

Classification of cancerous cells based on the one-class problem approach

Nabeel A. Murshed[‡], Flávio Bortolozzi[†] and Robert Sabourin^{††}

^{‡†}Centro Federal de Educação Tecnológica do Paraná (CEFET-PR) -
Pontificia Universidade Católica do Paraná (PUC-PR)

Av. Sete de Setembro 3165, Curitiba - Paraná 80230-901, Brasil.

^{‡†}email: murshed@dainf.cefetpr.br

^{††}École de Technologie Supérieure

4750 rue Henri-Julien, Montréal QC, H2T 2C8, Canada.

e-mail: sabourin@gpa.etsmtl.ca

ABSTRACT

One of the most important factors in reducing the effect of cancerous diseases is the early diagnosis, which requires a good and a robust method. With the advancement of computer technologies and digital image processing, the development of a computer-based system has become feasible. In this paper, we introduce a new approach for the detection of cancerous cells. This approach is based on the one-class problem approach, through which the classification system need only be trained with patterns of cancerous cells. This reduces the burden of the training task by about 50%. Based on this approach, a computer-based classification system is developed, based on the Fuzzy ARTMAP neural networks. Experimental results were performed using a set of 542 patterns taken from a sample of breast cancer. Results of the experiment show 98% correct identification of cancerous cells and 95% correct identification of non-cancerous cells.

Key Words

Classification of cancerous cells, two-class problem, one-class problem, Fuzzy ARTMAP Neural Networks.

1. INTRODUCTION

Cancer, the angel of death if one may call it, is considered the second cause of mortality among adults in various countries. Such a terrifying fact is due to the inability of modern medicine to find a cure for this disease or even understands its cause. However, a number of treatments exist which may reduce the severity of the cancer. The effectiveness of any cancer treatment is related to two factors: early discovery of the tumor and its severity.

Many exams are employed for the purpose of identifying a cancerous tumor. One common exam employed by most pathologists is the visual analysis of microscopic images of cells taken from the suspected tumor. Unfortunately, the high density of such microscopic images and their peculiarities make such analysis a difficult and a time consuming task. Not mentioning the possible errors that may be introduced into the results, as a consequence of human factors. In a work reported by Brug¹, three expert pathologists were asked to analyze 52 microscopic samples taken from tumors of various patients and to give their decision as to whether a particular tumor is cancerous or not. After a period of a month, the same pathologists were asked again to perform the same analysis on the same 52 samples, but, this time without telling them that they are analyzing the same samples again. From the two experiments two sets of results were collected. The first set, table 1¹, indicates the percentage of making the same prognostic on the same sample twice. For example, as it can be seen from table 1, 85% of the decisions made by pathologist 'A' were in agreement, whereas, 15% of his/her decisions were in conflict. The second set, table 2¹, indicates the percentage of correct prognostics made by the three pathologists.

As it can be seen from this table and according to Brug¹, 16% of the negative cases (non cancerous) were considered by the pathologists as being positive (cancerous); and 39% of the positive case were considered as being negative. The work reported by Brug¹ demonstrate that if a good prognostic is to be obtained, a robust analysis of the cytological images is required. Such analysis should assist the pathologist in making his/her decision with the minimum errors possible.

The work presented in this paper is not concerned with cancerous diseases nor it is concerned with introducing a new methods for curing them. The main concern of this work is the development of a computer-based classification method aimed at identifying cancerous cells. It is part of a long term research projected, initiated at the Laboratory de Sistemas Adaptativos (LSA) of the Centro Federal de Educação Tecnológica do Paraná (CEFET-PR), with the objective of developing a computer-based system for the analysis and classification of cancerous cells. In the next section we will describe our approach to the classification of cancerous cells. Sections 3 and 4 describe, respectively, the experimental results and discussion and conclusion.

Pathologist 'A'	Pathologist 'B'	Pathologist 'C'
85%	73%	70%

Table 1. Percentage of identical prognostics made on the same sample twice.

	Negative (Pathologist)	Positive (Pathologist)
Negative (Actual)	84%	16%
Positive (Actual)	39%	61%

Table 2. Percentage of correct/erroneous prognostics made by the three pathologists.

2. CLASSIFICATION OF CANCEROUS CELLS BASED ON THE ONE-CLASS PROBLEM APPROACH

2.1 Problem statement

A typical sample of a microscopic images of cells taken from cancerous tumor is shown in figure 1. The objects contained in such sample is divided into circular cells, artifacts and superposed cells. Circular cells are nucleus with quasi-circular shape, of medium size and connected contour. These cells are considered by the pathologist as being good cells for further analysis. Artifacts are particles of circular shape and relatively small size. The superposed cells are of random deformed shapes. Artifacts and superposed cells are not used for analysis. With this composition, the objective of the computer-bsead classification system is to analyze the cytological sample and separate the circular cells (cancerous cells) from the rest of objects contained in the image. Ideally, the system should recognize all the cancerous cells (Fig. 1b). However, in most practical cases a small percentage of classification error could be accepted (Fig. 1c), since the pathologist could perform a visual inspection or even use another alternative techniques.

2.2 The approach

In the field of object recognition by Neural Networks, one is faced with the following design criterion: How to build into the network a knowledge of the object(s) to be classified or recognized? For a problem similar to that of classifying cancerous cells described previously, the conventional approach to such criterion would be to form two sets of patterns: a set of patterns of the objects to be classified (positive patterns) and a set of patterns of other objects to be rejected (negative patterns), and then to train the network with patterns from both sets. This approach is conventionally known as a two-class problem approach.

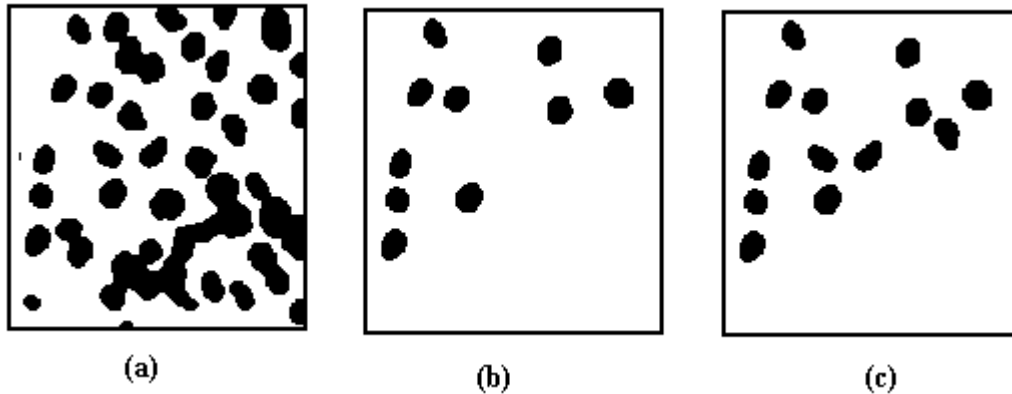


Figure 1. A typical sample of microscopic image of a breast cancer. a) Binary image obtained by applying Otsu's algorithm on the original gray-level image. b) The ideal result of the classification process. c) The practical case.

Let's assume that a child is to be trained to recognize apples. Following the two-class problem approach, one must form a training set that consists of two classes: a class of apples, ω_1 , and a class of other types of fruits (bananas, oranges, pears, etc.), ω_2 . Then, during training, the child will be shown some *prototypes from both classes* and will be told which prototype is an apple and which one is not an apple. Based on this approach, the child will *learn* the features in the class of apples (color, shape, etc.) as well as those of all other fruits in class ω_2 and will encode these information into his/her memory. When training has been successfully completed, the child will have had *knowledge of the features belonging to both classes*. Alternatively, the child may be trained with prototypes of apples only. In such a case, the training set will contain *prototypes from one class only*, the class of apples. During training, the child will be shown those prototypes and will be told that they are apples. Consequently, the child will *learn* the features belonging to the class of *apples only*. When training has been successfully completed, the child will have had *features belonging to one class only*.

In both approaches the child will produce the correct answer. But as far as the objective is concerned, one may ask: if the child is required to recognize apples, then why he/she needs to be trained to recognize other types of fruit. Certainly it will be more interesting if the child responds:

This fruit is not an apple but is a banana, an orange, or a pear.

But, this is an extra information that it is not needed.

When the child is shown an apple and is told that this is an apple, he/she associates the name 'apple' with the features of the apple (shape, color, etc.). These features provoke a certain neuronal activity patterns which are

associated with the word apple. Likewise, the features belonging to every other fruits provoke a neuronal activity patterns which are associated with the phrase 'not an apple'. In both cases the activity patterns along with their respective association are stored in memory. In the future when the child is presented with any given fruit, the features contained in this fruit will provoke an activity patterns which in turn will retrieve the correct association. This feature of our cognitive process, among others, permit us to recognize a class of patterns without the necessity of being trained with prototypes from other classes.

By the same analogy, one may argue that: if the objective of the pattern recognition system for the classification of cancerous cells is to decide whether a certain cell is cancerous or not cancerous, then the system should be trained with patterns of cancerous cells only. This approach is called a **one-class problem approach**² and is stated bellow:

"If the objective of the pattern recognition system is to recognize a certain class of similar objects, then the pattern recognition system should be trained with prototypes from this class only".

To implement a computer-based classification method, based on one-class problem approach, we've used the Fuzzy ARTMAP neural network as the classifier network. This is justified by the ability of the Fuzzy ARTMAP network to perform a comparison process prior to making its decision and to perform pattern association in way similar to that of the human being. These two features are necessary for implementing the one-class problem approach. The next subsection gives a brief description of the Fuzzy ARTMAP network .

2.3 Brief description of the Fuzzy ARTMAP neural network ³

The Fuzzy ARTMAP is a supervised neural network developed by Stephen Grossberg and his colleagues of Boston University. It incorporates two Fuzzy ART modules, ART_a and ART_b, linked together by a map field E^{ab} network, as illustrated in figure 2³, and is capable of associating an arbitrary set of input patterns with their corresponding classification patterns. The number of neurons in the map field (predictive neurons) equals the number of exemplar neurons in the ART_b module and are connected to them through non-adaptive 1-to-1 connections. The purpose of the predictive neurons is to connect two exemplar neurons, one in each module, to each other, if these neurons make the same prediction at the map field. During training, the network receives an arbitrary input pattern **I**^a and its correct association **I**^b. The objective of the network, thereafter, is to learn this association without any intervention from the outside world.

At each Fuzzy ART modules, ART_a and ART_b, categorical classes are formed at the output layers E^a and E^b, each of which is represented by an exemplar neuron in the respective layer. The number of categorical classes depends on the similarity between an input pattern and a category exemplar. The weight vectors, **w**_J^a and **w**_J^b, represent, respectively, the exemplar of each category in E_a and E_b. The similarity between a category exemplar and an input pattern, in each module, is determined by the following equation:

$$\frac{|\mathbf{I} \cap \mathbf{w}_J|}{\mathbf{I}} \geq \rho \quad (1)$$

During training when correct prediction occurs at the map field, the weight vector **w**_J^{ab} stores this information by locking the value of the weight connection w_J^{ab} , between an exemplar neuron in ART_a and a predictive neuron in the map field, at its current value, that is, $w_J^{ab} = 1$. This value stays for ever. If on the other hand, the exemplar neurons at each module activate two distinct prediction neurons at the map field, then

prediction is disconfirmed. As a consequence, an internal controller is activated to correct the prediction error made by ART_a at the map field. The internal controller, thereafter, increases the value of the ART_a vigilance parameter ρ^a by the minimum amount needed to select another previously learned exemplar neuron J . If no such exemplar neuron exists, then another uncommitted exemplar neuron will be assigned to represent the new pattern.

During test trial, and normal operation, input patterns I^a are presented without I^b , thereafter, an exemplar neuron is activated and a prediction is made at E^{ab} . The output x^{ab} of the map field is then taken as the predicted pattern, and is governed by the following equation⁴:

$$x^{ab} = \begin{cases} y^b \wedge w_j^{ab}, & \text{if the } J\text{th } E^a \text{ exemplar neuron is active and } E^b \text{ is active} \\ w_j^{ab}, & \text{if the } J\text{th } E^a \text{ exemplar neuron is active and } E^b \text{ is inactive} \\ y^b, & \text{if } E^a \text{ is inactive and } E^b \text{ is active} \\ 0, & \text{if } E^a \text{ is inactive and } E^b \text{ is inactive} \end{cases} \quad (2)$$

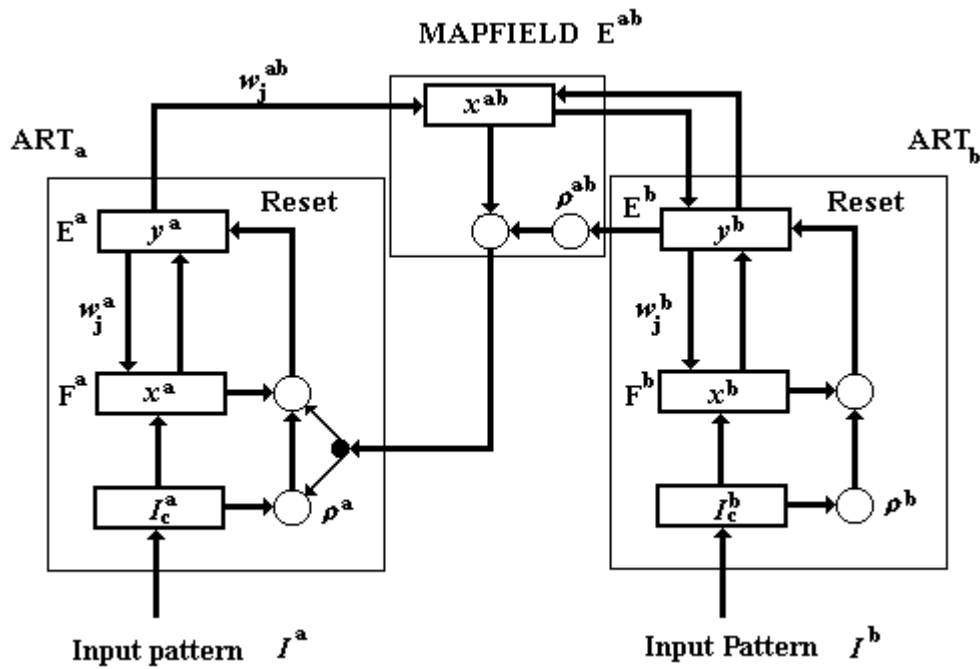


Figure 2. Architecture of the Fuzzy ARTMAP neural network³

It can be observed from equation 2 that if the Fuzzy ARTMAP is trained with patterns from one class only along with the respective association, 1 for example, then the network will learn this association for each pattern pair that are presented during the training process. During test, or normal operation, when an unknown pattern is presented to the network, the output, x^{ab} , of the map field would be either equal to the weight vector, w_j^{ab} , if the unknown input had activated a previously learned exemplar neuron in the ART_a module, or equal to 0, if the input exemplar had not activated any previously learned exemplar neuron. This feature of the Fuzzy ARTMAP neural network makes it ideal for implementing the one-class problem approach, as mentioned previously.

3. EXPERIMENTAL RESULTS

3.1 Definition of the experimental data

To evaluate the performance of the proposed classification method, two sets of data were collected from tumor of breast cancer. The first set contained 383 cancerous cells only, and the second set contained 159 cells, 54 of which is cancerous cells and 105 is composed of superposed and deformed cells. The first set were used for training only and the second set were used for evaluation. The artifacts, the circular cells of smaller size, were removed from the samples from the samples by applying a simple separation process based on the area. In this regards, objects of area bigger than 350 pixels were considered for the classification process.

From the set of selected objects (cancerous, superposed and deformed cells), 7 parameters were extracted. These parameters are: perimeter, area, factor of compactness, minimum radius, maximum radius, circularity and diameter. For the sake of clarity, the seven parameters will be labeled, respectively, as $x_0, x_1, x_2, x_3, x_4, x_5, x_6$. Typical values of these parameters from both data sets are listed in table 3.

x_0	x_1	x_2	x_3	x_4	x_5	x_6	class lable	
99.59800	655.00000	1.20517	9.21954	19.10497	3	0.48257	739.08837	1
103.74015	768.00000	1.11512	10.77032	20.09975		0.53584	866.59521	1
81.66907	508.00000	1.04482	10.19803	14.56022		0.70040	573.21661	1
80.08328	400.00000	1.27589	5.38516	16.15549		0.33333	451.35165	0
123.29651	896.00000	1.35015	8.54400	23.70653		0.36040	1011.02771	0

Table 3. Typical values of the parameters used in the experiment. The class labels 1 and 0 indicate, respectively, cancerous cells and superposed/deformed cells.

3.2 Parameter selection

A selection process was performed prior to the final experiment, the objective of which was to determine the most optimum parameters that would produce a better results. The selection process was performed by training and evaluating six Fuzzy ARTMAP networks, each with different set of parameters. The sets of parameters were: $x_0x_4, x_1x_6, x_2x_3, x_2x_5, x_2x_3x_5$ and $x_0x_2x_3x_4x_5$. Due to the high numerical differences between the parameters, such combination were necessary in order to make the input space as homogeneous as possible. The training and test sets for the parameter selection process were taken from the total data set as illustrated in figure 4.

The results of the evaluation for the selection process, in terms of the False Acceptance Rate (FAR) and False Rejection Rate (FRR), are shown in table 4. These errors describe, respectively, the percentage of negative cases accepted by the network as being cancerous cells, and the percentage of positive cases accepted as being non cancerous cells. Each value in the last raw is the average of the two errors. The first raw indicates the set of parameters used for training and evaluation. From table 4, it can be observed that the combination of parameters $x_2x_3, x_2x_3x_5$ and $x_0x_2x_3x_4x_5$ produced the lowest errors. These combinations were then used for the final experiment.

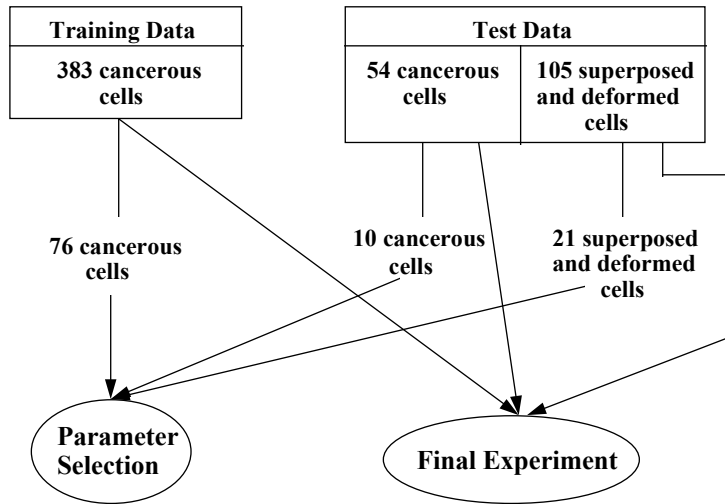


Figure 4. Definition of the experimental data.

	x_0x_4	x_1x_6	x_2x_3	x_2x_5	$x_2x_3x_5$	$x_0x_2x_3x_4x_5$
FAR (%)	14.29	33.33	4.76	14.29	4.76	4.76
FRR (%)	30.00	20.00	0.00	20.00	10.00	10.00
Et (%)	22.14	26.66	2.38	17.14	7.38	7.38

Table 4. Evaluation results for the selection process.

3.3 The final experiment

The structure of the final system is depicted in figure 5. Each Fuzzy ARTMAP network (FA) is trained and evaluated with one combination of parameters. The network parameters in both experiments were: $\beta = 1$, $\alpha = 0.01$ and $\rho = 0.99$. The final decision of the system is based on the majority decision rule stated as follows:

- ◆ Consider an unknown cell as being a cancerous cell if at least two of the three networks made the same positive decision and as being a non-cancerous, otherwise.

The intermediate classification results, made by each Fuzzy ARTMAP, are shown in table 5 and the final classification of, after the application of the majority decision rule are shown in table 6.

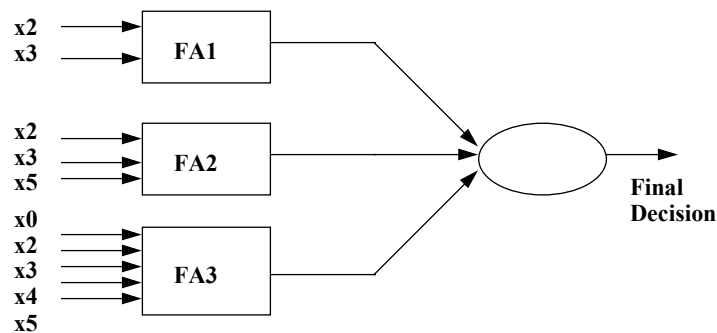


Figure 5. Architecture of the system for the classification of cancerous cells

	x_2x_3	$x_2x_3x_5$	$x_0x_2x_3x_4x_5$
FAR (%)	18.00	15.00	12.00
FRR (%)	3.70	7.41	5.55
Et (%)	10.85	11.21	8.78

Table 5. Intermediate classification results.

FAR (%)	FRR (%)	Et (%)
5.00	1.85	3.42

Table 6. Final classification results.

4. DISCUSSION AND CONCLUSION

In this paper we've presented a computer-based method for the classification of cancerous cells. The method is based on the one-class problem approach, for which the classifier network need be trained with patterns of cancerous cells only. The one-class problem approach was implemented using the Fuzzy ARTMAP neural network. This decision is justified by the ability of this network to classify patterns of the same class without being trained with patterns from other classes. The experimental results were performed with a data set of 542 patterns taken from a sample of breast cancer. For the purpose of classification, a set of seven parameters were used to represent each pattern and to form the input space. The input space was then divided into homogeneous subspaces. A selection process was then performed to select the most the optimum subspace(s) for performing the final classification. The final architecture of the classification was based on three Fuzzy ARTMAP networks, each of which performs the classification with different combinations of the features. Each Fuzzy ARTMAP network was then trained and tested separately using patterns of cancerous cells, for training, and patterns of both cancerous and superposed/deformed cells for test. All the experiments were performed with the Neural Works Simulator running on an IBM DX4 compatible PC. The final results of the system was then based on a majority decision rule, in which a given input pattern is classified as being a cancerous cell if at least two of the Fuzzy ARTMAP networks have agreed upon such decision, or as being a non-cancerous, otherwise.

As it can be observed from table 6, the final results are very promising. Though, the false acceptance error are somewhat higher than the false rejection. This error can be minimized by investigating the use of a better set of parameters and/or a better method for selecting the optimum parameters. The final decision of the acceptance of the error rates will strongly depend on the pathologist himself/herself.

5. REFERENCES

1. G. Brug. "Pattern recognition , image processing, related data analysis and expert systems integrated in medical microscopy," *Proc. Int. Conf. Pattern Recognition*, pp.286-293. 1988.
2. N. A. Murshed. "A natural approach to signature verification". **Master thesis**. Centro Federal de Educação Tecnológica do Paraná (CEFET-PR). 1995.
3. G. A. Carpenter, S. Grossberg, N. Markuzon, and J. H. Reynolds. "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps". *IEEE Tran. Neural Networks*. Vol. 3, No. 5, pp. 698-713, 1992.
4. N. A. Murshed, F. Bortolozzi and R. Sabourin, "Off-line signature verification, without a priori knowledge of class w2. A new approach," *Proc. Int. Conf. Image Analysis and Recognition*, Vol. I, pp. 191-196, Montreal, 1995.
5. N. Otsu. "A threshold selection method from gray-level histograms". *IEEE Trans. Syst. Man. Cybernetics*, Vol. SMC-9, NO. 1, pp. 62-66, 1979.
6. S. B. Filho, " Um quantificador da ploidia tumoral através da citofotometria". **Master thesis**. Centro Federal de Educação Tecnológica do Paraná (CEFET-PR). 1994.