An Analog Bayesian Classifier Implementation, for Thyroid Disease Detection, based on a Low-Power, Current-Mode Gaussian Function Circuit

Vassilis Alimisis, Georgios Gennis, Christos Dimas and Paul P. Sotiriadis

Department of Electrical and Computer Engineering
National Technical University of Athens, Greece
E-mail: alimisisv@gmail.com, giorgosyennis@gmail.com, chdim@central.ntua.gr, pps@ieee.org

2021 International Conference on Microelectronics (ICM) | 978-1-6654-0839-4/21/$31.00 ©2021 IEEE | DOI: 10.1109/ICM52667.2021.9664939

***Abstract*—The thyroid gland is a small organ that's located in the front of the neck, wrapped around the windpipe. Thyroid releases and controls hormones that help the metabolism work correctly. Metabolism plays a main role in many different systems throughout the human body. Thyroid disorder involves the abnormal production of thyroid hormones. In this regard, if a thyroid disease could be detected, patients could take a specific treatment and greatly reduce the symptoms. This work proposes a novel low power, low voltage (0.6V) analog architecture of a Bayesian classifier for thyroid disease detection. The architecture is based on a new Gaussian function circuit and the Lazzaro Winner-Take-All circuit. The proper operation of the analog classifier is verified using a real-world dataset. The proposed architecture is realized in TSMC 90nm CMOS process and was simulated using the Cadence IC Suite.**

***Index Terms*—Bayesian Classifier, thyroid disease detection, low-power design, analog VLSI implementation**

I. Introduction

The Internet of Things (IoT) is expanding rapidly, creating an environment of devices and sensors that in many cases will function entirely on batteries [1], [2]. Many consumer and industrial applications have systems which utilize IoT devices. Some of these devices are employed without online recharging capabilities. As a result hardware designers are relying on power management solutions to efficiently handle the power needed.

There is a new trend in which IoT applications are combined with Machine Learning (ML) algorithms in order to extract useful information from real-time measurements [3]. With the aim of real-time computation, a new domain is growing based on new computation methods (edge computing and analog computing). Edge computing [4] is a method in which data are processed at the periphery of the network, as close to the originating source as possible. Moreover, analog computing [5] is closer to the physical laws by which all computation is realized (which are continuous), analog circuits often use fewer devices (area efficiency) than corresponding digital circuits. Also, analog computing is characterised by implemented circuits operating in the sub-threshold region [6] which reduces the power consumption of a system.

Recently, an increasing number of research topics on wireless remote medical device for monitoring various physiological parameters about human diseases require portable mobile ability (wearable architectures) [7]. Motivated by the low-power and low-area requirements of analog computing for ML and IoT applications which also deals with the problems of digital bottleneck, we introduce a low-power, low-voltage area efficient and high-speed analog Bayesian classifier for thyroid disease detection. The proposed classifier is designed and verified on a real-world thyroid disease dataset [8].

The remainder of this paper is organized as follows. Section II refers to the background of this work. More specifically, the characteristics of a thyroid disease and a brief presentation of the mathematical bayesian model are provided. The proposed architecture and the basic building blocks of the proposed classifier are described in Section III. The proper behavior of the proposed Bayesian classifier is confirmed via a real-world thyroid disease dataset in Section IV. Some concluding remarks are given in Section V.

II. Background

A. Thyroid Disease

Thyroid is a butterfly-shaped gland in the front of the neck [8], [9]. More specifically, it is a very small organ which is located below the Adam's apple wrapped around the trachea (windpipe). The thyroid is part of the endocrine system, which is made up of glands that produce, store, and release hormones into the bloodstream so the hormones can reach the body's cells. Its main role is the regulation of a person's metabolism (has important roles to regulate numerous metabolic processes) by producing hormones. They also aid in muscle control, brain development, mood regulation and digestive function.

Thyroid disorders are conditions that affect this organ. A person can develop numerous problems if this organ abnormaly produces hormones [10]. The two most common types of thyroid disease are hyperthyroidism and hypothyroidism. Other abnormal conditions include thyroiditis, thyroid nodules, goiter, and thyroid cancer. After the detection, treatment options depend on the specific form of thyroid disease, and include medications, radioactive iodine, and sometimes surgery.

Hypothyroidism results from the thyroid gland producing an insufficient amount of thyroid hormones. Its main symptoms include tiredness and fatigue, trouble sleeping, depression, sensitivity to cold temperatures, dry skin and hair, difficulty in concentrating, frequent and heavy periods, and joint and muscle pain. On the other hand, hyperthyroidism describes excessive production of thyroid hormone, a less common condition than hypothyroidism. The main symptoms are anxiety, irritability or moodness, hyperactivity, sweating or sensitivity to high temperatures, hand trembling, hair loss, and missed or light menstrual periods.

B. Bayesian Model

A naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' theorem considering independence between the input features [11]. Despite this assumption, when combined with kernel density estimation, it can achieve high accuracy.

Using Bayes' theorem, the conditional probability of a vector input X originating from a class C_k can be expressed as:

$$p(C_k|X) = \frac{p(C_k)p(X|C_k)}{p(X)}, \quad (1)$$

where, $p(C_k)$ is the prior probability of class k, $p(X)$ is the evidence probability of the input X and $p(X|C_k)$ is the value of the probability density function (PDF) of class k for the input X. For a Gaussian PDF with a diagonal covariance matrix (as assumed by the Bayesian model), $p(X|C_k)$ is given by:

$$p(X|C_k) = \prod_{n=1}^{N} \frac{1}{\sqrt{(2\pi)\cdot\sigma_{kn}^2}}\, e^{-\frac{1}{2}\cdot\frac{(x_n-\mu_{kn})^2}{\sigma_{kn}^2}}. \quad (2)$$

Here, N is the number of features, μ_{kn} and σ_{kn} are the mean value and the variance corresponding to the n-th feature of class k, repsectively and x_n is the n-th feature of the input vector X.

The overall classifier determines the winning class by applying the argmax operator in the probabilities $p(C_k|X)$ for all the classes. In practice, the evidence probability is ignored and, therefore, the output of the classifier is described by:

$$y = \underset{k\in[1,K]}{\text{argmax}}\{p(C_k|X)\} = \underset{k\in[1,K]}{\text{argmax}}\{p(C_k)p(X|C_k)\}. \quad (3)$$

III. Proposed Architecture

The architecture of an analog Bayesian classifier is explained in this Section. For generality, the classifier is based on the mathematical analysis of the previous Section. The main building block of a Gaussian Bayesian Model is the Gaussian function. For that case, there are specific analog circuits, called Bump circuits, which implement the Gaussian function [12]. In this work, a novel current-mode Bump circuit, which uses a Lazzaro Winner-take-All (WTA) [13] neuron as a basic building block, is introduced.

The proposed Bump circuit, shown in Fig. 1, consists of two sub-circuits, a symmetric current correlator [14] and two Lazzaro WTA neuron cells. The neuron cells replace the differential block of a typical Bump circuit topology [15], [16]. This design achieves electronical tunability in all the characteristics of a Gaussian function (mean value, variance,

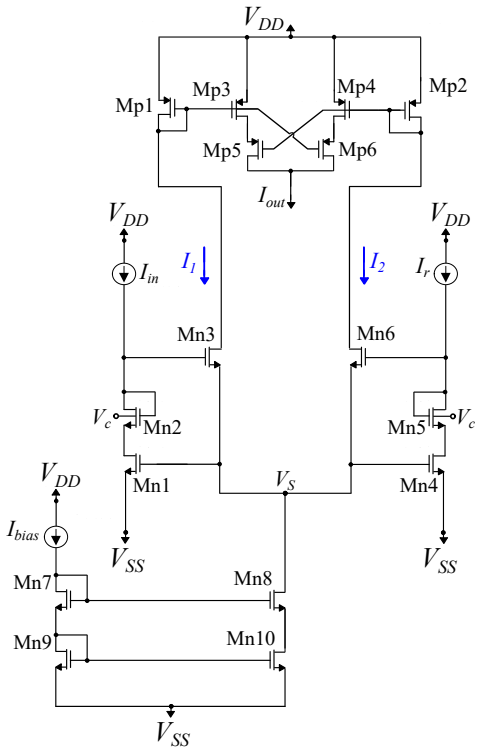

Fig. 1: Proposed Bump circuit.

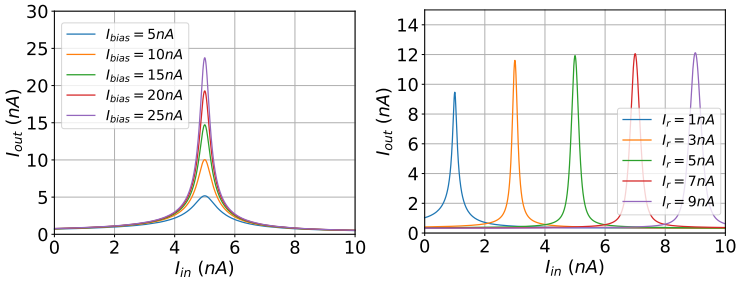

Fig. 2: Parametric analysis (left) over I_{bias}, for $I_r = 5nA$, $V_c = 0V$ and $M = 1$ (right) over I_r, for $I_{bias} = 12nA$ $V_c = 0.2V$ and $M = 1$.

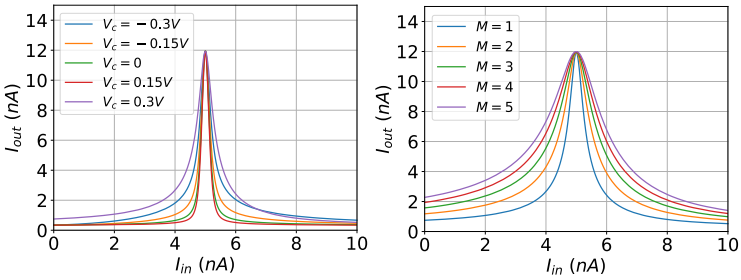

Fig. 3: Parametric analysis (left) over V_c, for $I_{bias} = 12nA$, $I_r = 5nA$ and $M = 1$ (right) over M, for $I_{bias} = 12nA$, $I_r = 5nA$ and $V_c = 0.3V$.

height). Specifically, the current parameter I_r, the voltage parameter V_c and the bias current I_{bias} control the mean value, the variance and the height, respectively. The variance is also tuned by the neuron (M_{n3}, M_{n6}) transistors' dimension ratio $M \cdot (\frac{W}{L})_{3,6}$. The appropriate curves are provided in Figs. 2 and 3. All transistors operate in the sub-threshold region and their dimensions are summarized in Table I. The power supply rails are set as $V_{DD} = -V_{SS} = 0.3V$.

TABLE I: MOS Transistors' Dimensions (Fig. 1).

Block	W/L (µm/µm)	Current Correlator	W/L (µm/µm)
M_{n1}-M_{n9}	0.4/1.6	M_{p1}-M_{p6}	0.4/1.6
M_{n10}	0.8/1.6	-	-

The mathematical Gaussian Bayes model is based on a Gaussian PDF. Specifically, in real-world applications this PDF is a multivariate Gaussian function. Based on (2) the multivariate Gaussian PDF is calculated by multiplying univariate Gaussian PDFs. By connecting two or more bump circuits in a cascaded form this multiplication is achieved [14]. In particular, the output current of the $(n$-1)-th Bump circuit is connected to the bias current of the n-th Bump. Only the first bump circuit is biased with a specific bias current (I_{bias}) representing the prior probability ($p(C_k)$) of the corresponding class k.

The second basic building block of the analog Bayes classifier is the WTA circuit, shown in Fig. 4 [13]. A WTA circuit is capable of performing the argmax operator. Specifically, given a set of N input signals and assuming that there is a single maximum among them, located at index j $\leq$ N, the output I_{out_j} has a non-zero value (winner), whereas the rest are zero. This behavior is achieved when all the transistors operate in the sub-threshold region. The transistors' dimensions are equal to $W/L = 0.4$µm$/1.6$µm. Based on (3) and utilizing the aforementioned building blocks, the proposed classifier is shown in Fig. 5.

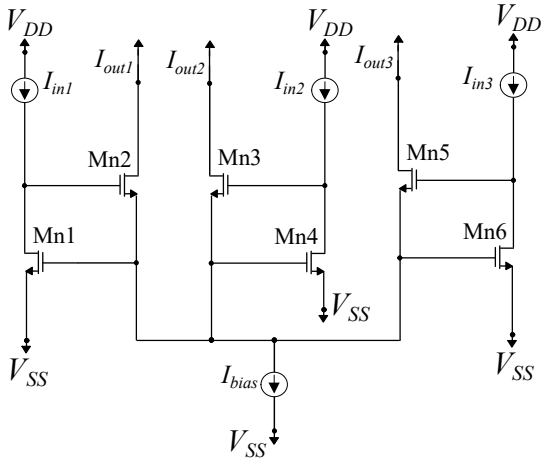

Fig. 4: Lazzaro WTA circuit with 3 neurons.

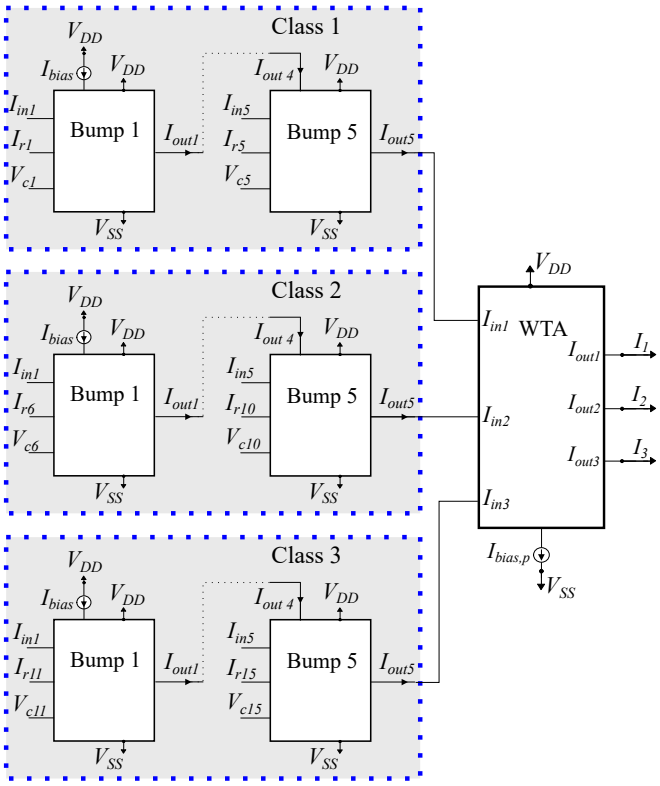

Fig. 5: Analog Bayesian classifier with 3 classes and 5-D inputs. (left) Multivariate Gaussian function circuits (right) WTA circuit.

IV. Application Example and Simulation Results

In this Section, the proper operation of the proposed Bayesian classifier is confirmed via a real-world thyroid disease detection problem. The proposed architecture has been designed in a TSMC $90nm$ CMOS process using the Cadence IC suite. The power supply rails for the whole classifier are set as $V_{DD} = -V_{SS} = 0.3V$. All the simulation results are conducted on the layout (post-layout simulations) presented in Fig. 6. The dataset is acquired from the University of California, Irvine (UCI) Machine Learning Repository [8] and contains (five) blood test metrics (related to thyroid) for patients with normal thyroid, hypothyroidism and hyperthyroidism. These metrics are feed directly to the classifier. The system's necessary parameters are provided by calculating the mean value, the variance and the prior probability of each class.

To highlight the gains of the proposed Bayesian classifier, two different tests are conducted. The first test compares the proposed architecture with a software-based one in terms of classification accuracy. Specifically, in Fig. 7 the classification accuracies over 20 different training test-cases are provided for both implementations. The results are also summarized in Table II. The second, tests the circuit's sensitivity via Monte Carlo analysis for $N = 100$ points. The Monte Calro histogram, shown in Fig. 8, has a mean value of $\mu_M = 0.916$ and a standard deviation of $\sigma_M = 0.023$. Both testcases confirm

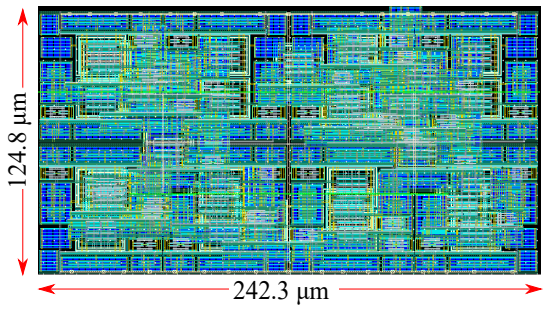

Fig. 6: Proposed Classifier's Layout.

the performance, the high accuracy and the desired sensitivity of the proposed classifier. The performance characteristics are summarized in Tables II and III.

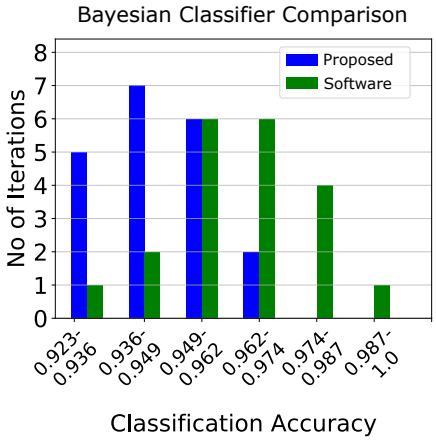

Fig. 7: Post-layout simulation results over 20 iterations.

TABLE II: Accuracy Results (over 20 iterations).

Method	Best	Worst	Mean	Std.
Software	1.000	0.9231	0.9639	0.0184
Proposed	0.9692	0.9231	0.9423	0.0149

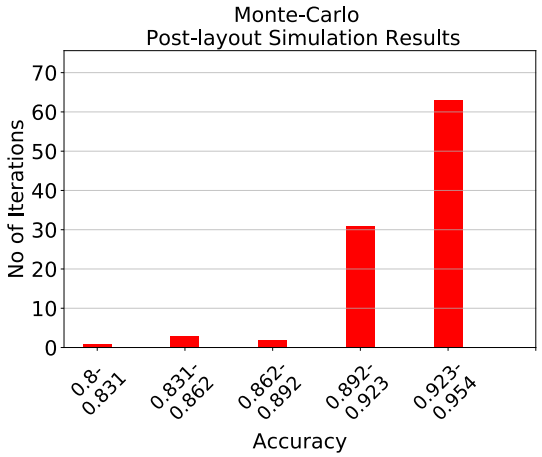

Fig. 8: Post-layout Monte Carlo simulation results.

TABLE III: Performance Summary

Technology	Power Supply	Power Consumption	Classification speed
$90nm$	$0.6V$	$365nW$	$170K \frac{\text{classifications}}{\text{second}}$

V. Conclusion

In this work, a novel architecture of an analog Bayesian classifier for thyroid disease prediction was proposed. The presented architecture consists of an alternative implementation for a Bump circuit and the Lazzaro WTA circuit. The proposed classifier can be used as a basic building block for the design of more complicated and accurate diagnosis systems. Based on the post-layout simulation results, the implemented architecture achieves 94.23% classification accuracy and reasonable sensitivity characteristics.

References

[1] J. Henkel, S. Pagani, H. Amrouch, L. Bauer, and F. Samie, "Ultra-low power and dependability for iot devices," *Invited Paper for IoT Technologies)*, pp. 978–3, 2017.

[2] K. Goebel, B. Saha, A. Saxena, J. R. Celaya, and J. P. Christophersen, "Prognostics in battery health management," *IEEE instrumentation & measurement magazine*, vol. 11, no. 4, pp. 33–40, 2008.

[3] Y. Chen, T. Chen, Z. Xu, N. Sun, and O. Temam, "Diannao family: energy-efficient hardware accelerators for machine learning," *Communications of the ACM*, vol. 59, no. 11, pp. 105–112, 2016.

[4] W. Shi and S. Dustdar, "The promise of edge computing," *Computer*, vol. 49, no. 5, pp. 78–81, 2016.

[5] W. Haensch, T. Gokmen, and R. Puri, "The next generation of deep learning hardware: Analog computing," *Proceedings of the IEEE*, vol. 107, no. 1, pp. 108–122, 2018.

[6] A. Wang, B. H. Calhoun, and A. P. Chandrakasan, *Sub-threshold design for ultra low-power systems*. Springer, 2006, vol. 95.

[7] L. Zhang, H. Ji, H. Huang, N. Yi, X. Shi, S. Xie, Y. Li, Z. Ye, P. Feng, T. Lin *et al.*, "Wearable circuits sintered at room temperature directly on the skin surface for health monitoring," *ACS Applied Materials & Interfaces*, vol. 12, no. 40, pp. 45 504–45 515, 2020.

[8] D. J. Newman, S. Hettich, C. L. Blake, and C. J. Merz, "Uci repository of machine learning databases, 1998," 1998. [Online]. Available: https://archive.ics.uci.edu/ml/machine-learning-databases/thyroid-disease/

[9] M. P. Vanderpump, "The epidemiology of thyroid disease." *British medical bulletin*, vol. 99, no. 1, 2011.

[10] A. Jabbar, A. Pingitore, S. H. Pearce, A. Zaman, G. Iervasi, and S. Razvi, "Thyroid hormones and cardiovascular disease," *Nature Reviews Cardiology*, vol. 14, no. 1, pp. 39–55, 2017.

[11] C. M. Bishop, "Pattern recognition," *Machine learning*, vol. 128, no. 9, 2006.

[12] V. Alimisis, M. Gourdouparis, G. Gennis, C. Dimas, and P. P. Sotiriadis, "Analog gaussian function circuit: Architectures, operating principles and applications," *Electronics*, vol. 10, no. 20, p. 2530, 2021.

[13] J. Lazzaro, S. Ryckebusch, M. A. Mahowald, and C. A. Mead, "Winner-take-all networks of o (n) complexity," 1988.

[14] V. Alimisis, M. Gourdouparis, C. Dimas, and P. P. Sotiriadis, "A 0.6 v, 3.3 nw, adjustable gaussian circuit for tunable kernel functions," in *2021 34th SBC/SBMicro/IEEE/ACM Symposium on Integrated Circuits and Systems Design (SBCCI)*. IEEE, 2021, pp. 1–6.

[15] ——, "Ultra-low power, low-voltage, fully-tunable, bulk-controlled bump circuit," in *2021 10th International Conference on Modern Circuits and Systems Technologies (MOCAST)*. IEEE, 2021, pp. 1–4.

[16] M. Gourdouparis, V. Alimisis, C. Dimas, and P. P. Sotiriadis, "An ultra-low power,±0.3 v supply, fully-tunable gaussian function circuit architecture for radial-basis functions analog hardware implementation," *AEU-International Journal of Electronics and Communications*, vol. 136, p. 153755, 2021.