Modified Oregonator: an Approach from the Complex Networks Theory

Oregonador Modificado: un Enfoque desde la Teoría de Redes Complejas

Jesús Andrés Arzola Flores, José Fernando Rojas Rodríguez, Esmeralda Vidal Robles

Benemérita Universidad Autónoma de Puebla

ABSTRACT

Within the framework of Systems Biology, this paper proposes the complex network theory as a fundamental tool for determining the most critical dynamic variables in complex biochemical mechanisms. The Belousov-Zhabotinsky reaction is proposed as a study model and as a complex bipartite network. By determining the structural property *authority*, the most relevant dynamic variables are specified, and a mathematical model of the Belousov-Zhabotinsky reaction is obtained. The bidirectional coupling of the proposed model was made with other models associated with biological processes, finding synchronization phenomena when varying the coupling parameter. The time series obtained from the numerical solution of the coupled models were used to construct their images using the *Gramian Angular Field* technique. In the end, a supervised learning tool is proposed for the classification of the type of coupling by analyzing the images, obtaining score percentages above 94%. The hereby proposed methodology could be extended to the experimental field in order to determine anomalies in the coupling and synchronization of different physiological oscillators.

KEYWORDS: Systems Biology; BZ Reaction; Complex Networks; Supervised Learning; Gramian Angular Field

RESUMEN

En el marco de la Biología de sistemas, se propone en el presente trabajo a la teoría de redes complejas como una herramienta fundamental para la determinación de las variables dinámicas más importantes en mecanismos bioquímicos complejos. Se emplea como modelo de estudio la reacción de Belousov-Zhabotinsky y se plantea como una red compleja bipartita. Mediante la determinación de la propiedad estructural *autoridad*, se determinan las variables dinámicas con mayor relevancia y se obtiene un modelo matemático de la reacción de Belousov-Zhabotinsky. Se realizó el acoplamiento bidireccional del modelo planteado con otros modelos asociados a procesos biológicos, encontrándose fenómenos de sincronización al variar el parámetro de acoplamiento. Las series de tiempo obtenidas de la solución numérica de los modelos acoplados se emplearon para construir sus respectivas imágenes mediante la técnica de *campo angular gramiano*. Finalmente, se propone una herramienta de aprendizaje supervisado para la clasificación del tipo de acoplamiento mediante el análisis de las imágenes, obteniéndose porcentajes de exactitud por encima del 94%. La metodología propuesta en el presente trabajo podría extenderse y trasladarse al campo experimental con la finalidad de determinar anomalías en el acoplamiento y sincronización de distintos osciladores fisiológicos.

PALABRAS CLAVE: Biología de sistemas; Reacción BZ; Redes Complejas; Aprendizaje supervisado; Campo angular gramiano

Corresponding author

TO: **Esmeralda Vidal Robles**

INSTITUTION: **Benemérita Universidad Autónoma de Puebla**

ADDRESS: **Av. San Claudio S/N, Col. Ciudad Universitaria, C. P. 72570, Puebla, Puebla, México**

E-MAIL: **esmeralda.vidal@correo.buap.mx**

INTRODUCTION

It is possible to study any physiological process through an intricate network of biochemical reaction mechanisms from the systems biology paradigm [1] [2] [3] [4], such mechanisms are responsible for regulating a wide variety of processes of utmost importance for life, through complex positive and negative feedback systems [5] [6]. A prominent example of negative feedback is the thyroid hormone regulatory mechanism [7] carried out by either thyroid cells or the leptin-insulin axis [8], which is strongly related to metabolic processes. Discrepancies in these feedback mechanisms can lead to multiple pathologies at the systemic level [7]. With the development of experimental tools for studying such complex systems, the need to understand the underlying dynamics of these regulatory mechanisms also arose, so biochemists undertook the task of studying the chemistry of these mechanisms to determine the kinetic parameters of the reactions participating in these processes. Meanwhile, biophysicists had to translate these biochemical processes to mathematical models using tools based on Dynamical Systems Theory (DST) and the discoveries made by biochemists [9] [10] [11].

B. Belousov was a Russian biophysicist who pioneered in the study of complex regulatory mechanisms present in biochemical processes at an experimental level. In 1950 he was given the task of proposing a reaction mechanism analogous to the Krebs cycle to study feedback processes [12] [13] [14]. Belousov conceived a chemical mixture made essentially of citric acid, bromate ions, and cerium ions in an acid medium under constant agitation [13]. During the reaction, Belousov observed that the mixture changed from being transparent to a yellow hue and vice versa. The phenomenon occurred time after time, being reminiscent of the Krebs cycle feedback processes. Nevertheless, his work was never officially published because reviewers concluded that the process was caused by mixture impurities, and that the phenomenon violated the natural laws of thermodynamics [13] [14]. Years later, Prigogine defined the basis of thermodynamics of irreversible processes [12] [13] [14].

Later on, another Russian biophysicist, A. Zhabotisky resumed Belousov's work. He replaced citric acid with malonic acid and cerium ions with iron ions to visualize chemical species concentration-oscillations obtaining a solution that shifted from red to blue and vice versa [12]. This chemical mechanism was thus named the Belousov-Zhabotinsky (BZ) reaction [12] [13] [14]. The work of Belousov and Zhabotinsky set the basis for the study of oscillatory biochemical processes [14].

In 1972, Field, Köros and Noyes proposed the first mathematical model in differential equations that described the underlying dynamics of the BZ reaction (FKN Model) and laid the groundwork for mathematical modeling of chemical mechanisms with oscillatory behaviors [15] [16]. This opened a vast field of study for the development of mathematical models in the area of systems biology [17]. The FKN model is extremely robust; however, it considers many chemical species as dynamic variables of the system, which makes it difficult to handle from an analytical point of view [15] [16]. To solve this problem, in 1974, Field and Noyes proposed a much simpler mechanism to describe reaction dynamics based on the FKN model, also known as *Oregonator* because it was created at the University of Oregon. However, said mathematical model was constructed from a reaction mechanism that consists of only five irreversible chemical reactions and that describes the BZ reaction qualitatively [18]. Years later, in 1990, Györgyi, Turányi, and Field proposed a detailed mechanism of the BZ reaction, that consisted of 80 chemical reactions and 27 chemical species and which is currently the most accepted reaction mechanism [19]. The BZ mechanism has two subsets, the first consists of a set of inorganic chemical reactions while the second one consists of organic chemical reactions. Just one year later, in his seminal

work *A history of chemical oscillations and waves*, Zhabotinsky presented a mathematical model in differential equations based on a subset of chemical reactions from the work of Györgyi et al., and mentioned that it is possible to explain BZ reaction dynamics with said subset [12]. At this point, the following questions arise: How reliable and valid is to reduce an extremely complex reaction mechanism to a subset of reactions? What criteria should be used to make this reduction? Could reducing the mechanism eliminate important system information? Is it possible to use mathematical tools to carry out this reduction properly without losing information? And if this is possible, can these mathematical techniques be used to study biochemical regulatory mechanisms to identify the most critical system variables?

To answer these questions, it is necessary to understand that DST is not the only mathematical tool used to study the complex mechanisms of biochemical regulation [20]. There are tools such as models of agents and cellular automata that have been used to explain at least qualitatively the phenomena that emerge from different physiological processes [20]. However, the tool that has attracted the most attention today is the Complex Networks Theory (CNT) based on graph theory [21] [22]. The CNT has been mainly developed by physicists, with the pioneering works of Barabasi et al., [23]. This theory has allowed us to understand the emergence of extremely complex behaviors in systems that involve a large number of variables because, through statistics, it enables us to study the structural properties of the networks to identify the variables or entities with greater relevance in the system of interest [24]. Multiple measures of centrality can be used to determine the level of importance of a variable in a complex network, such as degree, clustering, hub, *authority*, etc., [25]. In Biomedical research, CNT is widely used to study many phenomena including disease propagation, genetic regulation networks, protein-protein interaction networks, and identification

of possible therapeutic targets in complex biochemical reaction mechanisms [25]. In their excellent work, Costa et al. describe the CNT as a vital tool for Systems Biology [26].

It is well known that the emergence of chronic degenerative diseases (such as insulin resistance, diabetes mellitus II, cancer, cardiovascular diseases among others) result from the mismatch of a wide variety of physiological processes that are in turn regulated by an intricate network of biochemical reactions, which makes it difficult to study the interaction, coupling, and activation that may exist between different physiological processes [27] [28] [29] 30]. The CNT could facilitate the study of different biochemical processes related to each other taking these processes as complex networks where the participating chemical species can be considered nodes or vertices, and all possible physicochemical interactions between them as links or edges. It is possible to identify the chemical species with greater relevance by determining network centrality properties, and applying the standard chemical kinetics techniques (CK) [31] gives rise to mathematical models in differential equations that facilitate the study of coupling and synchronization phenomena between different physiological processes and their possible relationship with various pathologies (see Figure 1).

It is necessary to identify when a group of physiological processes is coupled or synchronized, which requires tools that permit identifying such phenomena. Thanks to the development of machine learning, it is possible to create models capable of learning to recognize or identify a series of patterns with high precision through the acquisition of experience (data) [32]. Machine learning can be divided in supervised and unsupervised learning and can be further classified into different combinations of these [32] [33]. In general, supervised learning consists of providing a series of input data with their respective label to a model so

that the model is generalized and subsequently allows the label to be predicted knowing only the input data or characteristics [32] [33]. On the other hand, unsupervised learning consists of providing only the input data or characteristics to the model without its label, to identify specific patterns of information [32] [33]. Each supervised learning model must go through a training, validation, and testing process to ensure generalization [32] [33]. Supervised learning models can be divided into two large groups: classification models and regression models; in the former, the variable is to be predicted as a qualitative or categorical variable. In regression models, the variable to be predicted is a quantitative variable, either continuous or discrete [32] [33]. Hereby, the following question arises: Is it possible to use a supervised learning model to predict the coupling or synchronization of biochemical processes represented by systems of differential equations?

To answer this question, it is necessary to ask another one: How can data be provided to the supervised learning model to achieve the prediction of a coupling state? The answer is that by modeling biochemical processes as systems of differential equations and coupling them unidirectionally or bidirectionally, it is possible to numerically solve these models, from which the change in the concentration of chemical species with respect to time is obtained. Therefore, it is possible to determine if, for any value of the coupling parameter, these chemical species are synchronized. The degree of synchronization can be determined by evaluating some nonlinear metric, such as the fractal dimension of the time series, evaluation of Lyapunov exponents, entropy, the study of the synchronization variety in the phase space of the variables under study, etc. [35] [36] [37] [38], then it is possible to use some of these metrics to determine, through a supervised learning model, whether these systems are coupled or not. It is also possible to obtain images from the numerical solution of the systems of differential equations; such is the case of *recurrence diagrams* [39]
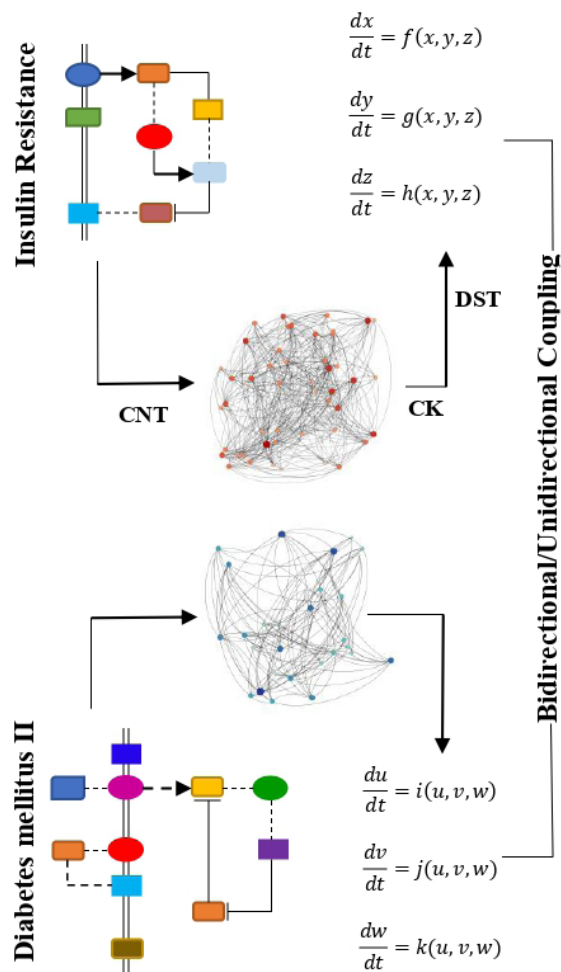

$$\frac{dx}{dt} = f(x, y, z)$$

$$\frac{dy}{dt} = g(x, y, z)$$

$$\frac{dz}{dt} = h(x, y, z)$$

$$\frac{du}{dt} = i(u, v, w)$$

$$\frac{dv}{dt} = j(u, v, w)$$

$$\frac{dw}{dt} = k(u, v, w)$$

FIGURE 1. **General diagram for obtaining a system of differential equations from a biochemical system. It is possible to build complex networks from biochemical reaction mechanisms and, through the network's structural properties, identify the most relevant variables to use CK standard tools to build a dynamic system that is representative of this process. Finally, coupling the systems of equations obtained employing a bidirectional or unidirectional coupling to study possible synchronization effects and their relationship with various pathologies. The block diagrams of the biochemical processes are only schematic representations of Insulin Resistance and Diabetes Mellitus II. The colors of the blocks and lines are also a graphic representation of the chemical species and the possible interactions present in these processes. However, they do not describe the actual biochemical process. For the construction of the network diagrams, the Gephi software was used (https://gephi.org/) [34].**

[40] or *Gramian angular field images* (*gaf*) [40], which can be used with a supervised learning model to determine if a group of variables, in this case, chemical species related to important biochemical processes, are coupled or not, and thereafter, to associate a possible state of synchronization with various pathologies. Currently, it is possible to monitor blood concentration of a variety of chemical species related to critical metabolic processes, which are associated with different pathologies [41] [42] [43] [44], making it possible to construct time series and use them to build a database of *gaf* images to train, validate and test a supervised learning model to identify possible synchronization states and their relationship with pathologies; however, clinical trials can be expensive and, as a result, conducting this type of study is not practical [41] [42] [43] [44]. On the other hand, using mathematical models that represent the dynamics of these variables and using their numerical solution to obtain their respective *gaf* images to train, validate and test the supervised learning model, offers a solution to this problem. The ability of each mathematical model to represent a biochemical process depends on its degree of complexity, i.e., whether the model involves not only metabolic biochemical processes but also epigenetic processes [45].

To analyze the possibility of using CNT in biochemical regulatory mechanisms for identifying the most critical variables of the system and building a system of differential equations that model it, the BZ reaction was used as a study model since it is currently the most complex oscillating reaction discovered so far and it has been used to study the biochemical feedback mechanisms present in the Krebs cycle [12] [13] [14] [15] [16] [17] [18] [19]. On the other hand, to emulate the effects of coupling and synchronization, the obtained model was bidirectionally coupled with other models associated with oscillatory biochemical processes, and its numerical solution was used to obtain the *gaf* images. Finally, a supervised learning tool was used to identify the type of models coupled using the *gaf* images.

MATERIALS AND METHODS

For this study, the general mechanism of the BZ reaction proposed by Györgyi et al. [19] was used. A network was built considering two types of nodes or vertices, the "chemical species" type nodes, and the "chemical reaction" type nodes. The link or edge between species and chemical reactions is given by the reaction rate constant of each of the reactions, which gives place to a "bipartite" network [20] [21] [22] [23] [24] [25]. Once the network was built, the structural property *authority* was evaluated using the algorithm proposed by Kleinberg, which was initially proposed to determine the level of importance and information flow of websites, to reveal the sites with the highest traffic in a virtual hyperlink environment [46]. In this work, the structural property *authority* was used to determine the importance of each of the chemical species involved in the BZ reaction mechanism. To build a system of nonlinear differential equations capable of describing the BZ reaction mechanism, it was assumed that the most relevant variables have the most significant flow of "chemical" information. Once the most critical chemical species in the reaction mechanism were determined under the *authority* criteria, the mathematical model was constructed using standard techniques of CK [31]. When the model was obtained, the effects of synchronization that could emerge due to its coupling with different chemical oscillators were studied to emulate the synchronization processes present in different biochemical systems in mammals [47]. To this end, a bidirectional coupling was carried out with three different models [48]. The first one was an identical model, the second one was the model proposed by Levefer also called Brusselator [49] [50], which shows interesting autocatalytic processes and, lastly, the model proposed by Selkov that describes the oscillatory behaviors present in glycolysis [51] [52]. For each case, the coupling parameter was varied and the numerical solution of the systems of coupled differential equations was obtained using the standard fourth-order Runge Kutta technique [53]. Once the time series of the numerical solu-

tion of the differential equations were obtained, the image of the time series of the bidirectionally coupled variables was constructed using the *gaf* technique [54], which was described extensively by Wang et al [55]. In general, this technique consists in representing a time series in a polar coordinate system, each X_i value of the time series ($\tilde{X}$) is scaled so that all values are in the interval [-1,1] or [0,1], subsequently, with the rescaled values the angular cosine is obtained (φ_i= arccos (x_i), -1 ≤ x_i ≤ 1, $x_i \in \tilde{X}$) and the coordinate r ($r= \frac{t_i}{N}$, $t_i \in \mathbb{N}$), where t_i corresponds to the time interval between each value of the time series and N is a constant factor [55]. Lastly, the addition/subtraction between each point is determined to identify possible temporal correlations within different time intervals [55]. It is possible to obtain two types of *gaf*, the sum (*gasf*= cos ($\varphi_i + \varphi_j$)) and the subtraction (gadf= sin ($\varphi_i + \varphi_j$)). These values are used to construct the *Grammar matrix* that is used to obtain the image (see Figure 2) [55]. This technique in combination with supervised learning tools has been used to study time series of electroencephalograms [56], electrocardiograms [57], signals obtained from biosensors [58], etc.

FIGURE 2. General diagram for obtaining the *gaf* images. The time series are rescaled, later these are represented in a polar coordinate system and finally, the *gaf* images are obtained.

After the *gasf* and *gadf* images were obtained, the supervised learning model was generated using the Orange Data Mining software (*https://orange.biolab.si/*) [58] [59], which is a visual toolbox where it is possible to build workflows that allow the use of widgets for analysis and data processing, supervised and unsupervised learning tools, data visualization and model evaluation. It also has extensions for text mining, spectroscopy, complex network analysis, time series, bioinformatics, and image analysis [59] [60] [61]. The transfer learning technique, which is an artificial intelligence technique that consists of pre-training a model with an extensive database and the experience gained from said training to apply it to another problem that may be completely different, was used to process the *gaf* images [61] [62] [63]. This technique is used in image processing as follows: deep convolutional neural networks (CNN) are used, which are pre-trained with a large number of images of all kinds, later, activations of the penultimate layer of the model (CNN codes) are used to represent the images with vectors (embedded), i.e., CNN is used as a feature extractor or descriptor, allowing supervised learning models to be used and thus obtaining high precision values in image classification [61] [62] [63]. In Orange, it is possible to embed images using different CNNs, including Google's Inception V3 neural network that has been pre-trained with the ImageNet database consisting of 1.2 million images. The neural network has 2048 nodes in its penultimate layer, so each image represents it with a vector of dimension 2048 [33] [64] (CNN by default in Orange). On the other hand, embedding images can also be done with CNN SqueezeNet, which is a much simpler network than Inception V3; nevertheless, it achieves a precision close to that of CNN AlexNet on the ImageNet database, this CNN represents each image as a vector of dimension 1000 [65]. It is also possible to embed with CNN's VGG16 and VGG19 proposed by the Visual Geometry Laboratory of the University of Oxford [66], in the same way, pre-trained with the ImageNet database, the CNN DeepLoc
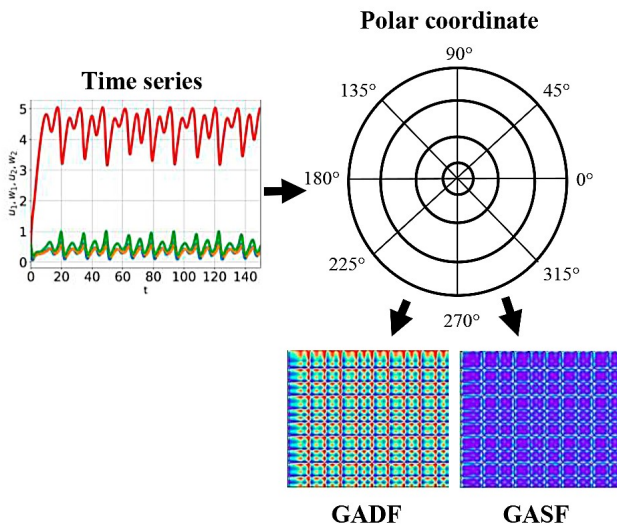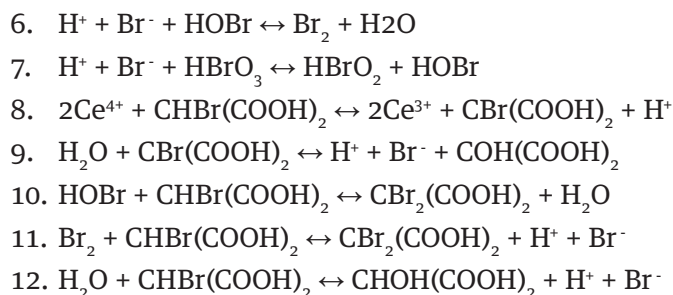
pre-trained with 21882 images [67] and CNN Painters, a pre-trained network with 79433 images [59] [60] [61]. For embedding the images, Orange sends them to an external server, except for CNN SquezeeNet, which is done locally, making embedding faster. In this work, the *gaf* images were embedded using CNN Google Inception V3 and SquezeeNet. Subsequently, a supervised learning model was trained as a classifier of the type of coupled oscillators using the *gaf* images. A logistic regression with Ridge penalty or regularization (L2 = 1) was used as a classification method. Six different evaluation techniques were used to train, validate and test the classifier: 1) the standard stratified holdout technique (70% training set / 30% test set), the training and testing subsets are repeated 10 times (A), 2) a 3-fold stratified cross-validation (B), 3) a 5-fold stratified cross-validation (C), 4) a 10-fold stratified cross-validation (D), 5) a 20-fold stratified cross-validation (E) and 6) the leave-one-out cross-validation (F), to identify the best evaluation technique and avoid overfitting [33] [68] [69]. The previous procedure was repeated implementing a principal component analysis (PCA) using a number of components such that they explain 95% of the total variance of the images (21 principal components for Google Inception V3 and nine principal components for SquezeeNet). After embedding the images with the CNN, the same six evaluation techniques were used to study the effect of reducing dimensions and to avoid overfitting [33] [68] [69]. For both procedures the classifier confusion matrix was obtained to determine the classification metrics: Classification Accuracy (CA), Precision (P), Recall (R), and F1-score (F1) [33] [70].
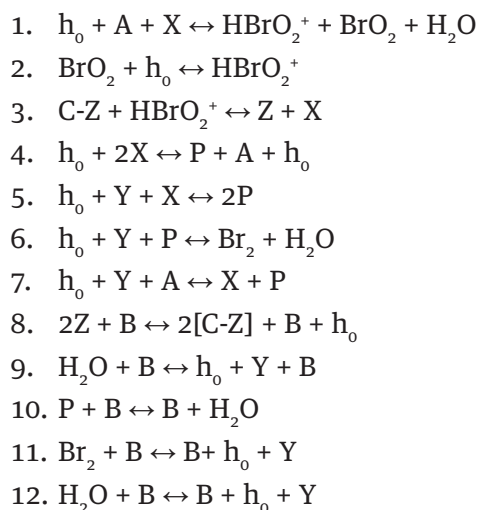
RESULTS AND DISCUSSION

Figure 3a is a diagram that represents the complex network for the general mechanism of the BZ reaction proposed by Györgi et al. The nodes represent each of the 27 chemical species and 80 chemical reactions participating in the reaction mechanism, and the links represent the reaction rate constant magnitude. The largest nodes, which are shown in Table 1, represent the nodes with the highest numerical value of *authority*. Figure 3b shows the network *authority* property histogram. Of the 107 nodes, 49 (i.e. 45.79%) have an *authority* value between 0 and 0.1 (none being actually zero), 40 have an *authority* value between 0.1 and 0.2 (37.38%), 9 nodes between 0.2 and 0.3 (8.41%), 3 nodes between 0.3-0.4 (2.8%), 3 nodes between 0.4 and 0.5 (2.8%), 2 nodes between 0.5 and 0.6 (1.87%), and one between 0.9 and 1 (0.93%). This last one is the most important node, considering the *authority* centrality criterion. In the histogram, it is easy to distinguish that there are few nodes with a high *authority* value, making it possible to identify the chemical species with greater importance, possibly due to power-law-like behavior [71]. Power laws are present in a wide variety of biological phenomena and are related to processes of a universal nature [72].

As expected, the chemical node or species with higher *authority* is the H^+ cation because the chemical reaction at the experimental level is carried out in the presence of sulfuric acid (H_2SO_4), which is a source of H^+ and allows the formation of bromous acid ($HBrO_2$), an essential variable for the chemical feedback mechanism that enables the existence of periodic behaviors [12] [13] [14]. Furthermore, we can note that H_2O is the second chemical species with higher *authority* since the precursor solutions for the chemical reaction use deionized water as a solvent [12]. The third most crucial chemical species is the carboxy radical (*COOH) inasmuch as the breakdown of malonic acid as a precursor agent results in short-lived chemical species [12] [19]. The presence of Br* radicals is mainly because during the progress of the reaction molecular bromine (Br_2) occurs due to the bromate anion precursor agent (BrO_3), said Br_2 results in the formation of bromide ions (Br) and subsequently to Br* radicals, which are fundamental chemical species for the feedback mechanism [12] [19]. Because the chemical reaction includes organic components, specifically malonic acid and its
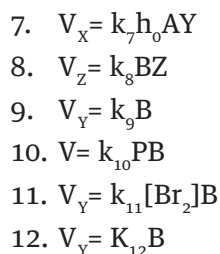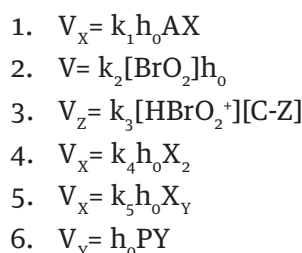
derivatives, it has carboxyl or carbonic acid groups that can be decomposed into carbon dioxide (CO_2), which is seen in the form of gas bubbles during the reaction [12] [13] [14].

a)

b)

FIGURE 3. **a) The complex network of the chemical mechanism of Györgyi et al [19]. The larger nodes correspond to those that have greater authority value; b) Histogram of structural property *authority*. It can be observed that most of the nodes have an *authority* value below 0.3 (91.58%), however, there are few nodes with *authority* values above 0.3 (8.42%), which can be considered as the most important nodes under the criterion of centrality *authority*. Gephi software was used for the construction of the network diagram (https://gephi.org/) [34].**

Lastly, we can see that the oxidized form of the Cerium metal (Ce^{4+}) is also among the ten nodes with the highest *authority* value, this results from said chemical species acting as a catalytic agent in the chemical reaction and that alongside its reduced form (Ce^{3+}), it gives rise to the typical color changes of the BZ reaction [12] [18]. With these results, it is possible to observe that applying CNT to the BZ reaction mechanism proposed by Györgyi et al. [19] permits identifying the chemical species with greater relevance under the criterion of *authority*. Such chemical species appear naturally in the mechanism described by Zhabotinsky, either in its ionic form or in the form of a precursor compound [12], so it is possible to approximate the BZ reaction to said subset of reactions. Still, it is essential to mention that it is possible to use other centrality criteria to identify the nodes with greater relevance [25].

TABLE 1. *Authority* values.

Node	Authority
H^+	1
H_2O	0.5435
*COOH	0.5086
Br^-	0.4598
Br*	0.4133
CO_2	0.4062
HOBr	0.3870
$HBrO_2$	0.3289
BrO_3^-	0.3030
Ce^{4+}	0.2939

The mechanism proposed by Zhabotisnky uses Iron (Fe) as a catalyst instead of Cerium (Ce) [12]. Below is the chemical mechanism studied by Zhabotinsky replacing Fe with Ce.

BZ Chemical model [12]

1. $H^+ + HBrO_3 + HBrO_2 \leftrightarrow HBrO_2^+ + BrO_2 + H_2O$
2. $BrO_2 + H^+ \leftrightarrow HBrO_2^+$
3. $Ce^{3+} + HBrO^{2+} \leftrightarrow Ce^{4+} + HBrO^2$
4. $H^+ + 2HBrO_2 \leftrightarrow HOBr + HBrO_3 + H^+$
5. $H^+ + Br + HBrO_2 \leftrightarrow 2HOBr$

6. $H^+ + Br^- + HOBr \leftrightarrow Br_2 + H2O$
7. $H^+ + Br^- + HBrO_3 \leftrightarrow HBrO_2 + HOBr$
8. $2Ce^{4+} + CHBr(COOH)_2 \leftrightarrow 2Ce^{3+} + CBr(COOH)_2 + H^+$
9. $H_2O + CBr(COOH)_2 \leftrightarrow H^+ + Br^- + COH(COOH)_2$
10. $HOBr + CHBr(COOH)_2 \leftrightarrow CBr_2(COOH)_2 + H_2O$
11. $Br_2 + CHBr(COOH)_2 \leftrightarrow CBr_2(COOH)_2 + H^+ + Br^-$
12. $H_2O + CHBr(COOH)_2 \leftrightarrow CHOH(COOH)_2 + H^+ + Br^-$

To make studying the reaction mechanism easier, it is possible to make the following variable change: A= [HBrO_3]; B= [Any organic species derived from malonic acid]; P= [HOBr]; X= [HBrO_2]; Y=[Br^-]; Z= [Ce^{4+}]; C= [Ce^{3+}]+[Ce^{4+}]; $h_0 \approx H^+$ [12]; therefore, the chemical mechanism is transformed into:

1. $h_0 + A + X \leftrightarrow HBrO_2^+ + BrO_2 + H_2O$
2. $BrO_2 + h_0 \leftrightarrow HBrO_2^+$
3. $C-Z + HBrO_2^+ \leftrightarrow Z + X$
4. $h_0 + 2X \leftrightarrow P + A + h_0$
5. $h_0 + Y + X \leftrightarrow 2P$
6. $h_0 + Y + P \leftrightarrow Br_2 + H_2O$
7. $h_0 + Y + A \leftrightarrow X + P$
8. $2Z + B \leftrightarrow 2[C-Z] + B + h_0$
9. $H_2O + B \leftrightarrow h_0 + Y + B$
10. $P + B \leftrightarrow B + H_2O$
11. $Br_2 + B \leftrightarrow B + h_0 + Y$
12. $H_2O + B \leftrightarrow B + h_0 + Y$

Considering species A, B and P as constants and chemical reactions as elementary (the power of variables X, Y, and Z is directly related to their stoichiometric coefficients), it is possible to use the standard techniques of CK to construct reaction speed laws [12] [31]:

Equations 1

1. $V_X = k_1 h_0 AX$
2. $V = k_2[BrO_2]h_0$
3. $V_Z = k_3[HBrO_2^+][C-Z]$
4. $V_X = k_4 h_0 X_2$
5. $V_X = k_5 h_0 X_Y$
6. $V_Y = h_0 PY$

7. $V_X = k_7 h_0 AY$
8. $V_Z = k_8 BZ$
9. $V_Y = k_9 B$
10. $V = k_{10} PB$
11. $V_Y = k_{11}[Br_2]B$
12. $V_Y = K_{12} B$

Where the variables X, Y, and Z are the variables that generally describe the chemical mechanism of the BZ reaction [12]. Using the law of mass action, the following system of nonlinear differential equations is obtained. It models the change of X, Y, and Z with respect to time and as a function of the chemical concentrations of the precursors [12]:

Equations 2

$$\frac{dX}{dt} = k_1 h_0 AX - k_5 h_0 XY + k_7 h_0 AY - 2k_4 h_0 X^2,$$

$$\frac{dY}{dt} = \frac{F k_8 k_9 BZ}{k_{-8} h_0 (C-Z) + k_9} - k_5 h_0 XY - k_7 h_0 AY + k_{12} B,$$

$$\frac{dZ}{dt} = 2k_1 h_0 AX - \frac{k_8 k_9 BZ}{k_{-8} h_0 (C - Z) + k_9}.$$

Where k_i are the reaction rate constants and F is a stoichiometric factor used as an adjustment parameter [12] [13]. It is possible to simplify the model by dividing the numerator and denominator of the term $\frac{k_8 k_9 BZ}{k_{-8} h_0 (C-Z) + k_9}$ by k_9, then doing a geometric series development and neglecting terms of greater order $k_8 BZ \frac{1}{\frac{k_8 h_0 (C-Z)+1}{k_9}} \approx k_8 BZ$ is obtained. Therefore, assuming that the concentration of H^+ remains constant and that $F = \frac{1}{2} f$, then the model proposed by Zhabotinsky can be reduced to [14]:

Equations 3

$$\frac{dX}{dt} = k_1 AX - k_5 XY + k_7 AY - 2k_4 X^2,$$

$$\frac{dY}{dt} = \frac{1}{2} f k_8 BZ - k_5 XY - k_7 AY + k_{12} B,$$

$$\frac{dZ}{dt} = 2k_1 AX - k_8 BZ.$$

The model described above is very similar to that proposed by Field and Noyes [13] [18], except for the term $k_{12}B$, which is related to the rate at which chemical species derived from malonic acid, specifically $CHBr(COOH)_2$ by its decomposition, it leads to the formation of new Br^- ions, which cannot be neglected from the mathematical model because they have a strong implication in the feedback process of the BZ reaction [12] [19] since the excessive production of these can lead to complete inhibition of oscillations. Also, these ions actively participate in the production of $HBrO_2$, which in turn facilitates the process of changing the oxidation state of the Ce catalyst.

Making the change of variable [13]:

$$u = \left[\frac{2k_4 X}{k_1 A}\right]; v = \left[\frac{k_5 Y}{k_1 A}\right]; w = \left[\frac{k_8 k_4 BZ}{(k_1 A)^2}\right]; t = k_8 B\tau.$$

So, Equations 3 in their dimensionless form are:

Equations 4

$$\frac{du}{d\tau} = \frac{qv - uv + u - u^2}{\varepsilon},$$

$$\frac{dv}{d\tau} = \frac{-qv - uv + fw^2 + \alpha}{\varepsilon'},$$

$$\frac{dw}{d\tau} = (u - w).$$

Where $\varepsilon = \frac{k_8 B}{k_1 A}$; $\varepsilon' = \frac{2k_4 k_8 B}{k_5 k_1 A}$; $q = \frac{2k_7 k_4}{k_5 k_1}$; $\alpha = \frac{k_{12}B}{k_1 A}$.

However, as in the model proposed by Field and Noyes, the parameter ε', is very small compared to ε [13] [18]; therefore, it is possible to consider that the variable v, remains in a stationary state, so the system of Equations 4 can be reduced to:

Equations 5

$$\varepsilon\frac{du}{dt} = u(1 - u) + (fw + \alpha)\frac{q - u}{q + u},$$

$$\frac{dw}{dt} = u - w.$$

We named this model *Modified Oregonator* (NOM). Amemiya et al. [73], and Krug et al. [74], associated the term to the sensitivity of the BZ reaction to the presence of oxygen and its photosensitivity. In contrast, we associate this term with the decomposition of the derivatives of the malonic acid that give rise to Br-reduction. The chemical mechanism proposed by Györgyi et al. was used as a starting point [19] identifying the most relevant chemical species through the CNT under the *authority* criterion, which approximates the reaction mechanism BZ, to the subset of reactions proposed by Zhabotisky [12].

To emulate the synchronization processes present in different physiological systems, the mathematical model obtained in this work was coupled with different oscillators, all of them dimensionless. The parameters of the models were selected according to the stability criteria to ensure the presence of oscillations [13] [18]. The coupling was linear bidirectional in all cases, and the coupling force is modulated by parameter k. The system of Equations 6 shows two coupled NOM oscillators:

Equations 6

$$\varepsilon_1\frac{du_1}{dt} = u_1(1 - u_1) + (f_1 w_1 + \alpha_1)\frac{q_1 - u_1}{q_1 + u_1} + k(u_2 - u_1),$$

$$\frac{dw_1}{dt} = u_1 - w_1,$$

$$\varepsilon_2\frac{du_2}{dt} = u_2(1 - u_2) + (f_2 w_2 + \alpha_2)\frac{q_2 - u_2}{q_2 + u_2} + k(u_1 - u_2),$$

$$\frac{dw_2}{dt} = u_2 - w_2.$$

The system of Equations 7, shows the NOM model coupled with the Brusselator model:

Equations 7

$$\varepsilon\frac{du_1}{dt} = u_1(1 - u_1) + (fw_1 + \alpha)\frac{q - u_1}{q + u_1} + k(u_2 - u_1),$$

$$\frac{dw_1}{dt} = u_1 - w_1.$$

$$\frac{du_2}{dt} = a - bu_2 + u_2^2 w_2 - u_2 + k(u_1 - u_2),$$

$$\frac{dw_2}{dt} = bu_2 - u_2^2 w_2.$$

And lastly, the NOM model was coupled with the Selkov glycolysis mathematical model (Equations 8),

Equations 8

$$\varepsilon \frac{du_1}{dt} = u_1(1 - u_1) + (fw_1 + \alpha)\frac{q - u_1}{q + u_1} + k(u_2 - u_1),$$

$$\frac{dw_1}{dt} = u_1 - w_1.$$

$$\frac{du_2}{dt} = v - \frac{u_2 w_2^\gamma}{w_2^\gamma u_2 + w_2^\gamma + 1} + k(u_1 - u_2),$$

$$\frac{dw_2}{dt} = \beta \left(\frac{u_2 w_2^\gamma}{w_2^\gamma u_2 + w_2^\gamma + 1} \right) - \eta w_2.$$

In Figure 4a, the numerical solution of the system of Equations 6 is shown. A full synchronization can be seen between variables u_1 and u_2, i.e., the oscillation rhythms are completely coupled [48] [75]. This type of behavior is common in biochemical processes carried out by cells that share the same microenvironment and can be stimulated either by a chemical or physical mechanism [76] [77]. On the other hand, in Figure 4b the gasf image of the variable u_1 can be seen, while in Figure 4c, the *gadf* image of the variable u_2 is shown. A characteristic pattern of bidirectional coupling between identical oscillators with periodic behavior can be seen in both figures.

In Figure 5a, the numerical solution for the system of Equations 7 is shown. Almost complete synchronization can be seen between the variables u_1, w_1 and u_2 [48] [75]. Anyway, this system of coupled differential equations is very sensitive to the value of the coupling parameter. Also, Figure 5b shows the *gasf* image of the variable u_1, while Figure 5c shows the gadf image of the variable u_2. In both figures, a characteristic pattern for a quasi-periodic system can be seen [78].
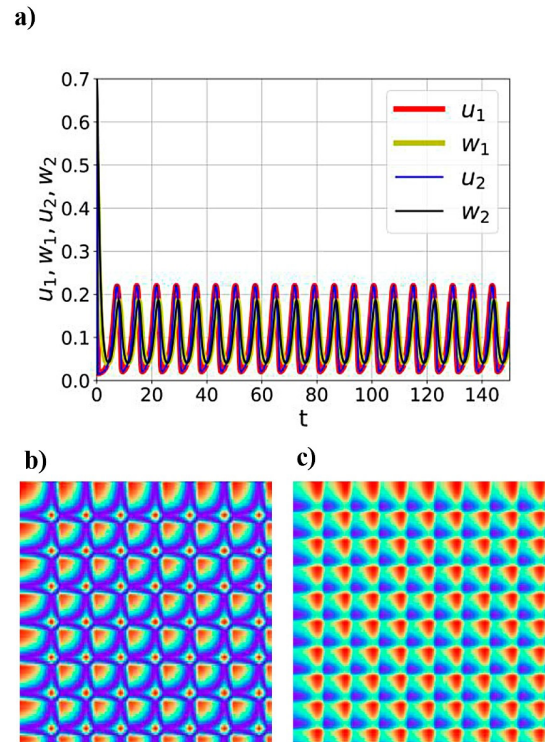

a)

b)   c)

FIGURE 4. a) Numerical solution of the system of Equations 6 with $\varepsilon_1 = \varepsilon_2 = 0.3$, $\alpha_1 = \alpha_2 = 0.03$, $q_1 = q_2 = 0.015$, $f_1 = f_2 = 1.0$ and $k = 1.0$. With $u_1(0) = 0.3$, $w_1(0) = 0.5$, $u_2(0) = 0.8$ and $w_2(0) = 0.7$, $tf = 150$ units of dimensionless time and $dt = 0.001$. An oscillatory behavior is shown for the 4 dynamic variables of the system and full synchronization is observed; b) *gasf* image of the variable u_1; c) *gadf* image of the variable u_2. For both images, a characteristic pattern of periodic systems can be seen.

In Figure 6a, the numerical solution of the system of Equations 8 is shown. The phase synchronization between the variables u_1, u_2 and w_1 [48] [75] can be seen. The NOM model oscillation rhythm causes variable u_2 to enter in-phase synchronization; however, the variable w_2 decays completely to zero. In the Selkov model, the variable u_2 is related to the concentration of ATP (adenosine triphosphate), while the variable w_2 is related to the concentration of ADP (adenosine diphosphate) [51] [52], so when said system is coupled with the NOM model, the synchronization process that could exist between the biochemical mechanism of glycolysis and the Krebs cycle is emulated.

Moreover, in Figure 6b the gasf image of the variable u_1 can be seen, while in Figure 6c the *gadf* image of the variable u_2 is shown. It can be distinguished that the *gasf* field shows the typical oscillatory behavior, while the image of the gadf field does not clearly show a distinct pattern.
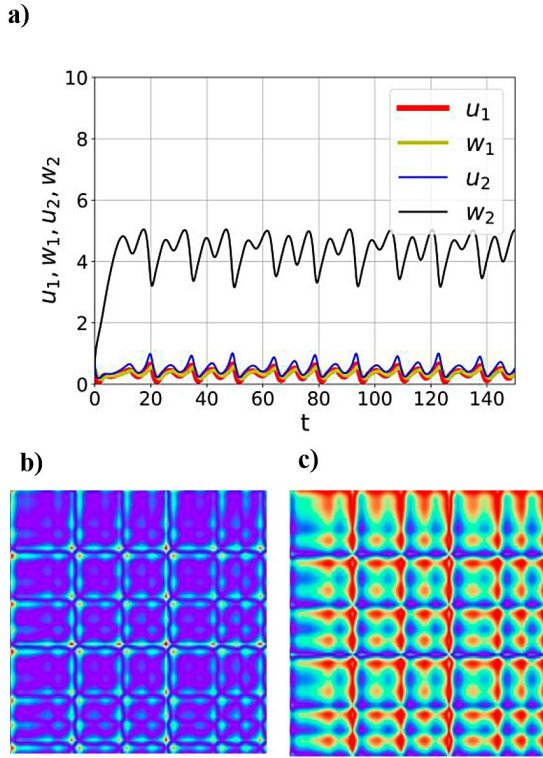
a)

b)   c)

FIGURE 5. **a) Numerical solution of the system of Equations 7 with $\varepsilon = 0.3$, $\alpha = 0.03$, $q = 0.015$, $f = 1.0$, $a = 1$, $b = 2.5$ and $k = 2.5$. With $u_1(0) = 0.3$, $w_1(0) = 0.5$, $u_2(0) = 0.8$ and $w_2(0) = 0.7$, $tf = 150$ units of dimensionless time and $df = 0.001$. A quasi-periodic behavior is shown for the 4 dynamic variables of the system and almost complete synchronization is observed between the variables u_1, u_2 and w_1; b) *gasf* image of the variable u_1; c) *gadf* image of the variable u_2. The images reflect the quasi-periodic behavior of the variables u_1 and u_2.**

The numerical solution of the equations without coupling can be found in the supplementary material, as well as their respective *gasf* and *gadf* images (see Figures S1, S2 and S3). The coupling of chemical oscillators to emulate the synchronization in biochemical processes, is of vital importance since it allows us to

understand the complex dynamics that underlie the feedback processes present in biological systems. Mismatches in the oscillation rhythms of physiological processes can lead to a wide variety of metabolic problems [7], so it is essential to know the type of coupling that can exist between oscillators (unidirectional, bidirectional, linear, nonlinear, etc., [48].) and whether these are coupled or not at any given time.
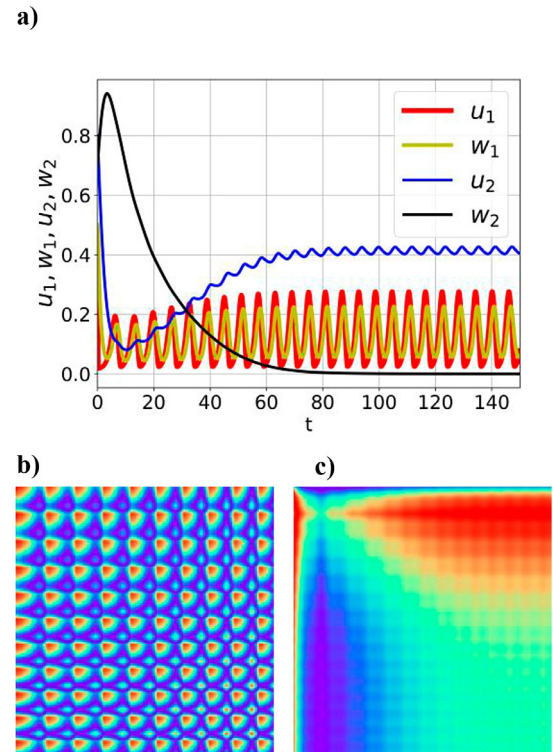
a)

b)   c)

FIGURE 6. **Figure 6. a) Numerical solution of the system of Equations 8 with $\varepsilon = 0.3$, $\alpha = 0.03$, $q = 0.015$, $f = 1.0$, $v = 0.0285$, $\eta = 0.1$, $\beta = 1.0$, $\gamma = 2$ and $k = 0.1$. With $u_1(0) = 0.3$, $w_1(0) = 0.5$, $u_2(0) = 0.8$ and $w_2(0) = 0.7$, $tf = 150$ units of dimensionless time and $df = 0.001$. Oscillatory behavior is distinguished for the variables u_1, u_2, and w_1, as well as phase synchronization between them. It is recognized that the variable w_2, for this case of the value of the parameters, tends to zero; b) *gasf* image of the variable u_1; c) *gadf* image of the variable u_2. The *gasf* image shows a typical pattern of a periodic oscillatory system. On the other hand, although the variable u_2 shows an oscillatory behavior, it is also coupled with the variable w_2, which is why a well-defined pattern in the *gadf* image is not clearly distinguished.**

To determine coupled oscillators type, the coupling parameter was varied between one and ten for the system of Equations 6, between one and ten for the system of Equations 7 and lastly, between 0.1 and 1 for the system of Equations 8 (because this system is highly sensitive to bidirectional coupling), to generate the images of the time series using the *gaf* technique. Once obtained, the images were embedded to subsequently train the supervised learning model using the procedure described in the materials and methods section (see Figure 7). Figure 8 shows the general Orange workflow for this procedure. In the supplementary material (Tables TS-1 and TS-2), the results of the evaluation of different models of supervised learning are shown (Multilayer Perceptron Neural Network (MLP) (Hyperparameters: Optimizer: L-BFGS-B, activation function: Logistics, regularization L2= 1, neurons in one hidden layer: 10, tol= 1E-4, max_iter= 200), K-Nearest-Neighbors (Hyperparameters: Number of nearby neighbors: 5, metric: Euclidean, wight= uniform) and Support Vector Machine (SVM) (Hyperparameters: Cost= 1.0, kernel: RBF, tol= 1E-3, iteration limit= 100)) [33] [68], which were used to classify the type of coupled oscillators using the gaf images, to compare them with those obtained with the logistic regression model. When training the different supervised learning models using the vectors provided by CNN SquezeeNet as image descriptors, the logistic regression model presented the highest CA and F1 values for all evaluation techniques, obtaining the maximum values in the methods A, D, E and F (CA= 0.983 and F1= 0.983). In contrast, the minimum values were obtained from methods B and C (CA= 0.966 and F1= 0.967) (see supplementary Tables ST-1). When using CNN Inception V3 to embed the images, similar values of CA and F1 were obtained for the logistic regression models, MLP and KNN for method A (CA= 0.961, F1= 0.961), methods B, C, D, E and F (CA= 0.950 and F1= 0.949-0.950). The SVM model presents the highest CA and F1 for methods B and C (CA= 0.966 and F1= 0.966). Notwithstanding, for methods D, E and F, the SVM model presents similar CA and F1 than the rest of the models (CA= 0.950 and F1= 0.949) (see supplementary Tables ST-1). Likewise, it is possible to observe that there are no major differences between the results obtained by CNN's Inception V3 and SquezeeNet. Godec et al., use transfer learning to embed biomedical images and mention that this technique allows them to obtain high precision values in classification models using small databases [61]. Godec et al., also found no major differences in the CA and F1 values of the logistic regression classifier when using either the CNN Inception V3 or the CNN SquezeeNet to embed the images [61].

It is worth mentioning that sometimes when the database used for training supervised learning models has a higher number of characteristics or descriptors compared to the sample size (as is often the case in biomedical databases), or is unbalanced, i.e., there is a greater amount of data from one class than from the rest, it is possible to overfit the model, leading to erroneous results [33] [61] [68] [69]. However, it is possible to prevent overfitting essentially through two procedures, the first is to use more data for training, decrease statistical bias and decrease the number of characteristics or descriptors, the second one is to limit the complexity of the model of supervised learning, employ regularization, either penalty L2 (Ridge), L1 (Lasso) or ElasticNet (L1 and L2 simultaneously) and use assessment techniques such as stratified cross-validation or leave-one-out cross-validation [33] [68] [79] [80] [81] [82] [83]. Therefore, we have also implemented a PCA after embedding the images with CNN's to study the effect of the reduction of dimensions in the classification of the type of coupled oscillators, using the same models of supervised learning and the same evaluation methods.

When performing the PCA implementation after embedding the images with CNN SquezeeNet using nine main components, which explain 95% of the total variance, the logistic regression model obtains the
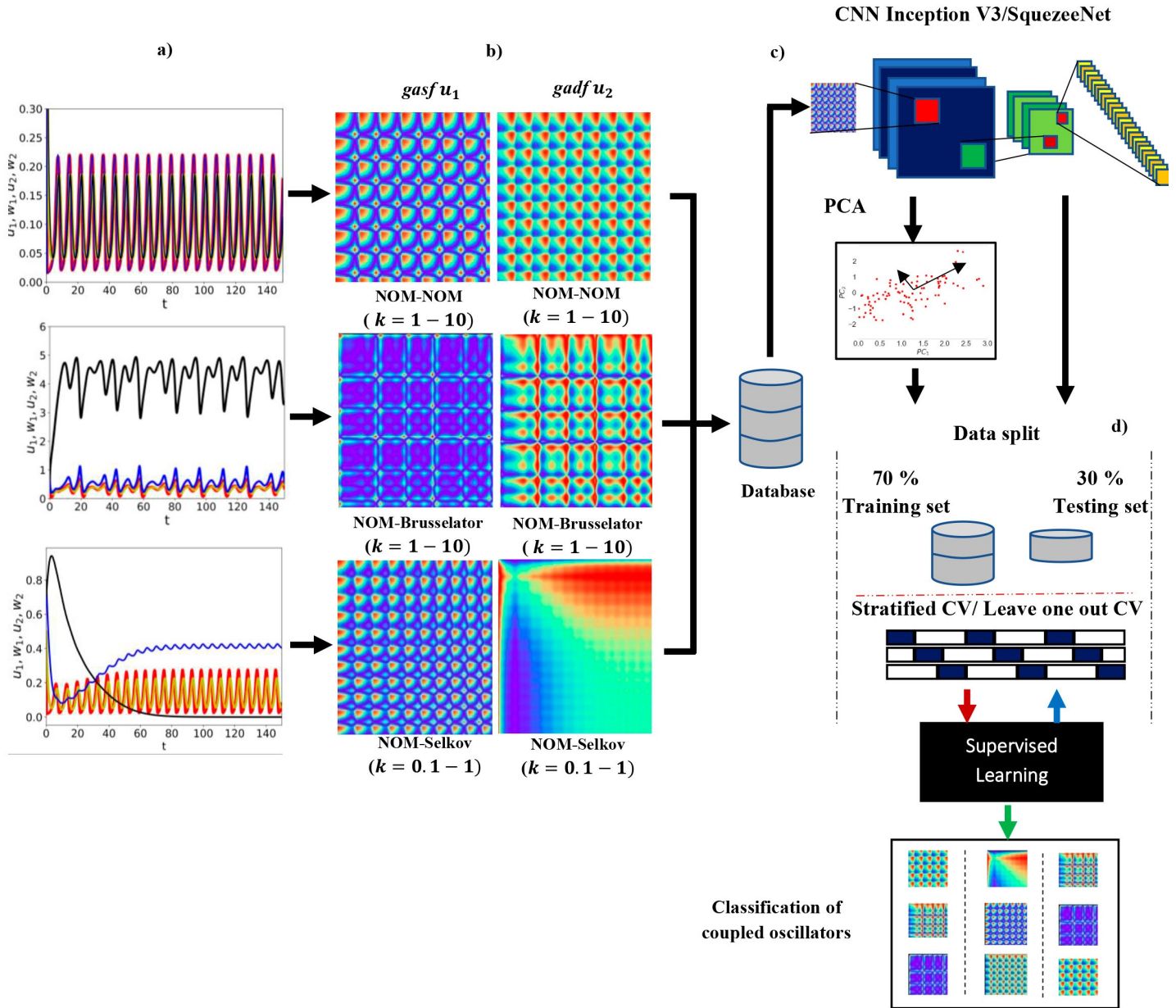
**FIGURE 7. General procedure to obtain the classification model of the type of coupled oscillators using the *gaf* images.
a) The numerical solution of the coupled oscillator systems is obtained by setting the parameters of the models
and varying the coupling parameter, b) subsequently *gaf* images are obtained from the time series of the numerical
solution of the coupled oscillators, c) after building the database, then embedding the images using Google's CNN Inception
V3 or SquezeeNet and using PCA for dimension reduction (optional). Lastly, d) the classification model is evaluated using
the techniques: stratified holdout (70% training set / 30% testing set), 3, 5, 10, 20-fold stratified cross-validation,
and leave-one-out cross-validation.**

highest CA and F1 value for method A (CA= 0.983 and F1= 0.983) nevertheless, it also shows the best classification metrics for methods C, D, E and F (CA= 0.966

and F1= 0.967), while for method B, the SVM model presents the best classification metrics with a CA= 0.966 and F1= 0.966 (see supplementary material
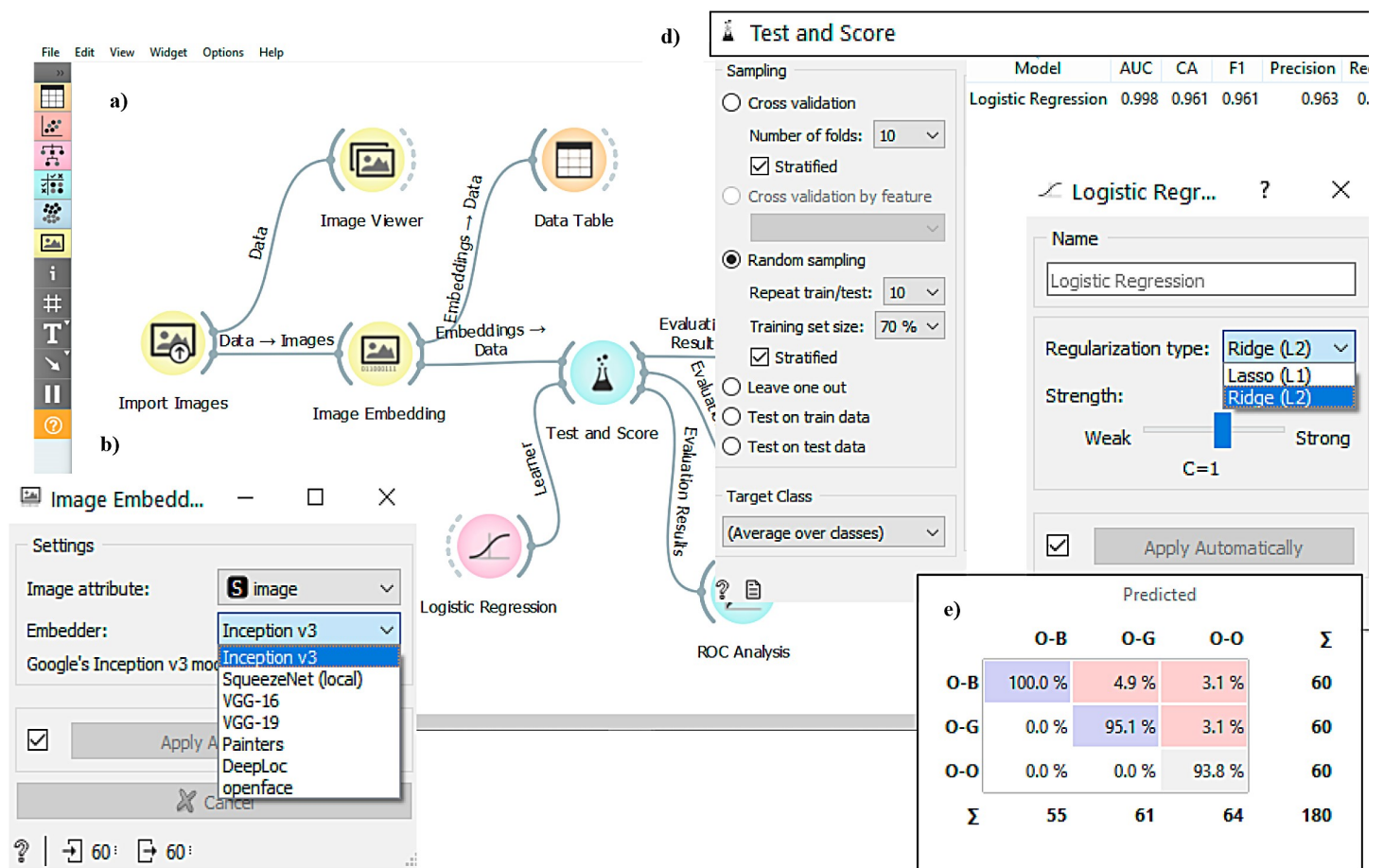
FIGURE 8. Orange workflow example for training the logistic regression model as a *gaf* image classifier. a) In the Orange interface, the images are imported first from a local folder using the "Import images" widget, then the images are embedded using the "Image embedding" widget and finally the training, validation and supervised learning model test. b) Window that shows the "Image Embedding" widget, which shows the CNNs that can be detected to extract the characteristics or descriptors of the images. The CNN of Google Inception V3 is shown by default. c) Window that shows the "Logistic Regression" widget. It is appreciated that it is possible to use the regularization of the sea Lasso (L1) or Ridge (L2). d) Window that shows the widget "Test and score." This widget shows the different evaluation techniques, which are: "k-fold cross-validation", "k-fold stratified cross-validation", "Random sampling", and "Leave-one-out cross-validation". Also, it shows the classification metric by class or average. e) Window that shows the "Confusion matrix" widget, which allows observing the confusion matrix of the supervised learning model. The widget allows displaying the confusion matrix in percentages or the number of images classified correctly and incorrectly.

Tables ST-2). Moreover, when a PCA is implemented after embedding with CNN Inception V3 using 21 main components, which explain 95% of the total variance, the logistic regression model shows the best classification metrics for the evaluation methods A, C, D, E and F, obtaining a value of CA= 0.961 and F1= 0.960 for method A, while for methods C, D, E and F, a CA= 0.950 and F1= 0.949 (see tables of supplementary material ST-2) were obtained. Regardless, the SVM model presents the highest values of CA and F1 for method B (CA= 0.966 and F1= 0.966). Having said that, it is possible to note that there are no major differences in the

classification metrics when using either the CNN SquezeeNet or the CNN Inception V3 to embed the images and, at the same time, there are no major differences in the metrics of classification when using the descriptors connected directly from CNN or when the main components are extracted from them. However, the computational time required for the evaluation of the models when using PCA is shorter. The supervised learning models present similar values in the classification metrics; nevertheless, the logistic regression model has the least complexity because it only uses one hyperparameter, which is used as a regularization or penalization [33] [67] [84] [85]. Therefore, logistic regression can be used as a classification model for the type of coupled oscillators. Table 2 shows the classification metrics for the logistic regression model for each of the evaluation methods, using descriptors obtained directly from CNN's as training data.

TABLE 2. Classification metrics for the logistic regression model using the characteristics obtained directly from the CNN for the evaluation.

Evaluation Technique	CNN							
	Inception V3				SquezeeNet			
A	CA	F1	P	R	CA	F1	P	R
B	0.961	0.961	0.963	0.961	0.983	0.983	0.984	0.983
C	0.950	0.949	0.953	0.950	0.966	0.967	0.969	0.966
D	0.950	0.949	0.953	0.950	0.966	0.967	0.969	0.966
E	0.950	0.949	0.953	0.950	0.983	0.983	0.984	0.983
F	0.950	0.949	0.953	0.950	0.983	0.983	0.984	0.983
A	0.950	0.949	0.953	0.950	0.983	0.983	0.984	0.983

Likewise, Table 3 shows the classification metrics for the same model, using the descriptors obtained from the PCA as training data. At this point, it is natural to ask what evaluation method should be used if all methods have similar ranking metrics. For this work, we chose method D because multiple experiments have been carried out that demonstrate that the best way to obtain high values in the metrics, be it classification or regression, is using stratified 10-fold

cross-validation, even when there is the possibility of computation to increase fold number in the evaluation of supervised learning models [33] [86]. In addition, as can be seen in Tables 2 and 3, for method D the same values of the classification metrics are obtained using the descriptors extracted directly from CNN Inception V3 and those obtained from the implementation of the PCA (CA= 950, F1= 0.949-0.950) for training. When using the descriptors extracted directly from CNN SquezeeNet for training, the classification metrics CA= 0.983, and F1= 983 were obtained, while those obtained due to the implementation of the PCA are CA= 0.967 and F1= 0.967, which means that there is no significant difference. In conclusion, there are no major differences between the use of the descriptors extracted from the implementation of the PCA after embedding the images with one or the other CNN. This shows that the dimensions reduction does not substantially affect the precision of the supervised learning model and, conversely, allows for a better generalization of it [61].

TABLE 3. Classification metrics of the logistic regression model using the characteristics obtained from the application of the PCA.

Evaluation Technique	CNN							
	Inception V3 (21 PC, explained variance: 95%)				SquezeeNet (9 PC, explained variance: 95%)			
A	CA	F1	P	R	CA	F1	P	R
B	0.961	0.961	0.963	0.961	0.983	0.983	0.984	0.983
C	0.950	0.950	0.953	0.950	0.950	0.951	0.957	0.950
D	0.950	0.950	0.953	0.950	0.967	0.967	0.970	0.967
E	0.950	0.950	0.953	0.950	0.967	0.967	0.970	0.967
F	0.950	0.950	0.953	0.950	0.967	0.967	0.970	0.967
A	0.950	0.950	0.953	0.950	0.967	0.967	0.970	0.967

Figure 9 shows the confusion matrix of the logistic regression model trained with the descriptors obtained from applying the PCA after being embedded with CNN Inception V3 and using the evaluation method D. 18 of the 20 images of the coupling of the NOM and Brusselator Oscillators (O-B) have been correctly clas-

sified, while two images have been erroneously classified, one as a coupling between the NOM and Silkov (O-G) oscillators and another as a coupling between the identical NOM (O-O) oscillators. Furthermore, 19 of the O-G coupling images have been correctly classified, while one has been erroneously classified as O-O coupling. All O-O coupling images have been correctly classified.

		Predicted		
		O-B	**O-G**	**O-O**	**Σ**
Actual	**O-B**	18	1	1	20
	O-G	0	19	1	20
	O-O	0	0	20	20
	Σ	18	20	22	60

FIGURE 9. Confusion matrix of the logistic regression method using the CNN Inception V3 as an image descriptor extractor and applying a reduction of dimensions using the PCA technique (21 PC, explained variance: 95%) (evaluation method D). It is observed that 18 of the 20 images of the O-B coupling have been classified correctly, while two have been erroneously classified, one as the O-G coupling and the other as the O-O coupling. On the other hand, the coupling shows that 19 of the 20 images of the O-G coupling have been correctly classified, while one has been classified as an O-O coupling. All images in the O-O coupling have been correctly classified.

Figure 10 shows the confusion matrix of the same model trained with the descriptors obtained from applying the PCA after being embedded with CNN SquezeeNet and using the evaluation method D. In the confusion matrix, it is possible to observe that 19 of the 20 images of the O-B coupling have been correctly classified, while one has been incorrectly classified as O-O coupling. Similarly, 1 of the 20 images of the O-G coupling has been incorrectly classified as O-O coupling. All images in the O-O coupling have been correctly classified. The decision to use one or the other CNN for embedding the images will depend on whether, as

users, we want our images to be sent to an external server for embedding. For privacy and security reasons, we prefer them to be embedded locally [61].

		Predicted		
		O-B	**O-G**	**O-O**	**Σ**
Actual	**O-B**	19	0	1	20
	O-G	0	19	1	20
	O-O	0	0	20	20
	Σ	19	19	22	60

FIGURE 10. Confusion matrix of the logistic regression method using as descriptor extractor of CNN SquezeeNet images and applying a reduction of dimensions using the PCA technique (9 PC, explained variance: 95%). It is observed that 19 of the 20 images of the O-B coupling have been correctly classified, while one has been classified as O-O coupling. While, 1 of the 20 images of the O-G coupling has been incorrectly classified as O-O coupling. All images in the O-O coupling have been classified correctly.

CONCLUSIONS

In the framework of Systems Biology, mathematical modeling of biochemical mechanisms involved in different physiological processes is of vital importance because it allows us to understand the non-linear dynamics that underlie these phenomena. This is why the use of mathematical tools and computational systems for the analysis of the complex feedback mechanisms present in living systems is necessary. The CNT is a mathematical tool that allows studying these mechanisms with a holistic approach and provides valuable information on each of the entities that make up the system [26] [87].

When determining the authority structural property of the complex network obtained from the BZ reaction mechanism proposed by Györgyi et al., and using it as the centrality criterion, the variables with the highest relevance were identified, i.e., those chemical species that have the greatest flow of information and that

could participate in the emergence of collective properties of the system. Identification of these variables led to the construction of a nonlinear system of differential equations similar to the reduction of the FKN model proposed by Field and Noyes (Oregonator) and which also explains the phenomenology of the BZ reaction. Hence, this result answers the question of using mathematical tools to reduce complex reaction mechanisms without losing generality. Therefore, it is possible to use this methodology in the study of nonlinear dynamics present in biochemical and physiological processes.

On top of that, by applying this methodology to biological systems, it is possible to translate any biochemical or physiological process to a mathematical model and study the phenomena of synchronization between different regulatory mechanisms [88] to decipher the complex dynamics that underlie living systems with a systemic approach.

The effect of coupling between oscillators of different nature can be clearly seen in the images obtained using the gaf technique, which can be used to train a supervised learning model to classify the type of coupled oscillators. The extraction of descriptors from gaf images through pre-trained CNNs (transfer learning) allows obtaining high precision values in the evaluation of different classification models; however, it is also possible to couple the pre-trained CNNs with the PCA to obtain high values in the classification metrics, comparable with the values of the metrics obtained by using only pre-trained CNNs as a descriptor extraction method. In particular, using the CNN's Inception V3 and SquezeeNet as extractors for descriptors of gaf images and obtaining the principal components of these descriptors, allows training classification models such as logistic regression and obtaining CA and F1-score values above 0.94 for different evaluation methods.

All things considered, the methodology proposed in this work can facilitate the determination of synchronization and desynchronization states in complex real biochemical and physiological mechanisms to recognize a possible correlation between these states and the emergence of different complex diseases.

AUTHOR CONTRIBUTIONS

J.A.A.F., carried out the theoretical-numerical calculation, J.F.R.R. and E.V.R. have supervised and reviewed all the calculations. All authors have contributed to the writing and corrections of the manuscript.

REFERENCES

[1] Draghici, S., Khatri, P., Tarca, A., Amin, K., Done, A., Voichita, C., Georgescu, C. and Romero, R. A systems biology approach for pathway level analysis. Genome Research. 2007;17(10):1537-1545. https://doi.org/10.1101/gr.6202607

[2] Chuang HY, Hofree M, Ideker T. A Decade of Systems Biology. Annual Review of Cell and Developmental Biology. 2010 Oct;26(1):721-44. https://doi.org/10.1146/annurev-cellbio-100109-104122

[3] Kell DB. Systems biology, metabolic modelling and metabolomics in drug discovery and development. Drug Discovery Today. 2006;11(23-24):1085-92. https://doi.org/10.1016/j.drudis.2006.10.004

[4] Emmert-Streib F, Dehmer M. Networks for systems biology: conceptual connection of data and function. IET Systems Biology. 2011 Jan;5(3):185-207. https://doi.org/10.1049/iet-syb.2010.0025

[5] Maeda YT, Sano M. Regulatory Dynamics of Synthetic Gene Networks with Positive Feedback. Journal of Molecular Biology. 2006;359(4):1107-24. https://doi.org/10.1016/j.jmb.2006.03.064

[6] Ferrell JE. Self-perpetuating states in signal transduction: positive feedback, double-negative feedback and bistability. Current Opinion in Cell Biology. 2002;14(2):140-8. https://doi.org/10.1016/S0955-0674(02)00314-9

[7] Mullur R, Liu Y-Y, Brent G. Thyroid Hormone Regulation of Metabolism. Physiological Reviews. 2014 Apr;94(2):355-382. https://doi.org/10.1152/physrev.00030.2013

[8] Porte D, Baskin DG, Schwartz MW. Leptin and Insulin Action in the Central Nervous System. Nutrition Reviews. 2002;60:S20-S29. https://doi.org/10.1301/002966402320634797

[9] Chen WW, Niepel M, Sorger PK. Classic and contemporary approaches to modeling biochemical reactions. Genes & Development. 2010;24(17):1861-75. https://doi.org/10.1101/gad.1945410

[10] Alves R, Antunes F, Salvador A. Tools for kinetic modeling of biochemical networks. Nature Biotechnology. 2006;24(6):667-672. https://doi.org/10.1038/nbt0606-667

[11] Radulescu O, Gorban AN, Zinovyev A, Noel V. Reduction of dynamical biochemical reactions networks in computational biology. Frontiers in Genetics. 2012;3:131. https://dx.doi.org/10.3389%2Ffgene.2012.00131

[12] Zhabotinsky AM. A history of chemical oscillations and waves. Chaos: An Interdisciplinary. Journal of Nonlinear Science. 1991;1(4):379-86. https://doi.org/10.1063/1.165848

[13] Sagues F, Epstein IR. Nonlinear Chemical Dynamics. Dalton Transactions. 2003;32(7):1201-1217. https://doi.org/10.1039/B210932H

[14] Epstein IR, Pojman JA, Steinbock O. Introduction: Self-organization in nonequilibrium chemical systems. Chaos: An Interdisciplinary Journal of Nonlinear Science. 2006;16(3):037101. https://doi.org/10.1063/1.2354477

[15] Noyes RM, Field R, Koros E. Oscillations in chemical systems. I. Detailed mechanism in a system showing temporal oscillations. Journal of the American Chemical Society. 1972;94(4):1394-5. https://doi.org/10.1021/ja00759a080

[16] Field RJ, Koros E, Noyes RM. Oscillations in chemical systems. II. Thorough analysis of temporal oscillation in the bromate-cerium-malonic acid system. Journal of the American Chemical Society. 1972;94(25):8649-64. https://doi.org/10.1021/ja00780a001

[17] Shanks, N. Modeling biological systems: the Belousov-Zhabotinsky reaction. Foundations of Chemistry. 2011; 3(1): 33-53.

[18] Field RJ, Noyes RM. Oscillations in chemical systems. IV. Limit cycle behavior in a model of a real chemical reaction. The Journal of Chemical Physics. 1974;60(5):1877-84. https://doi.org/10.1063/1.1681288

[19] Györgyi L, Turanyi T, Field RJ. Mechanistic details of the oscillatory Belousov-Zhabotinskii reaction. The Journal of Physical Chemistry. 1990;94(18):7162-70. https://doi.org/10.1021/j100381a039

[20] Sayama H. Introduction to the modeling and analysis of complex systems. New York: Open SUNY Textbooks, Milne Library, State University of New York at Geneseo; 2015. 478p.

[21] Lesne A. Complex Networks: from Graph Theory to Biology. Letters in Mathematical Physics. 2006;78(3):235-62. https://doi.org/10.1007/s11005-006-0123-1

[22] Mason O, Verwoerd M. Graph theory and networks in Biology. IET Systems Biology. 2007;1(2):89-119. https://doi.org/10.1049/iet-syb:20060038

[23] Albert R, Barabási A-L. Statistical mechanics of complex networks. Reviews of Modern Physics. 2002;74(1):47-97. https://doi.org/10.1103/RevModPhys.74.47

[24] Costa LDF, Rodrigues FA, Hilgetag CC, Kaiser M. Beyond the average: Detecting global singular nodes from local features in complex networks. EPL (Europhysics Letters). 2009;87(1):18008. https://doi.org/10.1209/0295-5075/87/18008

[25] Loscalzo J, Barabási Albert-László, Silverman EK. Network medicine: complex systems in human disease and therapeutics. Cambridge, MA: Harvard University Press; 2017. 448p.

[26] Costa LDF, Rodrigues FA, Cristino AS. Complex networks: the key to systems biology. Genetics and Molecular Biology. 2008;31(3):591-601. http://dx.doi.org/10.1590/S1415-47572008000400001

[27] Hornberg J, Bruggeman FJ, Westerhoff HV, Lankelma J. Cancer: A Systems Biology disease. Biosystems. 2006;83(2-3):81-90. https://doi.org/10.1016/j.biosystems.2005.05.014

[28] Cardon LR, Bell JI. Association study designs for complex diseases. Nature Reviews Genetics. 2001;2(2):91-99. https://doi.org/10.1038/35052543

[29] DeFronzo RA. Insulin Resistance, Hyperinsulinemia, and Coronary Artery Disease: A Complex Metabolic Web. Journal of Cardiovascular Pharmacology. 1992;20 (Suppl 1):S1-S16. https://doi.org/10.1097/00005344-199200111-00002

[30] Karalliedde J, Gnudi L. Diabetes mellitus, a complex and heterogeneous disease, and the role of insulin resistance as a determinant of diabetic kidney disease. Nephrology Dialysis Transplantation. 2016;31(2):206-13. https://doi.org/10.1093/ndt/gfu405

[31] Upadhyay SK. Chemical kinetics and reaction dynamics. New Delhi: Springer Netherlands; 2006. 256p. https://doi.org/10.1007/978-1-4020-4547-9

[32] Kotsiantis SB, Zaharakis ID, Pintelas PE. Machine learning: a review of classification and combining techniques. Artificial Intelligence Review. 2006;26:159-90. https://doi.org/10.1007/s10462-007-9052-3

[33] Berzal, F. Redes neuronales & deep learning. 1ra. ed. Granada, España: Edición Independiente; 2018. 753p.

[34] Bastian M, Heymann S, Jacomy M. Gephi: an open source software for exploring and manipulating networks. International AAAI Conference on Weblogs and Social Media. 2009;361-362.

[35] Moon FC. Chaotic and fractal dynamics: introduction for applied scientists and engineers. 2nd. ed. New York: Wiley; 1992. 528p.

[36] Afraimovich VS, Lin WW, Rulkov NF. Fractal dimension for poincaré recurrences as an indicator of synchronized chaotic regimes. International Journal of Bifurcation and Chaos. 2000;10(10):2323-2337. https://doi.org/10.1142/S0218127400001456

[37] Liu Z, Lai Y, Matías M. Universal scaling of Lyapunov exponents in coupled chaotic oscillators. Physical Review E. 2003;67(4): 045203. http://dx.doi.org/10.1103/PhysRevE.67.045203

[38] Porta A, Baselli G, Lombardi F, Montano N, Malliani A, Cerutti S. Conditional entropy approach for the evaluation of the coupling strength. Biological Cybernetics. 1999;81:119-129. https://doi.org/10.1007/s004220050549

[39] Eckmann, JP, Kamphorst SO, Ruelle D. Recurrence plots of dynamical systems. Europhysics Letter. 1987; 4(9):973-977. https://doi.org/10.1142/9789812833709_0030

[40] Hernández Sánchez S, Fernández Pozo R, Hernández Gómez LA. Deep Neural Networks for Driver Identification Using Accelerometer Signals from Smartphones. In: Abramowicz W, Corchuelo R. (eds) Business Information Systems. Lecture Notes in Business Information Processing. Switzerland: Springer, Cham; 2019, p. 206-220. https://doi.org/10.1007/978-3-030-20482-2_17

[41] Bode BW. Clinical Utility of the Continuous Glucose Monitoring System. Diabetes Technology & Therapeutics. 2000;2(Suppl 1):S35-41. https://doi.org/10.1089/15209150050214104

[42] Yoo EH, Lee SY. Glucose Biosensors: An Overview of Use in Clinical Practice. Sensors. 2010;10(5):4558-4576. https://doi.org/10.3390/s100504558

[43] Frost MC, Meyerhoff ME. Implantable chemical sensors for real-time clinical monitoring: progress and challenges. Current Opinion in Chemical Biology. 2002;6(5):633-641. https://doi.org/10.1016/s1367-5931(02)00371-x

[44] Frost MC, Meyerhoff ME. Real-Time Monitoring of Critical Care Analytes in the Bloodstream with Chemical Sensors: Progress and Challenges. Annual Review of Analytical Chemistry. 2015;8(1):171-192. https://doi.org/10.1146/annurev-anchem-071114-040443

[45] Coveney PV, Fowler PW. Modelling biological complexity: a physical scientist's perspective. Journal of The Royal Society Interface. 2005;2(4):267-280. https://doi.org/10.1098/rsif.2005.0045

[46] Kleinberg JM. Authoritative sources in a hyperlinked environment. Journal of the ACM. 1999;46(5): 604-632. https://doi.org/10.1145/324133.324140

[47] Aton SJ, Herzog ED. Come Together, Right...Now: Synchronization of Rhythms in a Mammalian Circadian Clock. Neuron. 2005;48(4): 531-534. https://dx.doi.org/10.1016%2Fj.neuron.2005.11.001

[48] Lorenzo González MN. Influencia del Ruido Gaussiano Correlacionado en la Sincronización de Sistemas Caóticos [Ph.D.'s thesis]. [Santiago de Compostela]:Universidad de Santiago de Compostela, 2000.161p. Spanish.

[49] Lefever R, Nicolis G, Borckmans P. The brusselator: it does oscillate all the same. Journal of the Chemical Society, Faraday Transactions 1: Physical Chemistry in Condensed Phases. 1988;84(4):1013-1023. https://doi.org/10.1039/F19888401013

[50] Vaidyanathan, S. Anti-synchronization of Brusselator chemical reaction systems via adaptive control. International Journal of ChemTech Research. 2015;8(6):759-68.

[51] Sel'kov EE. Self-Oscillations in glycolysis. 1. A Simple Kinetic Model. European Journal of Biochemistry. 1968;4(1):79-86. https://doi.org/10.1111/j.1432-1033.1968.tb00175.x

[52] Keener JP, Sneyd J. Mathematical physiology. 2nd. Ed. New York: Springer; 2009. 470p. https://doi.org/10.1007/978-0-387-75847-3

[53] Butcher J. Numerical methods for ordinary differential equations in the 20th century. In: Brezinski C, Wuytack, L. (eds). Numerical Analysis: Historical Developments in the 20th Century. Amsterdam: Elsevier; 2001.449-77p. https://doi.org/10.1016/B978-0-444-50617-7.50018-5

[54] Faouzi J, Janati H. pyts: a python package for time series classifi-cation. Journal of Machine Learning Research. 2020; 21(46):1–6.

[55] Wang Z, Oates T. Imaging time-series to improve classification and imputation. Proceedings of the 24th International Joint Conference on Artificial Intelligence. 2015;3939-45.

[56] Брагин, AB, Bragin A, Спицын ВГ, Spicyn, V.Electroencephalogram Analysis Based on Gramian Angular Field Transformation. Proceedings of the 29th International Conference on Computer Graphics and Vision. 2019;2485:273-75. https://doi.org/10.30987/graphicon-2019-2-273-275

[57] Damaševičius R, Maskeliūnas R, Woźniak M, Polap D. Visualization of physiologic signals based on Hjorth parameters and Gramian Angular Fields. 2018 IEEE 16th World Symposium on Applied Machine Intelligence and Informatics (SAMI). 2018; 91-6. https://doi.org/10.1109/sami.2018.8323992

[58] Qin Z, Zhang Y, Meng S, Qin Z, Choo K-KR. Imaging and fusing time series for wearable sensor-based human activity recognition. Information Fusion. 2020;53:80-7. https://doi.org/10.1016/j.inffus.2019.06.014

[59] Demsar J, Curk T, Erjavec A, Gorup C, Hocevar T, Milutinovic M, Mozina M, Polajnar M, Toplak M, Staric A, Stajdohar M, Umek L, Zagar L, Zbontar J, Zitnik M, Zupan B. Orange: Data Mining Toolbox in Python. Journal of Machine Learning Research. 2013; 14(35): 2349–2353.

[60] Demšar J, Zupan B. Orange: Data Mining Fruitful and Fun- A Hystorical Perspective. Informatica 2013;37:55-60.

[61] Godec P, Pančur M, Ilenič N, Čopar A, Stražar M, Erjavec A, Pretnar A, Demšar J, Starič A, Toplak M, Žagar L, Hartman J, Wang H, Bellazzi R, Petrovič U, Garagna S, Zuccotti M, Park D, Shaulsky G, Zupan B. Democratized image analytics by visual programming through integration of deep models and small-scale machine learning. Nature Communications. 2019;10(1):4551. https://doi.org/10.1038/s41467-019-12397-x

[62] Weiss K, Khoshgoftaar T, Wang D. A survey of transfer learning. Journal of Big Data. 2016;3(1):1345-1359. **https://doi.org/10.1186/s40537-016-0043-6**

[63] Tan C, Sun F, Kong T, Zhang W, Yang C, Liu C. A survey on deep transfer learning. In Kůrková V, Manolopoulos Y, Hammer B, Iliadis L, Maglogiannis I. International conference on artificial neural networks. Rhodes: Springer International Publishing. Cham; 2018. p. 270-279. **https://doi.org/10.1007/978-3-030-01424-7_27**

[64] Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR); Las Vegas: IEEE; 2016. 2818-2826p. **https://doi.org/10.1109/CVPR.2016.308**

[65] Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size [Internet]. arXiv.org; 2020. Available from: *https://arxiv.org/abs/1602.07360*

[66] Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition [Internet]. arXiv.org; 2020. Available from: *https://arxiv.org/abs/1409.1556v6*

[67] Almagro Armenteros JJ, Sønderby CK, Sønderby SK, Nielsen H, Winther O. DeepLoc: prediction of protein subcellular localization using deep learning. Bioinformatics. 2017;33(21):3387-3395. **https://doi.org/10.1093/bioinformatics/btx431**

[68] Müller AC, Guido S. Introduction to machine learning with Python a guide for data scientists. Beijing: OReilly; 2018. p. 251-303.

[69] Hawkins DM. The Problem of Overfitting. Journal of Chemical Information and Computer Sciences. 2004;44(1):1-12. **https://doi.org/10.1021/ci0342472**

[70] Sokolova M, Japkowicz N, Szpakowicz S. Beyond Accuracy, F-Score and ROC: A Family of Discriminant Measures for Performance Evaluation. In Sattar A, Kang B. (eds). AI 2006: Advances in Artificial Intelligence. Berlin: Springer; 2006. 1015-21p. **https://doi.org/10.1007/11941439_114**

[71] Pinto CMA, Mendes Lopes A, Tenreiro Machado JA. A review of power laws in real life phenomena. Communications in Nonlinear Science and Numerical Simulation. 2012;17(9):3558-78. **https://doi.org/10.1016/j.cnsns.2012.01.013**

[72] Gisiger T. Scale invariance in biology: coincidence or footprint of a universal mechanism? Biological Reviews of the Cambridge Philosophical Society. 2001;76(2):161-209. **https://doi.org/10.1017/s1464793101005607**

[73] Amemiya T, Kádár S, Kettunen P, Showalter K. Spiral Wave Formation in Three-Dimensional Excitable Media. Physical Review Letters. 1996;77(15):3244-7. **https://doi.org/10.1103/physrevlett.77.3244**

[74] Krug HJ, Pohlmann L, Kuhnert L. Analysis of the modified complete Oregonator accounting for oxygen sensitivity and photosensitivity of Belousov-Zhabotinskii systems. The Journal of Physical Chemistry. 1990;94(12):4862-6. **https://doi.org/10.1021/j100375a021**

[75] Ma J, Li F, Huang L, Jin W-Y. Complete synchronization, phase synchronization and parameters estimation in a realistic chaotic system. Communications in Nonlinear Science and Numerical Simulation. 2011;16(9):3770-3785. **https://doi.org/10.1016/j.cnsns.2010.12.030**

[76] Rajesh S, Sinha S, Sinha S. Synchronization in coupled cells with activator-inhibitor pathways. Physical Review E. 2007;75(1): 011906. **https://doi.org/10.1103/physreve.75.011906**

[77] Belykh I, Lange ED, Hasler M. Synchronization of Bursting Neurons: What Matters in the Network Topology. Physical Review Letters. 2005;94(18):1881011-1881014. **https://doi.org/10.1103/physrevlett.94.188101**

[78] Belhaq M, Houssni, M. Quasi-periodic oscillations, chaos and suppression of chaos in a nonlinear oscillator driven by parametric and external excitations. Nonlinear Dynamics. 1999;18(1):1-24. **https://doi.org/10.1023/A%3A1008315706651**

[79] Ng A. Feature selection, L 1 vs. L 2 regularization, and rotational invariance. In Brodley C (ed.). ICML '04: Proceedings of the twenty-first international conference on Machine learning. New York: Association for Computing Machinery; 2004. 78-86p. **https://doi.org/10.1145/1015330.1015435**

[80] Ng A. Preventing" overfitting" of cross-validation data. In Fisher DH (ed.). ICML '97: Proceedings of the Fourteenth International Conference on Machine Learning. San Francisco: Morgan Kaufmann Publishers Inc; 1997. 245-253p.

[81] Mao W, Mu X, Zheng Y, Yan G. Leave-one-out cross-validation-based model selection for multi-input multi-output support vector machine. Neural Computing and Applications. 2012;24(2):441-451. **https://doi.org/10.1007/s00521-012-1234-5**

[82] Grünauer A, Vincze M. Using Dimension Reduction to Improve the Classification of High-dimensional Data [Internet]. OAGM Workshop; 2020. Available from: *https://arxiv.org/pdf/1505.06907.pdf*

[83] Azure Machine Learning. Evitar el sobreajuste y los datos desequilibrados con el aprendizaje automático automatizado. Microsoft [Internet]. 2019. Available from: https://docs.microsoft.com/es-es/azure/machine-learning/concept-manage-ml-pitfalls

[84] Christodoulou E, Ma J, Collins GS, Steyerberg EW, Verbakel JY, Calster BV. A systematic review shows no performance benefit of machine learning over logistic regression for clinical prediction models. Journal of Clinical Epidemiology. 2019;110:12-22. **https://doi.org/10.1016/j.jclinepi.2019.02.004**

[85] Rajagopal S, Hareesha KS, Kundapur PP. Performance analysis of binary and multiclass models using azure machine learning. International Journal of Electrical and Computer Engineering (IJECE). 2020;10(1):978-986. **http://doi.org/10.11591/ijece.v10i1.pp978-986**

[86] Kohavi R. A study of cross-validation and bootstrap for accuracy estimation and model selection. IJCAI'95: Proceedings of the 14th international joint conference on Artificial intelligence. 1995;14(2):1137-1145

[87] Alducin Castillo J, Yanez Suárez O. Brust Carmona H. Electroencephalographic analysis of the functional conectivity in habituation by graphics theory. Revista Mexicana de Ingeniería Biomédica. 2016;37(3):181-200. **https://doi.org/10.17488/RMIB.37.3.3**

[88] Glass L. Synchronization and rhythmic processes in physiology. Nature. 2001;410(6825):277-84. **https://doi.org/10.1038/35065745**