World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:5, No:4, 2011

Dynamic Threshold Adjustment Approach For Neural Networks

Hamza A. Ali and Waleed A. J. Rasheed

Open Science Index, Computer and Information Engineering Vol:5, No:4, 2011 publications.waset.org/15808.pdf

***Abstract*—**The use of neural networks for recognition application is generally constrained by their inherent parameters inflexibility after the training phase. This means no adaptation is accommodated for input variations that have any influence on the network parameters. Attempts were made in this work to design a neural network that includes an additional mechanism that adjusts the threshold values according to the input pattern variations. The new approach is based on splitting the whole network into two subnets; main traditional net and a supportive net. The first deals with the required output of trained patterns with predefined settings, while the second tolerates output generation dynamically with tuning capability for any newly applied input. This tuning comes in the form of an adjustment to the threshold values. Two levels of supportive net were studied; one implements an extended additional layer with adjustable neuronal threshold setting mechanism, while the second implements an auxiliary net with traditional architecture performs dynamic adjustment to the threshold value of the main net that is constructed in dual-layer architecture. Experiment results and analysis of the proposed designs have given quite satisfactory conducts. The supportive layer approach achieved over 90% recognition rate, while the multiple network technique shows more effective and acceptable level of recognition. However, this is achieved at the price of network complexity and computation time. Recognition generalization may be also improved by accommodating capabilities involving all the innate structures in conjugation with Intelligence abilities with the needs of further advanced learning phases.
.

***Keywords*—**Classification, Recognition, Neural Networks, Pattern Recognition, Generalization.

I. Introduction

ARTIFICIAL Neural Network (ANN) is an information processing paradigm that mimics biological nervous systems, such as the brain. Any ANN is composed of a large number of highly interconnected processing elements (neurons) working together in order to solve specific problems. ANNs, like human, learn by examples. The learning process involves adjustments to the synaptic connections (weights) that exist between the neurons. They have the ability to derive meanings from complicated or imprecise data which can be used to extract patterns and detect trends that are too complex to be noticed by either humans or other computer techniques. A trained NN can be thought of as an "expert" in the category of information it has been given to analyze. This expert can then be used to provide projections given new situations of interest and answer "what if" questions [1-4]. Moreover, ANNs have the following advantages [5]:

H. A. ali is with the Computer Engineering Department, Basrah University, Basrah, IRAQ (phone: 00964-781-3 68928; email: alsewadi@hotmail.com).
W. A. J. Rasheed, with the IIE. ORG/SRF , JORDAN (phone: 00962-777-489311; email: wali20012001@yahoo.com).

1. Self-Organization: An ANN can create its own organization or representation of the information it receives during learning time.
2. Real Time Operation: ANN computations may be carried out in parallel, and special hardware devices are being designed and manufactured which take advantage of this capability.
3. Fault Tolerance via Redundant Information Coding: Network capabilities may be retained even with partial or major network damage.
4. Adaptive learning: The ability to learn how to do tasks based on the data given for training or initial experience.
5. Pattern recognition: ANN can identify input pattern even if they were not used for training. In this case, the network gives the output that corresponds to a taught input pattern that is least different from the given pattern.

A. Architecture of Neural Networks

NN may be built in various architectures as follows:

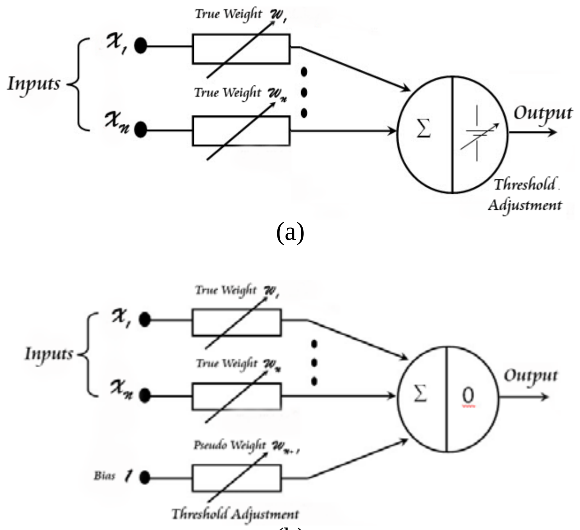
a. Feed-forward network: ANNs allow signals to travel one way only; from input to output, i.e. output of any layer does not affect the same layer. They tend to be straight forward networks that associate inputs with outputs. They are extensively used in pattern recognition.
b. Recurrent networks: ANNs can have signals traveling in both directions by introducing loops in the network referred to as feedback network. They are very powerful and can get extremely complicated. They are dynamic; their 'state' is changing continuously until they reach an equilibrium point and remain at this point until the input changes. Then a new equilibrium needs to be found. This architecture is also referred to as interactive or recurrent.
c. Competitive neural networks: ANNs that are termed "self-organizing neural nets" or "Kohonen network". A Kohonen network is a two-layered network, much like the Perceptron. But the output layer for a two-neurode input layer can be represented as a two-dimensional grid. The input values are continuous, typically normalized to any value between -1 and +1. Training of the Kohonen network does not involve comparing the actual output with a desired output. Instead, the input vector is compared with the weight vectors leading to the competitive layer. The neurode with a weight vector most closely matching the input vector is called the winning neurode.

B. Perceptrons and Parametric Adaptation

Perceptrons were the early neural nets suggested by Frank Rosenblatt [5]. They mimic the basic idea behind the mammalian visual system. They were mainly used in pattern recognition even though their capabilities extended a lot more. Moreover, Minsky and Papert [6] later described the limitations of single layer Perceptrons that led to significant decreased interest in NN. This state continued until the realization of multilevel Perceptrons with appropriate training

Open Science Index, Computer and Information Engineering Vol:5, No:4, 2011 publications.waset.org/15808.pdf

that managed to realize more complicated operations. These principles led to sophisticated architectures to appear along with their training techniques and algorithms. NN training basically is the process of adapting the connection scheme initiated with random values in a manner such that input/output association is enabled to be tagged to each other. Different training algorithms have been suggested and oriented to optimize many measures of speed, error compensations, upgrading generalization capabilities and suiting the nature of the structural elements. In general, most of the algorithms considered two adaptive parameters; weight and threshold values.



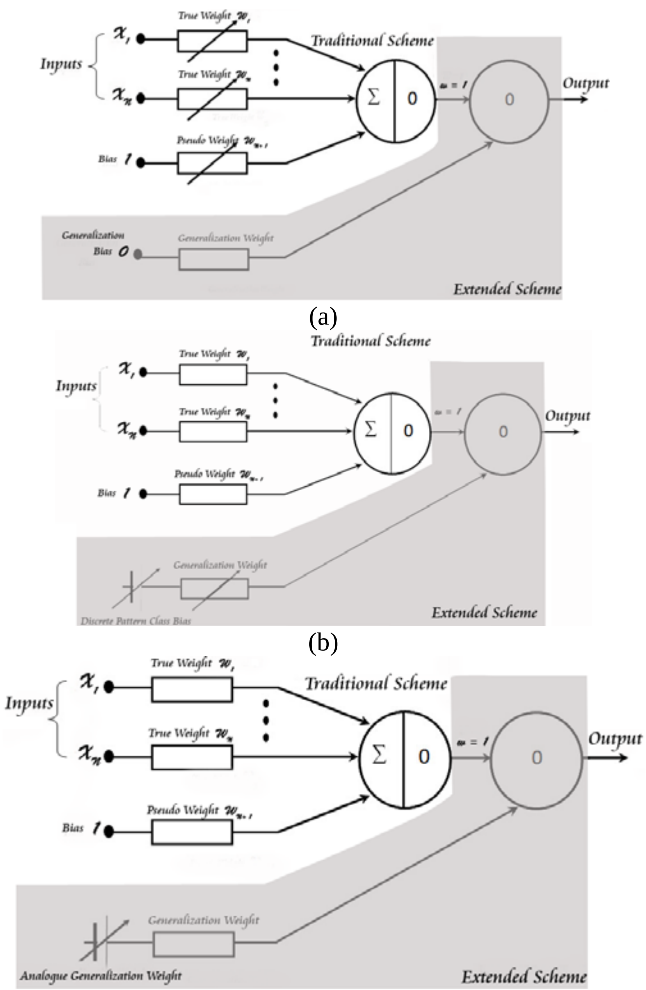
(a)

(b)

Fig 1 Perceptron threshold to weight representation

Due to the simplicity of computation, the latter (as in fig 1-a) has been represented as an extended pseudo connection weight [7] (as in fig 1-b). Hence, weights emerge as the decisive factor of adaptation or training in most of the developed models. The figure thus depicts the threshold to weight representation, where w_i, $i = 1, 2, 3,\ldots n$ are true connection weights, while w_{n+1} is pseudo weight added to accommodate the effect of threshold adjustment.

The architecture of the proposed schemes in this paper reconsiders the idea of integrating weights and threshold as common parametric attributes into their original nature, i.e. separate weight and threshold values. The weights are taken as association mean of traditional conduct, whereas threshold is regarded as an active parameter that continues to vary in accordance with input. This enables the overall structure of NN to respond dynamically even after the completion of traditional training. Furthermore, such considerations necessitate an additional phase of training that cover the requirement of generalization training as a complementary training phase. In this context, fig 2. clarifies three stages for utilizing the architecture of the proposed schemes. Fig 2-a depicts traditional training stage where the extended scheme (the generalization supported structure) is an idle structure, hence training is entirely traditional, resembling Pavlov principle, as the traditional scheme weights are the only affected parameters. Once the traditional training completed, all weights are fixed, allowing generalized training to commence as shown in fig 2-b. This phase is driven by applying discrete class bias values that discriminate each pattern class from others and tuning the generalization weights of the extended scheme. This training mode stands for the classification embedded process along the traditional association, as each pattern class is given a common bias value. Such capability in fact denotes two tasks of supervising the training process; target tagging and class definition of input. Finally, fig 2-c illustrates the implementation phase, where obviously all weights were fixed except threshold value tuning the output layer according to the input patterns.The extended scheme is modeled in the current work in two levels; single supportive layer and auxiliary net that tunes all layers of the traditional net. The contribution of this scheme can be attributed to the above detailed characteristics.



(a)

(b)

(c)

Fig 2 Perceptron adaptation modeling

C. Classification and Recognition

The most important performance criterion of NNs is their generalization ability [8]. Perceptrons work as linear classifiers as reported by Chuanyi and Sheng [9] can do a little better than

Open Science Index, Computer and Information Engineering Vol:5, No:4, 2011 publications.waset.org/15808.pdf

making random guesses. These classifiers, when combined through a majority vote, can result into good generalization performance and a fast training time. To improve classification and recognition using ANNs, many methods were developed. Ishibuchi and Nii [10] used fuzzification of input vector to avoid over fitting. Recently, a new algorithm [11] to improve the learning performance of neural network through results-feedback, called FBBP algorithm, presented by Wu and Wang, can improve NNs' generalization ability too. This FBBP-based algorithm is an inner-and-outer layer learning method in which weight value renewing plays the dominating role with the assistance of input renewing. It minimizes the error function of neural network through the dual functioning of weight value and input vector value tuning, where tuning of the input vector is similar to fuzz the input vector. This idea brings new inspiration that people had previously devoted large amounts of time for tuning weights of NNs in order to improve their performance (including the generalization ability), but lacked new ideas. Feng *et. al.* [12] suggested an approach that appropriately shrinks or magnifies input vector, thereby makes the generalization ability of NNs improved. This algorithm is called "Shrinking-Magnifying Approach" (SMA) that finds the appropriate shrinking-magnifying factor (SMF) and obtains a new neural network having better generalization ability. Ganchev *et. al.* [13] tackled generalized locally recurrent probabilistic neural networks GLRPNN, for text independent speaker verification. It is contrasted with that of Locally Recurrent PNNs, Diagonal Recurrent Neural Networks, Infinite Impulse Response and Finite Impulse Response MLP-based structures, as well as with Gaussian Mixture Models-based classifier. Although these methods can improve the classification and recognition ability of ANNs to some extent, however, the problem is generally still not completely solved. This can be attributed to the fact that the principle behavior of ANNs is of instance-based learning, i.e. they are expected to learn any relation using limited data but they should respond properly to inputs they have never seen before [14]. Therefore, it is impossible for ANNs to solve all the problems by learning from limited examples. Hence, research for developing new methods to improve classification and recognition ability of ANNs is still of great importance.

This paper presents two modifications to ANN structure with the aim of enhancing classification and recognition ability. They are both based on Pavlov and Piaget theorems [15, 16]. Basically the suggested structures incorporate an extra layer (s) added to traditional networks in order to introduce dynamical adjustment to the threshold values during implementation phase. Genetic Algorithm technique is adopted for the network training. The process involves two learning cycles; one deals with the traditional NN scheme while the other deals with the additional layer (s). The first stands for Pavlov learning assimilating capability and the second substantiates Piaget arguing through the accommodating capability. Different testing data have been used in a wide range of experiments. Adequate results of success are gained and that in turn used to approve the validity of the proposed model.

D. Background; Pavlov and Piaget Generalization Structural Interpretation

Throughout the intensive studies of human brain, neural networks appear as one of the successful and efficient abstracting models. These models prompted enormous interest of researches in psychology and physiology besides other related supporting applied sciences and medical investigations. The concrete basis, used to establish the main concept, is envisaged to lay on Pavlov theorem of conditional simple association [16]. This theorem has been conjugated with Hebb's theorem to simulate the weighting characteristics of the reticular formation of the in between cell connections of the nervous system, especially the synaptic junctions [17]. However, there were no literal interpretation to the natural processing carried out in the brain as a system with its associated behavior and constituents.

Based on the foregoing discussion and that of the psychological fundamentals, it could be stated that Pavlov theorem is faithfully interpreted and implemented with the traditional neural network models, but unfortunately to what relates Piaget's theorem, these networks failed to do so. It is known through the literature of the developed models, generalization is envisaged as an intuitive and as side effect of the connection schemes. While the significant deduction, as Piaget argued, generalization is an active learned process rather than being passive behavior of an association scheme. This might address the major obstacle stands behind improving the generalization capability of the traditional connection schemes where generalization enhancement had been attributed to data selection and net layering dimension scales as major trajectories of the efforts devoted for the developing purposes [18].

II. MATERIALS AND METHODS

Two modified models of the traditional artificial neural network ANN structure are suggested in this paper. They are introduced aiming to enhance the classification and recognition ability of ANN. They are both based on Pavlov and Piaget theorems [15,16] one involves a supportive layer as an extension to the ANN after the output layer and the other involves an auxiliary NN that support the traditional NN, as will be shown in the following.

A. Supportive Layer Scheme

This scheme involves dynamic response to data generation by simply incorporates the addition of an extra output layer with its own biasing neuron to the traditional neural network. This network consists of input layer, number of hidden layers and an output layer with biasing neuron and will be referred to hereafter as the atomic scheme. The extra layer is designed to differ from those of the common preceding ones in the connection layout by its neuronal threshold setting mechanism and control of its variations. The weights matrix of the

Open Science Index, Computer and Information Engineering Vol:5, No:4, 2011 publications.waset.org/15808.pdf

traditional layers is adjusted during the training phase but they are kept constant at the testing phase, whereas the additional layer keeps on changing its neuronal thresholds on both of the training and testing phases. Therefore, the static structure nature of the neural network is observed for the traditional network, i.e. the signals propagate from the input to the output layers via the hidden layers on fixed values of connection weights and threshold values as they imply the main data association. However, this model suggests neuronal threshold tuning of a specified layer in order to accommodate variations in the inputs that have not been seen before during training stage. Moreover, a convenient procedure has been adopted for training the whole network with the aid of Genetic Algorithm.
The schematic diagram of the proposed model is illustrated as shown in fig 3. It illustrates the traditional NN scheme followed by the added Supportive Layer. This layer extends signal propagation of the whole net in order to generate the output in two modes; the first deals with the required output of trained patterns with predefined settings, while the second tolerates output generation dynamically with tuning capability for any newly applied input.

A band selector neuron is incorporated in the last layer of the atomic scheme. This neuron is employed to initiate neuronal threshold tuning of the supportive layer. The output of this neuron is used as a bias to the supportive layer. Therefore, unlike all layers of the atomic scheme, wherein a bias input is adjusted and kept fixed afterwards, the supportive layer tunes its neuronal threshold values in accordance to the output of the band selector neuron continually. In fact, this output is made to be regulated as a function of the input of the whole model.

However, the major association attributes of the supporting layer denotes the weight values of the connections needed to link the band selector neuron to its neurons, and they are committed to the second cycle of model training. This cycle, definitely, will be commenced when the first cycle terminates and obtains the needed association in similar manner to that of the classical phase of training in traditional nets. The only difference here is that an extra output value is added to each pattern of the training set, as an additional argument representing band selector output. It must be noted here that neurons of the last layer in the atomic scheme are connected to their counterparts of the supporting layer with unity weight and in one-to-one configuration. Therefore, the complete model incorporates both, association and classification functions. However, output generation here is regarded as a result of adaptable propagation mean rather as being static mean. A proper mode can be described by the following expression.

$$\underset{\textit{(Normal operation)}\ \textit{(testing \& impl. phase)}}{\text{Output}} = \boldsymbol{f}\,(\underset{\textit{(Static attributes)}}{\text{weight scheme}},\ \underset{\textit{(Input dependent attributes)}}{\text{bias}}) \quad \ldots\ldots\ (1)$$

Where, bias works as a reference for the classification process and weight scheme works as a reference for the association structure.
Detailed training and testing of the supportive layer scheme is reported in [6].

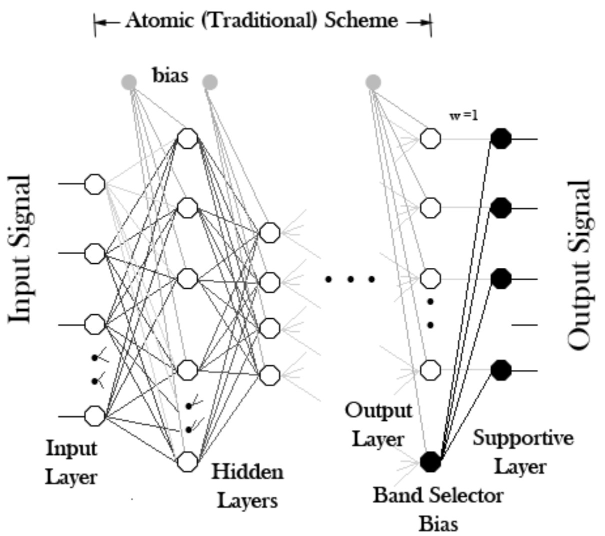


Fig 3 NN Scheme with supportive layer

a. Auxiliary Net Scheme

As an extension of the supportive layer scheme, a novel model of a dual network scheme or cooperative network is proposed. It features a full dynamic layering response instead of single layer. In this scheme, two cooperative networks are involved termed as main and auxiliary nets as shown in fig 4.

The main net accepts inputs and correspondingly generates associating output. While the auxiliary net acts as a driving net for the threshold settings of the main net. Such dual structure of neural networks addresses both pattern association and pattern classification. Hence, the main net stands for the association task whereas the auxiliary net carries out the classification task that supplements the main net. The auxiliary net acts as a real time threshold values generator corresponding to the applied input. These values are fed as bias settings to the main net in order to drift the activation functions of the scheme and hence, properly tuning the resultant output to be generated in agreement with the pattern class threshold space. Therefore, this design introduces the principle of dynamically tuning threshold values during testing and normal operations phases rather than depending only on gained network experience at the training phase. i.e. more accurate and interactive recognition possibilities are likely to be achieved.

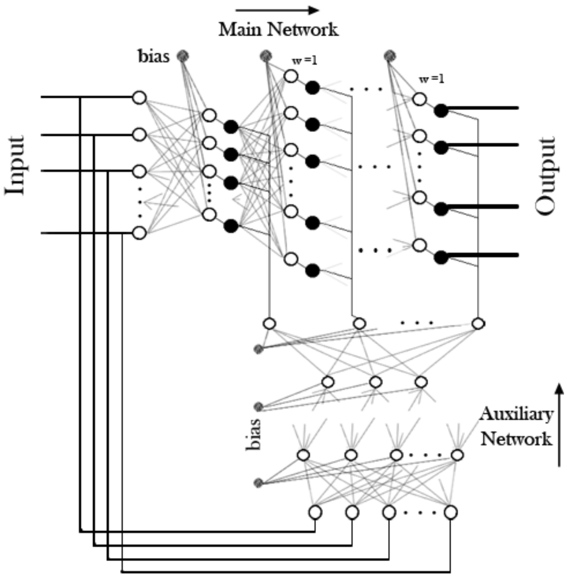


Fig 4 Overall layout of the auxiliary net model

The suggested mechanism forces the net response to follow input patterns in adapting the threshold in terms of a drifted threshold value instead of being a constant value during the data retrieval phase and hence it is made capable of involving two significant properties, namely supporting input patterns with structural classification parameters and making net operation dynamically responsive to the inputs instead of being static scheme.

III. Model description

A detailed diagram of the proposed model is shown in fig 4. It consists of two interlinked neural networks; main and auxiliary networks. The main network is a feed forward traditional structure with a dual-layer set. It accepts inputs and generates corresponding outputs. Each dual-layer consists of two sets of neurons except for the input layer which consists of one set only. The neurons of the first set are connected to the neurons of the second set through a full weight link (w = 1) with one to one configuration. Moreover, the first set adjusts its neuron thresholds by standard bias structure where a unity bias feeder is mounted and related connections are reticulated to each neuron. Whereas neurons of the second set are made such that their thresholds are adjusted by bias resources taken from the auxiliary network. Furthermore, the activation function of the second set in the dual-layer is chosen to be linear function whereas for the first set, any activation function can be implemented depending on the design requirements. The auxiliary net is a traditional scheme, taking its input signals as the same inputs of the main net, i.e. input patterns are simultaneously fed to the main and the auxiliary networks. The outputs of this net constitute a set of neurons, which are fed as bias resources to the second set of the dual-layers of the main net successively. Therefore, the number of output neurons in the auxiliary net equals to the number of the dual-layers excluding the input layer. Functionally as the model shows, there are two different responses characterizing each network independently. The main net works as associating network as it relates input-output patterns, while the auxiliary net generates a related output (set of thresholds) to the input patterns, which can be described as a classifier. Therefore, the complete model integrates both, association and classification functions cooperatively. This fact is interpreted along the following functional expressions:

$$\underset{\textit{(Normal operation) (testing \& impl. phase)}}{\text{Output}} = \boldsymbol{f}\,(\underset{\textit{(Fixed attributes)}}{\text{weight scheme}},\ \underset{\textit{(Input dependent attributes)}}{\text{bias}}) \,.\,.\,.\,.\,.\,.\,. \quad (2)$$

The bias here works again as a reference for the classification process resulting as output from the auxiliary net and main net weight scheme works as a reference for the association structure. It works to maintain that those two schemes are the outcome of two integrated training phases, a traditional training simulates Pavlov mode of learning and generalization training simulates Piaget mode of learning..

IV. Implementation

Genetic Algorithm technique is adopted to determine the overall connection structure parameters for the implementation of both schemes understudy. Moreover, although there are no anomalous restrictions to apply dedicated activation function or limit bounds to the input and output levels, it is found more applicable to use identity mode of activation function to the supportive layer (or the second layer of the dual-layer structure). This function offers efficient error compensation when output drifts are detected on the preceding layer of the traditional scheme, and thus it tends to recall the required output at the supportive layer (or the second layer of the dual-layer structure) responses throughout the training. The implementation of both schemes is considered in the following.

A.. Implementation of the Supportive Layer Scheme

A pre-organization is adopted to facilitate the training; patterns of the training set emergently are divided into two groups; the first ideally involves the most primitive pattern associations, while the second involves the patterns that are supposed to support the generalization capability. These patterns, in general, are extended by an extra argument in their related outputs. The value of this argument is given zero estimation to all patterns of the first group, and a random number to the patterns of the second group. Training can be characterized by two stages, namely the concerned structure of NN denotes the atomic structure involving the band selector expanding the last layer. The connection of the last layer of the atomic net to the supportive layer is one-to-one with the bias of each

Open Science Index, Computer and Information Engineering Vol:5, No:4, 2011 publications.waset.org/15808.pdf

Open Science Index, Computer and Information Engineering Vol:5, No:4, 2011 publications.waset.org/15808.pdf

neuron at this layer is derived from the band selector neuron resulting into the determination of the weights of the bias connections only.

B. Implementation of the Auxiliary net Scheme

This scheme is designed to implement two different modes of responses, classical response for simple association and a non classical for higher level classification. The first mode covers signal propagation from input to output along the main net. The first set of each dual-layer organization in the main net sums up its inputs applies the activation function and generates the outputs correspondingly. Meanwhile, the second set is switched into an idle state because no drifting for threshold is stimulated due to the full connection weighting linking each neuron to the preceding one. Therefore the output on the second set is simply identical to the output of the first set in each of the dual-layer configuration. In the second mode, tuning process for the activation function of the second set is stimulated to show anomalous mode of operation. When the first set of the dual-layer configuration sums up the inputs and applies its activation functions, it transfers the output to the second set. The second set is no longer being in idle mode, because it will drift the threshold in accordance with the generated bias injected from the auxiliary net via the generalization connections. Here, the model acts as an adaptive structure rather than being static. i.e. the threshold values are adaptively changing in accordance with input patterns. This dynamic threshold modification enhances adaptation of the designed network to any input drift away from the standard patterns.Before initiating the training phase, two separate tables are needed. Obviously the first table fulfills the main network training and the second is needed for the training of the auxiliary net. For the first table, pattern association (inputs and outputs) is divided into three sections. The first stands for standard pattern group, the second for non-standard pattern group and the third for performance measurement as a testing group. The second table is constructed independently in order to adjust the output of the association. In this table, the input patterns have the same number and take the same values of those of the first table, while the outputs number is assigned equals to the number of the dual-layers in the main net.

V. Results And Discussion

Probably the major problem which researchers confront in the course of testing any proposed neural network structure is the standardization issue of the compared schemes. Structural constituents of layering organization, neuronal compositions of each layer and the data of the underlying applications used are the main parameters addressed into this context. Anyway results could not be judged perfectly certain without any doubts. That is because of the absence of identical simulation programming coding, data representation and training algorithms. However, it is intended in this work to standardize the comparison parameters between the traditional nets and the presented structure as much as possible. Specifying same constituents with different examples and utilizing common data, which have been provided on Proben1 set [12], denotes all the possible trends that have been implemented to conduct the experimentation task. In this task, Genetic Algorithm is used as the training tool. Results of the two proposed schemes are summarized and discussed below.

A. Supportive layer scheme

Experiments on the supportive layer scheme involve wide range of application fields as shown in Table 1. It also lists the number of data items allocated for training and testing. The provided application data is usually divided into two sets constituting 80% and 20% ratios of the universe for training and testing purposes, respectively. Moreover, the 80% sample set is further been subdivided into two groups in order to cover the requirements of the first and second training stages of the proposed network.

TABLE I
EXPERIMENTATION APPLICATION FIELDS

Application	No. of inputs		No. of Outputs		No. of Patterns		
	Binary	Real	Binary	Real	Training	Generalization	testing
Cancer Diagnosis	-	9	2	-	350	175	174
Glass Types	-	9	6	-	107	54	53
Solar Flair	-	24	-	3	533	267	266
Majority Functions	7	-	1	-	64	32	32
Randon	-	6	-	2	60	15	25

For the cancer diagnosis example under study, various numbers of layers and neurons for the NN structure were implemented that results into different network specifications which can be summarized in Table II. For all the three different configurations listed in table 2, both the mean square error (MSE) and the recognition improvements are measured. The error measurements were conducted extensively for training and testing phases, both for the atomic net alone and the net with supportive layer. The details are listed in [6].

TABLE II
SUMMARY OF EXPERIMENTAL SPECIFICATIONS FOR THE CANCER DIAGNOSIS NN

No. of hidden layers	No. of generat-ions	Popula-tion size	Muta-tion Rate	Selected method	No. of hidden cells	Hidden & Output layers act. function	Overall Selection rate
1	600	50	2%	Rank	10	sigmoid	80%
2	800	50	5%	Rank	10, 5	Sigmoid	80%
3	1000	50	5%	Rank	14,10, 6	Sigmoid	80%

Open Science Index, Computer and Information Engineering Vol:5, No:4, 2011 publications.waset.org/15808.pdf

B. Auxiliary Net Scheme

The standard data given by Probin [19] were used for the experiment conducted on the proposed auxiliary NN structure. Similar to the case of the supportive layer scheme, different combinations of network elements such as number of hidden layers, number of neural cells in input, hidden and output layers were implemented aiming to reduce the mean square error of the overall network. Genetic algorithm technique is adopted for the network training with different mutations and number of employed generations. It is intended not to exaggerate the number of generations in order to notice the effect of generalization training and to what extent it can compensate for the error. This technique of intentional low adaptation of weights is followed by the authors in order to be able to magnify the effect of the generalization improvements caused by the support of auxiliary net despite the reduction in the number of generations.

It is noticed that in order to gain the same results with traditional structure and training scheme, the number of generations drastically exceeds this number, (practically it is > 3000). A number of 300 generations with 100 population size for the traditional net learning is chosen as initial training in all tests understudy while performance measurement of generalization is embodied via the integrating weight adaptation of auxiliary to main connecting weights (or generalization weights). And 20% mutation ratio is used to moderate the spreading of searching in the variable ace.

TABLE III
MEAN SQUARE ERROR MEASUREMENTS

Test	MSE_t	MSE_p	Recognition Improvement
1	2.7781	1.8529	33.33 %
2	0.1048	0.0510	51.34 %
3	0.0840	0.0756	10. 00 %
4	1.0640	0.5318	50.02 %
5	2.9203	1.2152	56.38 %

In our experiment, the calculated mean square error, MSE is computed for two different stages; MSE_t for main net only (running 300 generations), and MSE_p for the main and auxiliary net working together. The results are listed in Table 3, which show remarkable improvement in the network recognition capability. A recognition improvement of about 56% was noticed [20].

It is also noticed that the number of iterations in the experiment conducted is far less as compared with those reported using traditional neural network scheme only reported by Dayhoff [21]. This improvement is surely attributed to the incorporation of the auxiliary net with its classification enhancement behavior that positively affects the association action of the main net. This process obviously boosted the association action intelligence of the traditional neural network, theoretically Claimed by Caudill and Butler [22].

VI. CONCLUSION

The conducted experiments reflect the logical interpretation of the psychological postulates by modifying the traditional feed forward models with extension neuron structures utilizing Pavlov dependent scheme at their related training phases. This design resembles the assimilating human mental capabilities. Moreover, the extending neural net structures accommodate the enhanced capabilities that involve all innate structures in conjugation with Intelligence abilities but with a need for further advanced learning phases. i.e. it suggests neural network structure designs based on high level learning theorems for behavioral development. The procedure involves both assimilating capability of Pavlov and the accommodating capability of Piaget. Merging psychological learning concept with artificial neural network capabilities in order to improve intelligence representation and performance may be considered as a novel approach. To simulate the inherent behavior that exist in biological neural cells of dynamic knowledge adjustment during recognition process, the added auxiliary net in the proposed scheme facilitates continuous adjustment of threshold. This is suggested in order to accommodate variations in the input patterns away from the standard patterns. The design improvements can be attributed to the fact that the proposed model was developed based on both functional and behavioral philosophies.

REFERENCES

[1] A. Krogh and J. Vedelsby, "Neural Network Ensembles, Cross Validation, and Active Learning, Advances in Neural Information Processing Systems, G. Tesauro, D. Touretsky, and T. Leen, eds., 7, Cambridge, Mass.: MIT Press, 1995.

[2] R. V. Purohit and S. A. Imam, " Feature Extraction Pattern Recognition: A Review", MASAUM Journal Of Reviews and Surveys, 1 (1), September 2009. pp 27 – 45.

[3] B. Lerner, H. Guterman, M. Aladjem and I. Dinstein, "A Comparative Study of Neural Network Based Feature Extraction Paradigms", Pattern Recognition Letters, 20 (1), pp. 7- 14, 1999.

[4] B. Ripley, "Pattern Recognition and Neural Networks", Cambridge, Mass.: Cambridge Univ. Press, 1996.

[5] C. Stergiou and D. Siganos, "Neural Networks", http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol4/cs11/report.html.

[6] [6] W. A. J. Rasheed and H. A. Ali, "Generalization Aspect of Neural Networks on Upgrading Assimilation Structure into Accommodating Scheme", Journal of Computer Science, 5 (3), 2009, pp: 177-183.

[7] [7] R. Rojas, "Neural Networks: A Systematic Introduction", Springer-Verlag, Berlin Medleberg, New York, 1996.

[8] [8] M. T. Hagan, H. B. Demuth, M. Beale, "Neural Network Design", China Machine Press, CITIC Publishing House, Beijing, 2002.http://portal.acm.org/citation.cfm?id=249049.

[9] [9] Ji. Chuanyi and Ma, "Combinations of Weak Classifiers", IEEE Trans. On Neural Networks, 8 (1), 1997, pp: 32-42. DOI: 10.1109/72.554189.

[10] [10] H. Ishibuchi and M. Nii, "Fuzzification of input vector for improving the generalization ability of neural networks", The Int'l Joint Conference on Neural Networks, Anchorage, 4th-8th May, Alaska, 2, 1998, pp: 1153-1158. DOI: 10.1109/FUZZY.1998.686281.

[11] [11] Wu. Yan and S. Wang, "A New Algorithm to Improve the Learning Performance of Neural Network through Result-Feedback", Journal of Computer Research and Development (in Chinese), 41 (9), 2004, pp: 488-492.
DOI: 10.1109/WCICA.2004.1341928.

[12] [12] N. Feng, F. Wang and Y. Qiu, "Novel Approach for Promoting the Generalization Ability of Neural Networks", International Journal of

Signal Processing, Waset, 2 (2), 2006, pp: 131-135. www.waset.org/ijsp/v2/v2-2-20.pdf.

[13] [13] T. Ganchev, D. K. Tasoulis, M. N. Vrahatis, and N. Fakotakis, "Generalized Locally recurrent probabilistic neural networks for text independent speaker verification", Proceeding of the EuroSpeech, 3, 2007, pp: 1673–1676.
DOI: 10.1016/j.neucom.2006.05.012.

[14] [14] E. Inohira, T. Uoi and H. Yokoi, "Generalization Capability of Neural Networks for Generation of Coordinated Motion of Hybrid Prosthesis with a Healthy Arm", International Journal of Innovative Computing, Information and Control, 4 (2), 2008, pp: 471- 484. www.ijicic.org/07-044-1.pdf.

[15] [15] A. Engelbrecht, "Sensitivity Analysis for Selective Learning by Feed forward Neural Networks", Fundamental Informaticae, South Africa., 2001, http://portal.acm.org/citation.cfm?id=1219998.

[16] [16] J. Piaget, "The Origin of Intelligence in the Child", Routledge and Kegan Paul, Great Britain, 1979.

[17] [17] R. Nielson, "Neurocomputing. Addison Wesley Publishing", USA, 1989. http://portal.acm.org/citation.cfm?id=103996.

[18] [18] J. Robert, R. J. Sternberg and J. C. Kaufman, "The evolution of intelligence", Published by Lawrence Erlbaum Associates, 2002.
http://books.google.jo/books?id=xgQ1MKljVuIC&hl=en.

[19] [19] Proben1, "neural network bench mark Database Via FTP", from ftp.cs.cmu.edu [128.2.206.173] as /afs/cs/project/connect/bench/contrib./prechelt/probwn1.tar.gz or ftp.era.uka.de [129.13.10.90] as /pub/neuron/proben.tar.gz,1996.

[20] [20] W. A. J. Rasheed and H. A. Ali, "Cooperative Neural Network Generalization Model Incorporating Classification and Association", European Journal of Scientific Research, 36 (4), 2009, pp: 639-648.

[21] [21] J. Dayhoff, "Neural Network Architectures: An Introduction", Van Nostrand Reinhold, New York, 1990, ISBN 0-442-20744-1.

[22] [22] M. Caudill and C. Butler, "Naturally Intelligent Systems", MIT Press: Cambridge, Massachusetts, 1990.

Open Science Index, Computer and Information Engineering Vol:5, No:4, 2011 publications.waset.org/15808.pdf