

A Fuzzy ARTMAP-Based Classification System for Detecting 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) -
Pontifícia Universidade Católica do Paraná (PUC-PR)
Paraná-Brasil

email: murshed@dainf.cefetpr.br

^{††}École de Technologie Supérieure, Montréal-Canada.

Abstract. *This work investigates the use of Fuzzy ARTMAP neural network for detecting cancerous cells, based on the one-class problem approach. This approach is inspired by the way human beings perform pattern recognition. We all know that children and adults alike are capable of detecting patterns belonging to a certain class, by learning the features of these patterns only. Moreover, a child or an adult is capable of detecting an unknown pattern belonging to another class without an a priori knowledge of the features in these patterns. Based on this approach, a Fuzzy ARTMAPs-based system is developed for detecting cancerous cells by training the Fuzzy ARTMAPs with the features belonging to the class of cancerous cells only. This is different from the two-class problem approach which requires that the classifier must be trained with features from the class of cancerous cells and the class of non-cancerous cells. Experimental analysis were conducted using a set of 542 patterns taken from a sample of breast cancer. Training was performed with 383 cancerous cells. System performance was evaluated using 54 cancerous cells and 159 non-cancerous cells. Evaluation results show 98% correct identification of cancerous cells and 95% correct identification of non-cancerous cells.*

Key Words. Classification of cancerous cells, one-class problem, two-class problem, Fuzzy ARTMAP Neural Network.

1 Introduction

It is known that cancer is considered as the second cause of mortality among adults in various countries. As an example, according to the 1987 report of the American Cancer Society[1]:

" Breast cancer is still the most common form of cancer among American women, accounting for more than 30% of all cancers

in women. One in nine American women will develop breast cancer, about 180,000 this year alone. This year, about 46,000 American women will die because of breast cancer. Only lung cancer causes more cancer deaths among women. Breast cancer is mainly thought of as a woman's disease, but men can also develop breast cancer. About 300 men die of the disease each year. "

Such terrifying facts are due to the inability of modern medicine to find a cure for this disease or even understands its cause. However, according to the American Cancer Society: " the number of deaths due to breast cancer could be reduced if more women had mammograms as recommended beginning at age 40, had a yearly breast exam by a physician..." These precautions, according to the report, will help in discovering the cancer in its earliest, most curable stages. In the case of other types of cancers, microscopic analysis of tumors tissues is performed. In the mammogram or microscopic analysis, the pathologist visually inspects the images under test. The high density of some 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 Brug[2], 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 was cancerous or not. After a period of a month, the same pathologists were asked again to perform the same analyses on the same 52 samples, but, this time without telling them that they were analyzing the same samples again. From the two experiments two sets of results were collected. The first set, table 1[2], 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 2[2], indicates the percentage of correct prognostics made by the three

pathologists. As it can be seen from this table and according to Brug[2], 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 Brug[2] demonstrates that if a good prognostic is to be obtained, a robust analyses of the cytological images is required. Such robustness could be obtained by a computer-based classification system.

In this work we propose a Fuzzy ARTMAP-Based Classification System for detecting cancerous cells, based on the one-class problem approach. It is a part of a long term research project aimed at developing a computer-based system for the analysis and classification of cancerous cells.

In the next two sections we will describe the problem of detecting cancerous cells and how to solve it by the one-class problem approach.

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

Table 1. Percentage of identical prognostics made on the same sample twice[2].

	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].

2 Problem Definition

A typical sample of microscopic image of cancerous cells is shown in figure 1. The objects contained in this sample are 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

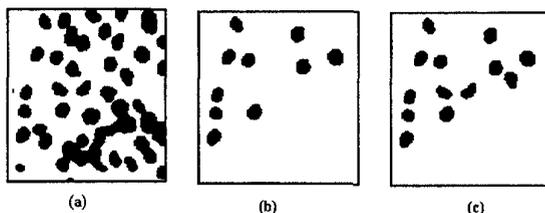


Figure 1. Example of microscopic binary images of breast cancer. a) Image to be analyzed by the classification system. b) Ideal classification results. c) Practical classification results.

and relatively small size. The superposed cells are of random deformed shapes. These Artifacts and superposed cells are not used for any analysis, so they can be eliminated from the sample.

With this composition, the objective of a computer-based classification system is to analyze the cytological sample and *identifies cancerous cells* (circular objects). Ideally, the system should be able to identify all 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 visual inspection or even use other alternative techniques.

3 Detecting Cancerous Cells Based One-Class Problem Approach

When developing a NN-based classification system, 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 identifying cancerous cells, as described above, one would adopt the two-class problem approach. This requires that the neural network be *trained* with patterns from two classes: the class of positive patterns (e.g. cancerous cells) and the class of negative patterns (e.g. non-cancerous cells). Negative patterns are included in the training data to teach the network not to confuse cancerous cells with non-cancerous cells. It is obvious that, when adopting the two-class problem approach and in order to achieve good classification results, one must obtain enough samples from both classes. The more sample we have, the better will be our knowledge of the patterns in both classes and, consequently, the better will be our design. Perfect! But, is it always possible to obtain enough samples of *negative patterns for training*? In our opinion, this is not possible. For example in the case of developing a computer-based signature verification system, obtaining a *good quantity* of signature forgeries for training is impossible. Another example is the case of developing a computer-based inspection system for detecting defects in biscuits. In such application, the defect is random and rare and, thus, obtaining enough samples of defect biscuits is not possible. How one would solve this problem? One possible solution is to use computer-generated patterns.

Another solution is to imitate the human approach to pattern recognition. We all know that children and adults alike are capable of identifying positive patterns (e.g., airplanes) without being taught with negative patterns (e.g., cars, building, etc.). Moreover, a child or an adult is capable of identifying a negative pattern (e.g., cars) as *not being an airplane* without an *a priori* knowledge of the class of cars. A child or an adult, of course, could be trained to recognize airplanes and cars as well. But, if the objective

is to recognize the class of airplanes, then there is no need to train the child or the adult with negative patterns. This characteristics of human being have given rise to the concept of *one-class problem approach* stated as[7]:

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

As an example, consider the case of developing a NN-based classifier to identify the letter 'A'. Based on the one-class problem approach, the task of building into the network, a knowledge of the letter 'A', becomes that of training the network with patterns from the letter 'A' only. After training, the network should be able to use the acquired knowledge to identify an unknown pattern as being an 'A' (positive patterns) or not an 'A' (negative patterns). It should be noted that when detecting negative pattern the network is not required to identify its type. If the network is required to do so type, then we have a two-class problem and the approach presented above is not valid.

The advantages of using the one-class problem approach is that network training is accomplished in a time shorter than that when using the two-class problem approach. This is obvious since the network, in the one-class approach, need to learn the features in one class only. Whereas, in the two class problem approach the network need to learn the features in two classes. The training phase could even get complicated when dealing with complex and large data sets. Another advantage of the one-class problem approach is that the system can be designed to operate in real time, when using the appropriate neural network.

To develop a computer-based classification system for detecting cancerous cells based on the one-class problem approach, we have used the Fuzzy ARTMAP neural network as the classifier network. This is justified by the ability of the Fuzzy ARTMAP to detect cancerous and non-cancerous cells based on its knowledge of cancerous cells only. This is because the Fuzzy performs a comparison process prior to making its decision, which is somewhat similar to that of human being.

To illustrate this feature of the Fuzzy ARTMAP, the following section gives a brief description of this network. A thorough description of the Fuzzy ARTMAP can be found in Carpenter et al[3]. Section 5 describes the system architecture and the experimental protocol. Discussion and conclusions are given in section 6.

4 The Fuzzy ARTMAP Neural Network

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 mapfield E^{ab} network, as illustrated in figure 2. The objective of the Fuzzy ARTMAP is to associate an arbitrary input pattern I^a with its correct association I^b. During training the network learns this association and encodes it in the weight vector w_{j^{ab}}. During test, and normal operation, input patterns I^a is presented without I^b, the network then predicts its association. The output x^{ab} of the mapfield is then taken as the predicted pattern, and is governed by the following equation[3]:

$$x^{ab} = \begin{cases} y^b \wedge w_j^{ab}, & \text{if the } j\text{th } E^a \text{ exemplar neuron is active and } E^b \text{ is active} \\ w_j^{ab}, & \text{if the } j\text{th } E^a \text{ exemplar neuron is active and } E^b \text{ is inactive} \\ y^b, & \text{if } E^a \text{ is inactive and } E^b \text{ is active} \\ 0, & \text{if } E^a \text{ is inactive and } E^b \text{ is inactive} \end{cases} \quad (1)$$

It can be observed from equation 1 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 mapfield would be either equal to the weight vector, w_{j^{ab}}, if the unknown input is found to be similar to a previously learned pattern, or equal to 0 otherwise. This feature of the Fuzzy ARTMAP neural network makes it ideal for implementing the one-class problem approach, as mentioned previously. Another interesting feature of the Fuzzy ARTMAP is the match tracking controller through which the network corrects any prediction error made by ART_a at the mapfield. This is accomplished by autonomously increasing the vigilance parameter ρ_a in search for another learned category exemplar.

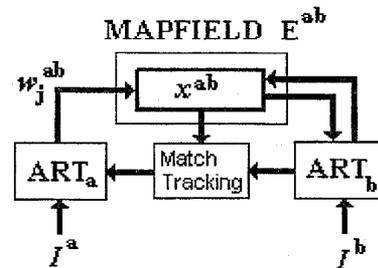


Figure 2. Block diagram of the Fuzzy ARTMAP NN.

5 System Description

5.1 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 was used for training and the second set was used for evaluation. The artifacts, the circular cells of smaller size, were removed from the samples by applying a simple separation process based on the area. That is, only objects of area bigger than 350 pixels were considered for the classification process.

5.2 Feature Vector

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. The equations for calculating these parameters from binary images can be found in [4]. 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 are: (99.59800, 655.00000, 1.20517, 9.21954, 19.104973, 0.48257, 739.08837) for the class of cancerous cells and (123.29651, 896.00000, 1.35015, 8.54400, 23.70653, 0.36040, 1011.02771) for the class of non-cancerous cells. As it can be observed from these samples, the feature space is not homogeneous. Hence, if the entire feature vector is applied to the network, the high-value features will have a stronger effect on NN learning than the small-value features will, even when a normalization process is performed. To render the feature space homogeneous, the feature vector was divided into 6 subvectors (FS0, FS1, FS2, FS3, FS4, FS5) each of which contains a different set of features. These sets are: x_0x_4 , x_1x_6 , x_2x_3 , x_2x_5 , $x_2x_3x_5$ and $x_0x_2x_3x_4x_5$. The division criteria was based on the similarity between feature values.

5.3 Selecting the Optimum Feature Subvectors

To determine the optimum feature subvectors that produces the best results, a selection process was performed by training and evaluating six Fuzzy ARTMAP networks, each with one feature subvector. The training and test sets were taken from the total data set as illustrated in table 3.

The results of the selection process for each sub-vector, in terms of the False Acceptance Rate (FAR) and False Rejection Rate (FRR), are shown in table 4. These errors

	Training Data 383 CC	Test Data 54 CC + 105 NCC
Selection Process	76	10 CC + 21 NCC
Final Experiment	383	54 CC + 105 NCC

Table 3. Data set description. CC and NCC indicate, respectively, cancerous cells and non-cancerous cells.

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 row is the average of the two errors. From table 4, it can be observed that the feature subvectors (FS2, FS4 and FS5) produced the lowest errors. These subvectors were then used for the final experiment.

	FS0	FS1	FS2	FS3	FS4	FS5
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.

5.4 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 feature subvector, as illustrated in figure 5. The network parameters in both experiments were: $\beta = 1$, $\alpha = 0.01$ and $\rho = 0.99$. These parameters were determined during the selection process. The final decision of the system is based on the majority decision rule stated as follows:

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

The intermediate and final classification results made, respectively, by each Fuzzy ARTMAP and by the system (after applying the majority decision rule), are shown in tables 5 and 6.

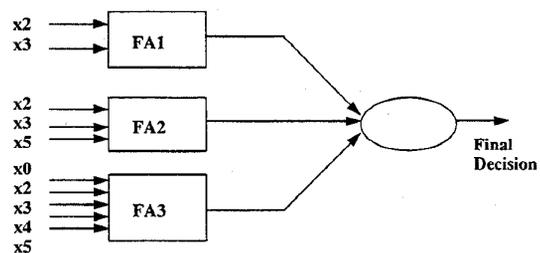


Figure 5. Architecture of the final classification system.

	FS2	FS4	FS5
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 after applying the majority decision rule.

6 Discussion and Conclusions

In this paper we've presented a Fuzzy ARTMAPs-based system for detecting cancerous cells. The approach to network training, in the presented system, was based on the one-class problem approach, in which the Fuzzy ARTMAPs were trained with patterns of cancerous cells only. This Fuzzy ARTMAP was selected because of its ability to identify positive and negative patterns (e.g., cancerous and non-cancerous cells) without being trained with negative patterns (non-cancerous cells). This is similar to the way human beings perform pattern recognition. Experiments were performed with a data set of 542 patterns taken from a sample of breast cancer. A set of seven parameters were selected to form the feature vector. Due to the high numerical differences between these parameters, the feature vector was divided into subvectors, each of which consisted of a set of parameters. A selection process was then performed to select the most optimum subvectors. The final architecture of the classification was based on three Fuzzy ARTMAPs, each of which was responsible for one set of parameters. Network training and testing was performed according to the data definition shown in table 3. 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 ARTMAPs have agreed upon such a decision, or as being a non-cancerous, otherwise. The intermediate and final results are shown in tables 5 and 6, respectively. For space limitations, we've decided in this work not to comment on the pre-processing techniques that were used to obtain the binary images shown in figure 1. Details of these techniques are given in Filho[5].

All the experiments were performed with the Neural Works Simulator running on an IBM DX4 compatible PC. We've found this simulator very practical for simulation purposes. However, it has some disadvantages. First it lacks some tools for visualizing network internal representation which is very important when dealing with images. Another disadvantage, is the lack of a timer module which permits the user to compare various networks, in terms of training and recall times. A third disadvantage, is related to the size of the generated C-code of a trained network. For a large network, the generated C-code is very huge and can not be compiled. It would be interesting if the simulator saves the connection weight into a file and generates the C-code to access this file. As with regard to using the Fuzzy ARTMAP, we've found it very flexible and easy to use.

6.1 Analysis of the Results

As it can be observed from table 5, the Fuzzy ARTMAPs performed reasonably well in identifying the

cancerous cells and rejecting the non-cancerous cells. The errors can be minimized by investigating the use of a better set of parameters and/or a better method for selecting the optimum parameters. However, the use of the majority decision rule is proved to be very effective, as it can be seen from table 6. This is because one set of parameters may not succeed to identify a particular cell, whereas, the other two sets may succeed.

6.2 Conclusions

From the results shown in tables 5 and 6, we may conclude that the one-class problem approach is very effective when applied to a pattern recognition problem similar to that of identifying cancerous cells; and that the Fuzzy ARTMAP neural network is an appropriate choice to implement this approach. However, we are very well aware that the system is still far from being perfect. Our future work will focus on obtaining a large database, optimizing the parameters selection process and comparing our results to that of some expert pathologists and of other systems.

7. References

- [1] The American Cancer Society. " Facts on breast cancer ", 1987.
- [2] G. Brug. "Pattern recognition , image processing, related data analyses and expert systems integrated in medical microscopy, " *Proc. Int. Conf. Pattern Recognition*, pp.286-293. 1988.
- [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] M. Coster and Y.L Chermant. *Precis d'analyse d'images*. 1. ed. França: Presses du CNRS, 1989.
- [5] S. B. Filho, " Um quantificador da ploidia tumoral através da citofotometria". **Master thesis**. Centro Federal de Educação Tecnológica do Paraná (CEFET-PR). 1994.
- [6] F. Giroud. "Cell nucleus pattern analysis : geometric and densitometric featuring, automatic cell phase identification". *Biol.Cell*, 44, p. 177 - 188, 1982.
- [7] N. A. Murshed. "A natural approach to signature verification". **Master thesis**. Centro Federal de Educação Tecnológica do Paraná (CEFET-PR). 1995.
- [8] 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 Analyses and Recognition*, Vol. I, pp. 191-196, Montreal, 1995.