

SWARM OPTIMIZED MODULAR NEURAL NETWORK BASED DIAGNOSTIC SYSTEM FOR BREAST CANCER DIAGNOSIS

Bipul Pandey¹ and Nitish Garg²

¹ Samsung Research & Development Institute, Delhi, India
pandeybipul@gmail.com

² Samsung Research & Development Institute, Delhi, India
nitishgarg88@gmail.com

Abstract

Artificial Neural Networks have long been considered a simple yet powerful and elegant paradigm for solving problems related to Pattern Recognition, Machine Learning and Knowledge Discovery. However, performance of traditional, monolithic neural network based systems, suffers when faced with complex problems which involve a large number of decision variables or dimensions. Also, performance of any such system depends on the architecture of the neural network involved. The architecture usually remains sub-optimal as human expertise is generally used to design the optimal architecture. In this paper, we describe how the twin paradigms of modularity and swarm intelligence based optimization could be successfully used to overcome these concerns. Here, instead of using a single monolithic expert, we use a modular neural network where several independent neural network experts individually work upon the inputs and give their outputs which is then integrated using an Integrator (here, a Fuzzy C-Means Integrator). Also, swarm intelligence has been used to determine the connections in each individual expert for achieving an optimized architecture for each expert. This approach has been used for the diagnosis of breast cancer disease. Experimental results show that the proposed approach gives a better diagnostic ability than those of other traditional methods used.

KEYWORDS

Pattern recognition, Breast Cancer Diagnosis, Modularity, Swarm Intelligence, Artificial Neural Network

1. INTRODUCTION

There has been a vast interest in the field of Artificial Neural Network for the past few decades. Researchers have appreciated the versatility of Artificial Neural Network in dealing with computational problems and this has led to concerted efforts aimed at exploiting the abilities of neural networks for problems in the fields as varied as Biometrics, Robotics, to Financial Forecasting, Data Security, Control Systems etc. [1-4]. Their ability to generalize and learn has made them preferable for problems where other traditional methods have failed to cut ice. Another aspect that has attracted the attention of researchers is their ability to adapt according to the input conditions. Due to this, their use in the fields of Speech Processing, Data Modeling and Prediction and others has enhanced.

An Artificial Neural Network can be considered as a collection of a large number of simple processing units called neurons. There are connections with each other called as weights. Along with biases, these weights or connections are responsible for the storage of knowledge present in the training dataset used. Such a nature of the network bestows upon it massive parallelism

leading to a high computation rate. [5] For simpler problems, even single layered neural networks give efficient solutions. To deal with increasing intricacy of problems, Neural Networks have to be bestowed with more layers with larger number of neurons. This leads to a much more complex network.

However, while dealing with very complex problems, usually with a large number of input decision variables, the performance of such monolithic networks suffer [6]. This is because the computation rate slows down with increasing complexity. Also, presence of a large number of decision variables makes proper training difficult. Another major concern regarding the performance of neural networks is the dependence of their performance on their architecture. Determining the architecture is usually a human intensive task which makes them susceptible to sub-optimality. These twin concerns are especially true for problems in Biometrics, Bioinformatics etc. where a large number of attributes are usually present to be used for classification or recognition.

Here, we have tried to deal with these concerns by introducing a novel approach that uses the twin paradigms of modularity and swarm intelligence to lead to a more optimal system for complex problem. The traditional monolithic network has been replaced with a modular neural network with several individual experts to deal independently with the input and their output is then combined by an Integrator. Swarm intelligence has been used to give an optimal set of connections or weights for each of the individual experts involved. The proposed approach has been used to deal with the specific objective of building an automated Breast Cancer Diagnostic System. The System comprises of a Modular Neural Network which comprises of six individual neural networks or experts. All of these individual experts are Feed-forward neural networks with single hidden layers. The output from these experts is combined via an Integrator. The Integrator used here is Fuzzy C-Means Integrator. During training, the connections are optimized with Particle Swarm Optimization. After this, the optimized set of weights so achieved is used as connections for the network. This is done for each of the individual neural expert. After the training is done, the optimized modular neural network thus obtained is used to classify input vectors from the breast cancer database as either cancerous or non-cancerous.

2. THE STATE OF ART

There has been a vast amount of research work involving the use of artificial neural network in solving issues involving a varied domain of fields like Robotic Navigation, Biomedical, and Biometrics etc. in the past many years. However, the use of such traditional monolithic neural approach has its share of pitfalls when faced with very complex problems. This is especially true for many problems in the field of Pattern Recognition and Classification. A Pattern can be described as “opposite of a chaos; it is an entity, vaguely defined, that could be given a name” [7]. Classification systems are those that use the inherent knowledge gained during learning to map input to the desired or correct class/output. Hence, Classification is nothing but learning the procedure that maps the input data [8]. Also, the basic function of a Pattern Recognition system is to classify the input patterns. This may be supervised or unsupervised depending on the input data and technique involved. The classes are defined using prior knowledge about the data related domain [9]. A few problems where Pattern Recognition approach has been immensely successful are recognition and categorization of sounds, images, texts and features [10].

An approach that has been used for solving the above mentioned concern regarding Neural Networks is the use of a Modular Neural Network instead of using a single traditional monolithic neural network. Presence of Modularity makes the network more robust, efficient with more effective and easy learning [6]. This modular approach has been used with immense success

especially in the field of Pattern Recognition. In [11], this approach of introducing modularity has been used for person recognition using Iris Biometric Measurement. The iris images have been first enhanced and then used for the recognition task. To deal with the number of attributes derived from the iris images, Modular Neural Network has been used as the recognition engine. Further, this approach of modularity has been tested for other biometric features like signature, fingerprint and face [12]. The approach has been found to be successful in both single computer and distributed environment. This use of modularity in biometrics has been taken a step further for recognition in multimodal environment [13]. Here, face and speech biometric has been fused to give a successful bi-modal recognition system based on this approach. To deal with the massive dimensionality present because of the use of two biometric, dimensionality has been introduced by using a number of separate experts whose outputs are joined by averaging operation. The resulting high efficiency implies that negligible loss of information by attribute division has been encountered.

Modular Neural Network has also seen usage in the field of financial forecasting. In [14], a non-parametric modular neural network (MNN) model has been used to price the S&P-500 European call options. It was found that this approach has a higher generalization capacity as compared to the traditional neural models used, leading to a much higher performance by this approach as compared to a host of traditional parametric and nonparametric models used. In [19], the approach has been tested for consumer prices in USA. Here, different neural models are implemented in the individual experts of the MNN and the best among the output is selected.

This approach has also seen usage in the field of Medical Diagnostics. In [15], breast cancer diagnosis has been tried using this approach. Features selected from stepwise LDA have been used by the ensemble classifier for detection. In several other efforts also, modularity approach has been found to be successful in providing a more efficient way for diagnosis of breast cancer [16, 17]. In [18], a two-level ensemble structure has been used for detection of lung cancer. The first level judges cells as normal or cancerous, while the second level further tests the cancerous listed cells from first level. The high recognition achieved highlights the importance of using the approach.

The second concern discussed about Artificial Neural Network earlier was regarding the sub-optimality of the network architecture present as a result of human-intensive nature of the task of architecture determination. Researchers have tried to answer this concern by using evolutionary algorithms for optimizing the architecture of the neural network being used. This optimization has been applied in two different ways: optimizing the modules, layers and neurons of the network and optimizing the connections among neurons (optimized training). In [20], evolution has been used to produce a modular neural network with the best architecture. This modular neural network is then used for multi-modal biometric recognition. The results obtained show the superiority of this approach over the traditional monolithic neural network and also offer significant improvement over the use of an un-optimized modular neural network. In [21], the optimization has been done both for achieving optimal topology (modules, layers and neurons) and optimal connections in the individual experts in the modular neural network. Hierarchical Genetic Algorithm has been used for the same. The resultant system thus obtained has been used for recognition of an individual voice and face. In [22], genetic algorithm has been used for tri-modal biometric system. Here, the modular neural network has been given three individual experts for the three biometrics i.e. iris, ear and voice recognition. The MNN has been optimized with genetic algorithm and the output from the three experts combined using an optimized fuzzy integrator.

3. METHODOLOGY

The basic objective here is to tackle the twin concerns of dealing with a large number of dimensions involved in complex problems and sub-optimality of architecture of the artificial neural network. This is achieved by using the paradigms of modularity and optimization by evolution. Instead of having a single monolithic expert, we devise a Modular Neural Network having a number of individual experts that independently compute their output from the inputs and an integrator that combines these outputs to give the final output. Further, to optimize the architecture of each of these experts, we use evolutionary approach to optimize the connections among the neurons in their networks while training.

3.1. Modular Neural Network

Modular Neural Networks are considered advancement over the traditional monolithic neural networks. They are also seen as the next step in the development of biologically inspired computation. Any Computational system that has two or more sub-systems that works upon same or different inputs independently is said to show modularity. As such, Modular Neural Networks can also be seen as an example of “Divide and Conquer” approach towards computation. Also, this brings them closer to how human brain works biologically. The human brain is known to have modularity comprised of different independently working modules. This thought to have occurred because of many changes that have affected the brain during evolution.

This property of modularity imbibes in them several advantages over the traditional monolithic neural network. Prominent among them are complexity reduction in model, scalability, robustness, and computational efficiency [6]. Another major benefit over monolithic neural models is the flexibility in design and implementation for complex problems. These properties lead to successful modelling of complex problems where the number of decision dimensions is very large. Several successful implementations show the efficiency achieved through the use of this paradigm in many problems [23, 24].

3.2. Optimization of the Architecture of Modular Neural Network

There has been considerable amount of interest in this issue in the last few years. The approach preferred for this is the use of evolutionary algorithms. They iteratively optimize the problem by optimizing each of the variables involved in the objective. This, in turn, optimizes the final objective of the effort [25]. Hence, they can be used to successfully optimize the architecture of the modular neural network. This optimization can be used for both the topology (number of modules, layers and neurons) and the connections (weights) among the neurons. Genetic Algorithms have emerged as a favorite among the evolutionary algorithms for this task [19, 21]. They are known to be strong optimization agents that lead to significantly more optimal structure for the neural networks. Hence, a combination of neural network and genetic algorithm is considered to provide far better intelligent systems as compared to traditional approach of relying on neural network alone [26]. However, they are known to suffer from certain deficiencies, notable among them are: difficulty in implementation, computationally demanding, and use of a large number of operators. Some new evolutionary techniques have thus found increasing usage in place of Genetic Algorithm for optimizing of neural network architectures. Prominent among them is Particle Swarm Optimization.

3.3. Particle Swarm Optimization

Like Genetic Algorithms, Particle Swarm Optimization (PSO) is a population based stochastic optimization technique [27]. They are designed to simulate the flocking of birds. In Particle Swarm Optimization, each single solution is fashioned as a “particle” in the search space. Each of

these particles has two parameters: position ($p[]$) in search space and velocity ($v[]$). The fitness of each particle is evaluated by a fitness function. First, the population of particles is randomly initialized. During iteration, the particles are known move in the search space with their velocities. At the end of the iteration, the population is updated by using two values. These are the “pbest” and the “gbest”. Pbest is the best solution (fitness) achieved by a particle so far. Gbest is the best solution obtained by any particle so far. These two values are used to update the velocity and positions with equations 1 & equation 2 [28].

$$v[] = v[] + c1 * \text{rand}() * (pbest[] - \text{present}[]) + c2 * \text{rand}() * (gbest[] - \text{present}[]) \quad (1)$$

$$p[] = p[] + v[] \quad (2)$$

Where $\text{rand}()$ is a random number between (0, 1) and $c1, c2$ are learning factors (usually $c1 = c2 = 2$).

The iteration is carried on till either the minimum error threshold is reached or maximum number of iterations is completed.

The pseudo code of the algorithm is as given below:

```

For each particle
    Initialize particle
END

Do
    For each particle
        Calculate fitness value
        If the fitness value is better than the best fitness value (pBest) in history
            set current value as the new pBest
        End
    Choose the particle with the best fitness value of all the particles as the gBest
    For each particle
        Calculate particle velocity according equation (1)
        Update particle position according equation (2)
    End
While maximum iterations or minimum error criteria is not attained
    
```

Swarm Optimization is known to show many advantages as compared to Genetic Algorithms (GA) in general [29]. PSO does not utilizes the large number of operators that genetic algorithm does. PSO does not have genetic operators like crossover and mutation. Particles update themselves with the internal velocity. As such, it is much easier to implement as compared to GA. So, the implementation is more efficient and the global optimum is thought to reach much faster as compared to GA. But, PSO is more prone to getting struck in local minimum as compared to GA.

The framework of the proposed approach is as shown in figure 1.

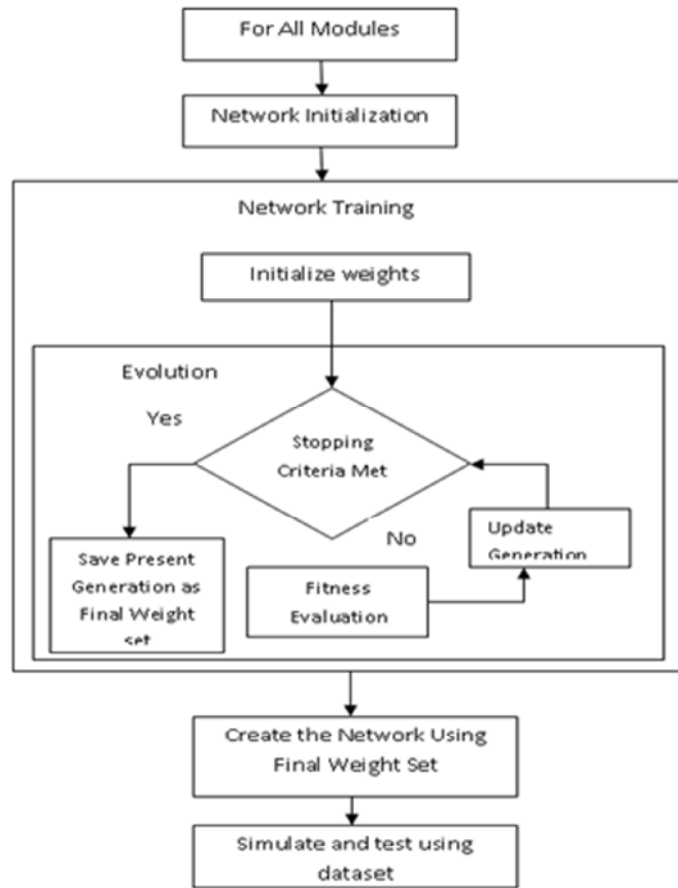


Figure 1. The Framework of the Proposed Approach

As the approach is to be applied for breast cancer diagnosis, the first step is to collect breast cancer data set from subjects. The purpose is to classify a tumor as either benign or malignant based on cell descriptions gathered by FNA image test. For our simulation, we have used the breast cancer data set from the UCI Machine Learning Repository. The data set is then divided into training and testing sets. 70% of the entire data set goes into the training set while 30% goes into the testing set. Each vector in the dataset has 30 decision attributes and one class attribute. The modular neural network to be used as the classification engine for the vectors from that dataset is now initialized. The number of individual experts (neural network modules) is kept as six. Each of the experts is then initialized as follows: Each module has one input layer with 30 nodes. One hidden layer is there with 30 nodes. The output layer has 2 nodes. Each individual neural network is a Feed-forward Neural Network. Now each of the experts is to be optimized. Here, we have relied on Particle Swarm Optimization for achieving optimized connections among the neurons. This is done by using PSO for training each of the experts. For each expert, first the weights are initialized randomly in the set $[-1, 1]$ to give a weight set equal to the number of the connections among the neurons of the expert. A number of such initialized weight sets is taken for initializing the population for PSO. Here, 25 such sets are taken as an initial population. Then, the stopping criterion is checked. Here, the only stopping criteria used is the maximum number of generations. Here, it is chosen as 100. Then the fitness is evaluated. The fitness function used here is the Root Mean Square Error (RMSE):

$$\text{RMSE} = (\sum (f(x_i) - y_i)^2 / 2)^{1/2} \quad (c)$$

Where $f(x_i)$ is the target and 'y_i' is the actual value and 'n' is the number of patterns used for training.

Then again the stopping criterion is checked. After the stopping criterion is achieved, the set with the least RMSE is taken as the set of weights for the individual neural network.

After training each of the individual experts in this way, the testing data set is used to testing the performance of the proposed framework.

4. EXPERIMENTAL RESULTS

The proposed approach is tested over the problem of breast cancer diagnosis. Here, we are given data vectors of attributes of different patients and we have to classify them as cancerous or non-cancerous. The dataset used here is the breast cancer data from the UCI Machine Learning Repository for this purpose (Wolberg, Mangasarian and Aha, 1992) [30]. The Breast Cancer data set used comprises of data vectors with a total of 30 input attributes. This database contains information about 569 patients out of which 212 have malignant tumor. The presence of malignant tumor point that the person involved has breast cancer. Attributes in the data set include radius mean of distances from center to points on the perimeter, texture means standard deviation of gray-scale values, smoothness means local variation in radius lengths, perimeter, area, smoothness (local variation in radius lengths) , compactness (perimeter² / area - 1.0), concavity (severity of concave portions of the contour), concave points (number of concave portions of the contour), symmetry and fractal dimension (coastline approximation - 1). These are measured for a total of 3 cells. The whole dataset has been divided into a training set and a testing set by taking 398 vectors as training data set (about 70% of the total data set) and rest as testing data set (about 30% of the total data set).

Matlab has been used as the platform for implementing the approach. The Modular Neural Network and PSO have been coded on the Matlab platform for implementation.

First, each of the six individual experts of the modular neural network is trained using the training dataset. The training is done by using Particle Swarm Optimization. After this, the testing set is used to classify the testing data vectors as cancerous or non- cancerous. This process is repeated fifteen times. Then, the average number of correctly identified data vectors and incorrectly identified data vectors is calculated for both training and testing data sets. This is then used to calculate mean training accuracy and mean testing accuracy. The results obtained are as listed in Table 1.

Table 1. Experimental Results Obtained from the Proposed Approach

S. No.	Property	Value
1.	Mean Training Accuracy	97.70 %
2.	Mean Testing Accuracy	97.92 %
4.	Mean Correctly Identified Instances (Training)	389
5.	Mean Incorrectly Identified Instances (Training)	9
6.	Mean Correctly Identified Instances (Testing)	167
7.	Mean Incorrectly Identified Instances (Testing)	4

Table 2. Comparison of Experimental Results Obtained from Various Approaches

S. No.	Algorithm	Training Accuracy	Testing Accuracy
1.	Proposed Approach	97.70 %	98.92 %
2.	MLP with BPA	97.10%	94.52%
3.	Fixed Architecture Evolutionary ANN	94.00%	95.27%
4.	Variable Architecture Evolutionary ANN	97.16%	95.00%
5.	Modular Neural Network	97.54%	95.60%

To compare the proposed approach, we also implemented four other popularly used approaches for breast cancer diagnosis in Matlab and applied them on the same dataset as used for the proposed approach. These are Multi-Layer Perceptron (MLP) with BPA training, Fixed Architecture Evolutionary ANN, Variable Architecture Evolutionary ANN, and Modular Neural Network. The results so obtained for each of these approaches along with that for the proposed approach are listed in Table 2 for comparison.

The Multi-Layer Perceptron (MLP) with BPA training gave a training accuracy of 97.10% while giving a testing accuracy of 94.52%. Evolutionary ANN with fixed architecture gave a training accuracy of 94.00% while giving a testing accuracy of 95.27%. Evolutionary ANN with variable architecture has provided an accuracy of 97.16% during training while providing an accuracy of 95.00% during testing. Modular Neural Network correctly identifies 97.54% of the training vectors while correctly identifying 95.60% of the testing vectors.

5. CONCLUSIONS

The high accuracy achieved by the proposed approach over the other traditionally used approaches proves that the proposed approach has been successful in dealing with the twin concerns of dimensionality and sub-optimality of the architecture of neural networks. The results also prove the significant advantage of using this approach in complex problems where decision variables are in large numbers.

This model could be extended to other such complex problems like multi-modal and single-modality biometric recognition, Robotic Navigation etc. Also, the evolution could be used for optimizing the number of modules, neurons, and layers of the modular neural network too. This

could lead to additional performance increase in terms of accuracy. Also, the PSO used could itself be optimized by heuristics prescribed in the literature. All this is planned to be done in future.

REFERENCES

- [1] Face Detection: A Survey; Erik Hjelmas, Boon Kee Low, *Computer Vision and Image Understanding* 83, pp: 236–274, 2001.
- [2] Artificial neural network models for forecasting and decision making; Tim Hill, Leorey Marquez, Marcus O'Connor, William Remus, *International Journal of Forecasting*, Volume 10, Issue 1, pp: 5-15, 1994.
- [3] Neural Networks and Statistical Models; Warren S. Sarle, *Proceedings of the Nineteenth Annual SAS Users Group International Conference*, 1994.
- [4] Neural Networks for Control; W. Thomas Miller, Richard S. Sutton, Paul J. Werbos, MIT Press, 1996.
- [5] An Introduction to Computing with Neural Nets; R. Lippman, *IEEE Transactions on Acoustic, Speech, and Signal Processing*, IEEE Signal Processing Society, Piscataway 4 (3), pp: 4–22, 1987.
- [6] “Biologically Inspired Modular Neural Networks”; F. Azam, PhD Dissertation, Virginia Tech. 2000.
- [7] Pattern Recognition: Human and Mechanical; S. Watanabe, Wiley, 1985.
- [8] Computer Systems That Learn: Classification and Prediction Methods from Statistics, Neural Networks, Machine Learning and Expert Systems; S. I. Weiss, C. Kulikowski, Morgan Kaufmann Publishers, 1991.
- [9] Data Mining: Concepts and Techniques; J. Han, M. Kamber, 2nd edition, Morgan Kaufmann Publishers, 2008.
- [10] Applied Pattern Recognition; D. Paulus and J. Hornegger, 2nd edition, Morgan Kaufmann Publishers, 1998.
- [11] Modular Neural Networks for Person Recognition Using the Contour Segmentation of the Human Iris Biometric Measurement; Fernando Gaxiola, Patricia Melin, Miguel López, *Soft Computing for Recognition Based on Biometrics*, Studies in Computational Intelligence, 2010, Volume 312/2010, pp: 137-153, 2010.
- [12] A Modular Neural Network with Fuzzy Response Integration for Person Identification Using Biometric Measures; Magdalena Serrano, Patricia Melin, *Evolutionary Design of Intelligent Systems in Modeling, Simulation and Control*, Studies in Computational Intelligence, 2009, Volume 257/2009, pp: 159-183, 2009.
- [13] Fusion of Speech and Face by Enhanced Modular Neural Network; Rahul Kala, Harsh Vazirani, Anupam Shukla, Ritu Tiwari, *Information Systems, Technology and Management, Communications in Computer and Information Science*, 2010, Volume 54, 6, pp: 363-372, 2010.
- [14] Option Pricing With Modular Neural Networks; N. Gradojevic, R. Gencay, D. Kukulj, *IEEE Transactions on Neural Networks*, Volume 20, Issue: 4, pp: 626 – 637, 2009.
- [15] Computer-aided classification of breast microcalcification clusters: Merging of features from image processing and radiologists; Y. Lo Joseph Y. Lo, Gavrielides Marios, K. Markey Mia, L. Jesneck Jonathan, *Proceedings of SPIE Vol. 5032*, 2003.
- [16] A Mixture of Experts Network Structure for Breast Cancer Diagnosis; Elif Derya Ubeyli, *Journal of Medical Systems*, Vol. 29, No. 5, 2005.
- [17] Improved Detection of Breast Cancer Nuclei Using Modular Neural Networks; F. Schnorrenber, N Tsopafsoulis, S. Iollios, M. Vossiliou, A. Adamou, K. Iyriacoui, *IEEE Engineering In Medicine And Biology*, 2000.
- [18] Lung Cancer Identification Based on Artificial Neural Network Ensembles; Zhi-Hua Zhou, Yuan Jiang, Yu-Bin Yang, Shi-Fu Chen, *Artificial Intelligence in Medicine*, Elsevier, 2002.
- [19] A hybrid modular neural network architecture with fuzzy Sugeno integration for time series forecasting; Patricia Melin, Alejandra Mancilla, Miguel Lopez, Olivia Mendoza, *Applied Soft Computing*, *Soft Computing for Time Series Prediction*, Volume 7, Issue 4, pp: 1217-1226, 2007.
- [20] Optimization of Modular Neural Networks with Interval Type-2 Fuzzy Logic Integration Using an Evolutionary Method with Application to Multimodal Biometry; Denisse Hidalgo, Patricia Melin, Guillermo Licea, *Bio-inspired Hybrid Intelligent Systems for Image Analysis and Pattern Recognition*, Studies in Computational Intelligence, Volume 256/2009, pp: 111-121, 2009.

- [21] Optimization of Modular Neural Network, Using Genetic Algorithms: The Case of Face and Voice Recognition; José M. Villegas, Alejandra Mancilla, Patricia Melin, *Soft Computing for Hybrid Intelligent Systems*, Studies in Computational Intelligence, Volume 154/2008, pp: 151-169, 2008.
- [22] Modular Neural Network with Fuzzy Integration and Its Optimization Using Genetic Algorithms for Human Recognition Based on Iris, Ear and Voice Biometrics; Daniela Sánchez, Patricia Melin, *Soft Computing for Recognition Based on Biometrics*, Studies in Computational Intelligence, Volume 312/2010, pp: 85-102, 2010.
- [23] Divide-and-Conquer Learning and Modular Perceptron Networks; .H. C. Fu, Y. P. Lee, C. C. Chiang, H. T. Pao, *IEEE Transactions on Neural Networks*, 12(2), pp: 250–263., 2001.
- [24] Automatic definition of modular neural networks; F. Gruau, *Adaptive Behavior*, 3(2), pp: 151–183, 1995.
- [25] *An Introduction to Genetic Algorithms*; M. Mitchell, Cambridge, Massachusetts: MIT Press, 1999.
- [26] *Evolving Artificial Neural Networks*; X. Yao, *Proceedings of the IEEE*, 87(9), pp: 1423-1447, 1999.
- [27] *Particle swarm optimization: An overview*; R. Poli, J. Kennedy, T. Blackwell, *Swarm Intelligence*, Volume 1, Number 1, pp: 33-57, 2007.
- [28] *Particle swarm optimization*; J. Kennedy, R. C. Eberhart, *Proceedings of IEEE int'l conf. on neural networks*, Vol. IV, pp. 1942-1948. IEEE service center, Piscataway, NJ, 1995.
- [29] *Comparison between Genetic Algorithms and Particle Swarm Optimization*, R. C. Eberhart, Y. Shi, *Evolutionary Programming VII*, Springer, 1998.
- [30] *UCI Machine Learning Repository* [<http://www.ics.uci.edu/~mllearn/MLRepository.html>] W. H. Wolberg, O. L. Mangasarian, D. W. Aha, University of Wisconsin Hospitals, 1992.