

# Improving Simplified Fuzzy ARTMAP Performance Using Genetic Algorithm for Brain Fingerprint Classification

# Ramaswamy Palaniappan<sup>1</sup>, Shankar M. Krishnan<sup>2</sup> and Chikkanan Eswaran<sup>3</sup>

<sup>1</sup>*Department of Computer Science, University of Essex,  
Colchester, United Kingdom  
rpalan@essex.ac.uk*

<sup>2</sup>*Biomedical Engineering Research Centre, Nanyang Technological University,  
Research TechnoPlaza, Singapore  
esmkrish@ntu.edu.sg*

<sup>3</sup>*Faculty of Information Technology, Multimedia University,  
Cyberjaya, Malaysia  
eswaran@mmu.edu.my*

## Abstract

*A genetic algorithm is proposed for ordering the input patterns during training for Simplified Fuzzy ARTMAP (SFA) classifier to improve the individual identification classification performance using brain fingerprints. The results indicate improved classification performance as compared to the existing methods for pattern ordering, namely voting strategy and min-max. As the ordering method is general, it could be used with any dataset to obtain improved classification performance when SFA is used.*

## 1. INTRODUCTION

Fuzzy ARTMAP (FA) [1] is an incremental neural network classifier, which has found use in numerous pattern recognition problems [2, 3]. Simplified Fuzzy ARTMAP (SFA) is a simpler version of FA, which is faster and performs equally well as FA [4, 5]. However, both the generic FA and its simplified SFA suffer from varying classification performance depending on the order of input patterns during training when it is used in the fast learning mode. Two methods have been proposed to overcome this problem, namely voting strategy [1] and min-max ordering [6].

In the first method, SFA is trained several times using training patterns presented in random order (i.e. permutations of the training patterns) and then the predicted classes of the test patterns are stored. Majority votes are used to determine the final class prediction for the test patterns [1]. It is also customary to state the average classification of test patterns from all the simulations in addition to the voting results.

To solve the problem of having to run many simulations, a single simulation method based on min-max clustering was proposed [6]. For a  $c$  - class problem, the method works by ordering the  $c$  training patterns that are maximally distant in the training feature space. Next, for the rest of the patterns,

the method orders training patterns that are minimally distant from these  $c$  patterns. Hence, it is known as min-max ordering.

In this study, a method that uses genetic algorithm (GA) [7] to select the presentation order of training patterns is proposed. The method works by using the selection, crossover and inversion operators in GA to select the presentation order of training patterns that maximises the SFA classification performance. Once the order is selected, only a single SFA training simulation (similar to min-max ordering) will be required for classification of test patterns. The performance of the proposed technique is compared with training patterns ordered by min-max and random ordering using brain fingerprints data (i.e. visual evoked potential (VEP) signals) to identify individuals.

## 2. METHODOLOGY

The methodology consisted of two distinct, yet related stages:

- Pattern ordering stage using either GA or min-max (as comparison);
- Performance testing stage with SFA training and testing for all the ordering methods (i.e. our proposed method, min-max and random).

The second stage was important as it showed the improvement in SFA performances when trained by training patterns ordered by GA as compared to min-max and random ordering.

The available dataset was split into 3 sets: datasets 1, 2 and 3. In the pattern ordering stage, GA was used with datasets 1 and 2. GA was simulated for 100 generations with 20 chromosomes using fitness given by SFA training and testing for each chromosome (hence SFA training and testing was conducted for 20 times in a generation).

SFA was trained with vigilance parameter (VP) value of 0. When this was completed, the presentation order of training patterns has been selected, and GA was not used anymore. Similarly, dataset 1 was used by the min-max method in the pattern ordering stage to order the presentation of training patterns. Once the order has been selected, the min-max method was not used anymore.

During performance testing stage, the presentation orders selected by GA and min-max methods were used. For random ordering, since there was no pattern ordering stage, the simulation was repeated 20 times with random permutations of the training patterns and voting strategy<sup>1</sup> as suggested in [1] was used to predict the final class of the test patterns.

In this second stage, we conducted classification experiments with dataset 1 for SFA training and dataset 3 for SFA testing. Dataset 2 was not used here (to be fair) as only GA has used this dataset before in the earlier stage. In other words, all the different ordering methods (proposed, min-max and random) were trained and tested with the same datasets.

Though SFA training and testing are involved in both stages for GA method, it should be noted that the SFA training and testing used in pattern ordering stage was different from the SFA training and testing used in performance testing stage. The SFA training (conducted for 2000 times) in pattern ordering stage was conducted with the aim to order the presentation of training patterns. But the SFA training in performance testing stage was conducted with the aim to test the classification performance of SFA with the selected presentation order of training patterns. So, in the second performance testing stage, for the presentation order selected by GA and min-max methods, only one SFA training and testing was completed. Note again that neither GA nor min-max methods was used in this second stage, only the presentation order selected earlier by GA or min-max methods was used. For random ordering, 20 SFA training and testing was conducted in the performance testing stage.

### 3. GA METHOD

The steps involved in the GA method are as follows:

#### A. Step 1: Initialisation

The number of bits in each chromosome was set to the number of patterns in dataset 1. Each bit was randomly set from integer values of 1 to this number, without repetition. Twenty similar chromosomes were generated, which represent the population.

#### B. Step 2: Fitness value

To calculate the fitness value of each chromosome, SFA was trained (using VP=0 to speed training and minimise overfitting) by patterns in dataset 1 by the presentation order given by the chromosome. The trained SFA was tested with VEP patterns from dataset 2 and the fitness value was the percentage of correctly classified VEP patterns over the total tested VEP patterns.

#### C. Step 3: Selection, Crossover and Inversion Operators

Two selection (reproduction) methods, namely, roulette wheel and tournament selection methods were used to select the chromosomes for the next generation. Half of the population (i.e. 10 chromosomes) were selected by each method. Here, the tournament selection method worked by selecting the best from 3 randomly chosen chromosomes. This step was repeated 10 times to obtain 10 chromosomes. Tournament selection would be useful to retain the chromosomes with high fitness values, but roulette wheel selection would be necessary to avoid premature convergence, i.e. to avoid GA from converging too quickly with suboptimal chromosomes. The roulette wheel method worked by selecting chromosomes with a higher probability of survival. In general, higher fitness chromosomes will have a higher chance of survival [7].

Two bits (i.e. genes) in a chromosome were randomly chosen and they were swapped if the crossover probability was not exceeded. This procedure was different from the common method of applying the crossover operation because of the nature of the problem, where every chromosome must be permutations of other chromosomes. This is also the reason why mutation operator was not used. Inversion operators were used to inverse the bits in the chromosomes. Here, a two-point inversion operator was used. Two points were randomly chosen and bits from a randomly chosen chromosome were inverted between these two points. The two points were chosen such that they lie within 3 point difference. This was to avoid too much disruption to the chromosome. Crossover and inversion operators were applied for certain number of times based on the probability,  $p$ , which was initially set at 0.9. The high initial probability was chosen because of the simple crossover and lack of mutation operator. The probability was gradually reduced with increasing number of generations using the formula;

$$p(n) = 0.9 * (1 - n / \text{max generation}) \quad (1)$$

where  $n$  would be the current generation.

#### D. Step 4

Steps 2 and 3 were repeated until a maximum generation number of 100 was reached. The overall best chromosome (with highest fitness value) was stored.

Since the best chromosomes were selected by the GA depend on the initial search space, GA simulation was repeated five times and the chromosome that had the fitness value closest to the average of the five best chromosomes' fitness values was stored. This chromosome represented the GA selected presentation order of training patterns for SFA. Figure 1 shows the steps involved in the GA method.

<sup>1</sup> Each random ordering will predict a certain class. The final output is based on the majority vote of the different classes.

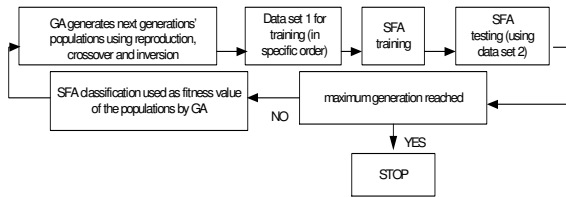


Fig. 1: GA method to select the presentation order

#### 4. VEP DATA

An experimental study was conducted to show the superior performance of the GA method compared to the random ordering and min-max methods. For this purpose, the data set used in an earlier work to identify individuals [8] was used.

The details of this data set will be briefly repeated here. VEP signals were extracted from 61 channels from 40 subjects while seeing a single black and white line picture. The pictures were common objects like a ball, a book, a car, etc. VEP signals from 40 trials with 1 second measurement intervals were stored. Gamma band (30-50 Hz) spectral power (GBSP) for each VEP signal was computed using zero-phase forward and reverse Butterworth filter and Parseval's time-frequency equivalence theorem. Each VEP pattern consists of GBSP features from 61 channels. These VEP patterns were classified into the 40 categories representing the different subjects.

The dataset consisting of 1600 VEP patterns was divided exclusively into 3 sets: datasets 1, 2 and 3. Datasets 2 and 3 each consisted of 13 VEP patterns from each subject, while dataset 1 consisted of 14 VEP patterns from each subject. Therefore, dataset 1 consisted of 560 VEP patterns, while datasets 2 and 3 consisted of 520 VEP patterns. Datasets 1 and 2 were used by the GA to select the presentation order of training patterns for the SFA.

The selections of patterns for each dataset were done randomly.

#### 5. RESULTS

The classification was carried out for VP values ranging from 0.1 to 0.9 (in steps of 0.1) but to save space, only the averaged results for classification performances, averaged training times (for a single pattern) and averaged SFA network sizes (based on number of Fuzzy ART clusters) are given in Table 1. Note that the SFA training times reported in Table 1 were the average time to train a single pattern in the performance testing stage.

Actually, the random-voting method would require 20 as many weights or 20 training times as shown in the Table. But in reality, either the training time or the weights would be 20 times more, but not both the training time and weights. This fact should be noted, though the average of both training time and weights are reported here. In addition, note that the SFA training time for random-average method was not from 20 SFA trainings but averaged from 20, which was done to approximate one SFA training time. Figure 2 shows the

classification performances for varying VP values using different ordering methods.

TABLE 1: SUMMARY OF AVERAGED RESULTS (VP= 0.1 TO 0.9) USING DIFFERENT ORDERING METHODS

Ordering method	Training time (s)	SFA size	Classification (%)
GA method	0.011134	94.10	93.48
Random – average	0.012028	98.96	89.17
Random – voting	0.012028	98.96	91.84
Min-max	0.011150	94.20	90.26

From Table 1 and Figure 2, it can be seen that the GA method gave superior classification performance over both random ordering and the min-max methods for all the VP values. It can also be seen that GA based presentation order of training patterns required lower training times and smaller SFA sizes when compared to the random ordering and min-max methods (true for all the VP values). The GA method was also advantageous over random ordering since it required only one simulation.

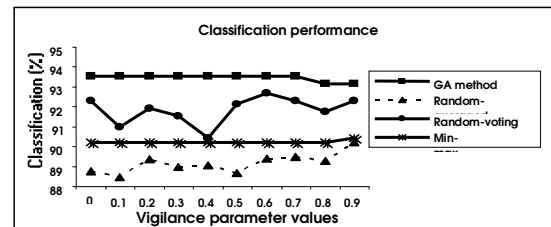


Fig. 2: Classification performances using different ordering methods

Another interesting fact that can be concluded from Figure 2 is that the VP values did not affect the classification performance of the GA method as significantly as they did for the random ordering and min-max methods. As such, if the GA method was used, the value of VP can be fixed at 0. The two main parameters that require tuning for SFA are presentation order of training patterns and VP. By using the GA method, the SFA does not require tuning of either of these parameters.

In the pattern ordering stage, our simulations indicated that GA selects the pattern order in much lesser time than min-max though the GA has to run 100 iterations! This difference becomes more evident as the numbers of training patterns become higher due to the increase of the min-max ordering complexity with increasing pattern size. However, exact time comparisons are not a matter of concern as there would be no such comparison possible with the random ordering method which does not have pattern ordering stage. Furthermore, the pattern ordering stage would

be generally conducted 'offline' and the important issues are actual performances addressed in the second stage (like training time, size, accuracy) once the presentation order of training patterns have been selected.

## 6. CONCLUSION

This paper has proposed the use of GA to select the presentation order of training patterns for SFA. The new method could also be applied to FA. The performances of the proposed method have been compared with the performances of the random ordering with a voting strategy and the min-max method for solving an individual classification problem using VEP signals. Though there are computational overheads for the proposed method, it was only during the pattern ordering stage and once ordered, the method performed the fastest.

It has been shown that SFA classification performances were better for the GA based method when compared to the random ordering and the min-max methods. Further the GA based method required lower training times and smaller SFA sizes when compared to the other methods. An additional advantage of the proposed method over random ordering method was that it required only a single simulation. The SFA classification performances, when training patterns were ordered by the GA method showed only a small variance for different VP values. SFA could be used with a VP of 0 for both the pattern ordering and performance testing stages. This means that the two parameters in SFA, namely, presentation order of training patterns and vigilance parameter do not require tuning for this method.

## REFERENCES

- [1] G. A. Carpenter, S. Grossberg, N. Markuzon, J. H. Reynolds, and D. B. Rosen "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Transactions on Neural Networks*, pp. 698-713, vol. 3, no.5, 1992.
- [2] R. Palaniappan, P. Raveendran, S. Nishida, and N. Saiwaki, "A new brain-computer interface design using fuzzy ARTMAP," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, pp.140-148, vol. 10, issue 3, 2002.
- [3] P. Raveendran, R. Palaniappan, and S. Omatu, "Fuzzy ARTMAP classification of invariant features derived using angle of rotation from a neural network," *Information Sciences: An International Journal*, pp.67-84, vol.130, 2000.
- [4] M. Vakil-Baghmisheh, and N. Pavesic, "A fast simplified fuzzy ARTMAP network," *Neural Processing Letters*, pp. 273-316, vol.17, 2003.
- [5] T. Kasuba, "Simplified fuzzy ARTMAP," *AI Expert*, pp. 19-25, vol. 8, no. 11, 1993.
- [6] I. Dagher, M. Georgiopoulos, G. L. Heileman, and G. Bebis, "An ordering algorithm for pattern presentation in fuzzy ARTMAP that tends to improve generalization performance," *IEEE Transactions on Neural Networks*, pp. 768-778, vol. 10, no. 4, 1999.
- [7] R. L. Haught and S. E. Haupt, *Practical Genetic Algorithm*, John Wiley & Sons, 1998.
- [8] R. Palaniappan, "Method to identify individuals using VEP signals and neural network," *IEE Proceedings - Science, Measurement and Technology Journal*, pp. 16-20, vol. 151, no. 1, January 2004.