Mixture Odor Classification using Fuzzy Neural Network and Its Optimization through Genetic Algorithm

BENYAMIN KUSUMOPUTRO
Computational Intelligence Research Laboratory
Faculty of Computer Science, University of Indonesia
Depok Campus, INDONESIA

Abstract: - This report presents an optimized fuzzy neural network through the use of genetic algorithms. Fuzzy neural networks are widely used as it can adaptively deal with measurement of error directly, however, this neural model creates a dilemma from the fact that both large and small networks exhibit a number of disadvantages. If the network size is too small, the error rate tends to increase due to the network might not be able to approximate enough the functional relationship between the input and the target output. While, if the size is too large, the network would not be able to generalize the input data that never been learned before. The developed optimized fuzzy neural system is then applied as the pattern classifier in the artificial odor recognition system. The performance of its recognition ability is explored and compared with that of the conventional Back-Propagation neural system and the un-optimized fuzzy neural. It is shown clearly that the average recognition rate of the GA-optimized fuzzy neural system has higher recognition capability compared with that of the other neural system.

Key words: Fuzzy neural networks, genetic algorithms, artificial odor recognition system

1 Introduction

Artificial odor discrimination system has been developed using an arrayed of quartz resonators as a sensor [1], and analyzed the measurement data through various neural networks as a pattern recognizer [2][3]. Experimental results show that the recognition rate of the developed system quite high if it is used for discriminating single-based odor [3][4], however, its recognition decreased considerably when a mixture odor has to be determined. Researchers have proposed approaches to incorporate fuzzy logic element into the neural network to improve its performance, especially to accelerate the speed of convergence and higher recognition rate [5]-[8]. The advantage of using fuzzy-neural system is that it can adaptively deal with fuzziness caused by measurement of error directly, and the resulting network topology can also perform fuzzy inference rules through analyzing the values of the connection weights.

Fuzzy neural networks are well known and widely used [6][7] and some work has been done on obtaining the proper network structure and the initial weights to reduce its training time [8]. However, this type of neural system, as same with that of multilayer perceptron, has a drawback due to its huge neural connections. The optimization of the network architecture is then be done by using discrete optimization method. In this approach, each network structure is assigned an evaluation value (e.g., an estimation of the generalization error), thus formulating a discrete optimization task. A variety of methods can be applied to this class of the problems, however, genetic algorithms [9][12], have gained a strong methodology as the evolutionary approach [13]-[16] by either removing the weak connections between neurons or removing the neurons that perform weak activations. In this paper we developed a GA-optimized fuzzy neural system and applied it as a pattern classifier to discriminate mixture of odors that could not be properly classified with the previous neural system. Analysis and comparison of the proposed neural system and the other two neural systems, e.g. Back-Propagation and un-optimized fuzzy neural will be explained.

2 Artificial Odor Recognition System

Artificial odor discrimination system is being considered as systems for the automated detection and classification of odors, vapors and gasses. The odor recognition system is consists of a quartz crystal microbalance as a sensory system, and a frequency counter for measuring the shifted frequency of the sensor as it absorbed the odorant molecule, and a computer to perform neural network analysis of the data and determined the odorant category. The sensory system consists of sixteen sensors of 20 MHz quartz resonator crystals each one is covered by a sensitive membrane, a frequency counter for measuring the characteristic frequency of the sensors and a computer performing the neural network. When odorant molecules are absorbed onto the membrane, the resonance frequency of the crystals will decrease significantly, and this phenomenon is called mass-loading effect [17]. As the responses of the sensors with different membrane to an input odor are also different slightly, the output pattern from the sensor array is also specific and can be used to identify the odor.

A chamber made of Corning Glass with a volume of 1300 ml is placed in a temperature-controlled bath. Sixteen AT cut quartz crystal microbalance sensors and its oscillation circuits are attached on the inner and outer sides of the chamber lid, respectively. The water bath with the chamber and its oscillation circuits is placed in a heat-insulated box to keep the temperature at 22^0C. After a sample is injected and evaporated in the chamber, the frequency shift is measured at the equilibrium point. Then the next sample is repeatedly injected in the same manner. When the odorant molecules are adsorbed onto the membrane, the characteristic-frequency of the sensor will reduce by a certain degree, and will recover to its characteristics-frequency after deadsorbtion procedure. A 16-bit frequency counter system is used to get a higher data accuracy, and the data is transferred to the computer for further analysis.

Since the shifted frequency is proportional to the total mass of the adsorbed odorant molecules, it is possible to use this mechanism as the fingerprint of the odor concern. To increase the accuracy of the recognition system, various types of membrane-coated sensors are necessary, which is arranged as an arrayed sensor. The shift of the frequency is given by [17]:

$$\Delta F = -2.3 \text{ x } 10^6 \text{ x } F^2 \text{ x } \frac{\Delta M}{A} \qquad (1)$$

where F denotes the characteristics frequency (MHz), ΔM the total mass of the absorbed molecule (g) and A the electrode area (cm^2).

3 Fuzzy Artificial Neural Networks

Research on fuzzy systems and neural networks has received considerable attention and obtained many successful application. Fuzzy systems have demonstrated to be well suited for dealing with ill-defined and uncertain systems, while neural networks are well known for its learning capability. Incorporated the fuzzy systems into artificial neural networks is then able to enhance the capability of intelligent systems to learn from experience and adapt to changes in an environment with uncertain or incomplete data.

Fuzzy–ANN is a type of multi-layers feed-forward neural network, however, this fuzzy–ANN has some basic difference compares to the conventional multi-layers feed-forward ANN. Instead of using conventional neuron that use the sum operation ($\sum$) over the multiplication of the input and weights in connection, two types of fuzzy neurons are implemented in the F-ANN. The fuzzy neurons are fuzzy AND-neuron and fuzzy OR-neuron where *t*-norms are used for AND operation and *t*-conorms for OR operation. The other difference is the neurons in the F-ANN used fuzzy type of data as their input and the output activation, so they can deal directly with the fuzziness of the data input from experiments.

The F-ANN will construct an inference system that consists of an antecedent part and a consequent part [18]-[21]. This system could be expressed in IF-THEN rule in the form of: *If* <u>*P is M*</u>, *then* <u>*Q is N*</u>, where <u>*P is M*</u> stands for the antecedent and <u>*Q is N*</u> denotes the consequent of the IF-THEN rule. The neural network topology that provides a logical construction of the IF-THEN rule is developed by using the Pedrycz's fuzzy set-based neurons [22][23] which consists of one input layer, several of AND-neurons as a hidden layer and OR-neurons as an output layer. Suppose *T* is used to denote *t*-norm function and *S* to denote *t*-conorm function, output activation of the AND-neuron with its weight *v* and input *x* can be expressed as:

$$z(t) = \underset{i=0}{\overset{n}{T}} [v_i(t) \text{ S } x_i(t)] \qquad (2)$$

while the output activation of the OR-neuron with its weight *w* and input *z* can be written as:

$$y(t) = \underset{i=0}{\overset{n}{S}} [w_i(t) \text{ T } z_i(t)] \qquad (3)$$

The commonly used operator for T is the min-operator, while for S is the max-operator.

The IF-THEN rule of the FANN is trained using back-propagation learning algorithm, which the weight updating process is done by using formula of:

$$W_{new} = W_{old} + \Delta W(t) \quad (4)$$

$$\Delta W(t) = -\alpha \frac{\partial E}{\partial w}(t) + \beta \Delta w(t-1) \quad (5)$$

As in the Back-Propagation learning rule, the calculation of the derivative error to its connection weight in this FANN should be done using the derivative of min and max operators with respect to its connection weight. These derivatives, however, are not linear in some conditions, and one approach to solve this problem is by using the Lukasiewicz's linearization formula for those derivatives, such that [21]:

$$\frac{\partial}{\partial w_j}\min(w_j, x_j) = \begin{cases} 1, & \text{if } w_j \leq x_j \\ 1 - w_j + x_j, & \text{if } w_j > x_j \end{cases} \quad (6)$$

$$\frac{\partial}{\partial f}\max(f(w_j), M_j^*) = \begin{cases} 1, & \text{if } f(w_j) \geq M_j^* \\ 1 + f(w_j) - M_j^*, & \text{if } f(w_j) < M_j^* \end{cases} \quad (7)$$

where $M_j^* = max_i\ (M_i)$, $M_i = min(w_i,x_i))$, $f(w_j) = min(w_j,x_j)$, x_j input signal, and w_j denotes the weight of the neuron.

Derivation of the error function for the OR-neuron, then, can be written as:

$$\frac{\partial E}{\partial w_{jk}} = (y_k - t_k)\frac{\partial}{\partial f}\{\max[f(w_{jk}), M_k^*]\}.\frac{\partial}{\partial w_{jk}}(\min(w_{jk}, z_{jk})) \quad (8)$$

while for the AND-neuron:

$$\frac{\partial E}{\partial v_{ij}} = \{\sum_{k=1}^{m}(y_k - t_k).\frac{\partial}{\partial f}\max[f(z_j), M_j^*].\frac{\partial}{\partial z_j}\min[z_j, w_{jk}]\} * \frac{\partial}{\partial g}\min[g(v_{ij}), M_i^{\#}].\frac{\partial}{\partial v_{ij}}\max(v_{ij}, x_i) \quad (9)$$

where $f(z_j) = min(w_{jk},z_j)$, $g(v_{ij}) = max(v_{ij},x_i)$, $M_i^{\#} = min[g(v_{ij})]$, and $M_j^*=max[f(z_j)]$.

Result of the FANN learning process also determines neuron with weak connection, which can be removed to find the ideal topology of network. In this paper, the removing weak connection for optimizing the network structure is done by evolving the process through Genetic Algorithms that will be explained in the next section.

4 Optimization of FANN through Genetic Algorithms

The genetic algorithms (GA) are searching algorithms that developed based on natural selection of genetics and evolution. The basic element processed by a GA is the string formed by concenating substrings, each of which is a binary coding of a parameter of the search space. Thus, each spring represents a point in the search space and hence a possible solution to the problem. Each string is decoded by an evaluator to obtain its objectivr function value. This value, which should be minimized by the GA is converted to a fitness value which determines the propability of the individual undergoing genetic operators. The population then evolves from generation to generation through the application of the genetic operators. The total number of strings included in a population is kept unchanged trough generations. A simple genetic algorithm that yields good results in many practical problems is composed of these operators: *reproduction*, *crossover* and *mutation* [24].

As previously described, the size of an ANN determines its performance. It is obviously known that if the size of the network is too small then the model will not be capable to represents the desired function. However, if the size is too big, the network will memorize all the examples by forming a large lookup table, but not be able to generalize well to the inputs that have not been learning before. As other statistical models, neural networks are subject to over-fitting when there are too many parameters (i.e., weights) in the model. Other additional problem is occurred, because the network size also determines the length of learning process. If there are m examples, and $|W|$ weights, each epoch takes $O(m|W|)$ calculation time.

The optimization of the FANN through GA is initially done by making a network with a rather big and complex structure. In their optimization process, GA will search the most optimal subset of the initial basic structure. Each subset structure will become an individual in the population to be processed, which is represented by an individual string. As the knowledge representation formed in FANN is kept in its connections, the GA optimization is directed in its FANN connection weights. Preliminary experimental result showed that optimization over networks connection weights work more efficient and accurate compare with that of networks hidden neurons. The optimization procedure of GA is implemented by initially encoding the problem and defining the objective function.

The objective function of the system to be optimized is done through its fitness value. To calculate the fitness value, each individual chromosome is decoded back to a FANN structure and be trained using backpropagation learning algorithm. Weight initialization of each structure is performed using Nguyen-Widrow method [33], and the network is trained until small error tolerance is accomplished or maximum epochs is reached. After the learning process is completed, the fitness value is calculated by:

*number_of_non_activated_connections / (error_rate * number_of_epochs).*

By using fitness value evaluated by the objective function, GA searches an individual best network structure with large number of non-activated connection, small error rate, and small number of epochs, by conducting evolution process through the operation of reproduction, crossover, and mutation.

5 Experiments on Artificial Odor Recognition System

The experiments are designed to elaborate the capability of the developed odor recognition system to recognize and determine the-unlearn of mixture odors. Three types of neural classifiers, i.e. Back-Propagation neural system, fuzzy neural system (FANN) and the GA-Optimized fuzzy neural, are conducted and compared its recognition capability to recognize the unlearn odor mixture.

The data used for the learning stage and its recognition tests are obtained from 10 experiments of each two-mixture of odors, where 100 data are taken from each sensor for each experiment, and the training/testing paradigm is determined to be 70%: 30%. The two-mixture of odors are prepared by mixing *50%* of aroma-based odor (citrus, canangga and rose) and *50%* of alcohol with various gradient concentrations, ranging from *0%* to *70%.* Table 1 shows in details the used sample odors including with percentage of the alcohol concentration.

Three groups of experiments are designed, simulating the degree of difficulties on recognizing the unlearn odors, i.e. 6 classes of two-mixture odor, 12 classes of two-mixture odor and 18 classes of the two-mixture odor. Table 2 shows that 12 classes experiments are conducted by combining every two of 6 classes experiments, while 18 classes experiments are conducted by combining all of the 6 classes experiment into only one experiment.

Table 1. Two-mixture odor with various gradient alcohol concentrations

No	Type of odor-mixture	Odor-mixture with various gradient alcohol concentration
1	**CnAlch**	CnA0%, CnA15%, CnA25%,
	Cannagga based odors	CnA35%, CnA45%, CnA70
2	**RoAlch**	CiA0%, CiA15%, CiA25%,
	Rose based odors	CiA35%, CiA45%, CiA70
3	**CiAlch**	CiA0%, CiA15%, CiA25%,
	Citrus based odors	CiA35%, CiA45%, CiA70

Table 2. Experimental design for recognizing the unlearn two-mixture of odor with different number of classes

Number of Classes	Data composition of the Two-Mixture odors		
6 classes	**RnAlch**	**CiAlch**	**CnAlch**
12 classes	**RnAlch + CiAlch**	**CiAlch + CnAlch**	**CnAlch + RnAlch**
18 classes	**RnAlch + CiAlch + CnAlch**		

Results of experiment on recognizing the unlearn odor within 6 classes of two-mixture odors are depicted in Table 3. It is shown that the average recognition rate of the Back-Propagation neural is about *83.9%,* while fuzzy neural is *99.1%,* respectively. The highest recognition rate, 100% could be achieved when the GA-optimized fuzzy neural is used.

Table 4 showed that the recognition rate on recognizing unlearn odor within 12 classes of two-mixture odors. As the number of classes increases, the recognition rates of all of the neural systems are decreased. Same with that of 6 classes of two-mixture odors, recognition rate of Back-Propagation neural system is the lowest, with only *54.1%* in average. This results show that the conventional Back-Propagation neural system could not be utilized when it is used to discriminate two-mixture of odors with higher number of classes. As also shown in this table, the un-optimized fuzzy neural system still shows higher recognition rate of average *81.7%,* while the optimized fuzzy neural shows an average of *91.9%.* It is clearly shown that both fuzzy neural system, the un-optimized and the optimized fuzzy neural, have stable recognition capability compare with that of using crips data.

Table 3. Comparison of recognition rate of neural systems to discriminate unlearn odor within 6 classes of two-mixture odors

	BP	Fuzzy NN	GA-Fuzzy NN
Ci Alch	68.0%	99.3%	100%
Cn Alch	98.2%	98.2%	100%
Ro Alch	85.6%	99.9%	100%
Average	83.9%	99.1%	100%

Table 4. Comparison of recognition rate of neural systems to discriminate unlearn odor within 12 classes of two-mixture odors

	BP	Fuzzy NN	GA-Fuzzy NN
Ci Alch + Cn Alch	50.5%	83.3%	92.3%
Ci Alch + Ro Alch	41.6%	79.4%	95.5%
Cn Alch + Ro Alch	67.3%	82.3%	87.9%
Average	53.1%	81.7%	91.9%

Table 5. Comparison of recognition rate of neural systems to discriminate unlearn odor within 6, 12 and 18 classes of two-mixture odors

	6 Classes	12 Classes	18 Classes
BP	83.9%	53.1%	38.1%
Fuzzy NN	99.1%	81.7%	61.1%
GA-Fuzzy NN	100%	91.9%	83.3%

However, when the fuzzy neural systems are applied for recognizing the unlearn odors within 18 classes of two-mixture odors, as it is clearly seen in Table 5, the recognition capability of the un-optimized fuzzy neural system decreases significantly, to only *61.1%*. This recognition rate is not enough to discriminate unlearn mixture odor properly. In contrary, the GA-optimized fuzzy neural system still shows its higher recognition rate of about *83.3%*. This results show that the GA-optimized fuzzy neural system is necessary when more difficult task should be encountered in discriminating two-mixture of odors.

6 CONCLUSION

In this paper, we present a method for optimization of the weight connections in Fuzzy NN structure through Genetic Algorithms. The size and topology of the neural network are optimized in term of its cost function, such as: its error rate, learning process cost, and generalization capability. The experiment results showed that this method has successfully found the nearly optimized fuzzy-neural network, which provide higher recognition rate even with lower computational cost. Compare with that of Back-Propagation method, fuzzy neural shows higher recognition rate, however, the optimized fuzzy neural has the highest recognition capability.

ACKNOWLEDGEMENT

The authors would like to express their gratitude to Ministry of Research and Technology of Indonesia for their support funding of this research through RUT X under contract number 14.06/SK/RUT/2003

References:

[1] T. Nakamoto, Y. Suzuki and T. Moriizumi, Study of VHF-band of QCM gas sensors, Sensors and Actuators, B 84, 2002 pp. 98-105

[2] T. Nakamoto and T. Moriizumi, Odor sensor using quartz resonator array and neural network pattern recognition, IEEE Proceeding of Ultrasonic Symposium, 1998, pp. 616-622

[3] B. Kusumoputro and M. Rivai, Discrimination of fragrance odor by arrayed quartz resonator and a neural network, *ICCIMA-98*, H. Selvaraj and B. Verma (Eds), Singapore: Word Scientific 1998, 264-270.

[4] B. Kusumoputro, Development of Self-Organized Network with a Supervised Training in Artificial Odor discrimination System, in Computational Intelligence for Modelling, Control & Automation, Masoud Mohammadian (Editor), IOS Press Netherlands (1999), 57 – 62.

[5] K. Hirota and W. Pedrycs, “Fuzzy logic neural networks: Design and computations,” Proc. Int. Joint Conf. Neural Networks, Singapore, Nov. 1991, vol. 3, pp.88-21

[6] S. Horikawa, T. Furuhashi and Y. Ushikawa, “On fuzzy modeling using fuzzy neural networks with the back propagation algorithm”, IEEE Trans. Neural Networks, vol. 3, no. 5, pp. 801 – 814, Sept. 1992

[7] H. Ishibushi, R. Fujioka and H. Tanaka, “ Neural networks that learn from fuzzy if-then rules”, IEEE Trans. Fuzzy Syst., vol. 1, pp. 85 – 97, May 1993

[8] J.R. Jang, “ Self-learning fuzzy controllers based on temporal back propagation” IEEE Trans. Neural Networks vol.3 no. 5 pp. 714-723, Sept. 1992

[9] H.R. Reed, “Pruning algorithm- A survey” IEEE Transaction on Neural Networks, vol.4, no.5, Sept. 1993, 740-747.

[10] B. Andree, “Constructive learning algorithms for feed-forward neural networks”, Ph.D. dissertation, Univ. Utrecht, The Netherlands, 1995.

[11] K. Mohraz and P. Protzel, “FlexNet-A Flexible neural network construction algorithm,” Proc. 4th European symposium on Artificial Neural Networks (ESANN’96), Brussels, 1996, pp.111-116.

[12]. T. Kwok and D. Yeung, “Constructive feedforward neural networks for regression problems: A survey,” Technical Report HKUST-CS95-43, Department of Computer Science, The Hongkong University of Science & Technology, 1995.

[13] J. Yan, N. Tokuda and J. Miyamichi, “A new constructive compound neural networks using fuzzy logic and genetic algorithm: 1 Application to artificial life,” IEICE Trans. Inf. & Syst., vol. E81-D, no. 12, 1998, pp.1507-1516.

[14] S. Nolfi, J. Elman and D. Parisi, “Learning and evolution in neural networks,” CRL Technical Report 9019, University of California, San Diego, 1990.

[15] K. Shimojima, Y. Hasegawa and T. Fukuda, “Unsupervised/supervised learning for RBF-fuzzy system-adaptive rules, membership functions and hierarchical structure by genetic algorithm,” Advances in Fuzzy Logic, Neural Networks and Genetic Algorithms, Springer-Verlag, 1994, pp. 127-147.

[16] G. F. Miller, P. M. Todd and S. U. Hedge, “Designing neural networks using genetic algorithms,” Proc.3rd International Conference on Genetic Algorithms, J.D. Schaffer, Ed., San Mateo, Calif., Morgan Kaufmann, 1989, pp. 379- 384.

[17] G. Sauerbrey, Verwedung von Schwingquaren Zur Wagung Dunner Shichten und Zur Mikrowagung, *Z. Phys.*, 155, 1959, 206-222.

[18] E. Sanchez, “Fuzzy logic knowledge systems and artificial neural networks in medicine and biology,” in An Introduction to Fuzzy Logic Aplications in Intelligent Systems, Norwell, MA: kluwer, 1989, ch. 11.

[19] R.R. Yager, “Using fuzzy logic to build neural networks,” Proc. IFSA’91, Brussels, Belgium, 1991, vol. On Artificial Intell., pp. 210-213.

[20] M. M. Gupta and D. H. Rao, “On the principles of fuzzy neural networks,” Fuzzy Sets and Systems, vol.61, no.1, pp.1-18, 1994.

[21] Z-Q. Liu and F. Yan, “Fuzzy Neural Network in Case- Based Diagnostic System,” *IEEE Trans. on Fuzzy Systems*, vol. 5, no.2, May 1997, 209-222.

[22] W. Pedrycs and A. F. Rocha, “Fuzzy-set based model of neurons and knowledge-based networks,” IEEE Trans. Fuzzy Syst., vol. 1, pp. 254-266, Nov. 1993.

[23] W. Pedrycs, “Fuzzy neural networks with reference neurons as pattern classifiers,” IEEE Trans. Neural Networks, vol.3, pp. 770-775, Sept. 1992.

[24] J. H. Holand, *Adaptation in Natural and Artificial Systems*, Cambridge, Massachusetts, MIT-Press, 1995.