

An Analog Integrated, Low-Power, Area-Efficient, Gilbert, Modulo-based Classifier with Application to Lung-Cancer Classification

Vassilis Alimisis, Nikolaos P. Eleftheriou, Savvas Leventikidis and Paul P. Sotiriadis

Department of Electrical and Computer Engineering
National Technical University of Athens, Greece

E-mail: alimisisv@gmail.com, eleftheriou_nikos@hotmail.com, savvas01@yahoo.gr, pps@ieee.org

Abstract—This study presents an alternative approach to develop low-power (744nW) analog classifiers capable of efficiently handling multiple input features while maintaining high levels of accuracy and minimizing power consumption. The proposed classifier relies on Voting and Bayes mathematical models, incorporating Gilbert two-signal four-quadrant multipliers and current comparators. The analog classifier is validated through testing with a real-world lung-cancer surgery dataset, achieving an accuracy of 75.45%. It predicts all testset samples of patients suffering from lung-cancer. Additionally, a comparison with related analog classifiers using the same dataset is conducted. The models are trained via a software-based implementation. The proposed architecture is realized using the TSMC 90nm CMOS process and simulated using the Cadence IC Suite.

Index Terms—Modulo-based classifier, Lung-cancer classification, low-power design, analog VLSI implementation

I. INTRODUCTION

The rapid expansion of the Internet of Things (IoT) has given rise to a variety of devices and sensors, many of which operate solely on batteries, making efficient power management crucial [1]. IoT devices find application in various consumer and industrial sectors, some of which lack online recharging capabilities. To address this, hardware designers are increasingly turning to innovative power management solutions.

A notable emerging trend involves integrating IoT applications with Machine Learning (ML) algorithms to extract valuable insights from real-time data [2]. In pursuit of real-time computation, a new domain is emerging, leveraging advanced computation methods like edge computing and analog computing. Edge computing [3] processes data as close to the source as possible, enhancing speed and efficiency. Analog computing [4] aligns more closely with the continuous nature of physical laws, often requiring fewer components compared to digital circuits. Additionally, analog circuits, by operating in the sub-threshold region [5], significantly reduce power consumption.

Recent advancements in wireless remote medical devices have sparked interest in monitoring various physiological parameters related to human health conditions with a focus on portability, particularly through wearable architectures [6].

Motivated by the demand for low-power and low-area solutions in analog computing for ML and IoT applications, an efficient, high-speed analog Bayesian classifier designed for lung-cancer classification is introduced. The proposed classifier has been rigorously tested on a real-world lung-cancer surgery dataset [7].

The remainder of this paper is organized as follows. Section II refers to a brief presentation of classifier's mathematical model. The proposed architecture and the basic building blocks of the proposed classifier are described in Section III. The proper behavior of the proposed classifier is confirmed via a real-world lung-cancer classification dataset and compared with the software-based implementation in Section IV. Section V provides a comparison study with related analog classifiers. Some concluding remarks are given in Section VI.

II. MATHEMATICAL MODELLING

The Naive Bayes classifier is a straightforward probabilistic classification method that applies Bayes' theorem while assuming independence between input features [8]. Even with this assumption, it can achieve impressive accuracy when combined with kernel density estimation. By employing Bayes' theorem, the conditional probability of a vector input X belonging to a class C_k is expressed as:

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

In this context, $p(C_k)$ represents the prior probability of class k , $p(X)$ denotes the evidence probability of the input X , and $p(X|C_k)$ signifies the value of the probability density function (PDF) of class k for the input X . Specifically, for a multivariate Gaussian PDF with a diagonal covariance matrix, as assumed by the Bayesian model, $p(X|C_k)$ is defined as:

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

In this context, N represents the number of features, leading to the generation of $N - d$ Gaussian functions. Parameters μ_{kn} and σ_{kn} denote the mean value and variance corresponding to the n -th feature of class k respectively, while x_n stands for the

n -th feature of the input vector X . The final decision for the winning class is taken by applying the argmax operator to the probabilities $p(C_k|X)$ for all classes. In practical application, the evidence probability is often disregarded, and the output of the classifier can be described as:

$$y = \operatorname{argmax}\{p(C_k|X)\} = \operatorname{argmax}\{p(C_k)p(X|C_k)\}. \quad (3)$$

for $k \in \{1, 2, \dots, K\}$.

In this study, to represent each sub-class using a single feature (a $1 - D$ cell), the mathematical model is articulated through a voting classifier [9], which can be approximated as:

$$y = \operatorname{mod}\{(C_1(X), C_2(X), C_3(X), \dots, C_K(X))\}. \quad (4)$$

In this context, $C_k(x)$ signifies the output of each 1-D Gilbert decision cell (GDC), essentially representing each sub-classifier. To further illustrate this concept, let's consider a scenario with five features involved in a binary classification task, each carrying equal weight. The functions $C_1(x)$, $C_2(x)$, $C_3(x)$ collectively yield the output for the first class (class 1), while $C_4(x)$, $C_5(x)$ produces the output for the second class (class 0). Consequently, the result is calculated as $y = \operatorname{mod}(1, 1, 1, 0, 0) = 1$ (indicating class 1).

III. PROPOSED ARCHITECTURE

In this section, the proposed architecture of the analog classifier along with its basic building blocks is presented. Since it can accommodate various numbers of classes and input dimensions, it is scalable and provides high versatility. Firstly, for the realization of GDC in equation (4), Gilbert two signal four-quadrant multipliers (Gilbert cells) [5] along with current comparators (Winner-take-all circuit) [10] are employed, as shown in Fig. 1. Transistors M_{n1} - M_{n4} and M_{n7} - M_{n10} implement the two Gilbert cells and transistors M_{n5} , M_{n6} , M_{n11} and M_{n12} implement the Winner-take-all (WTA) circuit. The GDC circuit operates in a translinear fashion, producing two decision output currents that signify the decisions for each feature in both classes. In this context, a higher current indicates the winning class.

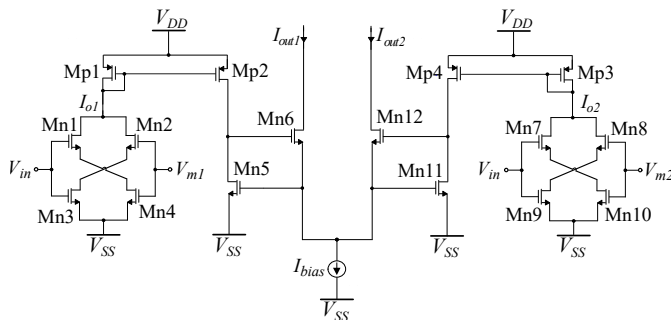


Fig. 1: The implementation of the GDC circuit. It consists of Gilbert two signal four-quadrant multipliers (Gilbert cells) along with current comparators. Here V_{in} is the input voltage and V_{m1} and V_{m2} are parameters describing the mean values of each function. The output current represents the decision according to a specific feature.

The architecture of the proposed classifier, as shown in Fig. 2, is designed for a classification problem involving

$N_{cla} = 2$ classes and $N_d = 16$ features (input dimensions). This illustration consists of 16 GDC circuits (input dimensions) and one WTA circuit (modulo implementation), shown in Fig. 2. Each GDC circuit describes the voting strength for each class regarding a specific feature. It produces two output currents, each one represent a feature's decision. All the currents related to one class are summed via current mirrors (CMs) to minimize potential distortions in calculations that might arise from undesired effects on the output currents of the GDC. The resulting output currents, distinguished by their high or low values, indicate the classifier's final prediction. All transistor dimensions are set to $W/L = 1.6\mu\text{m}/1.6\mu\text{m}$. The power supply rails set as $V_{DD} = -V_{SS} = 0.3$ V and all transistors operate in the sub-threshold region in order to achieve low power consumption.

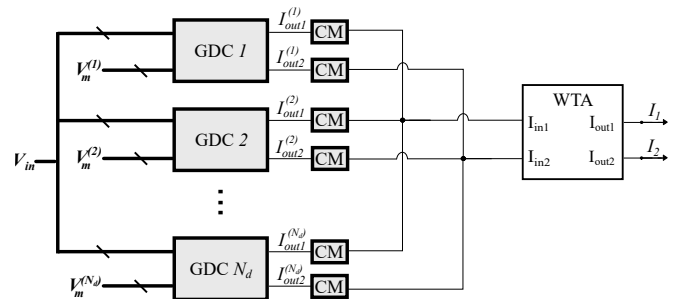


Fig. 2: The proposed classifier's top-level architecture.

IV. LUNG-CANCER DATASET AND SIMULATION RESULTS

In this section, the effectiveness of the proposed classifier using a real-world dataset related to lung-cancer [7] is challenged. The proposed architecture has been implemented in the TSMC 90nm CMOS process, employing the Cadence IC suite. All the simulation tests are conducted on the layout (post-layout simulations) illustrated in Fig. 3. This classification task revolves around lung-cancer, encompassing $N_{cla} = 2$ distinct classes and $N_d = 16$ inputs. As for the classifier's training, a software-based implementation is employed to tune the required parameters. All related metrics are directly fed into the hardware classifier. The necessary parameters for the system are computed by evaluating the mean value and prior probability of each class.

The Thoracic Surgery Data is a dataset available from the UCI Machine Learning Repository [7]. It encompasses clinical information about patients who underwent thoracic surgery for lung-cancer treatment. This dataset is valuable for research in medical and healthcare fields, as it includes attributes such as age, performance status, tumor size, and other relevant factors. The aim is to predict the survival status of patients after surgery based on these features. This dataset serves as a valuable resource for machine learning practitioners and researchers aiming to develop predictive models in the context of thoracic surgery outcomes.

To evaluate the proposed classifier's performance in terms of classification specificity and the circuit's behaviour under

Process-Voltage-Temperature (PVT) variations, two distinct tests were carried out on the layout. To account for random effects, the outcomes of 20 different training-testing iterations are depicted in Fig. 4. It predicts all the patients who have cancer but it has false-positive alarms for patients who have not cancer. As a result, it can be used as a wake-up engine for a digital back-end. Subsequently, the circuit's sensitivity on random variations is affirmed through a Monte Carlo analysis. Specifically, Fig. 5 displays the Monte Carlo Histogram for $N = 100$ data points.

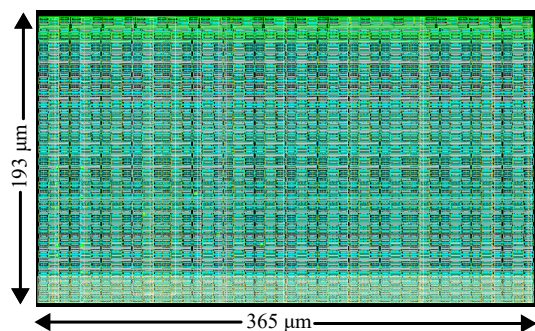


Fig. 3: Layout of the proposed classifier's architecture based on the design methodology (extra dummy transistors are used).

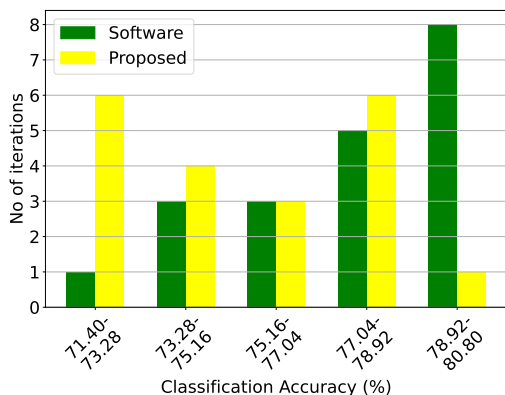


Fig. 4: Classification results of the proposed architecture and the equivalent software model on the lung-cancer classification dataset over 20 iterations.

V. COMPARISON STUDY AND DISCUSSION

In related literature, it is clear that the majority of analog classifiers are typically tailored to specific applications. It is a great challenge to compare different ML models or hardware implementations on the same application and deduce fair results. However, this challenge enables for adapting analog classifiers to serve a common application, thereby simplifying the process of evaluating performance that encompasses both ML models and alternative methodologies. Table I offers an overview of the performance comparison along a variety of related classifiers. Here, Gaussian Mixture Model (GMM) [11], Radial Basis Function [12], Long Short-Term Memory (LSTM) [13], K-means [14], Bayesian [17], ANN (Artificial

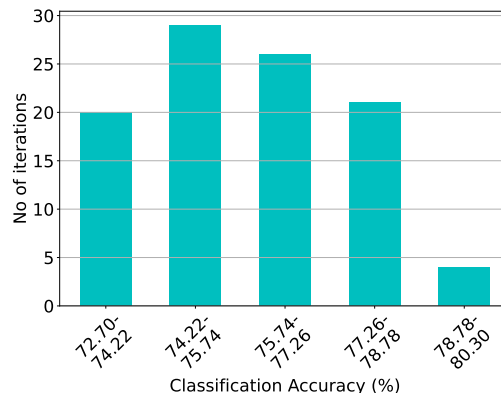


Fig. 5: Post-layout Monte-Carlo simulation results of the proposed architecture on the lung-cancer classification dataset with $\mu_M = 75.87\%$ and a standard deviation of $\sigma_M = 1.73\%$.

Neural Network) [15], Fuzzy [16], Support Vector Machine (SVM) [18], Threshold [19], Multilayer Perceptron (MLP) [20] and centroid-based [21] classifiers, all within the context of lung-cancer disease classification, are summarized.

The presented research introduces a solution, offering a balance between accuracy, power efficiency, and energy consumption per classification when compared to equivalent classifiers in the field. It's imperative to highlight that in this specific application, the design deal with a high input dimensionality. The proposed configuration holds a significant edge by obviating the necessity for Principal Component Analysis (PCA), allowing for the incorporation of all 16 input dimensions without any loss of crucial information. In contrast, many alternative topologies must reduce the dimensions to 11 to achieve optimal accuracy, representing a noteworthy constraint in prior similar studies [13]–[15], [20]. While the proposed classifier demonstrates its proficiency in accurately classifying a broader range of classes, we opt for a binary classification scenario to ensure a fair comparison. This adjustment facilitates a more meaningful assessment in relation to binary analog classifiers [16], [18], [19].

In terms of classification accuracy, the proposed architecture outperforms all its counterparts, except for MLP [20], LSTM [13] and K-means [14]. While these models achieve higher accuracy, they come at the cost of increased complexity and power consumption along with a larger silicon area due to their components number. On the other end of the spectrum, the Threshold classifier achieves the lowest power consumption in comparison with the other classifiers, albeit with a trade-off in accuracy and processing speed, attributed to its simple model design [19]. It's important to note that in biomedical applications of this kind, swift processing speed isn't of paramount importance, primarily due to their infrequent occurrence. Therefore, in the analysed approach, processing speed is decreased to enhance accuracy and optimize power consumption. Additionally, it touts lower energy consumption per classification compared to all classifiers, except for ANN [15], which achieves a lower classification accuracy.

TABLE I: Analog classifiers' comparison on the Lung-Cancer Disease Classification

	Classifier	Worst accuracy	Mean accuracy	Best accuracy	Power consumption	Processing speed	Energy per classification	No. of Dimensions
This work	Modulo	71.40%	75.45%	79.50%	744nW	320k $\frac{\text{classifications}}{\text{s}}$	$\frac{2.33 \text{ pJ}}{\text{classification}}$	16
[11]	GMM	68.40%	71.27%	73.80%	2.97 μ W	100K $\frac{\text{classifications}}{\text{s}}$	$\frac{29.70 \text{ pJ}}{\text{classification}}$	11
[12]	RBF	66.70%	70.41%	72.70%	27.87 μ W	200k $\frac{\text{classifications}}{\text{s}}$	$\frac{139.35 \text{ pJ}}{\text{classification}}$	11
[13]	LSTM	94.10%	97.54%	100.00%	22.54mW	870M $\frac{\text{classifications}}{\text{s}}$	$\frac{25.91 \text{ pJ}}{\text{classification}}$	16
[14]	K-means	88.30%	91.41%	95.10%	111.12 μ W	5M $\frac{\text{classifications}}{\text{s}}$	$\frac{22.22 \text{ pJ}}{\text{classification}}$	16
[15]	ANN	68.90%	72.43%	76.50%	2.63 μ W	14M $\frac{\text{classifications}}{\text{s}}$	$\frac{0.19 \text{ pJ}}{\text{classification}}$	16
[16]	Fuzzy	73.80%	78.65%	81.60%	3.67 μ W	4.55K $\frac{\text{classifications}}{\text{s}}$	$\frac{806.59 \text{ pJ}}{\text{classification}}$	11
[17]	Bayes	63.70%	68.72%	71.30%	1.79 μ W	100K $\frac{\text{classifications}}{\text{s}}$	$\frac{17.90 \text{ pJ}}{\text{classification}}$	11
[18]	SVM	70.10%	72.37%	74.70%	67.63 μ W	140K $\frac{\text{classifications}}{\text{s}}$	$\frac{483.07 \text{ pJ}}{\text{classification}}$	11
[19]	Threshold	67.60%	70.77%	75.90%	920nW	100K $\frac{\text{classifications}}{\text{s}}$	$\frac{9.20 \text{ pJ}}{\text{classification}}$	11
[20]	MLP	86.10%	87.56%	89.40%	354.18 μ W	930k $\frac{\text{classifications}}{\text{s}}$	$\frac{380.84 \text{ pJ}}{\text{classification}}$	16
[21]	Centroid	71.40%	73.87%	76.30%	2.98 μ W	100K $\frac{\text{classifications}}{\text{s}}$	$\frac{29.80 \text{ pJ}}{\text{classification}}$	11

VI. CONCLUSION

In this work, an alternative approach for a power-efficient (744nW), low voltage (0.6V), analog classifier for lung-cancer surgery classification was proposed. The presented architecture consists of Gilbert two-signal four-quadrant multipliers and current comparators. The circuit's parameters were adjusted through offline training of a Bayes software classifier. The post-layout simulation was conducted through a TSMC 90nm CMOS process and the results were assessed in comparison with both a software-based implementation and a variety of related analog classifiers. The realized architecture demonstrates a decent classification accuracy of 75.45% with notable sensitivity properties.

REFERENCES

- [1] K. Gulati, R. S. K. Boddu, D. Kapila, S. L. Bangare, N. Chandnani, and G. Saravanan, "A review paper on wireless sensor network techniques in internet of things (iot)," *Materials Today: Proceedings*, vol. 51, pp. 161–165, 2022.
- [2] J. P. Bharadiya, "Leveraging machine learning for enhanced business intelligence," *INTERNATIONAL JOURNAL OF COMPUTER SCIENCE AND TECHNOLOGY*, vol. 7, no. 1, pp. 1–19, 2023.
- [3] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE communications surveys & tutorials*, vol. 19, no. 4, pp. 2322–2358, 2017.
- [4] 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.
- [5] S.-C. Liu, *Analog VLSI: circuits and principles*. MIT press, 2002.
- [6] F. Custodio, F. J. Herrera, G. López, and J. I. Moreno, "A review on architectures and communications technologies for wearable health-monitoring systems," *Sensors*, vol. 12, no. 10, pp. 13907–13946, 2012.
- [7] [Online]. Available: <https://archive.ics.uci.edu/dataset/277/thoracic+surgery+data>
- [8] C. M. Bishop and N. M. Nasrabadi, *Pattern recognition and machine learning*. Springer, 2006, vol. 4, no. 4.
- [9] M. A. Khan, M. A. Khan Khattk, S. Latif, A. A. Shah, M. Ur Rehman, W. Boulila, M. Driss, and J. Ahmad, "Voting classifier-based intrusion detection for iot networks," in *Advances on Smart and Soft Computing: Proceedings of ICACIn 2021*. Springer, 2022, pp. 313–328.
- [10] J. Lazzaro, S. Ryckebusch, M. A. Mahowald, and C. A. Mead, "Winner-take-all networks of o (n) complexity," *Advances in neural information processing systems*, vol. 1, 1988.
- [11] V. Alimisis, G. Gennis, K. Touloupas, C. Dimas, M. Gourdouparis, and P. P. Sotiriadis, "Gaussian mixture model classifier analog integrated low-power implementation with applications in fault management detection," *Microelectronics Journal*, vol. 126, p. 105510, 2022.
- [12] S.-Y. Peng, P. E. Hasler, and D. V. Anderson, "An analog programmable multidimensional radial basis function based classifier," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 54, no. 10, pp. 2148–2158, 2007.
- [13] Z. Zhao, A. Srivastava, L. Peng, and Q. Chen, "Long short-term memory network design for analog computing," *ACM Journal on Emerging Technologies in Computing Systems (JETC)*, vol. 15, no. 1, pp. 1–27, 2019.
- [14] R. Zhang and T. Shibata, "An analog on-line-learning k-means processor employing fully parallel self-converging circuitry," *Analog Integrated Circuits and Signal Processing*, vol. 75, pp. 267–277, 2013.
- [15] S. T. Chandrasekaran, R. Hua, I. Banerjee, and A. Sanyal, "A fully-integrated analog machine learning classifier for breast cancer classification," *Electronics*, vol. 9, no. 3, p. 515, 2020.
- [16] E. Georgakilas, V. Alimisis, G. Gennis, C. Aletraris, C. Dimas, and P. P. Sotiriadis, "An ultra-low power fully-programmable analog general purpose type-2 fuzzy inference system," *AEU-International Journal of Electronics and Communications*, vol. 170, p. 154824, 2023.
- [17] V. Alimisis, G. Gennis, C. Dimas, and P. P. Sotiriadis, "An analog bayesian classifier implementation, for thyroid disease detection, based on a low-power, current-mode gaussian function circuit," in *2021 International conference on microelectronics (ICM)*. IEEE, 2021, pp. 153–156.
- [18] V. Alimisis, G. Gennis, M. Gourdouparis, C. Dimas, and P. P. Sotiriadis, "A low-power analog integrated implementation of the support vector machine algorithm with on-chip learning tested on a bearing fault application," *Sensors*, vol. 23, no. 8, p. 3978, 2023.
- [19] V. Alimisis, G. Gennis, E. Tsouvalas, C. Dimas, and P. P. Sotiriadis, "An analog, low-power threshold classifier tested on a bank note authentication dataset," in *2022 International Conference on Microelectronics (ICM)*. IEEE, 2022, pp. 66–69.
- [20] K. Lee, J. Park, and H.-J. Yoo, "A low-power, mixed-mode neural network classifier for robust scene classification," *Journal of Semiconductor Technology and Science*, vol. 19, no. 1, pp. 129–136, 2019.
- [21] V. Alimisis, V. Mouzakis, G. Gennis, E. Tsouvalas, C. Dimas, and P. P. Sotiriadis, "A hand gesture recognition circuit utilizing an analog voting classifier," *Electronics*, vol. 11, no. 23, p. 3915, 2022.