Accurate and Fast Off and Online Fuzzy ARTMAP-Based Image Classification With Application to Genetic Abnormality Diagnosis

Boaz Vigdor and Boaz Lerner

***Abstract*—We propose and investigate the fuzzy ARTMAP neural network in off and online classification of fluorescence in situ hybridization image signals enabling clinical diagnosis of numerical genetic abnormalities. We evaluate the classification task (detecting a several abnormalities separately or simultaneously), classifier paradigm (monolithic or hierarchical), ordering strategy for the training patterns (averaging or voting), training mode (for one epoch, with validation or until completion) and model sensitivity to parameters. We find the fuzzy ARTMAP accurate in accomplishing both tasks requiring only very few training epochs. Also, selecting a training ordering by voting is more precise than if averaging over orderings. If trained for only one epoch, the fuzzy ARTMAP provides fast, yet stable and accurate learning as well as insensitivity to model complexity. Early stop of training using a validation set reduces the fuzzy ARTMAP complexity as for other machine learning models but cannot improve accuracy beyond that achieved when training is completed. Compared to other machine learning models, the fuzzy ARTMAP does not loose but gain accuracy when overtrained, although increasing its number of categories. Learned incrementally, the fuzzy ARTMAP reaches its ultimate accuracy very fast obtaining most of its data representation capability and accuracy by using only a few examples. Finally, the fuzzy ARTMAP accuracy for this domain is comparable with those of the multilayer perceptron and support vector machine and superior to those of the naive Bayesian and linear classifiers.**

***Index Terms*—Fluorescence *in situ* hybridization (FISH), fuzzy ARTMAP neural network (NN), genetic abnormality diagnosis, image classification, off- and online learning.**

I. Introduction

In RECENT years, fluorescence *in situ* hybridization (FISH) has proved itself as a useful tool in the analysis of human chromosomes for clinical diagnosis and genetic research [1][2]. In contrast to conventional cytogenetics, FISH enables the detection of chromosomal numerical abnormalities during cell interphase. This analysis is vital in clinical applications requiring rapid results, where it is difficult to culture cells in vitro or stimulate cell proliferation. One of the most common applications of FISH is the enumeration of signals representing the inspected chromosomes within a population of cells, in a procedure known as signal (dot or spot) counting. Dot counting is used for detection and analysis of chromosomal numerical aberrations in e.g., prenatal inspection, diagnosis of tumors and haematopoietic neoplasia as well as for demonstration of disease-related chromosomal translocation [1]–[3]. However, one major limitation of the FISH technique in dot counting is that a large number of cells are needed to be scanned in order to get accurate estimation of chromosome distribution over cell population. Manual evaluation of cells by cytogenetic experts is laborious and time consuming, therefore it is only but natural to pursuit the automation of dot counting.

Previous research [4] proposed performing automatic dot counting based on signals classified as valid whereas signals classified as artifacts are removed from the analysis. This approach required the extraction of well-discriminating signal features and establishment of a highly accurate classifier. The research also studied different feature representations of FISH signals, as well as ways to select the most discriminative features appropriate to signal classification.

Extending the previous research to meet a broader range of requirements, we are interested in this investigation in finding and optimizing a classifier of signals demonstrating two genetic diseases, Down syndrome (trisomy 21, i.e., an excess of one copy of chromosome 21) and Patau syndrome (trisomy 13, i.e., an extra copy of chromosome 13). Since we wish the classifier to be adaptable to different cytogenetic applications, procedures and laboratories, as well as capable to discriminate high-dimensional signal representations accurately, we concentrate on neural networks (NNs) [5] that excel in classification problems. An earlier study [6] demonstrated the accuracy of a multilayer perceptron (MLP) NN in FISH signal classification. However, the MLP model required both large sample size and extended period for training. Frequently, these requirements are difficult to achieve in routine clinical inspection, especially in small laboratories and when the laboratory is required to provide a broad range of services in different applications or under time pressure. Therefore, we are interested here in providing the classifier the ability to learn rapidly, sometimes utilizing only a few examples. We would also like the classifier to have both possibilities of incremental (online) and batch (offline) learning in order to address all requirements of the cytogenetic laboratory, its natural growth in activity and the increasing number of FISH applications. Incremental learning may be also beneficial in order to accommodate any changes which have been accumulated not necessarily from the outset but following an initial period for which training was already completed successfully. These

Manuscript received April 23, 2005; revised November 17, 2005. This work was supported in part by the Paul Ivanier Center for Robotics and Production Management, Ben-Gurion University, Beer-Sheva, Israel.

The authors are with the Pattern Analysis and Machine Learning Laboratory, Department of Electrical and Computer Engineering Ben-Gurion University, Beer-Sheva 84105, Israel (e-mail: boaz@ee.bgu.ac.il).

Color versions of Figs. 2 and 4 are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TNN.2006.877532

changes may not justify retraining the already trained classifier on the whole of the data but only modifying classifier parameters based on recent examples. These qualities are advantageous for attaining fast pattern classification at any time and to assure plasticity in learning patterns of novel classes while keeping stability in recognizing patterns belonging to familiar classes. Finally, accomplishing all theses objectives should not come at the expense of achieving high classification accuracy.

A model that might achieve these objectives, as it is considered as one of the leading offline and incremental learning models to classification, is the fuzzy ARTMAP NN [7]. The fuzzy ARTMAP does extremely well in fast incremental supervised learning in a non-stationary environment. Moreover, the fuzzy ARTMAP allows learning new data without forgetting past data (tackling the so-called "plasticity-stability dilemma" [8]), which is crucial for achieving incremental learning. It is not always clear while experimenting with the MLP NN how to select a model (i.e., decide on the numbers of hidden layers and hidden neurons in each layer) with respect to the growing complexity of the data. The fuzzy ARTMAP architecture however expands its complexity following increase in data complexity and selects a model automatically during training. The fuzzy ARTMAP and its variants have been found to be accurate, fast learners as exemplified in performing various classification tasks, such as automatic target recognition based on radar range profiles [9], three-dimensional object understanding and prediction from a series of two-dimensional views [10], QRS-wave recognition [11], speaker-independent vowel recognition [12], medical diagnosis of breast cancer and heart disease [13], online handwritten recognition [14], classification of noisy signals [15], and discrimination of alcoholics from nonalcoholics [16].

In this work, we explore the fuzzy ARTMAP in off and online FISH signal classification allowing the automation of dot counting for genetic diagnosis of numerical abnormalities. We focus on studying and optimizing the classifier to the task and defer the implementation of dot counting based on the classification results to another study. We investigate the fuzzy ARTMAP in discriminating valid from artifact signals representing either a single or both Down and Patau syndromes, thus establishing two and four-class classification problems, respectively. We provide extensive examination of all main aspects of the fuzzy ARTMAP, starting from parameter sensitivity, ordering strategy for presenting training data to the classifier, training mode and selecting a paradigm for the different classification tasks. Finally, we compare the performance of the fuzzy ARTMAP in FISH signal classification to other machine learning and NN models, and show that the fuzzy ARTMAP achieves almost the highest accuracy and using lesser training resources (either a training period or sample size) than the other models. Together with the aforementioned characteristics, this makes the fuzzy ARTMAP a prominent candidate for a classification-based FISH dot-counter.

We briefly introduce FISH image acquisition, processing and analysis, including FISH signal representation in Section II. The main principles and dynamics of the fuzzy ARTMAP are summarized in Section III. Section IV presents experiments we performed using synthetic data in order to explore a several issues of the fuzzy ARTMAP, as well as the FISH data, followed by a thorough analysis of the experimental results. We further discuss the results in Section V.

II. Fish Image Analysis and Signal Representation

All stages of the biological procedure required for obtaining FISH images were described in [4]. Red and green signals, corresponding to chromosomes 21 and 13 respectively, were seen on blue DAPI stained nuclei. The analysis of these signals enabling the diagnosis of Down and Patau syndromes, respectively, was preliminary described in [17]. A total of 400 images were collected from five slides, stored in tagged image file format (TIFF) format and used in the signal classification experiments.

By analyzing each of the three color channels—red, green, and blue (RGB) of a FISH image separately—image processing and segmentation could be facilitated [4]. Multispectral FISH image analysis was beneficial not only to expedite preprocessing and segmentation, but also to yield color-based features that contribute to discriminative signal representation. Following image acquisition, the RGB color format was utilized because preprocessing, nuclei and signal segmentation, as well as the measurement of some features were performed more easily using this color format. However, as intensities of red and green signals, each measured in its own channel, were very similar to each other, the RGB format was not suitable for discriminating between signals of different colors. By contrast, signals of different colors represented by the hue parameter of the hue, saturation, intensity (HSI) color format [18] could be easily resolved due to their different hues [4].

Following segmentation, twelve features were measured for the signals and they are listed in Table I. The features include Area (a size measure) and Eccentricity (a shape measure), which were previously suggested [19] to describe FISH signals. In addition, we measured a number of spectral features. We computed at the specific color plane three RGB intensity-based measurements: the Total and Average Channel Intensities and the Channel Intensity Standard Deviation. We also computed four HSI hue-based measurements: Maximum Hue, Average Hue, Hue Standard Deviation, and Delta Hue. Delta Hue is the difference between the Maximum and Average Hue normalized by the Average Hue. This feature was added to the set after observing that the difference between values of the average and maximum hue for real signals was usually near zero, whereas for some kinds of artifacts (e.g., overlap of signals of two different colors) this difference was substantially large. Two additional features of the set are the two coordinates of the eigenvector corresponding to the largest eigenvalue of the red and green intensity components of the signal. The last feature is the Average Grey Intensity, i.e., average intensity over the three color channels. Previously [4], feature selection based on different qualitative and quantitative methodologies was applied to these twelve features in order to evaluate the contribution of each feature or combination of features to the classification process. However, since we focus in this study on the classifier, we skip feature selection and employ the features altogether.

Finally, we obtained 3008 signals each represented by the twelve features of Table I and a label associating the signal with its class as determined by a cytogeneticist expert. Table II shows

TABLE I
SET OF FISH SIGNAL FEATURES STUDIED IN THE WORK. TEXTURE INDICATES STANDARD DEVIATION OF CHANNEL INTENSITY (5) OR HUE (8). EIG. 1, 2 (10, 11) ARE ABBREVIATIONS FOR THE TWO COORDINATES OF THE EIGENVECTOR CORRESPONDING TO THE LARGEST EIGENVALUE OF THE RED AND GREEN INTENSITY COMPONENTS OF THE SIGNAL

Number	Feature	Number	Feature
1	Area	7	Average Hue
2	Eccentricity	8	Hue Texture
3	Total Channel Intensity	9	Delta Hue
4	Average Channel Intensity	10	Eig. 1
5	Texture	11	Eig. 2
6	Maximum Hue	12	Average Grey Intensity

TABLE II
SIGNAL DISTRIBUTION (NUMBER AND PERCENTAGE) AMONG THE FOUR CLASSES INVESTIGATED, CORRESPONDING TO REAL SIGNALS AND ARTIFACTS OF DOWN (RED) AND PATAU (GREEN) SYNDROMES

Class	"red real"	"red artifact"	"green real"	"green artifact"
# patterns	500	1220	552	736
%	16.62	40.56	18.35	24.47

the distribution of the 3008 signals between the four classes established by "real" (valid) and "artifact" signals of the two explored syndromes.

III. FUZZY ARTMAP PRINCIPLES AND DYNAMICS

The fuzzy ARTMAP NN for incremental supervised learning [7] incorporates two fuzzy adaptive resonance theory (ART) [8] modules denoted as ART_a and ART_b. These are linked by a map field module associating nodes (categories) from ART_a with nodes in ART_b. The fuzzy ART module [20] performs fast incremental unsupervised learning by clustering *M*-dimensional input patterns into categories, each forming a hyperrectangle region in the *M*-dimensional input space, in three stages: category choice, vigilance test, and learning. In the category choice stage, a choice function is calculated for the (complement-coded [7]) input $\mathbf{I}$ and each existing category w_j

$$T_j = \frac{||\mathbf{I} \wedge w_j||}{\alpha + ||w_j||} \tag{1}$$

where $\wedge$ is the fuzzy AND operation, $(\mathbf{X} \wedge \mathbf{Y})_i = min(x_i, y_i), \alpha > 0$ is a choice parameter[1] and the norm is L_1. The chosen category is the one achieving the highest value of the choice function. When a category *J* is chosen, a hypothesis test called vigilance test is performed in order to measure the category match to the input $\mathbf{I}$. If the match function exceeds the vigilance parameter $\rho \in [0, 1]$

$$\frac{||\mathbf{I} \wedge w_J||}{||\mathbf{I}||} \geq \rho \tag{2}$$

then the chosen category is said to win and learning is performed. Otherwise, the chosen category is removed from the search for this pattern. As a result, a new category maximizing the choice function (1) is chosen and the process continues until a chosen category satisfies the vigilance test (2). The vigilance parameter controls the similarity required between the chosen category and the input pattern in order to allow learning, where lowering the vigilance parameter provides broader generalization (large categories) and *vice versa*. If none of the existing categories meets the vigilance test, a new category is formed and learning is performed without a vigilance test. Either way, learning is accomplished by updating the weight vector[2] of the winning (or new) category according to

$$w_J^{\mathrm{new}} = \beta \left(\mathbf{I} \wedge w_J^{\mathrm{old}}\right) + (1 - \beta) w_J^{\mathrm{old}} \tag{3}$$

where $\beta \in (0, 1]$ is the learning rate and $\beta = 1$ defines fast learning.

In pattern recognition tasks, the input $\mathbf{I}_a$ to ART_a is the pattern and the input I_b to ART_b is the pattern label. As ART_b inputs are labels, ART_b vigilance parameter ρ_b is configured to one, so each label is clustered by a specific ART_b category. The map field includes a matrix of weights w^{ab} which maps ART_a categories to ART_b categories. The *J*th row vector of w^{ab} denotes the prediction of ART_b categories as a result of the *J*th winning category in ART_a. The map field is activated to produce the output

$$x_{ab} = y^b \wedge w_J^{ab} \tag{4}$$

where ART_b output, y^b, has Boolean coordinates

$$y_k^b = \begin{cases} 1, & \text{if the } k\text{th category wins in } \mathrm{ART}_b \\ 0, & \text{otherwise} \end{cases} \tag{5}$$

so $|x_{ab}|$ is the value of the weight that predicts the winning ART_b Kth category as a result of the winning ART_a Jth category. During the training phase, the map field performs a vigi-

[1]The interested reader can find an elaborated examination of the role of the choice parameter in [21].

[2]The initial weights are usually set at 1 [7].

lance test similarly to ART_a vigilance test, where if the match function exceeds the map field vigilance parameter $\rho_{ab} \in [0, 1]$

$$\frac{||x_{ab}||}{||y^b||} \geq \rho_{ab} \tag{6}$$

then resonance and learning occur. This test assures that the prediction of the correct class complies with the label represented by the winning ART_b Kth category. Else, a match tracking procedure is activated for finding a better category in ART_a. In this process, the map field raises ART_a vigilance parameter ρ_a

$$\rho_a = \frac{||\mathbf{I}_a \wedge w_J^a||}{||\mathbf{I}_a||} + \delta, \qquad 0 < \delta \ll 1. \tag{7}$$

This ensures that the current Jth category fails the vigilance test in ART_a and is removed from the competition. The search in ART_a proceeds until an ART_a category that predicts the correct ART_b category is chosen, otherwise a new category is created. When the Jth category upholds the map vigilance test (6), its association to the ART_b categories is adapted by the following learning rule:

$$w_J^{ab,\mathrm{new}} = \beta_{ab}(w_J^{ab,\mathrm{old}} \wedge \mathbf{y}^b) + (1 - \beta_{ab}) w_J^{ab,\mathrm{old}} \tag{8}$$

which is activated during resonance in the map field. In fast learning mode ($\beta_{ab} = 1$), the link between ART_a Jth category and ART_b Kth category is permanent, i.e., $w_{JK}^{ab} = 1$ for all data presentations. In the test phase, only ART_a is active, so the vigilance test in the map field is avoided. The class prediction is deduced from the map field weights of the winning ART_a category.

IV. Experimentation

A. The Methodology

We investigated the fuzzy ARTMAP classifier in discriminating FISH signals of two genetic syndromes. In all the experiments, the fuzzy ARTMAP was configured to fast learning ($\beta_a = \beta_{ab} = 1$) and the choice parameter was set to a small value $\alpha = 10^{-6}$ similarly to [9], [14]. Depending on the experiment, the ART_a vigilance parameter was either configured to a specific value or changed in [0, 1] in order to check possible influence of this parameter on the classification accuracy and number of categories.

Fig. 1 summarizes the experimental methodology that is detailed in four levels: the classification task, classification paradigm, ordering strategy and training mode. Three classification tasks were studied: Discriminating real from artifact red signals enabling Down syndrome detection ("red" in Fig. 1), discriminating real from artifact green signals enabling Patau syndrome detection ("green" in Fig. 1) and accomplishing both tasks simultaneously (i.e., a four-class classification task) enabling the detection of both syndromes ("4-class" in Fig. 1). We implemented two classification paradigms. The first was a monolithic classifier, i.e., a single fuzzy ARTMAP classifier that associated each signal with one of either the two or four classes. The second paradigm, implemented for the four-class classification task, was a two-stage hierarchical classifier. In the first stage, it classified the signals according to their color, i.e., red or green, and in the second stage as real or artifact of each of the color classes. The reasoning for this paradigm is that previous experience [6] has showed that decomposing the four-class problem into two two-class problems improves overall classifier performance.

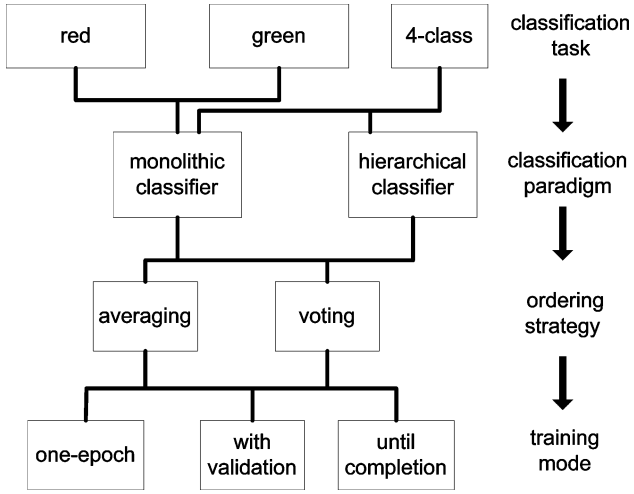


Fig. 1. Experimental methodology.

One troublesome issue of using fuzzy ARTMAP classifiers is the network inherent sensitivity to the presentation order of the training data [7], [22]. In order to overcome this problem, we implemented and compared two strategies named averaging and voting. In the averaging strategy, we duplicated the training set T_1 times and permuted each copy randomly. For each of these copies we conducted the experiment independently and averaged the classification accuracy over the experiments. Being a random variable of the data ordering, the true classification accuracy is approximated by the average accuracy over a several fuzzy ARTMAPs classifying each a different data ordering. Thus, the accuracy derived employing this strategy is a reference for the classifier accuracy and not a method for choosing a single classifier. In the voting strategy [7], the classifier was composed of T_2 fuzzy ARTMAP s each trained on a different permutation of the training set. In the test, a decision module received predictions of the T_2 fuzzy ARTMAP s associated with the different orderings and reported the majority label as the classifier decision. That is, both strategies are identical during the training stage but differ in the test stage. In addition, the voting strategy provides an ensemble classifier for the test. As the voting strategy had been found superior to the averaging strategy in a several applications [12], [13], it was interesting to compare both strategies while classifying the FISH image database. In the reported experiments, both T_1 and T_2 were equal to 5.

In addition, we implemented and studied three modes of training: "one-epoch," "until completion," and "with validation." In the one-epoch training mode, the fuzzy ARTMAP is trained using a single presentation of all the training patterns.

When training until completion, the classifier is repeatedly trained until predicting the training set perfectly.[3] In the training with validation mode, every training epoch the classifier accuracy is evaluated on a validation set and training is progressed until no further increase of this accuracy is reached. Training with validation is a popular approach in the machine learning community aiming at the elimination of overfitting in order to enhance generalization [5]. This approach also regards the usage of training until completion as erroneous as it is prone to overfitting [23]. On the other hand, the developers of fuzzy ARTMAP defined training until completion as the method of choice and it is used for fuzzy ARTMAP classifiers frequently [7], [23]. The same three training modes were also compared on artificial and "real-world" databases in [23], where the accuracy of training with validation was evaluated on a validation set after presentation of every 100 training patterns. For the artificial databases, training with validation increased accuracy significantly only in about one half of the inspected cases but reduced the number of categories in all cases compared to training until completion. On the real-world databases, no differences in either the accuracy or number of categories were found between the training modes.

In all our experiments, we employed a cross-validation procedure in order to estimate the classification accuracy. First, the randomly ordered database was divided into P equal size segments. The classifier was trained and tested P times, where in each time the test set was a specific segment and the training set a union of the remaining P-1 segments.[4] In this procedure, the patterns were utilized similarly and equally for training and test, while keeping disjoint test and training sets in each experiment. Test and training accuracy results were averaged over the P experiments. We used $P = 10$ segments, i.e., a CV-10 procedure.

We continue this section by first describing an experiment using synthetic data which is used to build some understanding of the fuzzy ARTMAP capability in classification and then presenting the main experimentation of the fuzzy ARTMAP in FISH signal classification.

B. Results for Synthetic Data

In order to get a preliminary insight into the way fuzzy ARTMAP classifies data, we have investigated the model accuracy and rise in the number of categories when trained to classify 250 patterns of each of two classes generated synthetically from two-dimensional Gaussian probability density functions. The two Gaussians had medium overlapping, equal *a-priori* probabilities, close means and covariance matrices of opposite orientations. The generated patterns and the Bayes' optimal decision boundaries are shown in Fig. 2. The Bayes' bound [24] of classification accuracy was estimated as 78.0%.

The fuzzy ARTMAP was trained using a particular data ordering, $\rho_a = 0.6$ and $\alpha = 10^{-6}$ for one epoch or until completion. As is shown in Fig. 3 (top), the categories formed in the first epoch of training cover most of the feature space. The additional

[3]Although until completion refers to 100% training accuracy, we stopped training when the model accuracy reached 99%.

[4]In training with validation mode, another segment was used for validation and $P - 2$ segments for training.

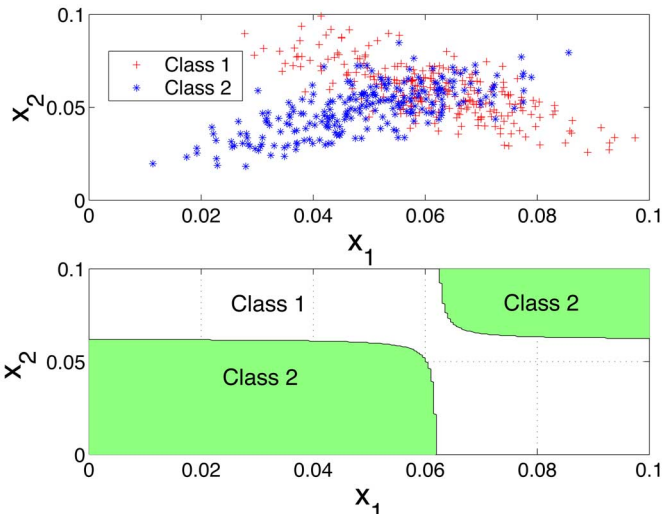


Fig. 2. Two classes having each 250 patterns generated randomly from a (top) two-dimensional Gaussian probability density function and the (bottom) corresponding Bayes' optimal decision boundaries for classifying the two classes.

categories, formed in the remaining epochs until training completion, reside mostly in the overlapping region between the two classes, as can be seen in Fig. 3 (bottom). When an existing category in the overlapping region that is associated with a specific class fails to correctly represent a pattern of another class, match tracking for this pattern begins. If no other existing category matches the pattern, a new category having a larger weight (2) and (7) and, thus, a smaller size [7] is formed. As long as these categories are not too small, and having that the training and test patterns are drawn from the same distribution, the new categories refine the decision boundaries and improve both training and test accuracies. This is demonstrated in Fig. 4 where the decision boundaries achieved following training for one epoch act as a layout for the more elaborated decision boundaries established in the remaining epochs. Overfitting the training patterns begins only when the new categories are too small, however these categories cannot deteriorate the test accuracy much if other categories for each class reside in the region. Repeated for 50 training orderings and averaged over these orderings, the training accuracy is raised from 78.2% for one epoch to 99.1% when training completes and the test accuracy from 69.3% to 71.2%. The average number of categories has increased from 23.9 to 72.3. That is, though expensive in categories leading to overfitting, the until completion training mode improves test accuracy compared with the one-epoch mode.

C. Results for the Cytogenetic Problem

1) Between the Methodologies: In the first batch of experiments with the cytogenetic database, we measured the fuzzy ARTMAP sensitivity to the classification task, classification paradigm, ordering strategy and training mode. We classified FISH signals of each disease separately (i.e., two two-class classification problems) using the monolithic paradigm as well as jointly (i.e., a four-class classification problem) using both the monolithic and hierarchical paradigms. In these experiments, we employed each of the three training modes—one-epoch,

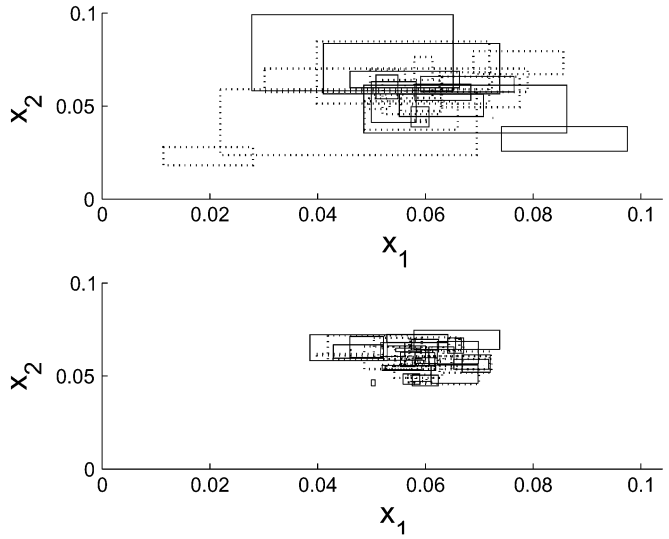


Fig. 3. Fuzzy ARTMAP categories formed while classifying the data shown in Fig. 2 at the end of the first training epoch (top) and after this epoch and until training completion (bottom). Solid and dashed line rectangles represent categories of Classes 1 and 2, respectively.

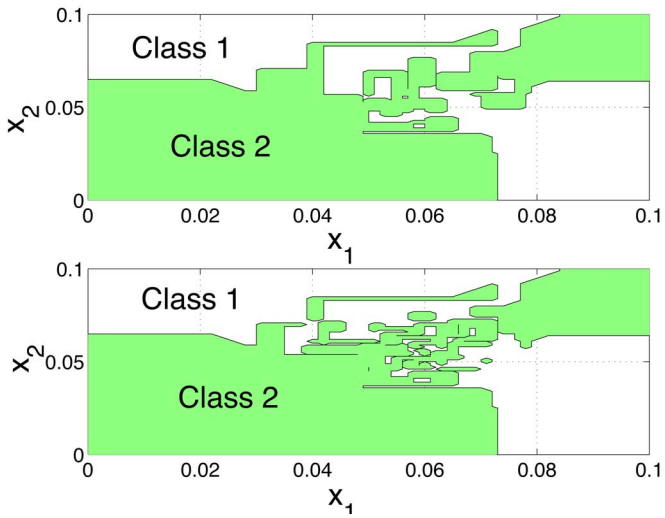


Fig. 4. Decision boundaries of the fuzzy ARTMAP classifying the data shown in Fig. 2 and having the categories of Fig. 3 at the end of the (top) first training epoch and at (bottom) training completion.

with validation and until completion, each of the ordering strategies—averaging and voting and three values of the ART_a vigilance parameter ρ_a, i.e., low (0.5), medium (0.7), and high (0.9). Performance was evaluated using the test classification accuracy as well as the numbers of categories and training epochs needed for achieving this accuracy.

Tables III–Table V show, respectively, for the three values of the vigilance parameter, the monolithic paradigm fuzzy ARTMAP performance for the averaging strategy, the three training modes and the three classification tasks. Similarly, Tables VI–VIII show the fuzzy ARTMAP performance using the voting strategy. Studying the figures in the tables, we summarize that both the numbers of categories and training epochs increase when training progresses from one-epoch, thorough with validation and until completion in all cases studied, when most of the categories are formed in the first epoch. The number of categories is also enlarged with the increase of ρ_a in all cases. As expected, the voting strategy yields higher test accuracy than the averaging strategy for all cases. Averaged over all classification tasks, training modes and values of the vigilance parameter, the voting strategy is accurate in 84.2% compared to 80.8% for the averaging strategy. The red signals are easier to classify than the green ones (85.8% accuracy versus 82.2% when averaging over all experiments), probably since they are more common in the database ($\sim 60\%$) and brighter in intensity. Simultaneous classification of signals of both syndromes yields slightly lower test accuracy (81.6% when averaging over all experiments) than for each of the separate syndromes. In agreement with the experiment for the synthetic database, experiments with the cytogenetic database show an advantage of the until completion training mode on the one-epoch and with validation modes for both the voting strategy (85.2% versus 84.4%, and 85.0%, respectively) and especially the averaging strategy (83.0% versus 80.0% and 81.8%, respectively). Test accuracies averaged over all cases for the three training modes are 84.1% (until completion), 82.2% (one-epoch) and 83.4% (with validation).

Extending the experiments of the four-class classification problem to the hierarchical strategy, Table IX shows a similar pattern of results to that shown by the monolithic paradigm. Averaged over all experiments, the accuracy of the hierarchical paradigm is slightly inferior to that of the monolithic paradigm (80.1% versus 81.6%).

2) Vigilance Sensitivity: In order to examine the sensitivity of the fuzzy ARTMAP to different values of the vigilance parameter ρ_a, we repeated the experiment with the red signals (Down syndrome), averaging strategy and the three training modes for increasing values of ρ_a. Figs. 5–7 show, respectively, for the one-epoch, with validation and until completion training modes, the training, validation (if applicable) and test classification accuracies for increasing values of ρ_a. In the one-epoch training mode and most of the range of ρ_a (Fig. 5), the gap between the training and test accuracies is relatively small (3–4%), but it grows as training continues (7–8% for training with validation (Fig. 6) and $\sim$13% for until completion (Fig. 7)). Accuracies for all training modes can be divided roughly into three ranges (except of course that training accuracy for the until completion mode is fixed inherently at (almost) 100% for all values of the vigilance parameter). In the first range where $0 \leq \rho_a \leq 0.2$, the accuracies (estimated means and standard deviations) are identical. In the second range ($0.2 < \rho_a < 0.7$) the accuracies are similar but not identical. In the third region ($0.7 \leq \rho_a < 1$), the training accuracy rises (usually) monotonically to 100% in one-epoch and with validation, while the test accuracy drops by approximately 4%, 3%, and 1% for training with validation, until completion and for one-epoch, respectively. In this region, the vigilance is very high, so more categories representing only very few signals (and even single signals for vigilance values near 1) are formed. Despite the drop in the fuzzy ARTMAP test accuracy for high vigilance values, this accuracy is high and remarkably stable for a wide range of vigilance values.

TABLE III
MEAN (STANDARD DEVIATION) VALUES OVER CV-10 OF THE TRAINING, TEST, AND VALIDATION (VAL.) (IF APPLICABLE) CLASSIFICATION ACCURACIES OF THE MONOLITHIC FUZZY ARTMAP PARADIGM UTILIZING THE AVERAGING STRATEGY AND $\rho_a = 0.5$ FOR THE THREE CLASSIFICATION TASKS. 1–3 REFER TO TRAINING MODES ONE-EPOCH, WITH VALIDATION AND UNTIL COMPLETION, RESPECTIVELY. ALSO INCLUDED, THE NUMBERS OF ART_a CATEGORIES AND TRAINING EPOCHS

Task	Mode	Training (%)	Val. (%)	Test (%)	Categories	Epochs
	1	86.2 (1.8)	-	78 (3.7)	113.1 (7.9)	1 (0)
4-class	2	94.4 (3.6)	81.4 (1.9)	80.2 (2.4)	155.9 (30.2)	3.4 (1)
	3	99.3 (0.3)	-	81.8 (1.8)	222.4 (10.8)	5.3 (0.7)
	1	88.3 (2.3)	-	83.9 (3.3)	45.7 (7)	1 (0)
red	2	91.6 (3.9)	85.7 (3.1)	84.1 (3.4)	56.2 (19)	3 (1.1)
	3	99.4 (0.3)	-	85.8 (2.8)	103.7 (7.2)	5.6 (0.7)
	1	84.8 (3.4)	-	78.1 (3.5)	42.5 (5.7)	1 (0)
green	2	90 (4.1)	82 (3.2)	81.3 (3.5)	50.7 (15.3)	2.7 (0.9)
	3	99.5 (0.3)	-	82.6 (3.2)	100.3 (7.5)	5.4 (0.7)

TABLE IV
MEAN (STANDARD DEVIATION) VALUES OVER CV-10 OF THE TRAINING, TEST, AND VALIDATION (VAL.) (IF APPLICABLE) CLASSIFICATION ACCURACIES OF THE MONOLITHIC FUZZY ARTMAP PARADIGM UTILIZING THE AVERAGING STRATEGY AND $\rho_a = 0.7$ FOR THE THREE CLASSIFICATION TASKS. 1–3 REFER TO TRAINING MODES ONE-EPOCH, WITH VALIDATION AND UNTIL COMPLETION, RESPECTIVELY. ALSO INCLUDED, THE NUMBERS OF ART_a CATEGORIES AND TRAINING EPOCHS

Task	Mode	Training (%)	Val. (%)	Test (%)	Categories	Epochs
	1	86.3 (2.2)	-	77.4 (3.6)	130.9 (8.5)	1 (0)
4-class	2	94.3 (3.9)	81.5 (2.2)	79.8 (2.4)	166 (31.7)	3.6 (1.2)
	3	99.3 (0.3)	-	81.6 (2.2)	233.1 (10.3)	5.2(0.6)
	1	87.6 (3.7)	-	82.5 (4.2)	53.2 (4.6)	1 (0)
red	2	92 (3.1)	85.9 (2.4)	84.6 (2.8)	62.8 (17.1)	2.8 (1)
	3	99.4 (0.3)	-	85.7 (2.6)	115.9 (8.1)	5.7 (0.9)
	1	84.8 (3.4)	-	79 (3.8)	53.3 (4.1)	1(0)
green	2	91.2 (4.5)	82.4 (3.6)	80.3 (4.1)	66 (17.6)	3 (1.2)
	3	99.5 (0.3)	-	81.6 (2.7)	111.3 (8)	5.4 (0.7)

TABLE V
MEAN (STANDARD DEVIATION) VALUES OVER CV-10 OF THE TRAINING, TEST AND VALIDATION (VAL.) (IF APPLICABLE) CLASSIFICATION ACCURACIES OF THE MONOLITHIC FUZZY ARTMAP PARADIGM UTILIZING THE AVERAGING STRATEGY AND $\rho_a = 0.9$ FOR THE THREE CLASSIFICATION TASKS. 1–3 REFER TO TRAINING MODES ONE-EPOCH, WITH VALIDATION AND UNTIL COMPLETION, RESPECTIVELY. ALSO INCLUDED, THE NUMBERS OF ART_a CATEGORIES AND TRAINING EPOCHS

Task	Mode	Training (%)	Val. (%)	Test (%)	Categories	Epochs
	1	90.8 (1.1)	-	79 (2.2)	205.5 (6.4)	1 (0)
4-class	2	96.4 (2.8)	81.5 (2.2)	80.2 (1.9)	272.4 (22.2)	3.4 (1)
	3	99.4 (0.3)	-	81 (1.9)	327.9 (12.3)	4.3 (0.5)
	1	91.7 (1.6)	-	82.6 (3.6)	135.4 (6)	1 (0)
red	2	96.3 (2.4)	85.1 (3.1)	84 (2.8)	138.3 (12.1)	3 (1)
	3	99.4 (0.3)	-	84.5 (2.9)	168.1 (6.9)	4.4 (0.6)
	1	89.9 (2.1)	-	79.2 (3.8)	122.5 (4.9)	1 (0)
green	2	95.4 (3.3)	82.8 (3.2)	81.4 (3.2)	125 (11.1)	3.08 (1.09)
	3	99.5 (0.3)	-	82.2 (3.3)	156.7 (7)	4.1 (0.6)

TABLE VI
MEAN (STANDARD DEVIATION) VALUES OVER CV-10 OF THE TRAINING, TEST AND VALIDATION (VAL.) (IF APPLICABLE) CLASSIFICATION ACCURACIES OF THE MONOLITHIC FUZZY ARTMAP PARADIGM UTILIZING THE VOTING STRATEGY AND $\rho_a = 0.5$ FOR THE THREE CLASSIFICATION TASKS. 1–3 REFER TO TRAINING MODES ONE-EPOCH, WITH VALIDATION AND UNTIL COMPLETION, RESPECTIVELY. ALSO INCLUDED, THE NUMBERS OF ART_a CATEGORIES AND TRAINING EPOCHS

Task	Mode	Training (%)	Val. (%)	Test (%)	Categories	Epochs
	1	86 (1.6)	-	82.8 (2.3)	109.7 (10.3)	1 (0)
4-class	2	92.8 (3.6)	80.6 (2.8)	83.6 (1.5)	160.7 (30.8)	2 (1.1)
	3	99.4 (0.3)	-	83.8 (1.3)	222.2 (12.5)	5.2 (0.7)
	1	88.1 (1.8)	-	87 (2.7)	42.3 (3.9)	1 (0)
red	2	92.9 (3.4)	85.8 (3)	87.8 (2)	68.6 (17.9)	2 (1.3)
	3	99.5 (0.3)	-	87.3 (1.8)	105.2 (8.9)	5.4 (0.7)
	1	84.8 (3)	-	84.3 (2.1)	41.5 (5)	1 (0)
green	2	91.8 (4)	82.4 (2.9)	84.73 (4.1)	68.4 (16.3)	2.1 (1.13)
	3	99.5 (0.9)	-	84.4 (2.8)	98.1 (5.6)	5.2 (0.8)

TABLE VII
MEAN (STANDARD DEVIATION) VALUES OVER CV-10 OF THE TRAINING, TEST AND VALIDATION (VAL.) (IF APPLICABLE) CLASSIFICATION ACCURACIES OF THE MONOLITHIC FUZZY ARTMAP PARADIGM UTILIZING THE VOTING STRATEGY AND $\rho_a = 0.7$ FOR THE THREE CLASSIFICATION TASKS. 1–3 REFER TO TRAINING MODES ONE-EPOCH, WITH VALIDATION AND UNTIL COMPLETION, RESPECTIVELY. ALSO INCLUDED, THE NUMBERS OF ART_a CATEGORIES AND TRAINING EPOCHS

Task	Mode	Training (%)	Val. (%)	Test (%)	Categories	Epochs
	1	86.2 (2.8)	-	82.6 (1.3)	124.9 (9)	1(0)
4-class	2	93.3 (2.9)	80.8 (2)	83 (1.2)	175.3 (24.8)	2.1 (1.1)
	3	99.4 (0.2)	-	84.5 (1.9)	236.3 (9.1)	5.3 (0.6)
	1	88.2 (2.5)	-	88.2 (3.4)	51.7 (4.6)	1 (0)
red	2	92.7 (3)	85.4 (2.6)	87.9 (2.3)	75.2 (15.6)	1.8 (0.9)
	3	99.5 (0.3)	-	87.5 (2.6)	114.9 (6.5)	5.6 (0.8)
	1	85.2 (3.1)	-	83.8 (3)	52.6 (4.1)	1 (0)
green	2	92.9 (3.4)	82.6 (3.2)	84.1 (1.6)	80.5 (15.3)	2.1 (1.2)
	3	99.5 (0.3)	-	84.3 (2.9)	112.6 (7)	5.6 (0.8)

TABLE VIII
MEAN (STANDARD DEVIATION) VALUES OVER CV-10 OF THE TRAINING, TEST AND VALIDATION (VAL.) (IF APPLICABLE) CLASSIFICATION ACCURACIES OF THE MONOLITHIC FUZZY ARTMAP PARADIGM UTILIZING THE VOTING STRATEGY AND $\rho_a = 0.9$ FOR THE THREE CLASSIFICATION TASKS. 1–3 REFER TO TRAINING MODES ONE-EPOCH, WITH VALIDATION AND UNTIL COMPLETION, RESPECTIVELY. ALSO INCLUDED, THE NUMBERS OF ART_a CATEGORIES AND TRAINING EPOCHS

Task	Mode	Training (%)	Val. (%)	Test (%)	Categories	Epochs
	1	90.4 (1.3)	-	82.7 (1.9)	256.7 (7.5)	1 (0)
4-class	2	95.7 (2.2)	80.6 (2.1)	83.6 (1.2)	282.5 (18.2)	2.4 (1.1)
	3	99.4 (0.3)	-	83.9 (1.8)	326.3 (10.2)	4.2 (0.5)
	1	91.8 (1.4)	-	86.1 (3.7)	134.6 (4.1)	1 (0)
red	2	96.1 (2.2)	85.1 (3)	87.1 (2.1)	145.7 (10.5)	2.1 (1)
	3	99.5 (0.2)	-	87.8 (2.1)	169.3 (7.5)	4.4 (0.6)
	1	90 (2.6)	-	82.4 (3.1)	122.6 (5.6)	1 (0)
green	2	95.6 (2.1)	82.1 (3.5)	83.5 (3.7)	133 (7.9)	2.1 (0.91)
	3	99.6 (0.3)	-	83.4 (4)	154.9 (6.6)	4.1 (0.7)

TABLE IX
MEAN (STANDARD DEVIATION) VALUES OVER CV-10 OF THE TRAINING AND TEST CLASSIFICATION ACCURACIES OF THE HIERARCHICAL FUZZY ARTMAP PARADIGM UTILIZING THE AVERAGING AND VOTING STRATEGIES, $\rho_a = 0.5, 0.7, 0.9$ AND CLASSIFYING THE TWO SYNDROMES SIMULTANEOUSLY. 1–3 REFER TO TRAINING MODES ONE-EPOCH, WITH VALIDATION AND UNTIL COMPLETION, RESPECTIVELY

Mode	Vigilance	Averaging Training (%)	Averaging Test (%)	Voting Training (%)	Voting Test (%)
1	0.5	81.5 (2.8)	74.5 (3.3)	84.9 (1.9)	81.3 (1.7)
2		92.5 (4.1)	78.7 (2.9)	93.4 (3.4)	82.0 (1.9)
3		99.5 (0.3)	80.6 (2.0)	99.4 (0.3)	81.8 (1.3)
1	0.7	82.7 (3.3)	74.8 (4.4)	84.8 (2.4)	81.9 (2.3)
2		93.3 (3.2)	79.0 (2.6)	92.4 (2.7)	81.2 (1.6)
3		99.4 (0.3)	80.4 (2.5)	99.4 (0.3)	82.1 (1.3)
1	0.9	90.0 (1.6)	77.0 (2.7)	90.2 (1.2)	82.7 (1.5)
2		96.2 (1.5)	79.4 (2.4)	95.2 (2.1)	82.5 (2.3)
3		99.4 (0.2)	80.1 (2.7)	99.4 (0.2)	83.1 (1.5)

In order that the fuzzy ARTMAP accuracy could be improved, the overlapping regions between classes (responsible for most of the errors) should affect the training process more heavily. This could be accomplished by either a larger training set (having signals in these regions) or by providing the model more flexibility in learning the existing signals within these regions. Since our data is not large enough, the fuzzy ARTMAP should gain flexibility from increasing either ρ_a, thereby allowing more categories to be formed, or the number of training epochs. The first case is more appropriate to the one-epoch training mode whereas the second case to training with validation and until completion. However, increasing both ρ_a (the number of categories) and training period yields the model over-flexibility which results in the undesirable phenomenon of overfitting the training set at the expense of inferior generalization to an independent (test) set, which is especially pronounced in training with validation and until completion in the third region of ρ_a. This is indeed manifested in the 4% and 3% drop in test accuracies for the with validation and until completion modes, respectively, compared to only 1% drop for the one-epoch mode when moving from small to large values of ρ_a.

3) Incremental Learning: We evaluated the fuzzy ARTMAP classifying signals of Down syndrome (the red signals) in incremental (online) learning ($\rho_a = 0.5$). For this purpose, we compiled a database composing of 500 red signals of each of the real and artifact classes, and employed randomly 900 of these signals in training while keeping 100 signals for the test. In each of 45 training iterations, we continued training the network using ten additional training signals from each class and tested the incrementally learned network on the same 100 test signals. In each iteration, we recorded the fuzzy ARTMAP number of categories and accuracy on both the training and test sets. We repeated this procedure until all 900 signals have been presented to the network, i.e., one training epoch, and continued this procedure for a several epochs and until no change of the training and test accuracies was noticed. We also repeated the whole experiment using CV-10 and 5 training orderings and averaged the fuzzy ARTMAP performance over all experiments. The number of categories as well as test and training accuracies are shown in Fig. 8 for the first eight epochs of the incremental learning of the fuzzy ARTMAP. The number of categories rises monotonically and smoothly to 64.4. The training accuracy for the first signals of each epoch is high and it reduces until the end of the epoch as more variants of the signals (noise) participate in the training set. However, from epoch to epoch, the slope of this reduction becomes more moderate until being flat at the seventh epoch, as the fuzzy ARTMAP learns to predict the training set perfectly. The corresponding test accuracy varies between 78.5% and 87.1% during the first epoch and then gradually converges to 85.4% after seven epochs.

V. DISCUSSION

The fuzzy ARTMAP was extensively experimented in off and online classification of FISH image signals needed for genetic abnormality diagnosis. The main aspects of the problem, that is, selecting a classifier paradigm (monolithic or hierarchical), ordering strategy (averaging or voting) and training mode (one-epoch, with validation or until completion) as well as evaluating parameter sensitivity (vigilance parameter) were examined for different classification tasks associated with the detection of either a single or two genetic syndromes simultaneously.

When comparing ordering strategies, the test accuracy can be interpreted as a random variable of the training set representation order. Therefore, the averaging strategy is a simple sample mean estimation of this random variable. The voting strategy, incorporating knowledge from different orderings of the training set, was originally suggested [7] in order to improve prediction accuracy. The results obtained in our study indicate that this incorporation of knowledge succeeds in diminishing prediction error and thereby the results supply additional evidence to the importance of the voting strategy in fuzzy ARTMAP training. Another ordering strategy [22], which is based on the max-min clustering method to identify a fixed ordering, is computationally efficient thus worthwhile for evaluation for the cytogenetic problem.

When assessing the experimental results for the training modes, we observe that training until completion produces higher accuracy than training with validation or for one-epoch.

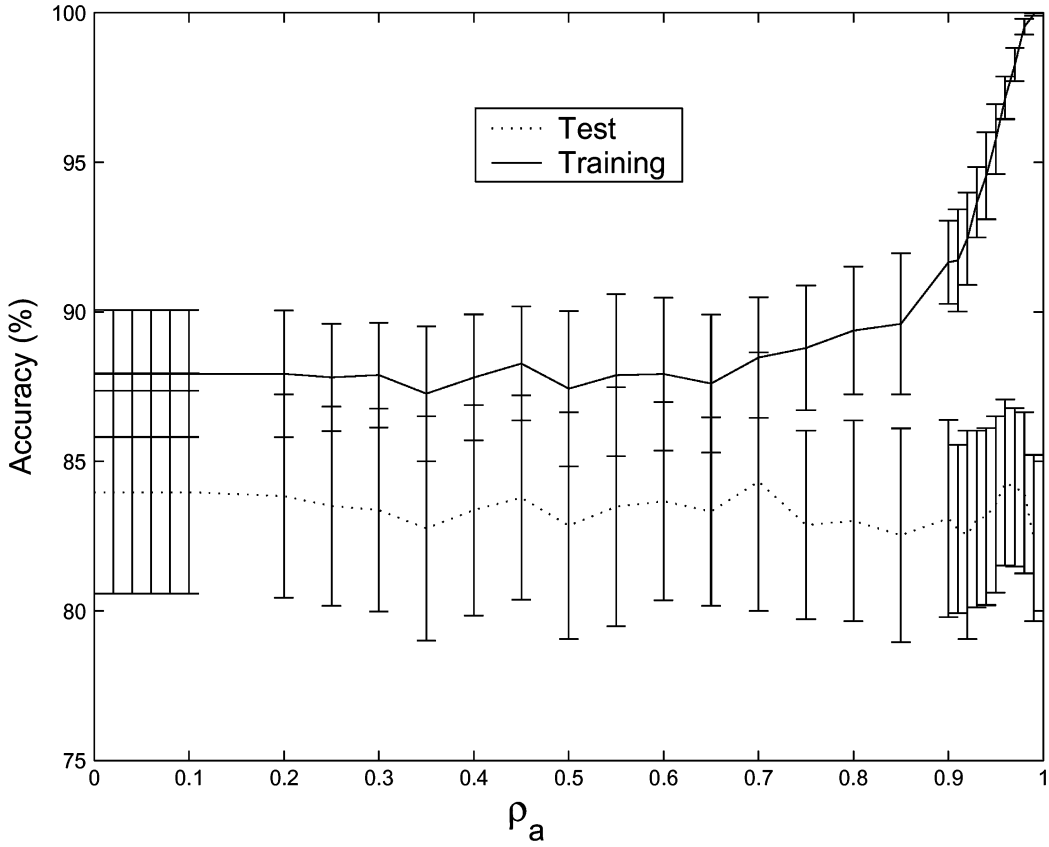


Fig. 5. Mean (±standard deviation) of the training and test accuracies of the fuzzy ARTMAP classifying Down syndrome signals using the averaging strategy for increasing values of the vigilance parameter, ρ_a, and the one-epoch training mode.

This observation highlights the fuzzy ARTMAP from other NNs, such as the MLP, and machine learning models, as training until completion usually leads these models to overfit the data. Overfitting in machine learning models should be avoided since the increase in training accuracy is at the expense of generalization to new data, and indeed different means to stop training before completion, such as the use of a validation set, are commonly applied. Interestingly, instead of loosing test accuracy when training progresses beyond a certain point, the fuzzy ARTMAP reaches and maintains high test accuracy until training is completed. Moreover, this conclusion is true for both off and online learning. The explanation to this phenomenon is that after approximately the first epoch, new categories are created in order to represent "noisy" training patterns, i.e., variants of patterns already learned by the existing categories. Each of these additional categories represents only a several training patterns or even a single pattern, and thus it has minor size and influence on the already-established representation. Therefore, the chances that these categories win the competition to represent next coming (or test) patterns is small. Although these additional categories are responsible to overfitting the training set (and increase of the training accuracy) they do not affect the test much, as most test patterns are likely to be represented by the large, previously established categories. However, those test patterns that are found to be represented and classified more accurately by the additional, rather than previously established, categories intensify the test accuracy. That is, this special kind of overtraining increases the computational complexity and memory requirements of the fuzzy ARTMAP classifier, and not only that it does not deteriorate the classifier generalization capability but it improves it. It will require further experimentation to judge whether the rise of performance when using training until completion is consistent with other than the cytogenetic domain.

Also very interesting to note is the contribution of a validation set to fuzzy ARTMAP learning. Compared to NNs and other machine learning models that gain both reduced complexity and improved accuracy from employing a validation set, the fuzzy ARTMAP only gains reduced complexity, i.e., a smaller number of categories. The accuracy of the model trained with validation has always been inferior to that when training until completion. This is mainly due to two reasons. First is that stopping training only because the accuracy on a validation set ceases to increase at some point does not infer that learning cannot continue and better representations be found as is explained previously. Second is the fact that drawing a validation set from (a not too large) training set reduces the number of patterns left for training and thus also the number of possible categories shaping the decision boundary. Similar conclusions can be drawn from analyzing the results of experiments held with other real-world databases [23]. Finally, as the test accuracy fluctuates intensively in the first epoch of the training (Fig. 8), and even if the highest accuracy is indeed reached on the validation set during this epoch, it is risky to count on validating the accuracy every constant number of patterns during the first epoch.

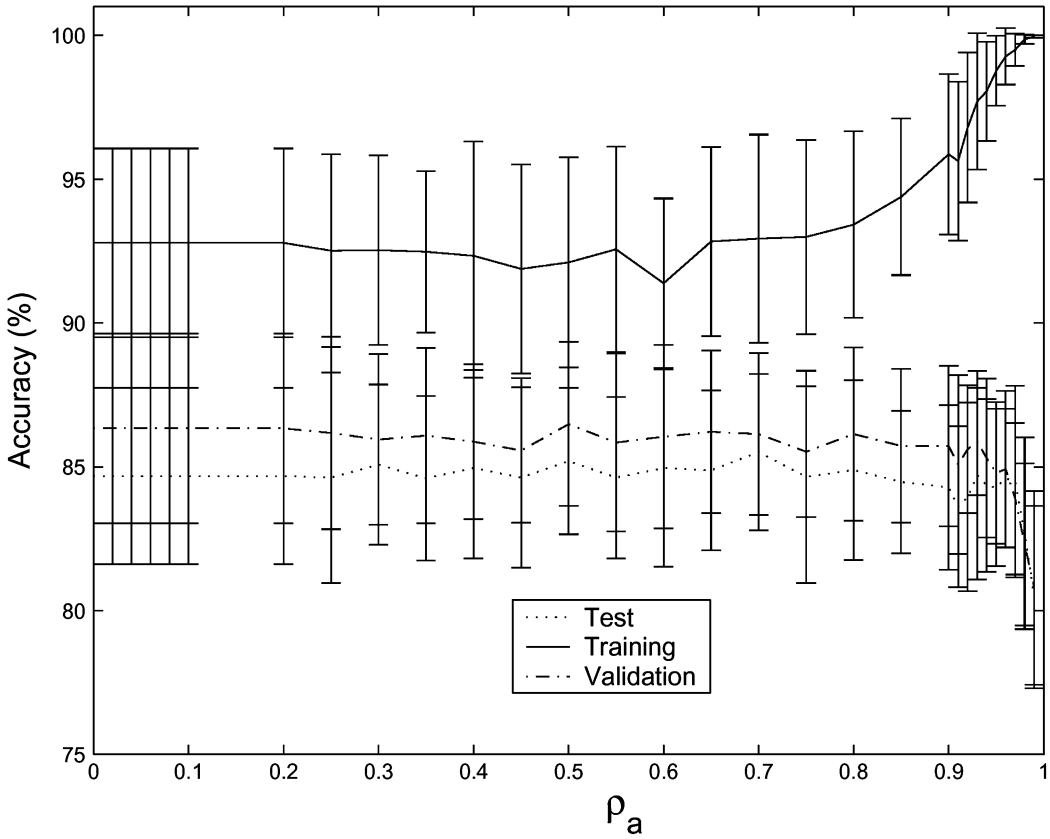


Fig. 6. Mean (± standard deviation) of the training, validation and test accuracies of the fuzzy ARTMAP classifying Down syndrome signals using the averaging strategy for increasing values of the vigilance parameter, ρ_a, and the with validation training mode.

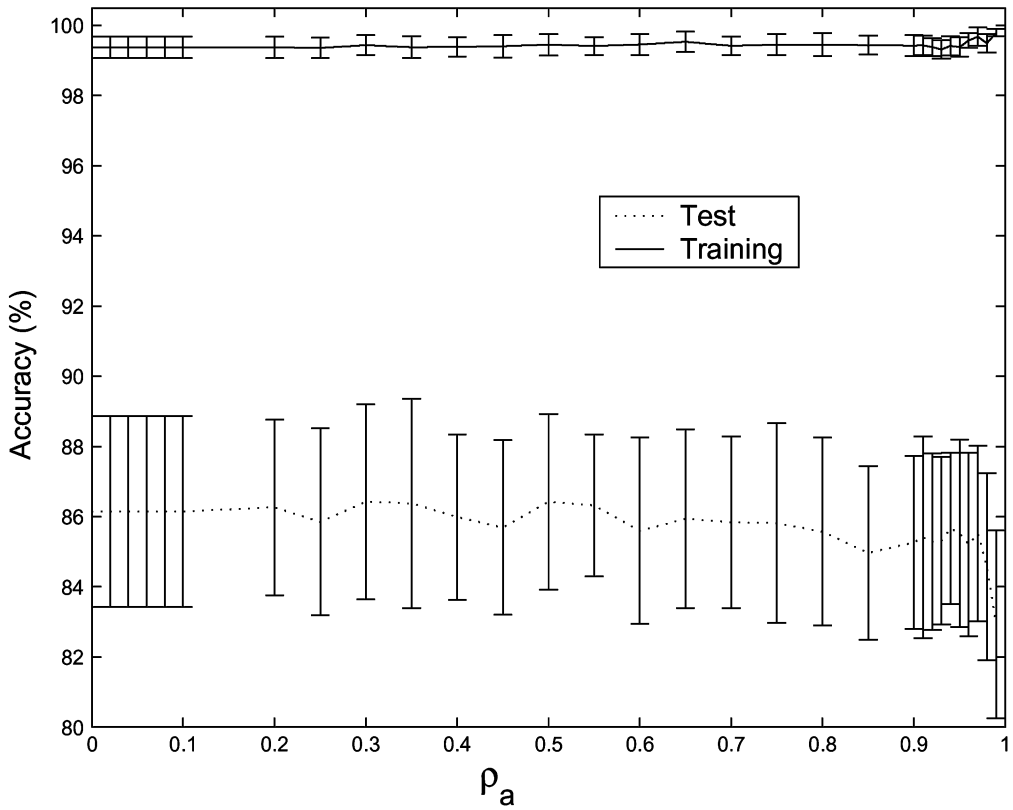


Fig. 7. Mean (± standard deviation) of the training and test accuracies of the fuzzy ARTMAP classifying Down syndrome signals using the averaging strategy for increasing values of the vigilance parameter, ρ_a, and the until completion training mode.

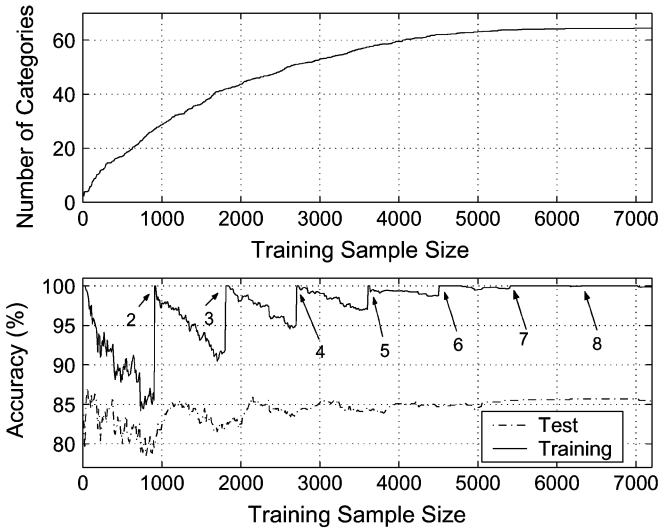


Fig. 8. Number of categories (top) as well as training and test classification accuracies (bottom) of the fuzzy ARTMAP that learns incrementally the red signals. Numbers at the bottom figure indicate the start of new training epochs.

TABLE X
ACCURACIES OF THE FUZZY ARTMAP AND OTHER STATE-OF-THE-ART MACHINE LEARNING PARADIGMS CLASSIFYING THE SAME FISH SIGNALS OF BOTH DOWN AND PATAU SYNDROMES

Classifier	Test accuracy (%)
Bayesian Neural Network (BNN)	87.1
Multi-Layer Perceptron (MLP) Neural Network	84.8
Fuzzy ARTMAP (until completion)	84.5
Support Vector Machine (SVM)	84.2
Fuzzy ARTMAP (one-epoch)	82.8
Linear Classifier	79.6
Naive Bayesian Classifier (NBC)	78.0

Although slightly inferior to the accuracy achieved when trained until completion, the high accuracy of the fuzzy ARTMAP when trained for one-epoch demonstrates its capability in fast offline learning. During and immediately after the first epoch, the fuzzy ARTMAP establishes most of the categories it will end with and accomplishes most of its learning. In our case, this indicates that data representation of the cytogenetic domain has been constructed almost completely during the first epoch of training. Changes in the number of categories and test accuracy after this first epoch are relatively small although the training accuracy keeps improving. In addition, training for one epoch is less sensitive to the value of the vigilance parameter and thus to overfitting. Similarly to offline learning, when learned incrementally the fuzzy ARTMAP builds its representation and accuracy mainly during the presentation of the first patterns before rapidly converges to a high classification accuracy.

Finally, Table X demonstrates that the accuracy of the fuzzy ARTMAP when trained until completion is either comparable to that of other state-of-the-art classifiers, such as the support vector machine [25] and the MLP NN [26], or superior to that of other classifiers, such as the naive Bayesian classifier and the linear classifier [26]. Still, the accuracy of the fuzzy ARTMAP, as well as those of all the above classifiers, are inferior to that of the Bayesian NN (BNN). Compared to the other classifiers, the BNN sophisticatedly gains its enhanced precision by combining a priori information with that acquired from the data. However, it depends on the availability of this information *a priori* and ways to introduce it to the model, as well as on methods (e.g., Markov chain Monte Carlo) to alleviate the computational difficulties when the model has a large number of parameters (network weights) [27]. When compared to the MLP, the most popular NN, the fuzzy ARTMAP NN trained until completion achieves comparable accuracy, however it requires less than six training epochs compared to hundreds of epochs that are required for the MLP [26]. If loosing some accuracy is acceptable ($\sim$2% for the cytogenetic data), then training (either on or offline) may end after one epoch or sooner employing the fuzzy ARTMAP. Both options make the fuzzy ARTMAP a remarkable alternative to the MLP NN.

We would summarize that experimenting with the fuzzy ARTMAP in FISH signal classification demonstrates a model which rapidly and accurately classifies signals in all tasks, both off and online, using all paradigms, model parameters, ordering strategies and training modes. The fuzzy ARTMAP shows a high degree of plasticity to the different objectives, however without loosing its stability and enhanced classification performance rendering itself an excellent candidate for genetic abnormality diagnosis and classification-based FISH dot counting. We further believe that most of the conclusions deduced in our study to the cytogenetic domain, such as, the superiority of the voting ordering strategy and training until completion mode can be generalized to other problems, but further experimentation is needed.

ACKNOWLEDGMENT

The authors would like to thank S. Dhanjal of Warwick University, U.K., for FISH preparation, as well as image acquisition and labeling.

REFERENCES

[1] D. Pinkel, J. Landegent, C. Collins, J. Fuscoe, R. Segraves, J. Lucas, and J. Gary, "Fluorescence in situ hybridization with human chromosome-specific libraries: Detection of trisomy 21 and translocations of chromosome 4," in *Proc. Natl. Acad. Sci.*, 1988, vol. 85, pp. 9138–9142.

[2] J. Nath and K. L. Johnson, "A review of fluorescence in situ hybridization (FISH): Current status and future prospects," *Biotech. Histochem.*, vol. 75, pp. 54–78, 2000.

[3] H. J. Tanke, R. J. Florijn, J. Wiegant, A. K. Raap, and J. Vrolijk, "CCD microscopy and image analysis of cells and chromosomes stained by fluorescence in situ hybridization," *Histochem. J.*, vol. 27, pp. 4–14, 1995.

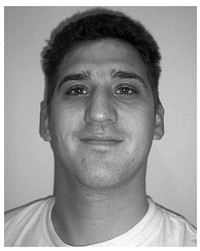
[4] B. Lerner, W. F. Clocksin, S. Dhanjal, M. A. Hultén, and C. M. Bishop, "Feature representation and signal classification in fluorescence in-situ hybridization image analysis," *IEEE Trans. Syst., Man, Cybern., A, Syst. Humans*, vol. 31, no. 6, pp. 655–665, Nov. 2001.

[5] C. M. Bishop, *Neural Networks for Pattern Recognition*. Oxford, U.K.: Clarendon, 1995.

[6] B. Lerner, W. F. Clocksin, S. Dhanjal, M. A. Hultén, and C. M. Bishop, "Automatic signal classification in fluorescence in-situ hybridization images," *Cytometry*, vol. 43, pp. 87–93, 2001.

[7] G. A. Carpenter, S. Grossberg, N. Markuzon, J. Reynolds, and D. Rosen, "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Trans. Neural Netw.*, vol. 3, no. 5, pp. 698–713, Sep. 1992.

[8] S. Grossberg, "Adaptive pattern recognition and universal encoding II: Feedback, expectation, olfaction, and illusions," *Biol. Cybern.*, vol. 23, pp. 187–202, 1976.
[9] M. A. Rubin, "Application of fuzzy ARTMAP and ART-EMAP to automatic target recognition using radar range profiles," *Neural Networks*, vol. 8, pp. 1109–1116, 1995.
[10] G. A. Carpenter and W. D. Ross, "ART-EMAP: A neural network architecture for object recognition by evidence accumulation," *IEEE Trans. Neural Netw.*, vol. 6, no. 4, pp. 805–818, Jul. 1995.
[11] Y. Suzuki, "Self-organizing QRS-wave recognition in ECG using neural networks," *IEEE Trans. Neural Netw.*, vol. 6, no. 6, pp. 1469–1477, Nov. 1995.
[12] J. R. Williamson, "Gaussian ARTMAP: A neural network for fast incremental learning of noisy multidimensional maps," *Neural Networks*, vol. 9, pp. 881–897, 1996.
[13] G. A. Carpenter and N. Markuzon, "ARTMAP-IC and medical diagnosis: Instance counting and inconsistent cases," *Neural Networks*, vol. 11, pp. 323–336, 1998.
[14] e.g. Sánchez, Y. A. Dimitriadis, J. M. Cano-Izquierdo, and J. Lopez-Coronado, "μ-ARTMAP: Use of mutual information for category reduction in fuzzy ARTMAP," *IEEE Trans. Neural Netw.*, vol. 13, no. 1, pp. 58–69, Jan. 2002.
[15] D. Charalampidis, T. Kasparis, and M. Georgiopoulos, "Classification of noisy signals using fuzzy ARTMAP neural networks," *IEEE Trans. Neural Netw.*, vol. 12, no. 5, pp. 1023–1036, Sep. 2001.
[16] R. Palaniappan, P. Raveendran, and S. Omatu, "VEP optimal channel selection using genetic algorithm for neural network classification of alcoholics," *IEEE Trans. Neural Netw.*, vol. 13, no. 2, pp. 486–491, Mar. 2002.
[17] B. Lerner and B. Vigdor, "An empirical study of fuzzy ARTMAP applied to cytogenetics," *Proc. 23rd IEEE Convention of Electrical Electronics Engineers in Israel*, pp. 301–304, 2004.
[18] Y. Ohta, *Knowledge-based Interpretation of Outdoor Natural Color Scenes*. London, U.K.: Pitman, 1985.
[19] H. Netten, L. J. van Vliet, H. Vrolijk, W. C. R. Sloos, H. J. Tanke, and I. T. Young, "Fluorescent dot counting in interphase cell nuclei," *Bioimag.*, vol. 4, pp. 93–106, 1996.
[20] G. A. Carpenter, S. Grossberg, and D. Rosen, "Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural Networks*, vol. 4, pp. 759–771, 1991.
[21] G. Heilman, G. Bebis, and M. Georgiopoulos, "Order of search in fuzzy ART and fuzzy ARTMAP: Effect of the choice parameter," *Neural Networks*, vol. 9, pp. 1541–1559, 1996.
[22] I. Dagher, M. Georgiopoulos, G. Heilman, and G. Bebis, "An ordering algorithm for pattern presentation in fuzzy ARTMAP that tends to improve generalization performnace," *IEEE Trans. Neural Netw.*, vol. 10, no. 4, pp. 768–778, Jul. 1999.
[23] A. Koufakou, M. Georgiopoulos, G. Anagnostopoulos, and T. Kasparis, "Cross-validation in fuzzy ARTMAP for large databases," *Neural Networks*, vol. 14, pp. 1279–1291, 2001.
[24] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001.
[25] A. David and B. Lerner, "Support vector machine-based image classification for genetic syndrome diagnosis," *Pattern Recogn. Lett.*, vol. 26, pp. 1029–1038, 2005.
[26] B. Lerner and N. D. Lawrence, "A comparison of state-of-the-art classification techniques with application to cytogenetics," *Neural Comput. Appl.*, vol. 10, pp. 39–47, 2001.
[27] R. M. Neal, *Bayesian Learning for Neural Networks*. New York: Springer-Verlag, 1996, Lecture Notes in Statistics 118.



Boaz Vigdor received the B.Sc. and M.Sc. degrees from Ben-Gurion University, Beer-Sheva, Israel, in 2002 and 2005, respectively. He is currently working toward the Ph.D. degree at Ben-Gurion University, Israel.

Boaz Lerner received the B.A. degree in physics and mathematics from the Hebrew University, Israel, in 1982, and the Ph.D. degree in computer engineering from Ben-Gurion University, Beer-Sheva, Israel, in 1996.

He performed research at the Neural Computing Research Group, Aston University, Birmingham, U.K., and the Computer Laboratory of the University of Cambridge, Cambridge, U.K. In 2000, he joined the Department of Electrical and Computer Engineering at Ben-Gurion University, Israel, where he is currently a Senior Lecturer. His current interests include machine learning approaches to data analysis, learning Bayesian networks, neural networks, and their application to "real-world" problems.