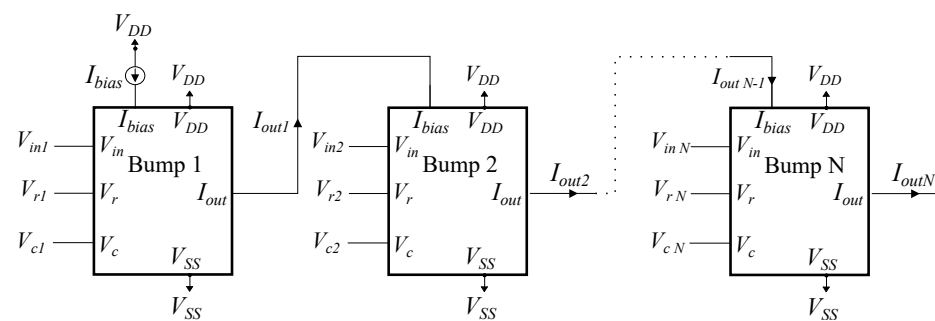

Figure 2. By arranging a series of *N* basic bump circuits in sequence, the final output of this arrangement provides the behavior of an N-dimensional Gaussian function. The individual parameters (V_r, V_c, I_{bias}) for each bump circuit are adjusted independently.

4.1. Gaussian Function Circuit

A broad spectrum of bump circuits has been developed for numerous applications [20]. However, for the purposes of this research, a modified version of the bump circuit (aspect ratio equal to 7) [65], as shown in Figure 3, has been employed to enhance the quality and resilience of the resulting Gaussian curve. Specifically, the adjusted circuit employs a symmetric current correlator (comprising transistors M_{p1}–M_{p6} in Figure 3) with a ratio of 2 instead of the non-symmetric version utilized in [41]. This modification is driven by the necessity for symmetric Gaussian curves when comparing two CLF prototypes. The adoption of a symmetric current correlator ensures that even minor currents maintain symmetry around the mean value, as illustrated in Figure 4. Moreover, with a ratio equal to 7, we achieve an increase in the linear region of the circuit (higher variance for the same V_c). Furthermore, to enhance mirroring performance even with low bias currents, a cascode current mirror consisting of transistors M_{n5}–M_{n10} (Figure 3) has been integrated. Details pertaining to the dimensions of the transistors are outlined in Table 1.
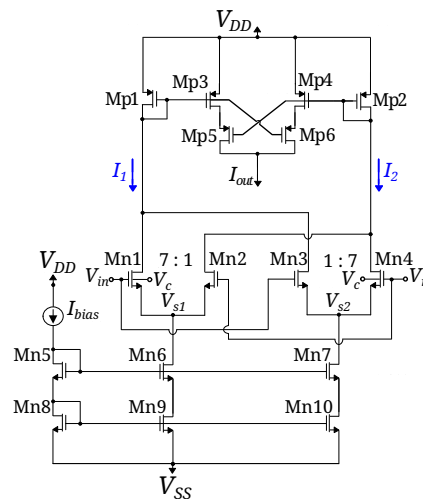

Figure 3. The utilized Gaussian function circuit is presented. The output current I_{out} resembles a Gaussian function controlled by the input voltage V_{in}. The parameter voltages V_r, V_c and the bias current I_{bias} control the Gaussian function's mean value, variance, and peak value, respectively.

Table 1. Bump circuit transistors' dimensions.

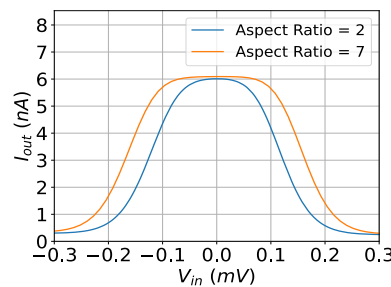| NMOS Differential Block	W/L (µm/µm)	Current Correlator	W/L (µm/µm)
M_{n1}, M_{n4}	2.8/0.4	M_{p1}, M_{p2}	1.6/1.6
M_{n2}, M_{n3}	0.4/0.4	M_{p3}, M_{p4}	0.4/1.6
$M_{n5}-M_{n8}$	0.4/1.6	M_{p5}, M_{p6}	0.4/1.6
M_{n9}, M_{n10}	1.6/1.6	-	-

Figure 4. The output current of the implemented bump circuit with respect to the aspect ratio of the input differential pair transistors. The simulation was conducted under $V_r = 0$ V, $V_c = 180$ mV and $I_{bias} = 6$ nA.

A multivariate Gaussian distribution as well as a corresponding multivariate Gaussian distance metric is formulated based on mathematical equations outlined in the background. In practical application, the serial connection of two or more bump circuits is akin to their multiplication [41]. The mean value and variance of each bump circuit are controlled by specific voltage parameters referred to as V_r and V_c. In this arrangement, the initial bump circuit incorporates an I_{bias}, which determines the peak of the Gaussian probability density function (PDF) and corresponds to the height of the Gaussian curve. Subsequent bump circuits are biased using the output current from the previous bump unit. An illustration of this cascade of bump circuits, facilitating the implementation of a multivariate distance function, can be observed in Figure 2. Nonetheless, a constraint of this design emerges as the count of bump cells within a cascaded implementation rises to accommodate high-dimensional data. In this scenario, the current scaling induced by I_{bias} does not exhibit a fully linear behavior. This deviation from linearity can be traced back to slight inaccuracies inherent in analog circuits. While these inaccuracies might have minimal impact on low-dimensional inputs, the cumulative effect becomes pronounced as more bump cells are linked in a series configuration. Consequently, the output current is notably influenced.

4.2. Winner-Take-All Circuit

The next circuit under consideration is the WTA circuit. To gain a comprehensive understanding of the modified WTA circuit implemented in this research, a concise analysis of the conventional Lazzaro WTA circuit is provided [21]. In the Lazzaro WTA circuit configuration, N_{cla} neurons are interconnected, sharing a common I_{bias} current as depicted in Figure 5. Each neuron corresponds to a distinct class and solely manages its input and output functions. Among these neurons, the one with the greatest input current generates a non-zero output equal to I_{bias} while the remaining neurons output zero. Instances involving similar input currents can result in multiple winners, a situation typically deemed undesirable in most classification scenarios.
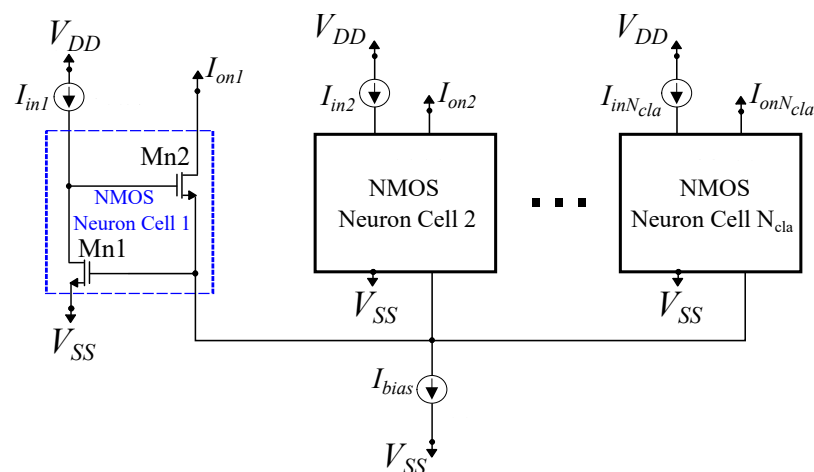

Figure 5. A N_{cla}-neuron standard Lazzaro NMOS Winner-Take-All (WTA) circuit.

A resolution to this challenge arises through the utilization of a cascaded WTA circuit, as depicted in Figure 6. The devised configuration integrates three WTA circuits interconnected in a cascaded manner [65]. The one-dimensional decision boundaries of the conventional Lazzaro WTA circuit and the proposed cascaded version are visualized in Figure 7. Notably, it becomes evident that the cascaded WTA circuit presents significantly steeper decision boundaries compared to the basic Lazzaro WTA circuit. As a result, the cascaded topology proves to be the appropriate choice for the essential argmax operation of the CLF. The dimensions of all transistors for the NMOS and PMOS neurons in Figure 6 are established at $W/L = 0.4\ \mu\text{m}/1.6\ \mu\text{m}$. The preference for extended transistors is guided by the requirement for reduced noise and enhanced linearity to effectively execute the argmax operator.
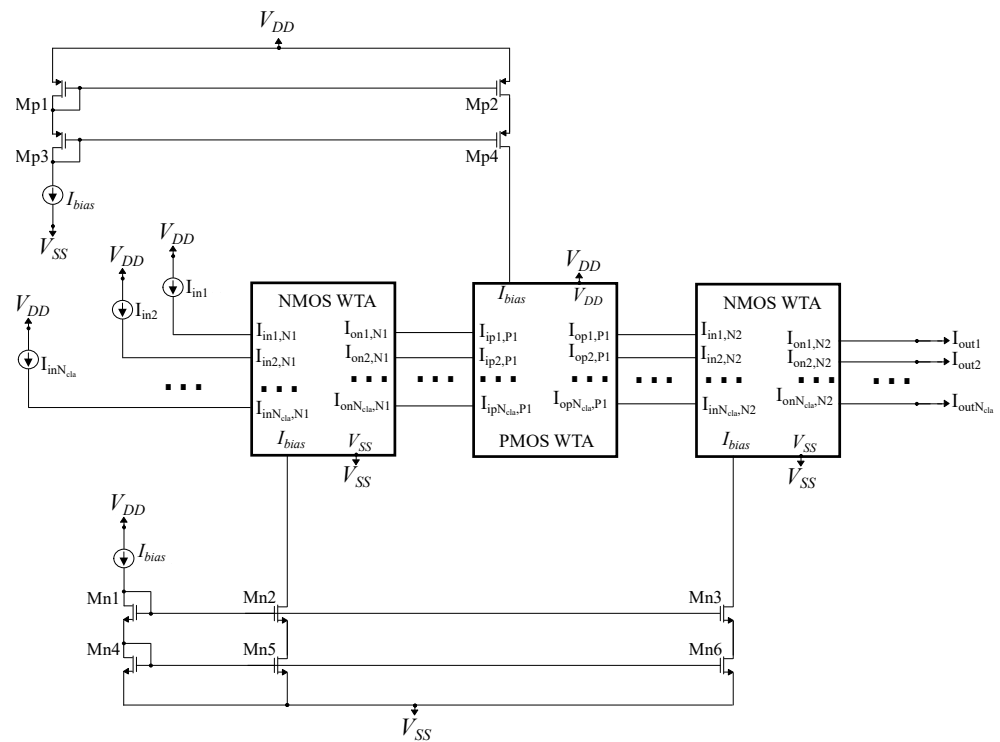
Figure 6. A cascaded NMOS-PMOS-NMOS WTA circuit. It is utilized to improve the performance of the standard WTA circuit.
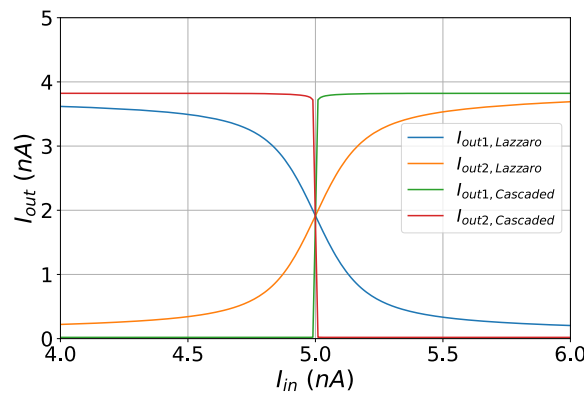

Figure 7. Decision boundaries of the standard and the cascaded WTA circuit.

5. Application Examples and Simulation Results

In this section, the proposed design methodology is tested on two real-world datasets: (a) the epilepsy seizure prediction problem from the CHB-MIT Scalp EEG database [12,13] and (b) thyroid disease detection from the University of California, Irvine (UCI) Machine Learning Repository [9] to confirm its proper operation. The design procedure has been designed using the Cadence IC suite in a TSMC 90 nm CMOS process. All simulation results are conducted on a single layout (post-layout simulations), which is shown in Figure 8. This layout, which integrates bell-shaped CLFs, has been meticulously designed with a primary emphasis on achieving area efficiency. Composed of three CLF cells (classes), each consisting of three sub-cells and with a total of five input dimensions, this layout is adept at accommodating all desirable bell-shaped CLFs (both datasets).
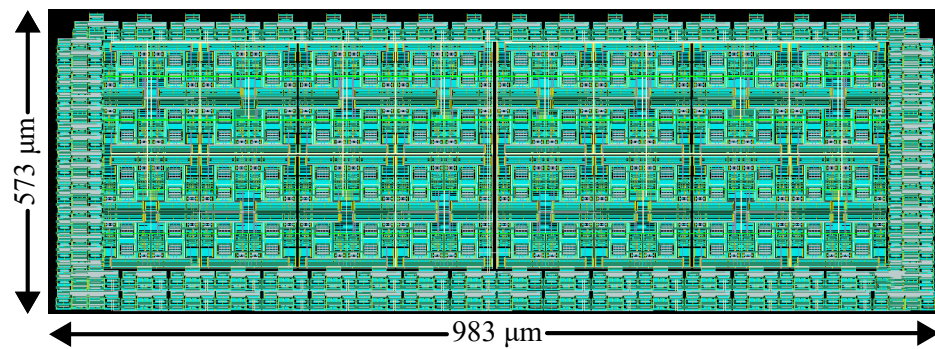
Figure 8. Layout related to the general design methodology. It combines all the implemented CLFs and extra switches in order to select the appropriate one.

The epilepsy seizure dataset [12,13] encompasses EEG signals obtained from children grappling with intractable epilepsy. This dataset has been meticulously labeled by expert physicians with the ictal periods being definitively identified. In this context, the scope of pre-ictal and post-ictal spans an hour before and an hour after the seizure correspondingly. Data instances that do not correspond to the ictal, pre-ictal, or post-ictal intervals are categorized as inter-ictal occurrences. For classification purposes, the system leverages four distinct features; the peak-to-peak voltage and energy percentages in the alpha, the first half of the gamma, and the second half of the gamma frequency bands [66]. These features can be effectively extracted from the raw EEG signals using analog feature extraction methodologies [6,67].

Prior to the operational deployment of the circuit, the system's essential parameters are ascertained through software-based training. The overarching objective of this classifier is to adeptly discriminate between the pre-ictal and inter-ictal periods. To function effectively as a low-power front-end wake-up circuit, it is imperative that the circuit accurately predicts all potential seizures while simultaneously minimizing the occurrence of false positive alarms. This dataset is a binary class problem with four features.

The second dataset is sourced from the University of California, Irvine (UCI) Machine Learning Repository [9]. It comprises blood test metrics associated with thyroid conditions, specifically normal thyroid function, hypothyroidism, and hyperthyroidism. These metrics are directly fed into the classifier for analysis. To establish the classifier's operational parameters key metrics, including the mean value, variance, and prior probability of each class are calculated. This dataset is a three-class problem with five features.

To underscore the advantages introduced by the proposed design methodology a comprehensive examination is undertaken through two distinct tests for each CLF. In the initial test, a comparative assessment of classification accuracy is performed among the hardware implementation and a software-based one. To mitigate the influence of stochastic variability stemming from the training algorithm, a total of 20 independent software-based training iterations are executed to derive the essential CLF parameters. Importantly, across all iterations, all CLFs are tested with the same parameters to ensure equitable evaluation. In the second test, a Monte Carlo simulation with $N = 100$ data points is executed to evaluate the proposed architecture's sensitivity behavior. The CLF's parameters under consideration in this instance are chosen to be one of the 20 candidates previously established in the first test. Regarding the epileptic seizure prediction all the CLFs successfully predict all 17 seizures (100% sensitivity) of the test set.

5.1. GMM CLF Implementation and Simulation Results

In this subsection, the implemented GMM-based CLF and its simulation results for both datasets are provided. Based on the proposed design methodology and the simulation result of the software implementation, the generic GMM-based CLF is composed of $N_{Cla} = 3$ classes, $N_{Clu} = 2$ clusters per class, and $N_d = 5$ input dimensions. The high-level

architecture of this CLF is depicted in Figure 9. More specifically, each class is composed of two 5-D Gaussian function circuits, which correspond to the two clusters and two current mirrors that are used to add the output currents of each cluster. In Figures 10 (thyroid) and 11 (epilepsy), the classification accuracy for both implementations (hardware/software) is presented, encompassing a total of 20 distinct training test cases (first case) for both datasets. The results are also summarized in Tables 2 and 3. Regarding the sensitivity, the Monte Carlo histogram shown in Figure 12 has a mean value of $\mu_M = 94.85\%$ and a standard deviation of $\sigma_M = 2.31\%$.
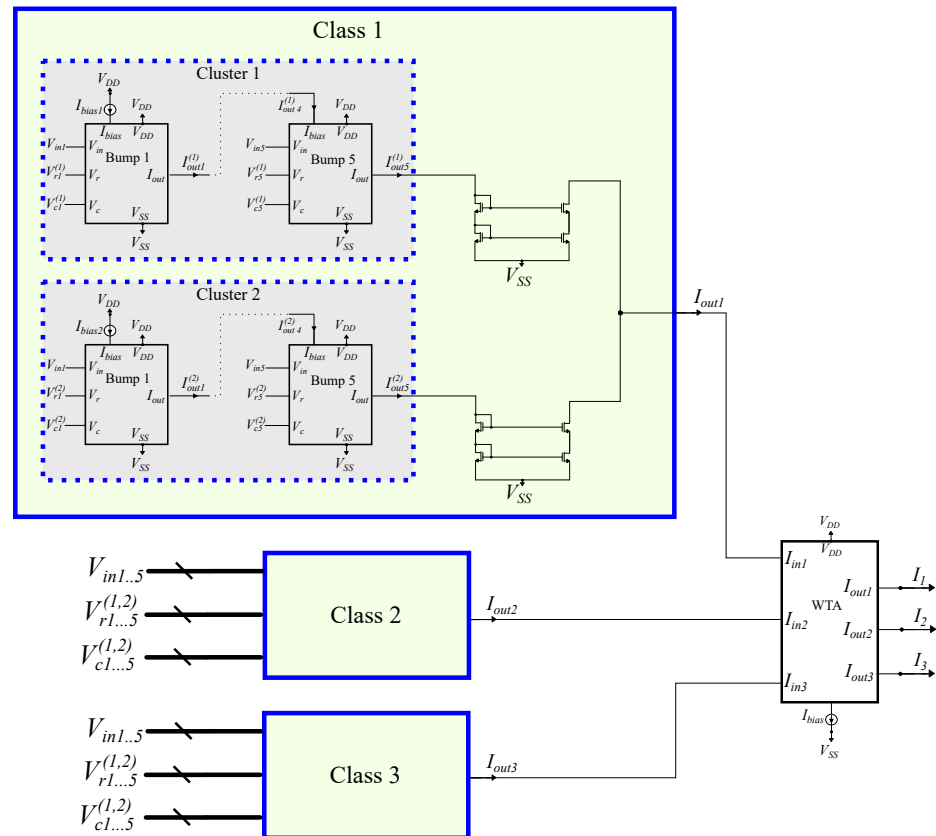

Figure 9. An analog GMM-based classifier comprises 3 classes, 2 clusters per class, and 5 input dimensions.
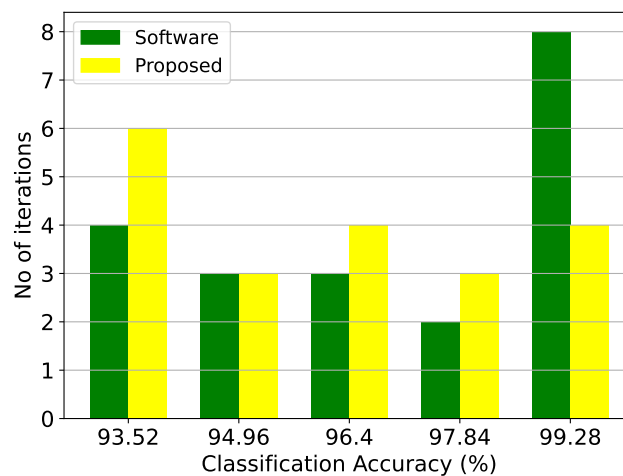

Figure 10. Classification results of the GMM architecture and the equivalent software model on the thyroid disease detection dataset over 20 iterations.
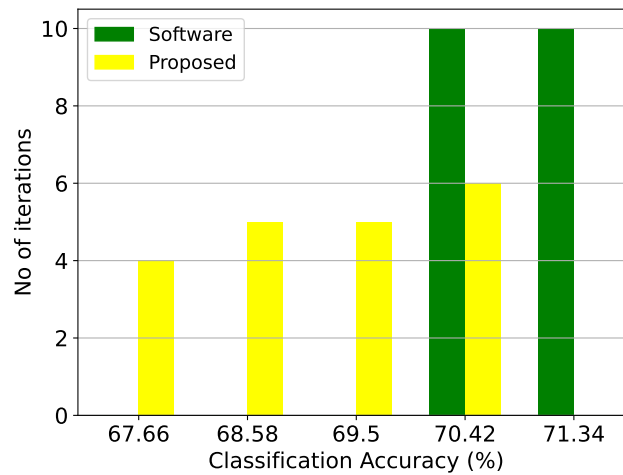
Figure 11. Classification results of the GMM architecture and the equivalent software model on the epileptic seizure prediction dataset over 20 iterations.
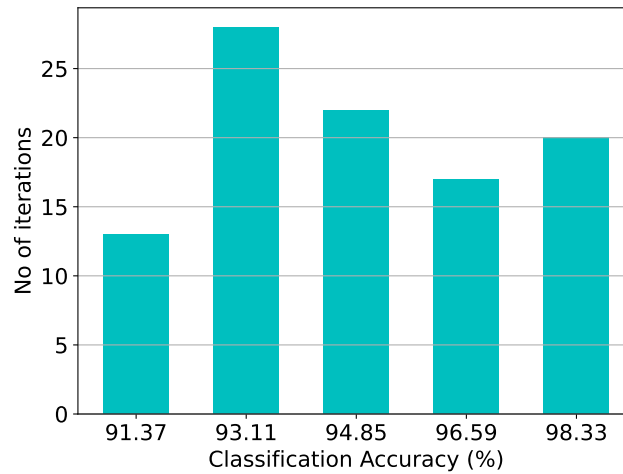

Figure 12. Post-layout Monte Carlo simulation results of the GMM architecture on the thyroid disease detection dataset with $\mu_M = 94.85\%$ and $\sigma_M = 2.31\%$.

Table 2. GMM-based CLF's accuracy for thyroid disease detection dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	100.00%	93.20%	97.35%	2.36%
Hardware	99.60%	92.80%	96.20%	2.07%

Table 3. GMM-based CLF's accuracy for epileptic seizure prediction dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	71.80%	70.00%	70.90%	0.54%
Hardware	70.60%	67.20%	69.20%	1.06%

5.2. Radial Basis Function CLF Implementation and Simulation Results

In this subsection, the implemented radial basis function (RBF) CLF and its simulation results for both datasets are provided. Based on the proposed design methodology and the simulation result of the software implementation, the generic RBF CLF is composed of $N_{Cla} = 3$ classes, $N_{Clu} = 3$ clusters per class, and $N_d = 5$ input dimensions. The high-level architecture of this CLF is depicted in Figure 13. More specifically, each class comprises three 5-D Gaussian function circuits which correspond to the three clusters and three current mirrors that are used to add

the output currents of each cluster. In Figures 14 (thyroid) and 15 (epilepsy), the classification accuracies for both implementations (hardware/software) are presented, encompassing a total of 20 distinct training test cases (first case) for both datasets. The results are also summarized in Tables 4 and 5. Regarding the sensitivity, the Monte Carlo histogram shown in Figure 16 has a mean value of $\mu_M = 93.83\%$ and a standard deviation of $\sigma_M = 2.33\%$.
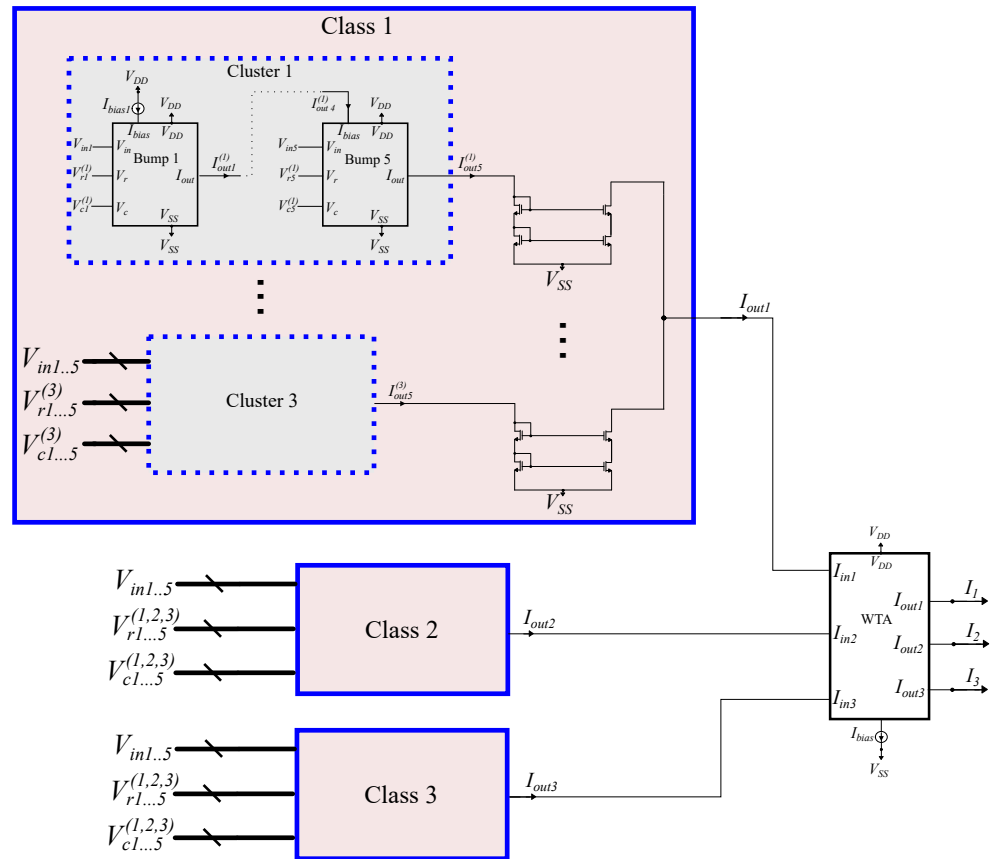

Figure 13. An analog RBF-based classifier comprises 3 classes, 3 clusters per class, and 5 input dimensions.
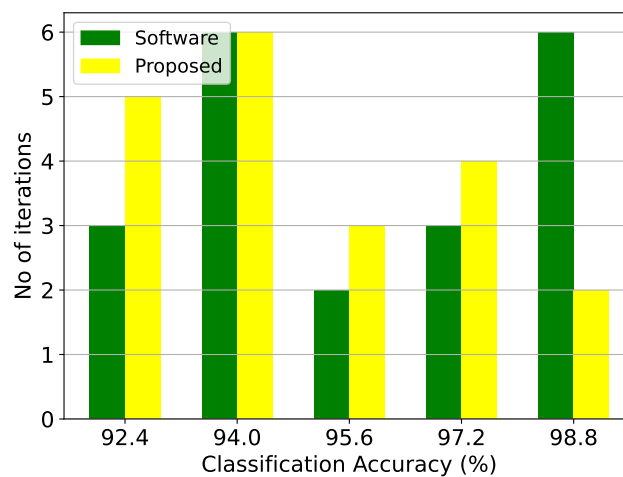

Figure 14. Classification results of the RBF architecture and the equivalent software model on the thyroid disease detection dataset over 20 iterations.
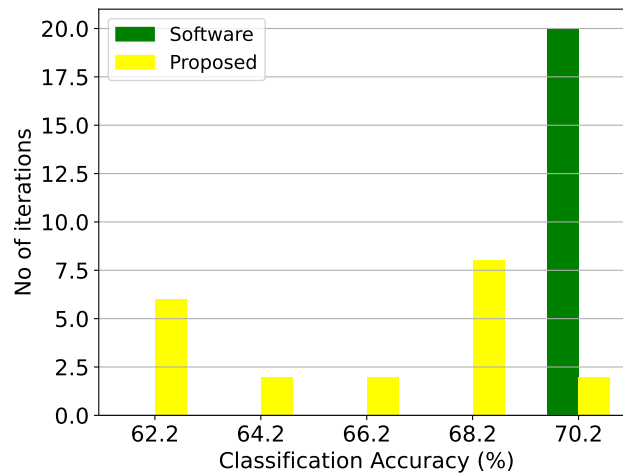
Figure 15. Classification results of the RBF architecture and the equivalent software model on the epileptic seizure prediction dataset over 20 iterations.
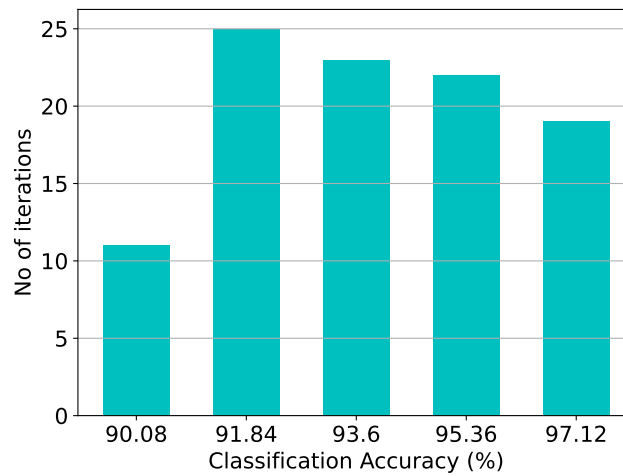

Figure 16. Post-layout Monte Carlo simulation results of the RBF architecture on the thyroid disease detection dataset with $\mu_M = 93.83\%$ and $\sigma_M = 2.33\%$.

Table 4. RBF CLF's accuracy for thyroid disease detection dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	99.60%	92.00%	95.55%	2.69%
Hardware	98.40%	91.60%	94.10%	2.16%

Table 5. RBF CLF's accuracy for epileptic seizure prediction dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	71.20%	69.30%	70.20%	0.55%
Hardware	69.20%	61.20%	66.75%	2.53%

5.3. Bayes CLF Implementation and Simulation Results

In this subsection, the implemented Bayes CLF and its simulation results for both datasets are provided. Based on the proposed design methodology and the simulation result of the software implementation, the generic RBF CLF is composed of $N_{Cla} = 3$ classes, $N_{Clu} = 1$ clusters per class, and $N_d = 5$ input dimensions. The high-level architecture of this CLF is depicted in Figure 17. To elaborate further, each class is composed of 5-D Gaussian function circuits corresponding to individual clusters. Subsequently, the resultant output current is fed into one

of the inputs of the WTA circuit. In Figures 18 (thyroid) and 19 (epilepsy), the classification accuracy for both implementations (hardware/software) are presented, encompassing a total of 20 distinct training test cases (first case) for both datasets. The results are also summarized in Tables 6 and 7. Regarding the sensitivity, the Monte Carlo histogram shown in Figure 20 has a mean value of $\mu_M = 94.49\%$ and a standard deviation of $\sigma_M = 2.29\%$.

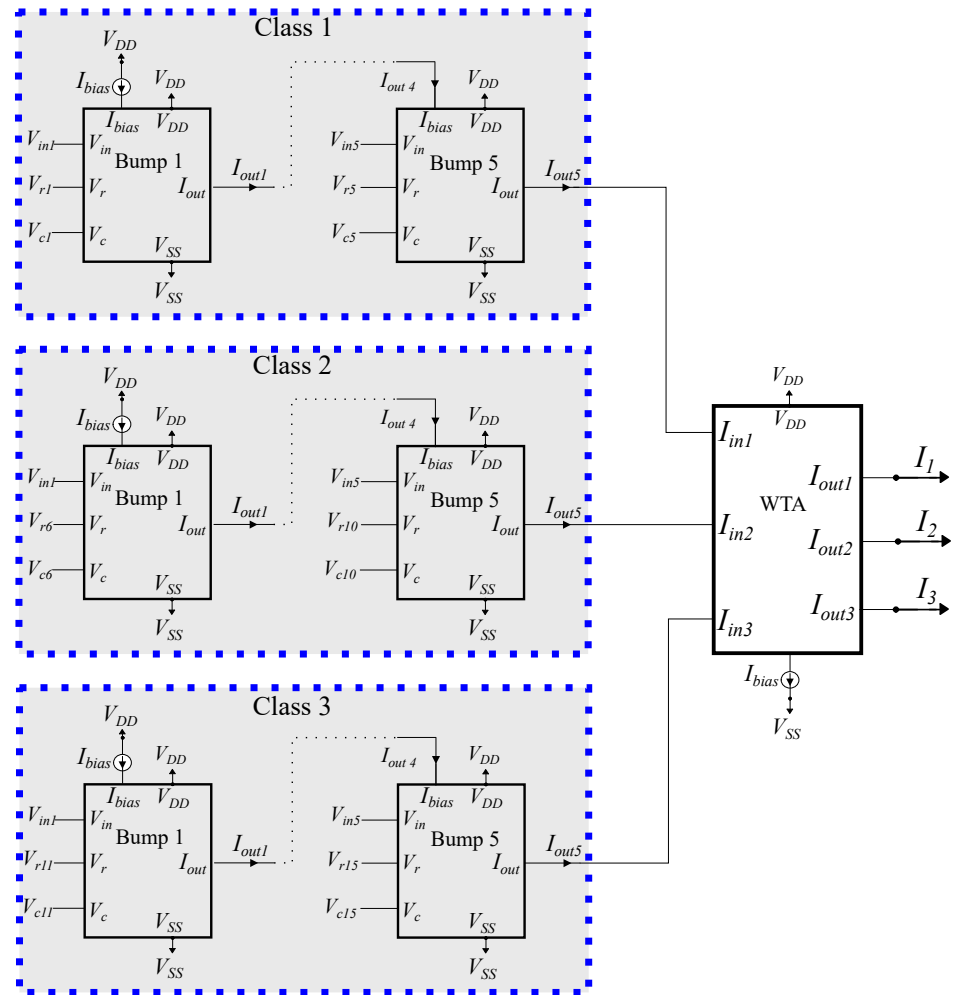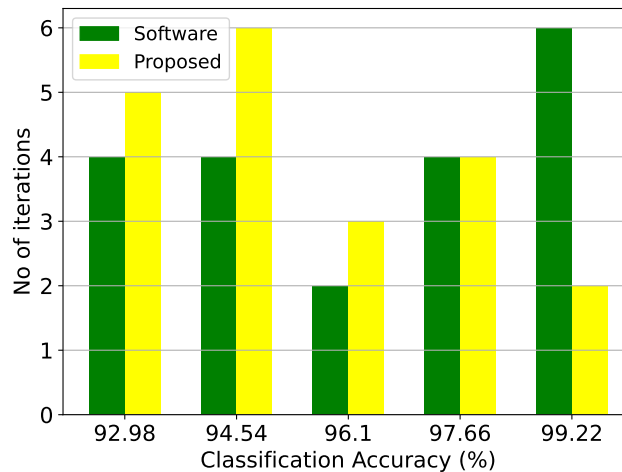**Figure 17.** An analog Bayesian classifier comprises 3 classes and 5 input dimensions.

Figure 18. Classification results of the Bayesian architecture and the equivalent software model on the thyroid disease detection dataset over 20 iterations.
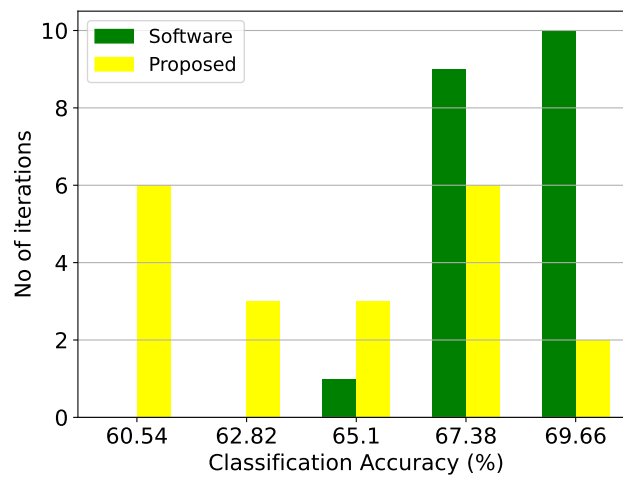
Figure 19. Classification results of the Bayesian architecture and the equivalent software model on the epileptic seizure prediction dataset over 20 iterations.
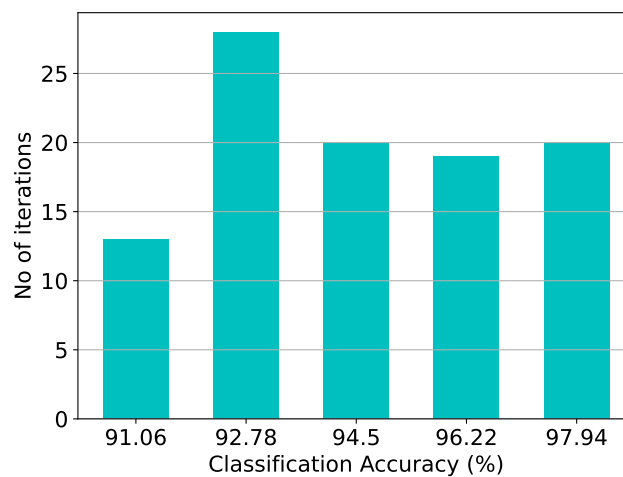

Figure 20. Post-layout Monte Carlo simulation results of the Bayesian architecture on the thyroid disease detection dataset with $\mu_M = 94.49\%$ and $\sigma_M = 2.29\%$.

Table 6. Bayes CLF's accuracy for thyroid disease detection dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	100.00%	92.70%	96.85%	2.41%
Hardware	99.00%	92.20%	94.70%	2.22%

Table 7. Bayes CLF's accuracy for epileptic seizure prediction dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	70.80%	66.20%	68.80%	1.34%
Hardware	68.80%	59.40%	64.80%	3.27%

5.4. Threshold CLF Implementation and Simulation Results

In this subsection, the implemented threshold CLF and its simulation results for both datasets are provided. Based on the proposed design methodology and the simulation result of the software implementation, the generic RBF CLF is composed of $N_{Cla} = 1$ class, $N_{Clu} = 1$ clusters per classm and $N_d = 5$ input dimensions. The second class is a threshold current, which is selected as the decision boundary of each classification problem. A threshold CLF operates in a binary manner, enabling us to ascertain, using the thyroid dataset, whether a patient is healthy or not. The high-level architecture of this CLF is depicted

in Figure 21. To elaborate further, each class is composed of 5-D Gaussian function circuits, corresponding to individual clusters. Subsequently, the resultant output current is fed into one of the inputs of the WTA circuit. In Figures 22 (thyroid) and 23 (epilepsy), the classification accuracy for both implementations (hardware/software) are presented, encompassing a total of 20 distinct training test cases (first case), for both datasets. The results are also summarized in Tables 8 and 9. Regarding the sensitivity, the Monte Carlo histogram shown in Figure 24 has a mean value of $\mu_M = 93.12\%$ and a standard deviation of $\sigma_M = 2.67\%$.

Figure 21. An analog Threshold classifier comprises 2 classes and 5 input dimensions. The second class is the decision boundary for the threshold implementation.
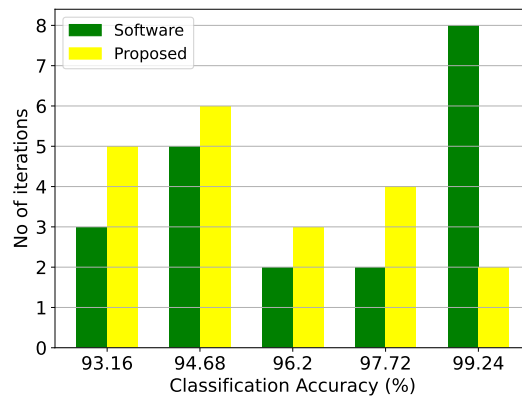
Figure 22. Classification results of the threshold architecture and the equivalent software model on the thyroid disease detection dataset over 20 iterations.
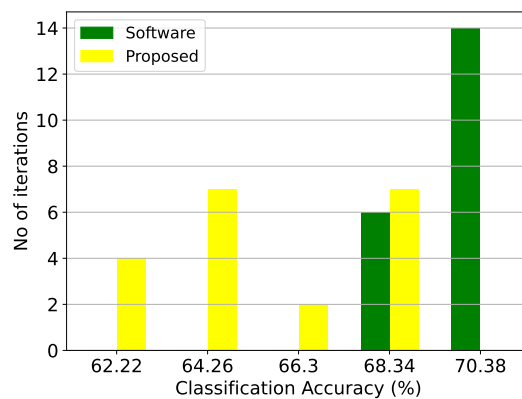
Figure 23. Classification results of the threshold architecture and the equivalent software model on the epileptic seizure prediction dataset over 20 iterations.
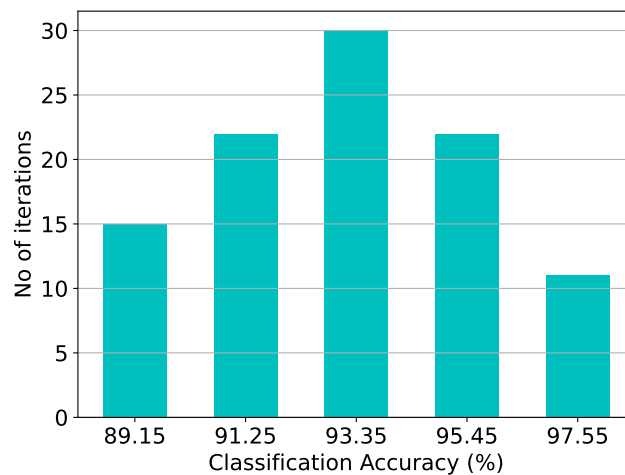
Figure 24. Post-layout Monte Carlo simulation results of the threshold architecture on the thyroid disease detection dataset with $\mu_M = 93.12\%$ and $\sigma_M = 2.67\%$.

Table 8. Threshold CLF's accuracy for thyroid disease detection dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	100.00%	93.00%	97.15%	2.48%
Hardware	99.20%	92.40%	94.90%	2.13%

Table 9. Threshold CLF's accuracy for epileptic seizure prediction dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	71.40%	68.80%	70.10%	0.91%
Hardware	69.10%	61.20%	65.00%	2.54 %

5.5. Multiple Centroid CLF Implementation and Simulation Results

In this subsection, the implemented centroid CLF and its simulation results for both datasets are provided. Based on the proposed design methodology and the simulation result of the software implementation, the generic RBF CLF is composed of $N_{Cla} = 1$ class, $N_{Clu} = 2$ centroids for Class 1 and $N_{Clu} = 1$ for Class 2 and $N_d = 5$ input dimensions. The first class is related to abnormal thyroid conditions (two centroids per class), and the second class describes the healthy patient. A centroid CLF operates in a binary manner, enabling us to ascertain, using the thyroid dataset, whether a patient is healthy or not. The high-level architecture of this CLF is depicted in Figure 25. To elaborate further, each class is composed of 5-D Gaussian function circuits, corresponding to individual clusters. The resultant output current is fed into one of the inputs of the WTA circuit. In Figures 26 (thyroid) and 27 (epilepsy), the classification accuracy for both implementations (hardware/software) are presented, encompassing a total of 20 distinct training test cases (first case), for both datasets. The results are also summarized in Tables 10 and 11. Regarding the sensitivity, the Monte Carlo histogram, shown in Figure 28, has a mean value of $\mu_M = 95.55\%$ and a standard deviation of $\sigma_M = 1.93\%$.
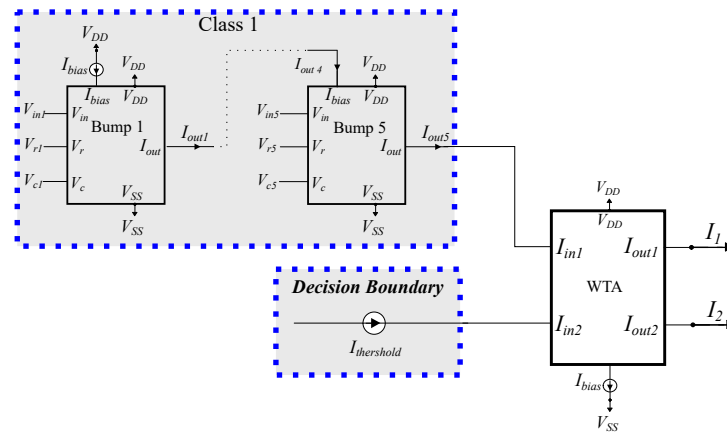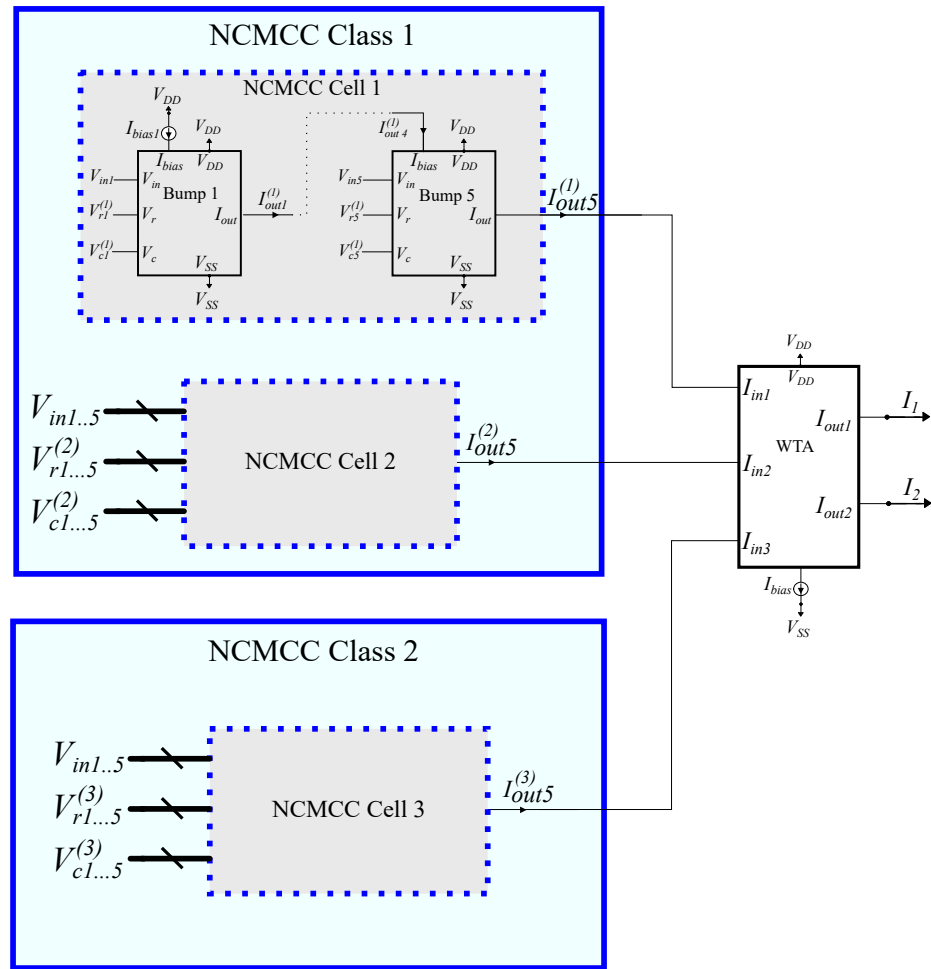
Figure 25. An analog centroid classifier comprises 2 classes, 2 centroids in the first class, 1 centroid in the second class, and 5 input dimensions.
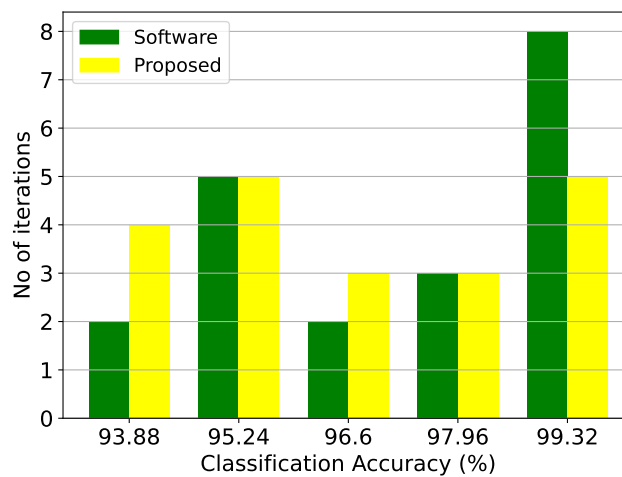

Figure 26. Classification results of the centroid architecture and the equivalent software model on the thyroid disease detection dataset over 20 iterations.
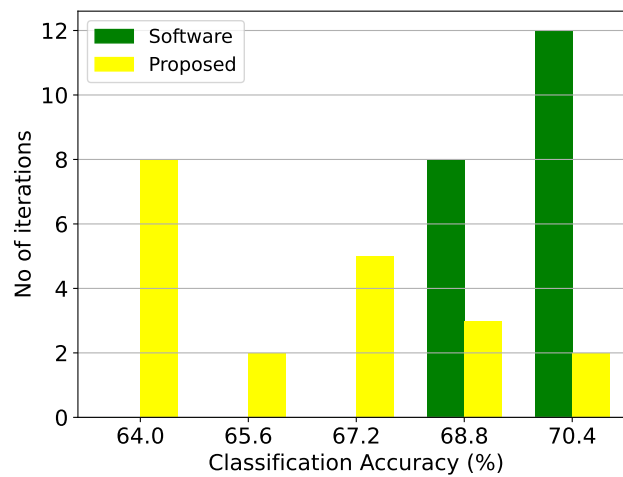
Figure 27. Classification results of the centroid architecture and the equivalent software model on the epileptic seizure prediction dataset over 20 iterations.
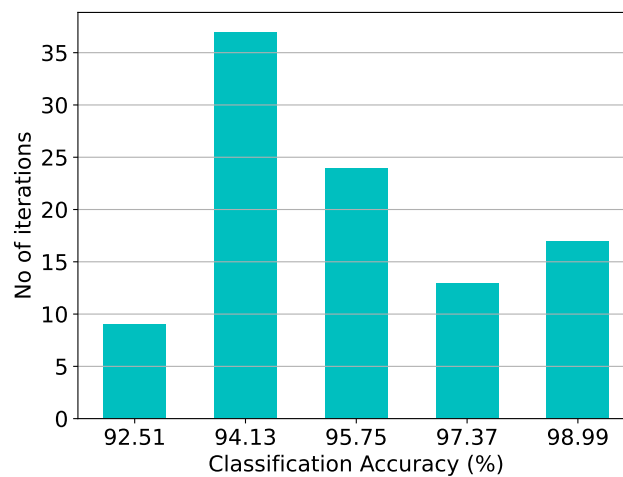
Figure 28. Post-layout Monte Carlo simulation results of the centroid architecture on the thyroid disease detection dataset with $\mu_M = 95.55\%$ and $\sigma_M = 1.93\%$.

Table 10. Centroid-based CLF's accuracy for thyroid disease detection dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	100.00%	94.20%	98.00%	2.23%
Hardware	100.00%	93.20%	96.60%	2.31%

Table 11. Centroid-based CLF's accuracy for epileptic seizure prediction dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	71.20%	68.80%	70.00%	0.78%
Hardware	69.50%	63.20%	65.80%	2.11%

For the implementation of the argmax operator the Lazzaro WTA circuit will be used as the basic building block [21]. The overall design is shown in Figure 29. Based on Equation (6), the similarity of the input vector with the three centroids is compared to determine the highest one, using a three-neuron WTA circuit. To decrease the linear region in which the WTA circuit produces multiple winners a second three-neuron WTA circuit is connected in a cascaded format [65]. The output of the second WTA is three currents in a digital one-hot-vector format. To convert these outputs into a representation

that corresponds to the two-class application we calculate two currents $I_{in1,N2}$ and $I_{in2,N2}$ as follows:

$$I_{in1,N2} = I_{op1,P1} + I_{op2,P1},\tag{7}$$

$$I_{in2,N2} = I_{op3,P1}.\tag{8}$$

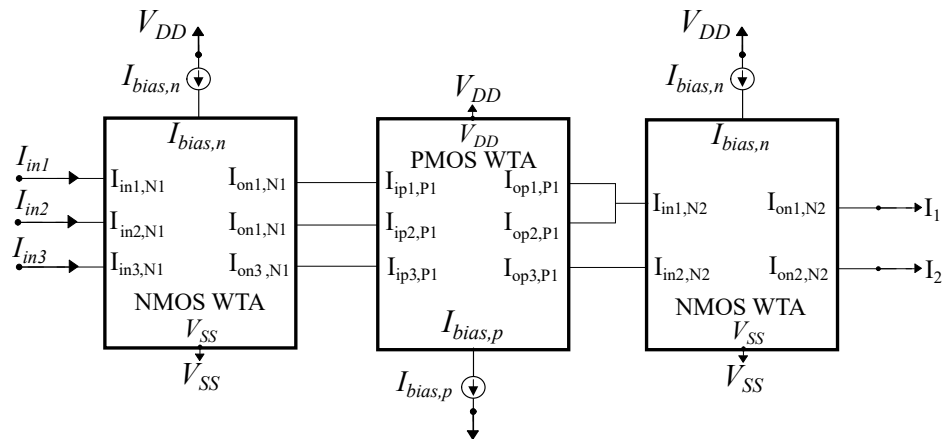**Figure 29.** The cascaded WTA circuit composed of two 3-neuron WTA circuits and one 2-neuron WTA circuit in cascaded connection. The outputs of the second WTA circuit are three currents in a digital one-hot representation; the two currents are summed up together and subsequently fed as input to the third WTA circuit along with the third output of the second WTA.

Then, the one additional two-neuron WTA circuit is added to further increase the quality of the output currents that indicate the winning class.

5.6. Support Vector Machine CLF Implementation and Simulation Results

The classification block of the support vector machine (SVM) comprises *M* RBF cells, *M* switches, and a WTA circuit. The test samples, which are five-dimensional vectors, are introduced synchronously into the classification block guided by an external clock signal. In each clock cycle, the *M* RBF cells individually compute the RBF kernel function for the current test vector. This calculation is based on the learning samples utilized in the training procedure. Notably, the RBF cells within the classification block are biased by replicating the adjusters' output currents from the learning block.

To predict the outcome of the classifier, it becomes necessary to determine the *sign* of the sum in the hardware-friendly SVM's decision rule. Rather than summing all currents and evaluating the overall sign, we separately aggregate positive and negative currents. This differentiation is achieved by employing switches where positive (or negative) currents correspond to input learning samples with positive (or negative) labels. The comparison between negative and positive values is facilitated through a current-mode circuit, the WTA circuit. The output of the WTA circuit encodes the classifier's prediction using a one-hot-vector format ($[I_{out1}, I_{out2}]$). A WTA circuit is used instead of a comparator due to the fact that information processing in the system is performed mainly in current mode.

An SVM CLF operates in a binary manner, enabling us to ascertain whether a patient is healthy or not using the thyroid dataset. The high-level architecture of this CLF is depicted in Figure 30. In Figures 31 (thyroid) and 32 (epilepsy), the classification accuracy for both implementations (hardware/software) is presented, encompassing a total of 20 distinct training test cases (first case) for both datasets. The results are also summarized in Tables 12 and 13. Regarding the sensitivity, the Monte Carlo histogram shown in Figure 33 has a mean value of $\mu_M = 95.16\%$ and a standard deviation of $\sigma_M = 0.75\%$.
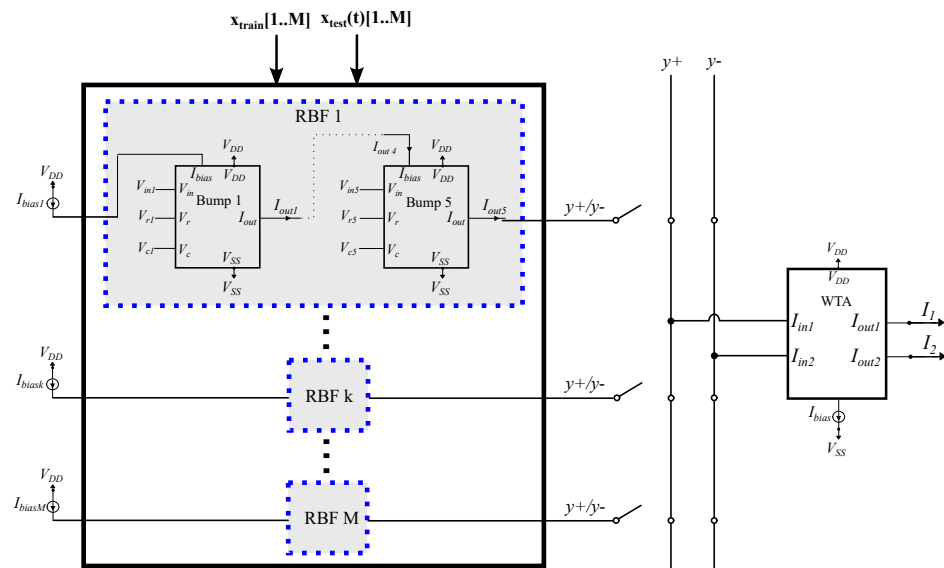
Figure 30. An analog SVM classifier consists of M RBF cells and 2 classes. The RBF cells receive the input dimensions and generate the suitable RBF patterns utilizing the learned parameters. These derived RBF patterns correspond to the Support Vectors in the model. To convey the polarity of the Support Vectors to the classification block switches are employed. A WTA approach is implemented for contrasting the positive and negative magnitudes.
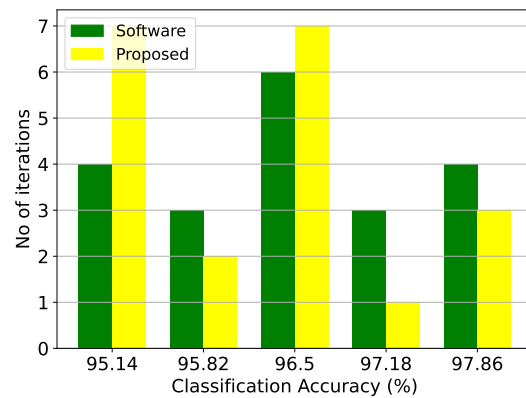
Figure 31. Classification results of the SVM architecture and the equivalent software model on the thyroid disease detection dataset over 20 iterations.
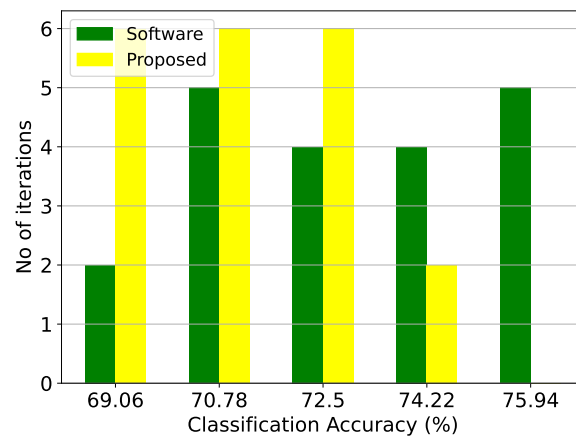
Figure 32. Classification results of the SVM architecture and the equivalent software model on the epileptic seizure prediction dataset over 20 iterations.
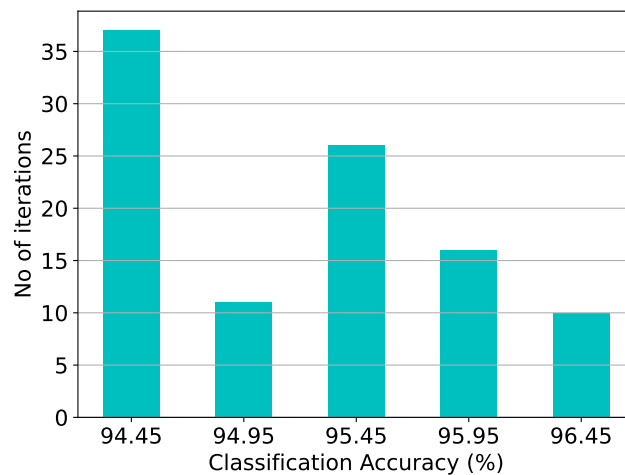
Figure 33. Post-layout Monte Carlo simulation results of the SVM architecture on the thyroid disease detection dataset with $\mu_M = 95.16\%$ and $\sigma_M = 0.75\%$.

Table 12. SVM CLF's accuracy for thyroid disease detection dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	98.20%	95.20%	96.40%	1.01%
Hardware	98.00%	94.80%	96.20%	1.05%

Table 13. SVM CLF's accuracy for epileptic seizure prediction dataset (over 20 iterations).

Method	Best	Worst	Mean	Variance
Software	76.80%	69.20%	73.00%	2.31%
Hardware	74.20%	68.20%	71.05%	1.76%

## 6. Performance Summary and Discussion

This section aims to present a comparative analysis of the analog classifiers implemented earlier in this work. Table 14 summarizes the main performance indexes referring to the thyroid disease detection dataset for a GMM, an RBF, a Bayesian, a threshold, a centroid, and an SVM classifier. Table 15 summarizes the main performance indexes referring to the epileptic seizure prediction dataset.

Table 14. Implemented analog classifiers' performance comparison for thyroid disease detection dataset.

Classifier	Best	Worst	Mean	Power Consumption	Processing Speed	Energy per Classification
GMM	99.60%	92.80%	96.20%	1.31 µW	$112\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{11.7\text{ pJ}}{\text{classification}}$
RBF	98.40%	91.60%	94.10%	2.48 µW	$100\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{24.8\text{ pJ}}{\text{classification}}$
Bayes	99.00%	92.20%	94.70%	421 nW	$130\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{3.2\text{ pJ}}{\text{classification}}$
Threshold	99.20%	93.00%	94.90%	**247 nW**	$130\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{\textbf{1.9 pJ}}{\textbf{classification}}$
Centroid	**100.00%**	93.20%	**96.60%**	2.57 µW	$112\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{22.9\text{ pJ}}{\text{classification}}$
SVM	98.00%	**94.80%**	96.20%	44.7 µW	$\textbf{140 K}\frac{\textbf{classifications}}{\textbf{s}}$	$\frac{319.3\text{ pJ}}{\text{classification}}$

In terms of best and mean classification accuracy, the most efficient design turns out to be the centroid classifier (best: 100%, mean: 96.60%). In contrast, regarding the worst-case classification accuracy, the highest score is achieved by the SVM classifier (94.80%). This design also overwhelms the others in terms of robustness, since it effectively minimizes

the range of variation between the highest and lowest values of classification accuracy (having the lowest standard deviation of accuracy). In addition, this classifier surpasses the others in terms of processing speed $\left(140\text{K}\frac{\text{classifications}}{\text{s}}\right)$, a quality that holds significant importance when dealing with real-time scenarios. In order to achieve the aforementioned performance, power consumption is sacrificed. However, the overall consumption of this design also encompasses hardware training, which comprises various components. Pertaining to the power consumption of the implemented classifiers, the best option proves to be the threshold implementation (247 nW), which also outperforms the rest of the designs in terms of energy per classification $\left(\frac{1.9\text{ pJ}}{\text{classification}}\right)$.

Table 15. Implemented analog classifiers' performance comparison for epileptic seizure prediction dataset.

Classifier	Best	Worst	Mean	Power Consumption	Processing Speed	Energy per Classification
GMM	70.60%	67.20%	69.20%	180 nW	$112\text{ K}\frac{\text{classifications}}{\text{s}}$	$\frac{1.6\text{ pJ}}{\text{classification}}$
RBF	69.20%	61.20%	66.75%	231 nW	$100\text{ K}\frac{\text{classifications}}{\text{s}}$	$\frac{2.31\text{ pJ}}{\text{classification}}$
Bayes	68.80%	59.40%	64.80%	123 nW	$130\text{ K}\frac{\text{classifications}}{\text{s}}$	$\mathbf{\frac{0.9\text{ pJ}}{\text{classification}}}$
Threshold	69.10%	61.20%	65.00%	**111 nW**	$130\text{ K}\frac{\text{classifications}}{\text{s}}$	$\frac{0.8\text{ pJ}}{\text{classification}}$
Centroid	69.50%	63.20%	65.80%	355 nW	$112\text{ K}\frac{\text{classifications}}{\text{s}}$	$\frac{3.2\text{ pJ}}{\text{classification}}$
SVM	**74.20%**	**68.20%**	**71.05%**	3.24 μW	$\mathbf{140\text{ K}\frac{\text{classifications}}{\text{s}}}$	$\frac{23.1\text{ pJ}}{\text{classification}}$

Similar results regarding the comparative advantages between the various classifiers occur from the epileptic seizure prediction dataset. In this case, the SVM classifier achieves the best accuracy because of the complexity of the implemented ML algorithm. However, this is a trade-off between complexity/accuracy and power consumption, since the most power-efficient design is the threshold classifier but with significantly lower accuracy. The necessary characteristic is sensitivity in seizure prediction (successively detects all 17 epileptic seizures) where all classifiers achieve 100%.

Within contemporary literature, it is clear that a considerable number of analog classifiers are typically tailored for specific applications. This situation presents a difficulty when aiming to conduct an impartial comparison between diverse implementations. However, this challenge enables us to adapt analog classifiers to serve a universal application, thereby facilitating an evaluation that includes both machine learning models and alternative approaches. It is worth noting that Table 16 provides a comprehensive summary of the performance metrics from this research alongside related classifiers, all within the context of epileptic seizure prediction. The presented design methodology introduces a compelling solution by striking a balance between accuracy, power efficiency, and energy consumption per classification in comparison to analog classifiers. While the other models attain greater accuracy, they also entail increased complexity, requiring higher power consumption and a larger hardware footprint due to their augmented component number. It can be used as a wake-up circuit design methodology in comparison with other classifiers that consume μW or mW. It is imperative to highlight that, within this particular application, the design methodology effectively manages all seizures.

Considering another crucial aspect of design involves finding the right balance between the power consumption and the specificity of the wake-up circuit. As the specificity of the wake-up circuit is enhanced, the overall power usage of the digital circuit, typically higher than its analog counterpart, diminishes. Nevertheless, achieving elevated specificity values demands a rise in the intricacy of the analog circuitry. In particular, enhancing the classifier's performance necessitates better-performing data acquisition devices, increased integration of analog feature extraction circuits, and larger analog memory units to store the classifier's parameters. These enhancements collectively contribute to escalated power consumption. In practical terms, caution is advised when amplifying the power consump-

tion of the analog front end; a classification system incorporating a power-hungry analog classifier that alternates with a digital counterpart might end up consuming more power than a fully digital system.

Table 16. Analog classifiers' comparison on the epileptic seizure prediction dataset.

Classifier	Best	Worst	Mean	Power Consumption	Processing Speed	Energy per Classification
GMM	70.60%	67.20%	69.20%	180 nW	$112\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{1.6\,\text{pJ}}{\text{classification}}$
RBF	69.20%	61.20%	66.75%	231 nW	$100\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{2.31\,\text{pJ}}{\text{classification}}$
Bayes	68.80%	59.40%	64.80%	123 nW	$130\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{0.9\,\text{pJ}}{\text{classification}}$
Threshold	69.10%	61.20%	65.00%	**111 nW**	$130\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{\textbf{0.8\,pJ}}{\textbf{classification}}$
Centroid	69.50%	63.20%	65.80%	355 nW	$112\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{3.2\,\text{pJ}}{\text{classification}}$
SVM	74.20%	68.20%	71.05%	3.24 µW	$140\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{23.1\,\text{pJ}}{\text{classification}}$
RBF [27]	67.80%	59.90%	64.33%	5.54 µW	$170\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{32.59\,\text{pJ}}{\text{classification}}$
RBF-NN [29]	69.10%	62.30%	67.41%	870 nW	$270\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{3.22\,\text{pJ}}{\text{classification}}$
SVM [30]	72.60%	69.40%	71.76%	141.7 µW	$870\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{162.87\,\text{pJ}}{\text{classification}}$
MLP [38]	85.70%	82.10%	83.47%	677.43 µW	$930\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{728.42\,\text{pJ}}{\text{classification}}$
K-means [43]	82.40%	76.30%	81.31%	101.32 µW	$5\,\text{M}\frac{\text{classifications}}{\text{s}}$	$\frac{20.26\,\text{pJ}}{\text{classification}}$
Fuzzy [45]	77.30%	69.20%	74.59%	761 nW	$4.55\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{167.25\,\text{pJ}}{\text{classification}}$
SVR [51]	81.30%	76.30%	78.42%	81.41 µW	$870\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{93.57\,\text{pJ}}{\text{classification}}$
LSTM [59]	99.20%	94.30%	97.51%	33.41 mW	$870\,\text{K}\frac{\text{classifications}}{\text{s}}$	$\frac{38.40\,\text{pJ}}{\text{classification}}$

7. Conclusions

This study introduces an innovative general methodology for adjustable analog integrated CLFs rooted in the principles of Gaussian function. Through the strategic manipulation and utilization of the Gaussian function and Winner-Take-All (WTA) circuits, it becomes possible to fabricate bell-shaped CLFs tailored to address a wide spectrum of scenarios, encompassing varying class quantities, cluster/centroid configurations and data dimensions. To illustrate the adaptability and effectiveness of this approach, the proposed design methodology is employed in the analysis of two distinct real-world datasets specifically curated for the purpose of thyroid disorder diagnosis and epileptic seizure prediction. The critical parameters controlling the functionality of this method are established through offline training of a bell-shaped CLF utilizing software-based methods. Thorough examination and comparisons of the classification outcomes within these scenarios serve to emphasize the effective functioning of the proposed methodology and provide validation for the implemented adjustments. The proposed design methodology can be used as a basic tool for the design of more complicated and accurate diagnosis systems.

Author Contributions: Investigation, V.A., N.P.E. and A.K.; Writing—original draft preparation, V.A., N.P.E. and A.K.; Writing—review and editing, V.A., N.P.E., A.K., G.G., C.D. and P.P.S. All authors have read and agreed to the published version of the manuscript.

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

References

1. Henkel, J.; Pagani, S.; Amrouch, H.; Bauer, L.; Samie, F. Ultra-low power and dependability for IoT devices (Invited paper for IoT technologies). In Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE), Lausanne, Switzerland, 27–31 March 2017; pp. 954–959.
2. Goebel, K.; Saha, B.; Saxena, A.; Celaya, J.R.; Christophersen, J.P. Prognostics in battery health management. *IEEE Instrum. Meas. Mag.* **2008**, *11*, 33–40. [CrossRef]
3. Alioto, M. *Enabling the Internet of Things: From Integrated Circuits to Integrated Systems*; Springer: Cham, Switzerland, 2017.
4. Haensch, W.; Gokmen, T.; Puri, R. The next generation of deep learning hardware: Analog computing. *Proc. IEEE* **2018**, *107*, 108–122. [CrossRef]
5. Wang, A.; Calhoun, B.H.; Chandrakasan, A.P. *Sub-Threshold Design for Ultra Low-Power Systems*; Springer: New York, NY, USA, 2006; Volume 95.
6. Zhang, Y.; Mirchandani, N.; Onabajo, M.; Shrivastava, A. RSSI amplifier design for a feature extraction technique to detect seizures with analog computing. In Proceedings of the 2020 IEEE International Symposium on Circuits and Systems (ISCAS), Seville, Spain, 12–14 October 2020; pp. 1–5.
7. Aktas, F.; Ceken, C.; Erdemli, Y.E. IoT-based healthcare framework for biomedical applications. *J. Med. Biol. Eng.* **2018**, *38*, 966–979. [CrossRef]
8. Muthu, B.; Sivaparthipan, C.; Manogaran, G.; Sundarasekar, R.; Kadry, S.; Shanthini, A.; Dasel, A. IOT based wearable sensor for diseases prediction and symptom analysis in healthcare sector. *Peer Peer Netw. Appl.* **2020**, *13*, 2123–2134. [CrossRef]
9. Quinlan, R. *Thyroid Disease*; UCI Machine Learning Repository: Irvine, CA, USA, 1987. [CrossRef]
10. Vanderpump, M.P. The epidemiology of thyroid disease. *Br. Med. Bull.* **2011**, *99*, 39–51. [CrossRef] [PubMed]
11. Jabbar, A.; Pingitore, A.; Pearce, S.H.; Zaman, A.; Iervasi, G.; Razvi, S. Thyroid hormones and cardiovascular disease. *Nat. Rev. Cardiol.* **2017**, *14*, 39–55. [CrossRef] [PubMed]
12. Goldberger, A.L.; Amaral, L.A.; Glass, L.; Hausdorff, J.M.; Ivanov, P.C.; Mark, R.G.; Mietus, J.E.; Moody, G.B.; Peng, C.K.; Stanley, H.E. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* **2000**, *101*, e215–e220. [CrossRef]
13. CHB-MIT Scalp EEG Database. Available online: https://physionet.org/content/chbmit/1.0.0/ (accessed on 10 July 2023) .
14. Karoly, P.J.; Rao, V.R.; Gregg, N.M.; Worrell, G.A.; Bernard, C.; Cook, M.J.; Baud, M.O. Cycles in epilepsy. *Nat. Rev. Neurol.* **2021**, *17*, 267–284. [CrossRef]
15. World Health Organization. *Atlas: Epilepsy Care in the World*; World Health Organization: Geneva, Switzerland, 2005.
16. Tsiouris, K.M.; Pezoulas, V.C.; Zervakis, M.; Konitsiotis, S.; Koutsouris, D.D.; Fotiadis, D.I. A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. *Comput. Biol. Med.* **2018**, *99*, 24–37. [CrossRef]
17. Banerjee, A.; Maity, S.; Mastrangelo, C.H. Nanostructures for biosensing, with a brief overview on cancer detection, IoT, and the role of machine learning in smart biosensors. *Sensors* **2021**, *21*, 1253. [CrossRef]
18. Sharma, A.; Polley, A.; Lee, S.B.; Narayanan, S.; Li, W.; Sculley, T.; Ramaswamy, S. A Sub-60 μA Multimodal Smart Biosensing SoC With >80-dB SNR, 35-μA Photoplethysmography Signal Chain. *IEEE J. Solid-State Circuits* **2017**, *52*, 1021–1033. [CrossRef]
19. Bishop, C.M.; Nasrabadi, N.M. *Pattern Recognition and Machine Learning*; Springer: Cham, Switzerland, 2006; Volume 4.
20. Alimisis, V.; Gourdouparis, M.; Gennis, G.; Dimas, C.; Sotiriadis, P.P. Analog gaussian function circuit: Architectures, operating principles and applications. *Electronics* **2021**, *10*, 2530. [CrossRef]
21. Lazzaro, J.; Ryckebusch, S.; Mahowald, M.A.; Mead, C.A. Winner-take-all networks of O (n) complexity. *Adv. Neural Inf. Process. Syst.* **1988**, *1*, 703–711.
22. Chi, P.; Li, S.; Xu, C.; Zhang, T.; Zhao, J.; Liu, Y.; Wang, Y.; Xie, Y. Prime: A novel processing-in-memory architecture for neural network computation in reram-based main memory. *ACM SIGARCH Comput. Archit. News* **2016**, *44*, 27–39. [CrossRef]
23. Shawahna, A.; Sait, S.M.; El-Maleh, A. FPGA-based accelerators of deep learning networks for learning and classification: A review. *IEEE Access* **2018**, *7*, 7823–7859. [CrossRef]
24. Mohri, M.; Rostamizadeh, A.; Talwalkar, A. *Foundations of Machine Learning*; MIT Press: Cambridge, MA, USA, 2018.
25. Liu, S.C.; Kramer, J.; Indiveri, G.; Delbrück, T.; Douglas, R. *Analog VLSI: Circuits and Principles*; MIT Press: Cambridge, MA, USA, 2002.
26. Jabri, M.; Coggins, R.J.; Flower, B.G. *Adaptive Analog VLSI Neural Systems*; Springer Science & Business Media: New York, NY, USA, 1996.
27. Peng, S.Y.; Hasler, P.E.; Anderson, D.V. An analog programmable multidimensional radial basis function based classifier. *IEEE Trans. Circuits Syst. Regul. Pap.* **2007**, *54*, 2148–2158. [CrossRef]
28. Dorzhigulov, A.; James, A.P. Generalized bell-shaped membership function generation circuit for memristive neural networks. In Proceedings of the 2019 IEEE International Symposium on Circuits and Systems (ISCAS), Sapporo, Japan, 26–29 May 2019; pp. 1–5.
29. Mohamed, A.R.; Qi, L.; Li, Y.; Wang, G. A generic nano-watt power fully tunable 1-d gaussian kernel circuit for artificial neural network. *IEEE Trans. Circuits Syst. II Express Briefs* **2020**, *67*, 1529–1533. [CrossRef]
30. Kang, K.; Shibata, T. An on-chip-trainable Gaussian-kernel analog support vector machine. *IEEE Trans. Circuits Syst. Regul. Pap.* **2009**, *57*, 1513–1524. [CrossRef]
31. Lee, K.; Park, J.; Kim, G.; Hong, I.; Yoo, H.J. A multi-modal and tunable Radial-Basis-Funtion circuit with supply and temperature compensation. In Proceedings of the 2013 IEEE International Symposium on Circuits and Systems (ISCAS), Beijing, China, 19–23 May 2013; pp. 1608–1611.

32. Watkins, S.S.; Chau, P.M.; Tawel, R. A radial basis function neurocomputer implemented with analog VLSI circuits. In Proceedings of the [Proceedings 1992] IJCNN International Joint Conference on Neural Networks, Baltimore, MD, USA, 7–11 June 1992; Volume 2, pp. 607–612.

33. Verleysen, M.; Thissen, P.; Voz, J.L.; Madrenas, J. An analog processor architecture for a neural network classifier. *IEEE Micro* **1994**, *14*, 16–28. [CrossRef]

34. De Oliveira, J.; Oki, N. An analog implementation of radial basis neural networks (RBNN) using BiCMOS technology. In Proceedings of the 44th IEEE 2001 Midwest Symposium on Circuits and Systems, MWSCAS 2001 (Cat. No. 01CH37257), Dayton, OH, USA, 14–17 August 2001; Volume 2, pp. 705–708.

35. Collins, S.; Marshall, G.F.; Brown, D. An analogue Radial Basis Function circuit using a compact Euclidean Distance calculator. In Proceedings of the IEEE International Symposium on Circuits and Systems-ISCAS'94, London, UK, 30 May–2 June 1994; Volume 6, pp. 233–236.

36. Anderson, J.; Platt, J.; Kirk, D.B. An analog VLSI chip for radial basis functions. *Adv. Neural Inf. Process. Syst.* **1992**, *5*, 765–772.

37. Hsieh, Y.T.; Anjum, K.; Pompili, D. Ultra-low Power Analog Recurrent Neural Network Design Approximation for Wireless Health Monitoring. In Proceedings of the 2022 IEEE 19th International Conference on Mobile Ad Hoc and Smart Systems (MASS), Denver, CO, USA, 19–23 October 2022; pp. 211–219. [CrossRef]

38. Lee, K.; Park, J.; Yoo, H.J. A low-power, mixed-mode neural network classifier for robust scene classification. *J. Semicond. Technol. Sci.* **2019**, *19*, 129–136. [CrossRef]

39. Cevikhas, I.; Ogrenci, A.; Dundar, G.; Balkur, S. VLSI implementation of GRBF (Gaussian radial basis function) networks. In Proceedings of the 2000 IEEE International Symposium on Circuits and Systems (ISCAS), Geneva, Switzerland, 28–31 May 2000; Volume 3, pp. 646–649.

40. Alimisis, V.; Gennis, G.; Dimas, C.; Gourdouparis, M.; Sotiriadis, P.P. An ultra low power analog integrated radial basis function classifier for smart IoT systems. *Analog. Integr. Circuits Signal Process.* **2022**, *112*, 225–236. [CrossRef]

41. Alimisis, V.; Gennis, G.; Touloupas, K.; Dimas, C.; Gourdouparis, M.; Sotiriadis, P.P. Gaussian Mixture Model classifier analog integrated low-power implementation with applications in fault management detection. *Microelectron. J.* **2022**, *126*, 105510. [CrossRef]

42. Alimisis, V.; Gennis, G.; Dimas, C.; Sotiriadis, P.P. An analog Bayesian classifier implementation, for thyroid disease detection, based on a low-power, current-mode gaussian function circuit. In Proceedings of the 2021 International Conference on Microelectronics (ICM), New Cairo City, Egypt, 19–22 December 2021; pp. 153–156.

43. Zhang, R.; Shibata, T. An analog on-line-learning K-means processor employing fully parallel self-converging circuitry. *Analog. Integr. Circuits Signal Process.* **2013**, *75*, 267–277. [CrossRef]

44. Alimisis, V.; Mouzakis, V.; Gennis, G.; Tsouvalas, E.; Dimas, C.; Sotiriadis, P.P. A Hand Gesture Recognition Circuit Utilizing an Analog Voting Classifier. *Electronics* **2022**, *11*, 3915. [CrossRef]

45. Georgakilas, E.; Alimisis, V.; Gennis, G.; Aletraris, C.; Dimas, C.; Sotiriadis, P.P. An ultra-low power fully-programmable analog general purpose type-2 fuzzy inference system. *AEU-Int. J. Electron. Commun.* **2023**, *170*, 154824. [CrossRef]

46. Alimisis, V.; Gennis, G.; Tsouvalas, E.; Dimas, C.; Sotiriadis, P.P. An Analog, Low-Power Threshold Classifier tested on a Bank Note Authentication Dataset. In Proceedings of the 2022 International Conference on Microelectronics (ICM), Casablanca, Morocco, 4–7 December 2022; pp. 66–69.

47. Alimisis, V.; Mouzakis, V.; Gennis, G.; Tsouvalas, E.; Sotiriadis, P.P. An Analog Nearest Class with Multiple Centroids Classifier Implementation, for Depth of Anesthesia Monitoring. In Proceedings of the 2022 International Conference on Smart Systems and Power Management (IC2SPM), Beirut, Lebanon, 10–12 November 2022; pp. 176–181.

48. Peng, S.Y.; Minch, B.A.; Hasler, P. Analog VLSI implementation of support vector machine learning and classification. In Proceedings of the 2008 IEEE International Symposium on Circuits and Systems (ISCAS), Seattle, WA, USA, 18–21 May 2008; pp 860–863.

49. Zhang, R.; Shibata, T. Fully parallel self-learning analog support vector machine employing compact gaussian generation circuits. *Jpn. J. Appl. Phys.* **2012**, *51*, 04DE10. [CrossRef]

50. Alimisis, V.; Gennis, G.; Gourdouparis, M.; Dimas, C.; Sotiriadis, P.P. A Low-Power Analog Integrated Implementation of the Support Vector Machine Algorithm with On-Chip Learning Tested on a Bearing Fault Application. *Sensors* **2023**, *23*, 3978. [CrossRef] [PubMed]

51. Zhang, R.; Uetake, N.; Nakada, T.; Nakashima, Y. Design of programmable analog calculation unit by implementing support vector regression for approximate computing. *IEEE Micro* **2018**, *38*, 73–82. [CrossRef]

52. Zhang, R.; Shibata, T. A vlsi hardware implementation study of svdd algorithm using analog gaussian-cell array for on-chip learning. In Proceedings of the 2012 13th International Workshop on Cellular Nanoscale Networks and their Applications, Turin, Italy, 29–31 August 2012; pp. 1–6.

53. Yamasaki, T.; Shibata, T. Analog soft-pattern-matching classifier using floating-gate MOS technology. *IEEE Trans. Neural Netw.* **2003**, *14*, 1257–1265. [CrossRef]

54. Yamasaki, T.; Yamamoto, K.; Shibata, T. Analog pattern classifier with flexible matching circuitry based on principal-axis-projection vector representation. In Proceedings of the 27th European Solid-State Circuits Conference, Villach, Austria, 18–20 September 2001; pp. 197–200.

55. Hasler, P.; Smith, P.; Duffy, C.; Gordon, C.; Dugger, J.; Anderson, D. A floating-gate vector-quantizer. In Proceedings of the 2002 45th Midwest Symposium on Circuits and Systems, 2002. MWSCAS-2002, Tulsa, OK, USA, 4–7 August 2002; Volume 1, pp. I–196.

56. Cauwenberghs, G.; Pedroni, V. A charge-based CMOS parallel analog vector quantizer. *Adv. Neural Inf. Process. Syst.* **1994**, *7*, 779–786.

57. Lu, J.; Young, S.; Arel, I.; Holleman, J. A 1 tops/w analog deep machine-learning engine with floating-gate storage in 0.13 μm cmos. *IEEE J. Solid-State Circuits* **2014**, *50*, 270–281. [CrossRef]

58. Yamasaki, T.; Shibata, T. An analog similarity evaluation circuit featuring variable functional forms. In Proceedings of the ISCAS 2001, The 2001 IEEE International Symposium on Circuits and Systems (Cat. No. 01CH37196), Sydney, NSW, Australia, 6–9 May 2001; Volume 3, pp. 561–564.

59. Zhao, Z.; Srivastava, A.; Peng, L.; Chen, Q. Long short-term memory network design for analog computing. *ACM J. Emerg. Technol. Comput. Syst. JETC* **2019**, *15*, 1–27. [CrossRef]

60. Odame, K.; Nyamukuru, M. Analog LSTM for Keyword Spotting. In Proceedings of the 2022 IEEE 4th International Conference on Artificial Intelligence Circuits and Systems (AICAS), Incheon, Republic of Korea, 13–15 June 2022; pp. 375–378. [CrossRef]

61. Tsai, H.; Ambrogio, S.; Mackin, C.; Narayanan, P.; Shelby, R.M.; Rocki, K.; Chen, A.; Burr, G.W. Inference of Long-Short Term Memory networks at software-equivalent accuracy using 2.5M analog Phase Change Memory devices. In Proceedings of the 2019 Symposium on VLSI Technology, Kyoto, Japan, 9–14 June 2019; pp. T82–T83. [CrossRef]

62. Adam, K.; Smagulova, K.; James, A.P. Memristive LSTM network hardware architecture for time-series predictive modeling problems. In Proceedings of the 2018 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), Chengdu, China, 26–30 October 2018; pp. 459–462. [CrossRef]

63. Li, F.; Chang, C.H.; Siek, L. A compact current mode neuron circuit with Gaussian taper learning capability. In Proceedings of the 2009 IEEE International Symposium on Circuits and Systems, Taipei, Taiwan, 24–27 May 2009; pp. 2129–2132.

64. Reynolds, D.A. Gaussian mixture models. *Encycl. Biom.* **2009**, *741*, 659–663.

65. Alimisis, V.; Gennis, G.; Touloupas, K.; Dimas, C.; Uzunoglu, N.; Sotiriadis, P.P. Nanopower Integrated Gaussian Mixture Model Classifier for Epileptic Seizure Prediction. *Bioengineering* **2022**, *9*, 160. [CrossRef] [PubMed]

66. Miller, R. Theory of the normal waking EEG: From single neurones to waveforms in the alpha, beta and gamma frequency ranges. *Int. J. Psychophysiol.* **2007**, *64*, 18–23. [CrossRef] [PubMed]

67. Chen, M.; Boric-Lubecke, O.; Lubecke, V.M. 0.5-μm CMOS Implementation of Analog Heart-Rate Extraction with a Robust Peak Detector. *IEEE Trans. Instrum. Meas.* **2008**, *57*, 690–698. [CrossRef]