RESEARCH ARTICLE

A New Artificial Immune System Based on Continuous Learning for Pattern Recognition

Um Novo Sistema Imunológico Artificial Baseado no Aprendizado Contínuo para Reconhecimento de Padrões

Simone S. F Souza [1], Fernando P. A. Lima [2]*, Fábio R. Chavarette[3]

Abstract: This paper presents a novel approach for pattern recognition based on continuous training inspired by the biological immune system operation. The main objective of this paper is to present a method capable of continually learn, i.e., being able to address new types of patterns without the need to restart the training process (artificial immune system with incremental learning). It is a useful method for solving problems involving a permanent knowledge extraction, e.g., 3D facial expression recognition, whose quality of the solutions is strongly dependent on a continuous training process. In this context, two artificial immune algorithms are employed: (1) the negative selection algorithm, which is responsible for the pattern recognition process and (2) the clonal selection algorithm, which is responsible for the learning process. The main application of this method is in assisting in decision-making on problems related to pattern recognition process. To evaluate and validate the efficiency of this method, the system has been tested on handwritten character recognition, which is a classic problem in the literature. The results show efficiency, accuracy and robustness of the proposed methodology.
Keywords: Pattern Recognition — Continuous Learning — Artificial Immune Systems — Negative Selection Algorithm — Clonal Selection Algorithm

Resumo: Este artigo apresenta uma nova abordagem para reconhecimento de padrões com base em treinamento contínuo inspirado na operação do sistema imunológico biológico. O principal objetivo deste artigo é apresentar um método capaz de aprender continuamente, ou seja, ser capaz de abordar novos tipos de padrões sem a necessidade de reiniciar o processo de treinamento (sistema imunológico artificial com aprendizado incremental). É um método útil para resolver problemas que envolvem uma extração permanente de conhecimento, por exemplo, reconhecimento de expressão facial 3D, cuja qualidade das soluções depende fortemente de um processo de treinamento contínuo. Nesse contexto, dois algoritmos imunes artificiais são empregados: (1) o algoritmo de seleção negativa, responsável pelo processo de reconhecimento de padrões e (2) o algoritmo de seleção clonal, responsável pelo processo de aprendizado. A principal aplicação desse método é auxiliar na tomada de decisões sobre problemas relacionados ao processo de reconhecimento de padrões. Para avaliar e validar a eficiência desse método, o sistema foi testado no reconhecimento de caracteres manuscritos, que é um problema clássico na literatura. Os resultados mostram eficiência, precisão e robustez da metodologia proposta.
Palavras-Chave: Reconhecimento de Padrões — Aprendizado Contínuo — Sistemas Imunológicos Artificiais — Algoritmo de Seleção Negativa — Algoritmo de Seleção Clonal

[1]*State University of Mato Grosso (UNEMAT), Campus of Tangará da Serra, Rodovia MT-358, Km 07, Jardim Aeroporto, 78300000 - Tangará da Serra, MT - Brasil*
[2]*Federal Institute of Science and Technology Education of Mato Grosso (IFMT), Advanced Campus of Tangará da Serra, Rua 28, 980 N, Vila Horizonte, 78300000 - Tangará da Serra, MT - Brasil*
[3]*Mathematical Department, Faculty of Engineering of Ilha Solteira (FEIS), UNESP, Universidade Estadual Paulista Júlio de Mesquita Filho, Av. Brasil, 56, PO Box 31, 15385-000, Ilha Solteira, SP - Brasil*
*Corresponding author: fernando.lima@tga.ifmt.edu.br

1. Introduction

Pattern recognition consists of a classification or categorization task where the designer defines the classes (supervised classification) or these classes are determined based on the similarities of the patterns (non-supervised classification).

Pattern recognition has been increasingly researched lately due to its several applications that are challenging and demand

hard computational efforts. The availability of advanced computational resources transforms the use and elaboration of methods for pattern recognition analysis. A pattern recognition system contains three main steps as follows:

1. Data acquisition (i.e., feature extraction) and preprocessing (selection of the most discriminative characteristics);
2. Data representation;
3. Decision-making (building a classifier).

The principal contribution in pattern recognition that is present in the specialized body of literature is the building of classification methods, which are based on intelligent systems like fuzzy logic [1], artificial neural networks [2], artificial immune systems [3], among others.

Generally, the pattern recognition methods based on intelligence have a training strategy in obtaining knowledge. However, these strategies (learning or training algorithms) reinitialize the training process every time the process is executed and, principally, when a new pattern is observed by the system. Such methods are efficient and are able to solve the current problems. However, they do not learn constantly, and when a new pattern is available, all the training processes are executed again, which is a task that cannot be considered intelligent in a strict sense. A modern and intelligent pattern recognition system must learn continuously, have the capacity to innovate and update new patterns, learn with the experience, and be in constant improvement.

Developing a system with such capacities is a complex task and is seldom researched in the literature. However, still there is no available an efficient methodology, so, the use of the artificial immune systems is an alternative solution. In the specialized literature, can be cited the papers [4], [5], [6], [7], [8], proposing approach in this regard. Thus, this paper proposes a new approach for pattern recognition, providing the continuous training on pattern recognition system.

This paper presents a new approach to develop a pattern recognition system with continuous training based on the Biological Immune System (BIS). To create this system, two different immune systems are used, including the Negative Selection Algorithm (NSA), [9] and the Clonal Selection Algorithm (CLONALG), [10]. The artificial immune algorithm, which is inspired by the BIS, is a promising technique in Intelligent Computing (IC). The BIS computationally reproduces the principal characteristics and the propriety functionalities and abilities. It has a stable, reliable, and adaptive architecture, allowing the continuous inclusion of the training module. This property allows the system to introduce new experiences and knowledge without the need to reinitialize the immune memory of the system. This is one advantage of using Artificial Immune Systems, which allow the possibility of continuous training unlike in other techniques such as using Artificial Neural Networks (ANNs). In some ANNs, to improve the training, it is necessary to reinitialize the process, consequently destroying the knowledge previously acquired,

except for a few architectures that have the plasticity propriety as in the neural networks of the ART (Adaptive Resonance Theory) family [11].

To evaluate and validate the efficiency of this new approach, a classical recognition problem of manuscript characters is used. It is compared with a conventional recognition system (NSA) to analyze the performance of the proposal.

This paper is organized as follows. Section II discusses the BIS, section III focuses on the NSA, and section IV tackles the CLONALG. The proposed methodology is presented in section V, and the applications and results are discussed in section VI. The conclusions are presented in section VII.

2. Biological Immune System

The BIS is the principal defence of the living organisms against infection agents that enter the human body. In this case, the BIS must instantaneously act against the infection agents and identify them to protect the human body (evidence of a disease). There are two kinds of defences: the innate immune system and the adaptive immune system.

The innate immune system is the first line of defence and responds faster. It is characterized by dendritic cells (APC, Antigen Presenting Cells) and phagocytes (granulocytes, macrophages, etc.), which are responsible in ingesting strange particles to the organism and other kinds of defences such as physical (skin) and chemical barriers [3].

The adaptive immune system is the second line of defence and is able to recognize microorganisms like virus, bacteria, fungus, protozoa, helminthic, and some kinds of worms. The adaptive immune system realizes the training process of the infecting agents at first contact with the antigen. It is also responsible in creating immune memory cells from the first exposure to the infecting agent to accelerate the response to the same kind of infecting agent in future exposure [12].

2.1 Operation of the Pattern Recognition System and Training of the BIS

The biological organism, which is unique in humans as being composed of several cells and molecules that work in harmony, aims answers to strange agents that are called antigens.

The BIS has fundamental characteristics such as pattern recognition and training and neutralization of the infecting agent. All these steps represent an immune answer that can be illustrated in a simplified form as shown in Figure 1.

Step (I) represents the BIS exposure to an antigen (infecting agent). Basically, the BIS can react in two ways: the innate immune response, which begins at step (II), and the adaptive immune response, which begins at step (VII). If the process begins at the innate immune system (step (II)), the infecting agent is considered unknown by the BIS; therefore, it is necessary to identify and inform the infecting cell of the organism to forward the adaptive immune system begins an adaptive answer (identify and generate the antibodies that will neutralize the antigen). If the process starts at the adaptive immune system (step (VII)), the antigen is known by
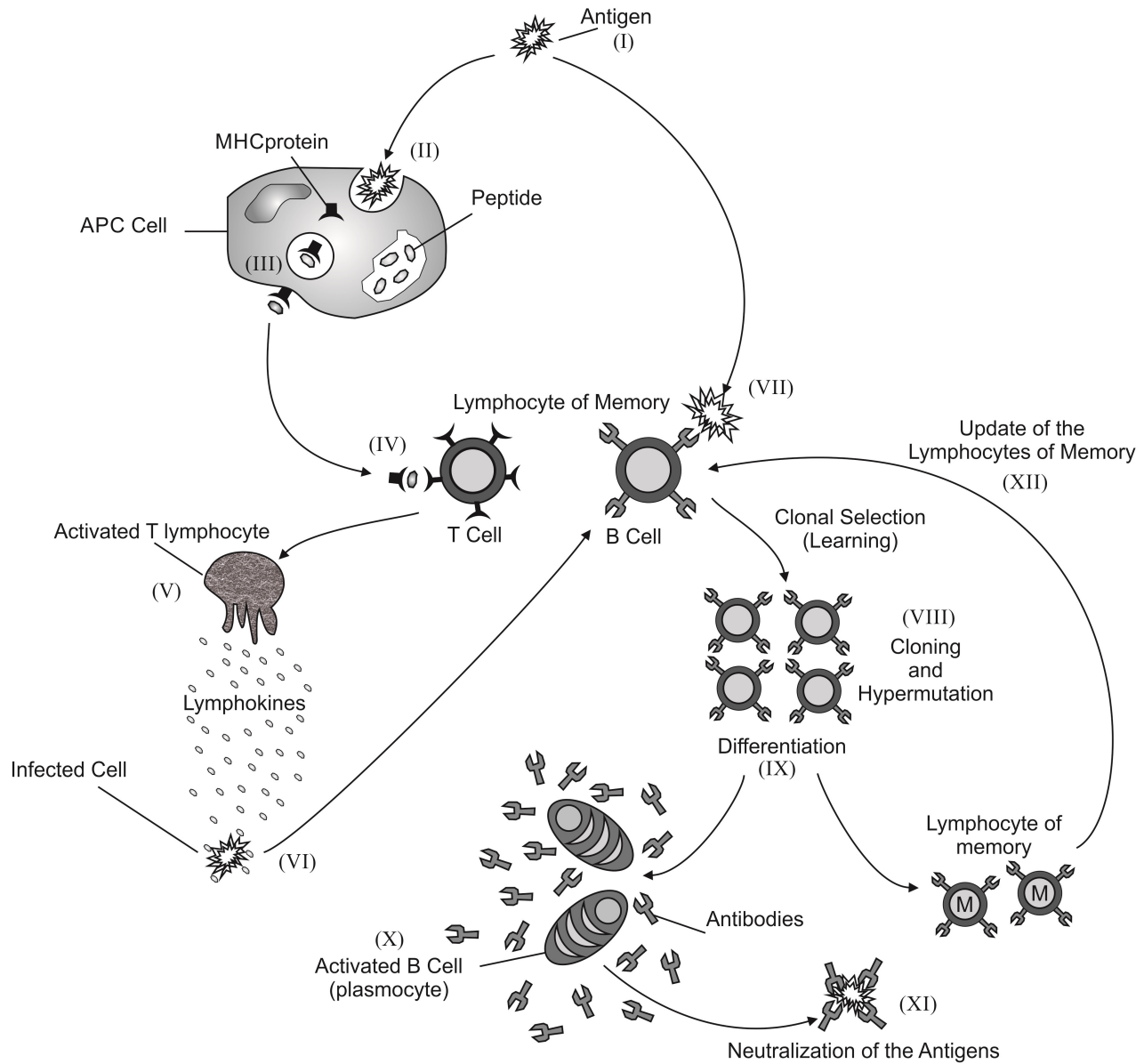
Figure 1. Operation of the BIS

the BIS, and the detection process is realized by the memory lymphocytes presented in the organism.

At step (II), the innate immune process is initialized when an antigen (infecting agent) is ingested by the dendritic cell (APC). At this phase, the antigens, which are fragmented in antigenic peptides, digest the antigens. At step (III), the fragmented peptides are linked to the MHC (Major Histocompatibility Complex) and presented to the surface of the dendritic cells. At step (IV), the T lymphocytes (which have receiving molecules in the surface) are able to recognize and identify different MHC/peptide antigens processed by the dendritic cells, that is, when there is connection (combination) that occurs in the recognition of the antigen and, thus, the lymphocyte state is activated. This step represents the dis-

crimination proper/no proper process realized by the organism distinguishing the proper cells from the infecting agents [13], [12]. Based on this pattern recognition principium of the NSA is proposed [9]. After the detection of the antigen at step (V), the T lymphocytes are activated, secreting chemical signals (lymphokine) that notify other components of the BIS that an antigen is found, or simply that the cell is infected.

At step (VI) when an infected cell is identified, the B lymphocytes can recognize the infected cell and initialize the training process, which is the so called clonal selection principium (step (VII)). In this case, the B lymphocytes are able to recognize the free antigens identified on the organism without the need to ingest and digest the APC, and thus, recognition happens. When the B cells recognize an antigen

(as for signalization or by memory cells), the clonal selection process begins, which is also known as training process (step (VIII)). At this step, the lymphocytes are subjected to a clonal and hypermutation process of generating a set of B lymphocytes on the organism. There is a differentiated process (step (IX)) where the B lymphocytes with highest affinity are separated to pertain to the memory set, and the B lymphocytes with low affinity are activated and transformed into plasmocyte-secreting antibodies (step (X)). At step (XI), the antigens detected by the BIS are neutralized by the generated antibodies, leading to the destruction of the danger (disease).

At step (XII), the lymphocytes that are separated on the differentiation process are transformed into memory lymphocytes and pertain to the set of memory lymphocytes substituting the old lymphocytes with low affinity. The new memory lymphocytes circulate on the organism, assuring an efficient and faster response to future exposure to the same type of antigen (infecting agent). It is emphasized that all the processes are realized with cooperation of the BIS cells where each one is responsible by a simple function and realizes a complex work [13], [12].

3. Negative Selection Algorithm (NSA)

The NSA, which was proposed in [9] to detect changes in the systems, is based on the negative selection of T lymphocytes over time. This process works on the discrimination of proper versus non-proper cells. The algorithm is executed in two phases, as follows [3], [14]:

1. Censor

   (a) Define a set of proper chains (S) to be protected;

   (b) Generate random chains, and evaluate the affinity (match) between each chain and the proper chains. If the affinity is greater than a predefined value, then the chain is rejected. Otherwise, the chain is filed into a detector set (R).

2. Monitor

   (a) Given a set of chains to be protected (protected chains), evaluate the affinity between each chain and the detector set. If the affinity is superior to a predefined value, then a non-proper element is identified.

The censor phase of the NSA primarily consists of generating a detector set from the data that were randomly chosen and verifying which data can recognise a non-proper pattern. The detectors are similar to mature T cells, which can recognise pathogenic agents [3].

The monitoring phase consists of monitoring a system to identify a change in the behaviour; thus, this phase classifies the change using the detector set that has been created in the censor phase. The censor phase occurs offline, and the monitoring phase occurs in real time [3], [15].

Figures 2 and 3 present the flowcharts of the censor and the monitoring phases of the negative selection algorithm.
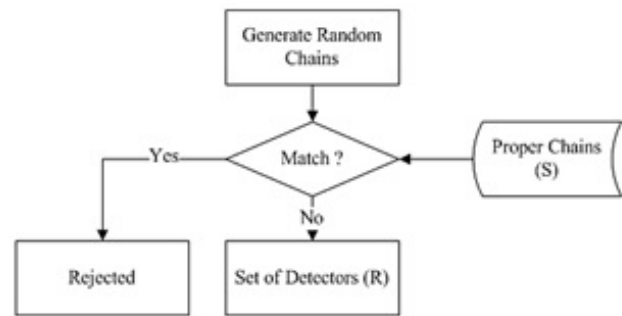

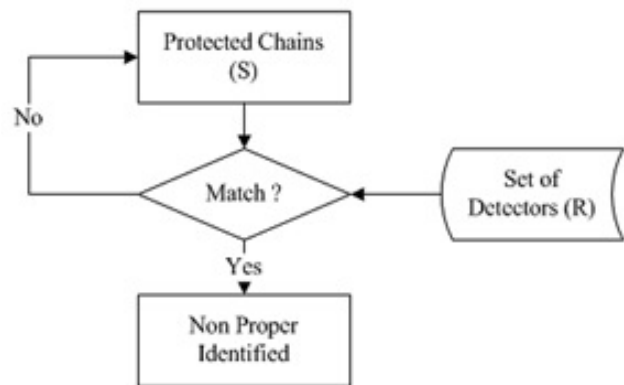**Figure 2.** Flowchart of the censor phase of the NSA.

Figure 3. Flowchart of the monitoring phase of the NSA.

The antigen (Ag) is the signal to be analysed in the negative selection algorithm and can be represented by expression (1) [12], [15].

$$Ag = Ag_1, Ag_2, Ag_3, Ag_4, ...Ag_n \qquad (1)$$

The detectors represent the antibodies (Ab) and are expressed according to expression (2) [12], [15].

$$Ab = Ab_1, Ab_2, Ab_3, Ab_4, ...Ab_n \qquad (2)$$

where L is the dimension of the space of the antigen and the antibody.

3.1 Matching Criterion

To evaluate the affinity with the chains and to prove that they are similar, the matching criterion is used, which has the same meaning as the combination. The matching can be perfect or partial [14].

The matching is perfect when the two analysed chains have the same value in every position, and the matching is partial when the patterns have only one identical position value to confirm the matching [16]. This quantity is known

as the affinity rate. The affinity rate represents that there is a similar grade for matching to occur between two analysed chains [14]. In the paper [16], the affinity rate is defined according to the following equation:

$$TAf = \frac{An}{At}100 \qquad (3)$$

TAf the affinity rate; *An* is the quantity of normal rates in the problem (proper rates); and *At* is the total number of chains in the problem (proper and non-proper chains).

Equation (3) allows the precise calculation of the affinity rate for the proposed problem and represents the statistical analysis with the samples of the problem. Expression (4) represents the method for quantifying the total affinity with the analysed patterns [16], [14]:

$$Af_t = \frac{\sum_{i=1}^{L} P_c}{L}100 \qquad (4)$$

where Af_t of the affinity with the patterns analysed; L is total quantity of positions; P_c is matched position; and, $\sum_{i=1}^{L}$ sum of the matched positions.

Thus, if Af_t is greater than TAf, then the combination or matching with the patterns occurs, that is, the patterns are considered to be equal or similar. Otherwise, there is no matching with the patterns.

4. Clonal Selection Algorithm

The clonal selection algorithm (CLONALG) is originally proposed in [10]. Different from what is proposed in [9], this algorithm considers two central characteristics of the clonal selection principium: the maturation and the proportional selection to the affinity. There are two versions of this algorithm that are available in the literature: one is to solve machine training problems and pattern recognition and the other is to solve optimization problems [3].

The CLONALG for pattern recognition and machine training is described according to the following steps [12], [3]:

1. Initialization: Generate a random population $(Ab = Ab_{\{M\}} + Ab_{\{R\}})$ with L lymphocytes for each antigen (Ag_i). L is given by $M + R$;
2. Affinity evaluation: Each antigen (Ag_i) is presented to every lymphocyte of the population (Ab) in a process to evaluate the affinity. An affinity vector f is then determined;
3. Selection: The n lymphocytes with the highest affinity f in relation to (Ag_i) are selected to compose a subpopulation $(Ab\{n\})$;
4. Cloning: The n lymphocytes selected proliferate (cloning) proportionally to the affinities to the antigen (Ag_i), generating a C clone population. The higher the affinity f is, the highest the quantity of clones will be in each one of the n selected lymphocytes;

5. Hypermutation: The C clone population is then subjected to the affinity maturation process to generate a new population C^*, where each lymphocyte will mutate on a rate inversely proportional to the affinity f. The higher the affinity f is, the less the mutation rate will be for each lymphocyte;
6. Affinity evaluation: Determine the affinity f^* with the set C^* of mutated clones and the antigen (Ag_i);
7. Re-selection: From the mature C^* population, re-select the n best matured lymphocytes to form the subpopulation $(Ab_{\{n\}})$. From this subpopulation, choose the best lymphocytes to get into the memory set $(Ab_{\{M\}})$. One lymphocyte gets into the memory set when it presents high affinity rates, and it can substitute some memory lymphocytes;
8. Metadynamics: Substitute d antibodies from $(Ab_{\{R\}})$ by $(Ab_{\{d\}})$ new individuals, thus inducing diversity on the repertoire. The antibodies with lesser affinities are substituted.
9. Repeat steps 2 to 8 until the stop criterion is satisfied.

At the end of the iterative process, the memory set $(Ab_{\{M\}})$ has M lymphocytes with the highest affinity rates in relation to the antigen (Ag_i). This memory set can be used by the NSA to detect and classify the antigen learnt during the clonal selection process.

It is emphasized that for training problems, step 8 (metadynamics) is not executed, making the parameter $d = 0$.

The quantity Nc of clones generated in step 4, for each lymphocyte i, is given by equation (5), [3]:

$$N_c^i = round\frac{\beta N}{i} \qquad (5)$$

where: β a multiplicative factor between 0 and 1, N_c^i is the total quantity of lymphocytes of the Ab population, and $round(.)$ is the rounding operator to the closest integer. The mutation rate (α) of each clone is defined by equation (6), [3]:

$$\alpha = exp(-\rho f^n) \qquad (6)$$

where: ρ is a damping control parameter of the exponential function and f^n is the normalized value of the affinity f, which is calculated based on equation (7):

$$f^n = \frac{f}{f_{max}} \qquad (7)$$

This way, each clone is mutated according to [17]:

$$m = round(\alpha * N(0,1)) \qquad (8)$$

where: m is the quantity of mutations for each clone, α is the mutation rate, and $N(0,1)$ is a random Gaussian variable with mean of zero and pattern deviation $\sigma = 1$.

5. Proposed Methodology

This section presents the proposed methodology to solve the recognition problem of manuscript characters with a pattern recognition algorithm with continuous training. The algorithm without the continuous training is also presented, that is, the NSA algorithm.

The database of the manuscript characters used to validate the method and the preprocessing technique of the images are then presented.

5.1 Database of the Manuscript Characters

A database containing numerical manuscript characters is built with 200 samples, with 20 samples from each type of pattern (numbers 0 to 9). Some individuals wrote the numbers in a table, and this table is digitized as shown in Figure 4. Each manuscript character is extracted with the dimension of 32x32 pixels and filed in BMP format [18].

Figure 4. Database of the manuscript characters.

5.2 Image Preprocessing

After every image has been extracted from the database, a preprocessing using an image processing technique is carried out. This process is important to avoid noise and to intensify the patterns. Thus, the original images are inversely transformed into binary, where the white color (zero) becomes black (one) and vice versa, as illustrated in Figures 5 (a) and 5 (b). The images are then expanded using a process called dilation, as shown in Figure 5 (c) [18].
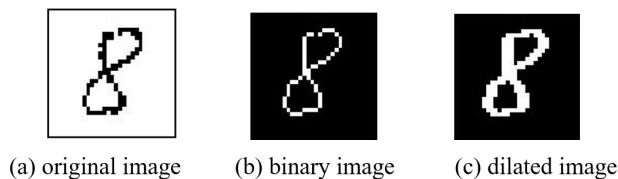
(a) original image   (b) binary image   (c) dilated image

Figure 5. Image preprocessing.

5.3 Manuscript Character Recognition System Without the Continuous Training

The manuscript character recognition system without the continuous training is based on the NSA proposed in [9]. This system is composed of three modules: (1) image preprocessing (as shown in Figure 5); (2) censor of the NSA (offline);

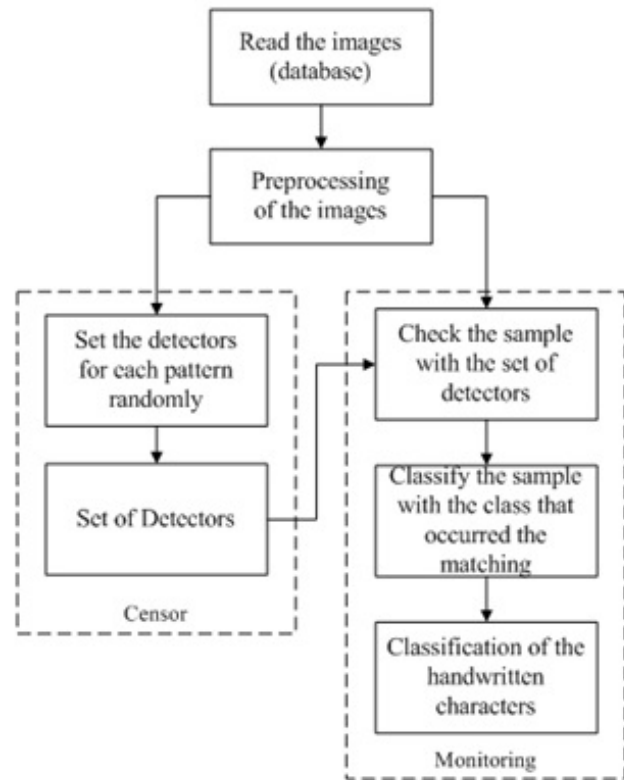and (3) monitoring of the NSA (on-line). The flowchart for this algorithm is presented in Figure 6).

Figure 6. Flowchart of the system.

Preprocessing is executed to improve the quality of the images and increase the characteristics of each pattern.

A σ quantity of detector for each pattern is randomly defined at the censor module of the NSA. This set of detectors is used on the monitoring process to verify the affinity and the matching in relation to the patterns under analysis.

The matching is verified during the monitoring of the NSA that evaluates the quality of the similar pixels with the patterns (quantifying the affinity using equation (4)). Matching occurs when the affinity with two patterns is superior to the defined affinity rate. When matching occurs with the patterns, the pattern that matches the detector is automatically recognized and classified as pertaining to the detector pattern. If the pattern is not identified by any known class, there is a classification error, that is, the pattern is considered unknown.

5.4 Manuscript Character Recognition System With Continuous Training

The manuscript character recognition system with continuous training proposed in this section is based on the pattern recognition process and training realized by the BIS as presented in 2.1.

The biological process presented is formulated and computationally reproduced using the NSA proposed in [9] and the CLONALG proposed in [10].

The NSA realizes the pattern recognition process, and the CLONALG is used as a form of training.

The pattern recognition system with continuous training is composed of six modules, where the preprocessing of the images is as shown in Figure (5): (1) censor of the NSA; (2) novelty detection module; (3) training module; (4) monitoring of the NSA; and (5) knowledge updating module. The flowchart for this algorithm is presented in Figure 7).

As shown in Figure (7), the reading of the images is done first (database). The preprocessing of the images is executed to improve the characteristics of each pattern.

The censor of the NSA is executed in an off-line process. At this time, detectors are randomly defined for the patterns. The quantity of detectors to be defined for each pattern is a parameter and depends on the system operator. The detector set of the NSA is used on the novelty detection module and monitoring of the NSA.

After the execution of the off-line module, the monitoring is executed on-line. Thus, the test set (antigens) is obtained from the processed images. This set is formed with every pattern available in the database. A pattern to be analyzed is then randomly chosen.

When a pattern is chosen to be analyzed, the novelty detection module is executed, where a comparison with the pattern under analysis and the detector set of the NSA are realized. The matching of the pattern with a detector is verified. If matching occurs, the pattern under analysis is considered known by the system, and there is no novelty. Otherwise, the pattern is considered a novelty and tagged as unknown by the system.

When the system identifies a novelty, the continuous training module is activated to learn about the new antigen (pattern under analysis), and the training process occurs, which is realized by the CLONALG [19].

The CLONALG aims to generate a memory set (knowledge) from an unknown pattern (antigen). This process follows the steps presented in section 4. Initially, a lymphocyte population is defined, which is randomly generated. A 32x32 bit matrix is created to randomly generate a lymphocyte, and it is randomly filled with binary values to generate a pattern.

Equation (4)) is used to quantify the affinity with the lymphocytes of the population (Ab) and the antigen. The n best lymphocytes with higher affinity values in relation to the antigen for the cloning and hypermutation process are selected. The quantity of clones is calculated using equation (5)) , and the quantity of mutations to be realized on the clones depends on equations (6)) - (8)) . Mutation is an evolutionary process aiming to realize little modification on the structure of the lymphocytes, such that the affinity in relation to the antigen is increased.

Mutation using this methodology is simply realized, where for each mutation, a matrix position (image) is randomly chosen and the value of the position is changed. If the position value is zero (white), it becomes black and vice versa. After maturation of the lymphocytes, the n best maturated clones

are re-selected to be introduced again. The selected lymphocytes substitute the worst lymphocytes of the population. The best lymphocytes of the population are then separated for the memory set [19]. The process is repeated until the stop criterion is satisfied. The stop criterion used in this work is when every lymphocyte of the memory set attains an affinity at a minimum of 90% with the antigen.

Figure 8 presents an illustration of the training process realized by the CLONALG training the patterns zero (0) and three (3). Figure 8 (a) illustrates the memory set on the first iteration after the patterns are randomly generated. Figure (8) (b) shows that the memory set is realized after 50 iterations. Figures (8) (c), (8) (d), and (8) (e) illustrate the memory set after 100 iterations, 150 iterations, and 200 iterations, respectively. Figure (8) (f) shows the memory set after the stop criterion is satisfied.

When the training process ended, the CLONALG presents a memory set (knowledge) in relation to the unknown antigen. This knowledge set is incorporated to the detector set of the NSA providing knowledge to the system such that in a new analysis of the same kind of antigen the system is able to perform the recognition. This process is called the continuous training [19].

When the novelty detection module does not identify any novelty, that is, when the NSA already knows the antigen, the process runs according to the conventional algorithm. Thus, the pattern under analysis must be compared with the detector set of the NSA, and the matching and the affinity to classify the pattern must be evaluated according to the class that matches.

After the pattern under analysis is classified, the system ends the monitoring process of the NSA. There is also the updating knowledge module, which evaluates if it is necessary to update the detector set of the NSA (knowledge) to improve the detection process. In this case, the criterion is the matching with the detectors and the antigen pattern. If all the detectors match the antigen pattern, it is not necessary to update the detector set of the NSA for the classified pattern.

If there are detectors of the NSA that do not match the classified pattern, the updating knowledge process occurs. When the NSA cannot recognize the pattern with every detector, it is necessary to update the detector set of the NSA to improve the recognition process and use the antigen to learn (obtain knowledge). If it is necessary to update the knowledge, the training process realized by the CLONALG is executed according to what has been described previously, which is the same as the process carried out in the continuous training module.

After the training process is realized and the memory set of the antigen pattern is generated, the best lymphocyte of the memory set must be chosen and the detector set of the NSA must be updated. The updating is carried out by substituting a detector that does not match with the antigen pattern. This detector is substituted by the best lymphocyte of the memory set generated on the training process realized by the CLONALG.
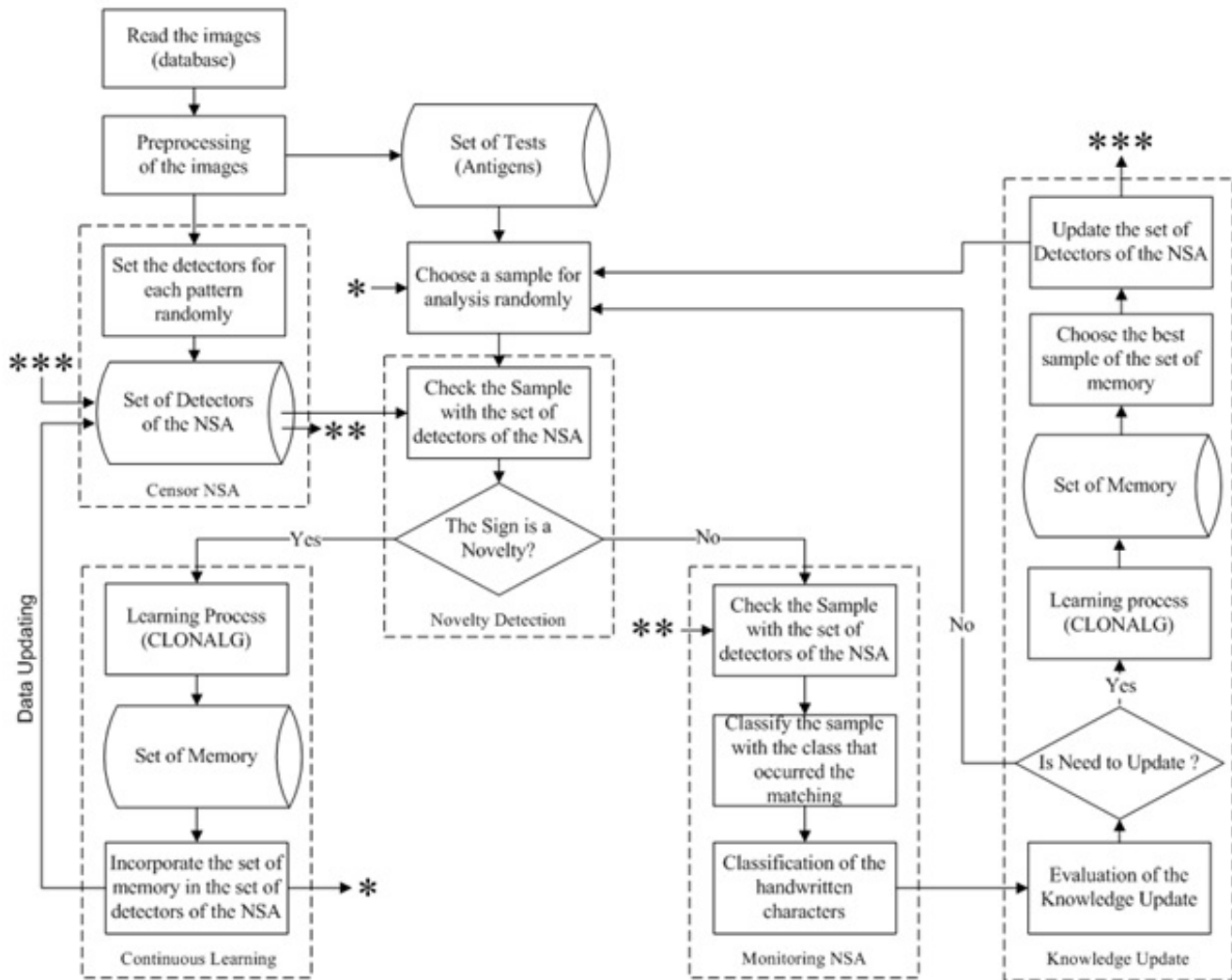
Figure 7. Pattern recognition process with continuous training.

The knowledge updating only occurs when it is necessary to reinforce the knowledge of the NSA, that is, when a detector does not satisfy the affinity criterion, which is not matching with the antigen. Thus, this detector is substituted by a new detector generated in a training process called knowledge updating.

The process illustrated in Figure (7) is repeated until every antigen is analyzed.

6. Applications and Results

To evaluate and validate the efficiency of this method, the system has been tested on handwritten character recognition, which is a classic problem in the literature.

This section presents the results obtained after applying the proposed methodology on the database of the manuscript characters. The algorithm is developed in MATLAB ®[20]. The evaluation of the methodology considers the efficiency, precision, and robustness on the recognition process of manuscript characters with the two algorithms presented.

Table 1 presents the parameters used on the algorithms with and without the continuous training.

Table 1. Parameters.

Parameters	NSA without the continuous training	NSA with the continuous training
TAf	85.00%	85.00%
N	-	15
M	-	6
R	-	9
n	-	3
β	-	0.5
ρ	-	5
d	-	0

The parameters presented in Table 1 were obtained using an empirical test.

Figure 8. Training process realized by the CLONALG.

Table 2. Results for the system without the continuous training.

Manuscript character	Tested patterns	Matched patterns	Match (%)
Zero (0)	20	20	100.00
One (1)	20	0	0.00
Two (2)	20	20	100.00
Three (3)	20	19	95.00
Four (4)	20	0	0.00
Five (5)	20	19	95.00
Six (6)	20	18	90.00
Seven (7)	20	19	95.00
Eight (8)	20	20	100.00
Nine (9)	20	0	0.00
Total	200	135	67.50

Table 3. Results for the continuous training system

Manuscript character	Tested patterns	Matched patterns	Match (%)
Zero (0)	20	20	100.00
One (1)	20	20	100.00
Two (2)	20	20	100.00
Three (3)	20	20	100.00
Four (4)	20	19	95.00
Five (5)	20	20	100.00
Six (6)	20	19	95.00
Seven (7)	20	20	100.00
Eight (8)	20	20	100.00
Nine (9)	20	19	95.00
Total	200	197	98.50

A quantity of detectors equal to 30% or 6 detectors is adopted for each pattern (digits 0 to 9) according to what is presented in [18], which is the configuration offering the best efficiency of the NSA.

The censor phase in both systems (with and without the continuous training) is realized by excluding three patterns, such as one (1), four (4), and nine (9), to evaluate the performance of the novelty detection module.

Doing this assures that the algorithms initiate the on-line monitoring process without the previous knowledge about these patterns. The option of excluding these patterns of the censor process (obtaining the knowledge) is due to the patterns that present the greatest index of errors according to [18].

According to the configurations and parameters previously presented, the pattern recognition algorithm without the continuous training obtains the results shown on Table 2 .

Table 2 shows that the recognition system cannot identify the patterns that are excluded from the censor process, that is, without previous knowledge about the patterns, the system is not able to identify these types of patterns.

Table 3 presents the results obtained with the pattern recognition system with the continuous training. The configurations and parameters are the same as those used for the conventional algorithm.

Table 3 shows that the system with the continuous training is able to learn about unknown patterns and identify them for further analysis. It is also observed that the knowledge updating module contributes in increasing the matching rate even when the system already has a previous knowledge about the pattern. This is due to the updating of the detector set of the NSA on the online monitoring, providing reinforcement and improvement on the knowledge.

Comparing Tables 2 and 3 shows that there is an improvement on detecting manuscript characters, that is, the pattern recognition with the continuous training is able to evolve, learns with experience, and begins to recognize patterns that are not previously recognized. It is also observed that the patterns excluded from the censor process are not recognized by the first system presenting 0% of matching. After using the continuous training, those patterns are identified with a matching rate of 100% for pattern one (1), 95% for pattern four (4), and 95% for pattern nine (9).

Finally, it is emphasized that it is possible to control the quality of the lymphocytes generated on the memory set of the training process realized by the CLONALG. To end the training process, it is necessary for the lymphocytes of the memory set and the antigen (pattern to be learnt) to have affinity rate of at least 90%. With this criterion, it is possible to assure that the memory lymphocytes have high affinity

values, providing efficiency to the recognition pattern process.

6.1 Comparative study

In this section presents a short comparison of the results obtained with the specialized literature in the problem of manuscript pattern recognition. The table 4 presents the comparative study.

Table 4. Comparative study.

Reference	Technique Used	Match (%)
[21]	Multi-layer Perceptron	73.61
[22]	Multi-layer Perceptron	81.80
[23]	Multiscale neural networks training	93.50
[18]	Negative Selection Algorithm	94.00
[21]	Neural Networks with Radial Bases	95.12
[8]	Online learning asymmetric	97.86
This Paper	Algorithm with continuous learning	98.50

As observed in the table 4, the proposed method presents performs well when compared to other methods available in the literature.

6.2 Advantages and disadvantages of the proposed methodology

After carrying out the tests and obtaining the results, we identified the strengths and weaknesses of the proposed methodology, these are presented in this section.

Strengths:

- The proposed methodology presented excellent performance;
- The algorithm require low processing time;
- The algorithm are robust and efficient;
- The proposed methodology learning continuously, having the capacity to innovate and update new patterns, learn with the experience, and be in constant improvement.

Weakness:

- The proposed methodology have several parameters that must be set in advance.

7. Conclusions

This work presents a new approach to develop a pattern recognition system with continuous training based on the BIS. Two artificial immune systems are used as tools in pattern recognition: the NSA and the CLONALG. The proposed methodology is applied to the recognition problem of manuscript characters, which is a classic problem in the literature on evaluating efficiency. A comparison with the conventional algorithm, that is, without the continuous training, is carried out. Three patterns are excluded from the censor process for both systems in evaluating the continuous training. The conventional algorithm presents 67.5% of matches in all patterns tested, while the algorithm with the continuous training presents 98.5% of matches in all patterns. It is emphasized that the continuous training module can continuously learn to identify unknown patterns. Efficiency, precision, and robustness of the continuous training algorithm are better than those of the conventional one. These advantages are very important for several applications. Finally, the principal advantage of this methodology in relation to others, such as the Artificial Neural Networks, is the capability to continuously improve and reinforce the knowledge, where training is permanently realized without the need to reinitiate the training process. This method is applied to solve a problem of manuscript pattern recognition. Moreover, this method can also be very useful in cases where pattern recognition is complex, such as in 3D facial expression recognition [24]. 3D facial recognition is a process whose quality of the solutions is strongly dependent on a continuous training process.

Acknowledgements

Author contributions

All authors contributed to the development of this article, highlighting the following items: The authors proposed a new artificial immunological algorithm with continuous learning for pattern recognition; The continuous training module can continuously learn to identify unknown patterns; The efficiency, precision and robustness of the continuous training algorithm are better than those of the conventional one. Finally, the method proposed by the authors is able to continuously improve and reinforce knowledge, where training is carried out permanently without the need to restart the training process. Simone S. F Souza and Fernando P. A. Lima contributed by developing the proposed method and writing the paper. Fábio R. Chavarette contributed with the guidance.

References

[1] ZADEH, L. A. Fuzzy sets. *Information and Control*, v. 8, p. 338–353, 1995.

[2] HAYKIN, S. *Neural networks: a comprehensive foundation*. USA, New Jersey: Prentice-Hall, Upper Saddle, 1994. 842 p.

[3] CASTRO, L. N. de. *Immune engineering: development and application of computational tools inspired by artificial immune systems*. Tese (Doutorado) — UNICAMP - Universidade Estadual de Campinas, Campinas, SP, Brasil, 2001.

[4] BARROS, A. C. et al. Detection and classification of the voltage disturbances in electrical distribution systems using

a neural network artmap euclidian modified with continuous training. In: *Proceedings of the Brazilian Symposium of Intelligent Automation*. Fortaleza, Brazil: [s.n.], 2013. p. 1–6.

[5] JAVED, O.; ALI, S.; SHAH, M. Online detection and classification of moving objects using progressively improving detectors. In: *Proceedings of the CVPR*. [S.l.: s.n.], 2005. p. 696–701.

[6] MARCHIORI, S. C. et al. Neural Network Based on Adaptive Resonance Theory With Continuous Training For Multi-Configuration Transient Stability Analysis of Electric Power Systems. *Applied Soft Computing*, v. 11, p. 706–715, 2011.

[7] OZA, N. C.; RUSSELL, S. Online bagging and boosting. In: *Proceedings of the 8th International Workshop on Artificial Intelligence and Statistics*. Florida, USA: [s.n.], 2001. p. 105–112.

[8] PHAM, M. T.; CHAM, T. J. Online learning asymmetric boosted classifiers for object detection. In: *Proceedings of the CVPR*. [S.l.: s.n.], 2007. p. 1–6.

[9] FORREST, S. A.; PERELSON, A. L.; CHERUKURI, R. Self-non-self discrimination in a computer. In: *Proceedings of the IEEE Symposium on Research in Security and Privacy*. Oakland: [s.n.], 1994. p. 202–212.

[10] CASTRO, L. N. de; ZUBEN, F. J. V. The clonal selection algorithm with engineering applications. In: *Proceedings of the GECCO, Workshop on Artificial Immune Systems and Their Applications*. Las Vegas: [s.n.], 1993. p. 36–39.

[11] CARPENTER, G. A. et al. Fuzzy artmap: A neural network for incremental supervised learning of analogy multidimensional maps. *IEEE Transactions on Neural Network*, v. 3, p. 689–713, 1992.

[12] CASTRO, L. N. de; TIMMIS, J. *Artificial immune systems: a new computational intelligence approach*. New York: Springer, 2002. 357 p.

[13] DASGUPTA, D. *Artificial immune systems and their applications*. New York: Springer-Verlag, 1998.

[14] LIMA, F. P. A.; LOTUFO, A. D. P.; MINUSSI, C. R. Disturbance detection for optimal database storage in electrical distribution systems using artificial immune systems with negative selection. *Electric Power System Research*, v. 109, p. 54–62, 2014.

[15] LIMA, F. P. A.; LOTUFO, A. D. P.; MINUSSI, C. R. Artificial immune systems applied to voltage disturbance diagnosis in distribution electrical systems. In: *Proceedings of the IEEE Powertech*. Grenoble, France: [s.n.], 2013. p. 1–6.

[16] BRADLEY, D. W.; TYRRELL, A. M. Immunotronics novel finite-state-machine architectures with built-in self-test using self-nonself differentiation. *IEEE Transactions on Evolutionary Computation*, v. 6, p. 227–238, 2002.

[17] FRANÇA, F. O. de; ZUBEN, F. J. V.; CASTRO, L. N. de. An artificial immune network for multimodal function

optimization on dynamic environments. In: *Proceedings of the GECCO*. Washington, DC, USA: [s.n.], 2005. p. 289–296.

[18] LIMA, F. P. A. et al. Manuscripts character recognition using artificial immune systems. In: *Proceedings of the Congress of Applied and Computational Mathematic*. Curitiba, PR, Brazil: [s.n.], 2014. p. 1–6.

[19] LIMA, F. P. A.; MINUSSI, C. R. Voltage disturbance analysis of electrical power systems based on artificial immune system with continuous learning. In: *Internal Publication, Electrical Engineering Department, UNESP, Univ Estadual Paulista, Ilha Solteira (SP)*. Ilha Solteira, SP, Brazil: [s.n.], 2014. p. 1–12.

[20] MATLAB 7.8 version. [S.l.], 2011.

[21] PERWEJ, Y.; CHATURVEDI, A. Machine recognition of hand written characters using neural networks. *International Journal of Computer Applications*, v. 14, p. 6–9, 2011.

[22] JÚNIOR, J. J. O. et al. Evaluating NN and HMM classifiers for handwritten word recognition. In: *Proceedings of the XV Brazilian Symposium on Computer Graphics and Image Processing*. [S.l.: s.n.], 2002. p. 210–217.

[23] GUANAPATHY, V.; LIEW, K. L. Handwritten characters recognition using multiscale neural networks training technique. *Word Academy of Science, Engineering and Technology*, v. 15, p. 32–37, 2008.

[24] ZHAO, W. W.; CHELLAPPA, R. ace processing: advanced modeling and methods. In: *Academic Press*. San Diego, USA: [s.n.], 2009.