

Fuzzy ARTMAP Network with Evolutionary Learning

P. Ramuhalli, R. Polikar, L. Udpa and S. S. Udpa

Materials Assessment Research Group
Dept. of Electrical and Computer Engineering
Iowa State University, Ames, IA 50011

ABSTRACT

Neural networks, particularly the multilayer perceptron, have been used extensively in automated signal classification systems with classification accuracy as the figure of merit. Three important issues that can enhance the utility of these systems are (i) incremental learning, (ii) confidence or reliability measures and (iii) performance improvement through continual learning. This paper investigates these issues using a Fuzzy ARTMAP network. A hypothesis testing based algorithm is developed for computing reliability measures, which are fed back to the network for retraining and performance improvement. Implementation results on ultrasonic data are presented.

1. INTRODUCTION

Automatic signal classification (ASC) systems are used extensively for signal interpretation in non-destructive evaluation (NDE) applications. Their popularity stems from the fact that they are capable of rapid analysis of large amounts of data. In addition, they offer more accurate and consistent data interpretation and allow storage of expert knowledge.

The overall signal classification strategy is a two-step process namely (i) preprocessing (ii) classification. The preprocessing step is used to prepare the data for input to a classification system. The preprocessing typically performs noise filtering, data compression and feature extraction. The output of this step is the lower dimensional feature vector, which contains the class discriminatory information. The feature vector is input into a classification algorithm that outputs the class information. A training step is necessary to ensure that the system learns to distinguish between the different classes.

The criterion of interest in design of ASC systems is the accuracy of signal interpretation. Three other critical issues that are often ignored are:

- Incremental Learning: the ability to learn new information without forgetting knowledge already learned.
- Confidence Measures: a measure of the reliability of classification decision.
- Experiential/Evolutionary Learning: ability to learn continually and improve with time. This allows the system to react to changes in the data and helps to keep the system current.

This paper presents a neural network based ASC system that addresses the three issues described. Fuzzy ARTMAP networks [1] are used as the classifier of choice, largely because these

networks are capable of incremental learning [2]. In addition, these networks are capable of fast learning and provide guaranteed convergence.

Confidence levels can indicate lack of adequate training data in a particular region of the input space, or over fitting of the input-output mapping function. Given any random variable x and its probability density function, the probability

$$P[c_1 \leq x \leq c_2] = \gamma = 1 - \delta \quad (1)$$

defines the confidence interval for x corresponding to a confidence coefficient γ where δ is the confidence level and $[c_1, c_2]$ is the confidence interval [3]. In the context of classification, for a given interval, confidence can be defined as the probability of making a correct decision. Thus, if \mathbf{x} is a pattern vector from a class A, then the confidence associated with the classification of \mathbf{x} in A can be measured by the posterior probability $P(A|\mathbf{x})$. This conditional probability can be determined using standard hypothesis testing or Bayesian methods.

Previously developed methods for computing confidence include the work by Mitra and Pal [4], and Shroff [5]. The first method [4] uses fuzzy logic to compute membership values for the input attributes. The second approach uses regression analysis [5] and is applicable to networks that use the output error for training. Confidence measures developed previously for Fuzzy ARTMAP networks use fuzzy inference systems to compute membership functions for the input data [6, 7].

The organization of this paper is as follows. The next section provides an introduction to hypothesis testing and its application to estimating confidence measures for the ARTMAP network. Section 3 describes the proposed algorithm for performance improvement. The database used for validation and discussions of results are provided in section 4. Finally, section 5 presents conclusions and suggests directions for future investigation.

2. CONFIDENCE MEASURES

2.1 Hypothesis Testing

Hypothesis testing is a well-established statistical procedure for parameter estimation [3]. In general, given a random variable x and its distribution $F_X(x, \theta)$ parameterized by θ , the null hypothesis $H_0: \theta = \theta_0$ is tested against the alternative hypothesis $H_1: \theta \neq \theta_0$. The purpose of hypothesis testing is to establish whether the evidence is sufficient to reject the null hypothesis.

Two types of errors are possible in hypothesis testing. Defining the critical region D_c as the region in the sample space where the

density $f_X(x, \theta_0)$ of input x under hypothesis H_0 is negligible, the two types of error are defined as

Type I error: H_0 is true and $x \in D_c$, denoted by α and is given by $\alpha = P[x \in D_c | H_0]$ (2)

Type II error: H_0 is false and $x \notin D_c$, denoted by $\beta(\theta)$ and is given by

$$\beta(\theta) = P[x \notin D_c | H_1] \quad (3)$$

Hypothesis testing can be used for estimating confidence measures since the difference $1-\alpha$ is the confidence in rejecting the null hypothesis.

2.2 Application to neural networks

The method of hypothesis testing can be applied to the output activations of networks [8] to generate confidences. In a neural network, each output node goes high to indicate that the input belongs to the class represented by the node. In the case of an input x and class of node j , we have

$$\begin{aligned} H_0 : x \notin \text{class } j \\ H_1 : x \in \text{class } j \end{aligned} \quad (4)$$

The assumption for applying hypothesis testing in this case is that the output activation under the null hypothesis is lower than the activation under the alternative hypothesis. The cumulative distribution function F of activations under the null hypothesis is obtained as

$$F_X(x | H_0) = P[X \leq x] \quad (5)$$

where X is the activation of the output node. Suppose the output node reaches an activation x_0 . The probability of a type I error is then given by

$$\alpha = 1 - F_X(x_0) = P[X > x_0] \quad (6)$$

The probability of rejecting the null hypothesis is therefore given by

$$C_{alt} = 1 - \alpha \quad (7)$$

The same procedure is repeated to find the confidence in rejecting the alternative hypothesis. The activation of inputs under H_1 is used to compute the probability of rejecting the alternative hypothesis. The above procedure can be implemented only when data has been adequately sampled under both the hypotheses. For networks with multiple output nodes, hypothesis testing can be applied to the winning node to determine the confidence in the decision.

Two issues need to be taken into consideration when using this procedure for computing confidences. First, in most applications, the null hypothesis (as defined here) produces a lower activation than the alternative hypothesis. However, this assumption of lower activation under the null hypothesis may not hold in all cases. In this case, the above method can be used with the definitions of the null and alternative hypothesis reversed.

Second, the method gives the probability of rejecting a particular hypothesis. It does not give the probability of accepting the hypothesis (Note that lack of rejecting a hypothesis does not mean acceptance). This situation may arise if there are more than two hypotheses to be examined, for instance, in a three-class problem. In such a case, hypothesis testing can only give the probability of rejecting each hypothesis. However, if there are only two possibilities (the input either belongs to a class or it

doesn't), rejecting a hypothesis is equivalent to accepting the other hypothesis. Thus, the confidence of accepting the alternative hypothesis in the above description is the same as rejecting the null hypothesis, as given by (7).

2.3 Hypothesis Testing in Fuzzy ARTMAP

Hypothesis testing can be used to compute the confidence in Fuzzy ARTMAP classifications. Implementation of this algorithm requires redefining the output of the ART_a network in the Fuzzy ARTMAP algorithm to be

$$y_j^a = T_j^a = \left| I \wedge \mathbf{w}_j^a \right| / \left(\alpha + \left| \mathbf{w}_j^a \right| \right) \quad (8)$$

where y_j^a is the output of node j in F_2^a , I is the input to the network, \mathbf{w}_j^a represents the weights for node j in F_2^a , α is a tiebreak parameter and \wedge denotes the fuzzy AND (minimum) operation. Define null hypothesis H_0 : pattern I does not belong to category j and alternative hypothesis H_1 : I belongs to category j . y_j^a will be higher under H_1 than under H_0 . The algorithm for deriving confidence estimates is as follows.

Training:

1. Train network using the Fuzzy ARTMAP algorithm. Let S_j represent patterns clustered to node j in F_2^a . $S_j = \{I | I \rightarrow \text{node } j\}$, $j = 1, 2, \dots, N_a$ (9)

2. Use S_j to estimate the cumulative distribution function $F_{y,j}$ of activations y_j^a for each node j in F_2^a .

$$y_j^a = \left| I \wedge \mathbf{w}_j^a \right| / \left(\alpha + \left| \mathbf{w}_j^a \right| \right), \quad I \in S_j, \quad j = 1, \dots, N_a \quad (10)$$

$$F_{Y,j}(y_j^a | H_0) = P\left[Y \leq y_j^a \right], \quad j = 1, 2, \dots, N_a \quad (11)$$

Save the distributions.

Testing for pattern x :

1. Use the Fuzzy ARTMAP algorithm to classify the pattern. Obtain the category J in F_2^a to which the pattern is clustered and the output y_J from (10).

2. Use the stored cumulative distribution function for node J , $F_{Y,J}(y_J^a)$, to compute the probability of type I error.

$$\alpha = 1 - F_{Y,J}(y_J^a) = P[Y > y_J] \quad (12)$$

The confidence in the decision is $1-\alpha$

The assumption is that the training data has been adequately sampled under the null hypothesis. A similar technique can be applied to compute the confidence in the null hypothesis if required.

3. PERFORMANCE IMPROVEMENT

A second aspect of automatic signal classification systems is the concept of experiential learning, whereby the system evolves with time. Evolution with time can be achieved by using the Fuzzy ARTMAP algorithm in a feedback configuration where the network is retrained with patterns classified with low levels of confidence. The feedback algorithm is summarized in Figure 1. First, the trained network is used to classify the patterns in the test database and the confidence measures are computed for each classification. A selection criterion is then used to select all patterns with ambiguous classifications. A classification decision

is flagged as ambiguous if confidence is less than a threshold τ_1 or the ratio of confidence in H_1 to confidence in H_0 is less than a second threshold τ_2 . This ensures that patterns with a high confidence in rejecting both the null and alternative hypotheses are also called ambiguous. In such a case, operator intervention is required and the network is retrained so that it can recognize the patterns if encountered in future.

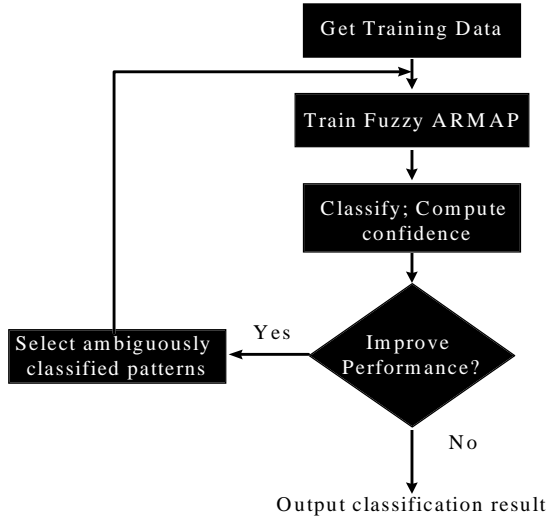


Figure 1. Flowchart for continual learning.

The input-output patterns are learned using the fast learning option given by [1]

$$\mathbf{w}_j^{new} = \beta(\mathbf{I} \wedge \mathbf{w}_j^{old}) + (1 - \beta)\mathbf{w}_j^{old} \quad (13)$$

with $\beta = 1$.

4. RESULTS

The algorithms discussed above were implemented on data obtained from ultrasonic inspection of piping welds in nuclear power plants (Figure 2). The problem involves classifying the ultrasonic waveforms into one of three classes – cracks, counterbores and weld roots. The database comprised one-dimensional A-scans from the three different classes obtained at an inspection frequency of 2.25 MHz and sampled at a frequency of 10 MHz. The database was divided into three parts – a training set of 135 signals (45 signals from each of the three classes), a test set of 120 signals (40 from each class) and a validation set of 72 signals (33 from cracks, 22 from counterbores and 17 from weld roots). The training data was used for initial training of the network while the test set was used for selecting patterns for retraining the network. The validation set was used as a “blind test database” and no signal from the validation set was used in training the network. The 256 point long A-scans were transformed using the discrete wavelet transform (DWT) and a set of 150 DWT coefficients were used as input to the Fuzzy ARTMAP network.

For computation of reliability measures, the Fuzzy ARTMAP network was trained using a vigilance parameter equal to 0.95 and α equal to 0.001. The network learned all the signals in one

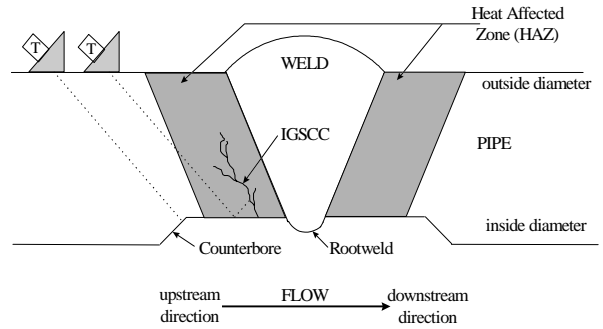


Figure 2. Inspection geometry for ultrasonic weld inspection.

epoch. Three subsequent iterations of retraining with ambiguously classified data were implemented. Quantitative measures of performance improvement are defined as (i) percentage of correct classifications and (ii) reliability of correct classifications. A sample of the results obtained with the validation set in four iterations is summarized in Table 1. The full table of results can be obtained on the Internet at [9]. In addition to the confidence measures, the table also summarizes the classification given by the network for each signal along with the true class. Result set I shows examples of four possible cases - (i) correct classification with high confidence (patterns 2, 35 and 58), (ii) correct classification with low confidence (patterns 17 and 41), (iii) incorrect classification with low confidence in the decision (patterns 10 and 53) and (iv) incorrect classification with high confidence (patterns 9 and 64).

Results of iterations II, III and IV show the improvement in the confidence levels as the network is retrained with ambiguously classified signals. The results indicate that the confidence in the decision for correctly classified signals increases as the network learns more data (signals 2, 17 and 58). In cases where the pattern was incorrectly classified initially, the network learns the correct classification and improves the degree of confidence in the decision (signals 10 and 36). In a few cases however, the signal is misclassified and the confidence in the decision increases over time (signal 14). This is usually due to overlap between the different classes.

The results of feedback learning are summarized in Table 2 and it is seen here that the number of signals correctly classified with a high confidence in the decision increases with time while those in the other three categories decrease over time. This indicates that the network is able to correct itself and evolve with time.

5. CONCLUSIONS

An automated signal classification system using the Fuzzy ARTMAP network has been described. An algorithm for estimating the reliability of the network decision based on hypothesis testing concepts is developed. Confidence measures are used to select data for retraining the network to improve its performance. Results on ultrasonic data obtained from the inspection of welds in nuclear power plant piping indicate that the network is capable of detecting its own misclassifications and retraining with them to improve its performance.

6. REFERENCES

- [1] G. A. Carpenter et al "Fuzzy ARTMAP: A neural network architecture for incremental supervised learning of analog multidimensional maps," *IEEE Trans. Neural Networks*, Vol. 3, No. 5, pp. 698-713, September 1992.
- [2] P. Ramuhalli, L. Udpa, S. S. Udpa and J. Spanner, "ARTMAP networks for classification of ultrasonic weld inspection signals", in *Review of Progress in QNDE*, Vol. 17, eds. D. O. Thompson and D. E. Chimenti, (Plenum, New York, 1997), pp. 751-757.
- [3] A. Papoulis, *Probability, random variables and stochastic processes*, McGraw Hill, New York, 1991.
- [4] S. Mitra and S. K. Pal, "Fuzzy multi-layer perceptron, inferencing and rule generation", in *IEEE Trans. Neural Networks*, vol. 6, number 1 (1995), pp. 51-63.
- [5] S. B. Shroff, *Estimation of reliability measures for artificial neural networks*, M. S. Thesis, Iowa State University, Ames, Iowa, 1997.
- [6] P. Ramuhalli, *Automatic signal classification systems using Fuzzy ARTMAP networks*, M. S. Thesis, Iowa State University, Ames, IA 1998.
- [7] P. Ramuhalli, L. Udpa and S. S. Udpa, "Use of reliability measures to improve the performance of Fuzzy ARTMAP networks," to appear in *Proc. Int'l. Joint Conf. Neural Networks*, Washington D. C., July 1999.
- [8] T. Masters, *Practical neural network recipes in C++*, Academic Press, Boston, MA, 1993.
- [9] <http://www.public.iastate.edu/~rpradeep/Articles/ICASSP2000/Results/table1.html>

Table 1. Confidence measures after four rounds of retraining (C: Crack, B: Counterbore, R: Weld Root).

		I			II			III			IV		
Ascan Index	True Class	Actual Call	Confidence in H1	Confidence in H0	Actual Call	Confidence in H1	Confidence in H0	Actual Call	Confidence in H1	Confidence in H0	Actual Call	Confidence in H1	Confidence in H0
2	C	C	0.989	0.080	C	0.991	0.085	C	0.992	0.090	C	0.993	0.093
17	C	C	0.059	0.576	C	0.994	0.553	C	0.994	0.584	C	0.995	0.611
35	B	B	0.989	0.022	B	0.991	0.036	B	0.992	0.048	B	0.993	0.056
41	B	B	0.027	0.331	B	0.032	0.313	B	0.037	0.314	B	0.040	0.288
58	R	R	0.989	0.022	R	0.991	0.037	R	0.992	0.05	R	0.993	0.060
10	C	B	0.057	0.335	B	0.057	0.317	C	0.994	0.278	C	0.995	0.222
36	B	C	0.308	0.832	B	0.021	0.263	B	0.027	0.270	B	0.121	0.818
14	C	B	0.346	0.750	B	0.055	0.325	B	0.082	0.337	B	0.141	0.838
9	C	B	0.989	0.328	B	0.991	0.311	B	0.992	0.309	B	0.993	0.284
53	B	C	0.187	0.641	B	0.034	0.292	B	0.064	0.307	C	0.042	0.510
64	R	B	0.771	1.000	R	0.910	0.441	C	0.878	0.997	C	0.861	0.999

Table 2. Summary of results for four iterations of retraining.

	I	II	III	IV
Correct classification, High confidence	23	29	40	42
Correct classification, Low confidence	26	16	17	13
Incorrect classification, High confidence	8	15	10	8
Incorrect classification, Low confidence	15	12	5	9