Estimation of Wavenet Optimistic Values by using Genetic algorithm to Recognize Colored images

Dr.Entidhar Mhawes Zghair
Foundation of Technical Education
Technical Instructors Training Institute
Lecturer

Abstract:-

The purpose of this work is to use Wave net and genetic algorithms for the prediction of the presented an original initialization procedure for the parameters of feed-forward wavelet networks, prior to training by gradient-based techniques.

Genetic algorithm global optimization techniques can be providing near optimal value for learning rate, translation, dilation, and hidden nodes. Then the net work can be learning in short learning time (reduce %50 from number of iteration) with desired accuracy is more than the desired accuracy without using genetic algorithm.

Then it uses the genetic algorithm method to determine a set of best wave net whose translation and dilation parameters with optimal value of learning rate. The hidden nodes are used as initial values for subsequence training. The results show high accuracy in classification and recognize of color medical images convert as gray level of applying of (WN).

Keywords: - Genetic algorithm ,Neural Network ,Back Propagation Algorithm, Wavelet

1. Introduction:

Intelligent systems cover a wide range of technologies related to hard sciences, such as modeling and control theory, and soft sciences, such as the artificial intelligence (AI). Intelligent systems, including neural wavelet networks (NWN), and wavelet techniques, utilize the concepts of biological systems and human cognitive capabilities. The major drawbacks of these architectures are the curse of dimensionality, such as the requirement of too many parameters in NWNs, the use of large number of wavelets, and the long training times, etc. These problems can be overcome with network structures, combined two or all these systems [1]

Wavelets are mathematical functions that cut up data into different Frequency components, and then study each component with a resolution matched to its scale. The fundamental idea behind wavelets is to analyze the signal at different scales or resolutions, which is called multiresolution. Wavelets are a class of functions used to localize a given signal in both space and scaling domains. A family of wavelets can be constructed from a mother wavelet. Compared to Windowed Fourier analysis, a mother wavelet is stretched or compressed to change the size of the window. In this way, big wavelets give an approximate image of the signal, while smaller and smaller wavelets zoom in on details. Therefore, wavelets automatically adapt to both the high-frequency and the low-frequency components of a signal by different sizes of windows. Any small change in the wavelet representation produces a correspondingly small change in the original signal, which means local mistakes will not influence the entire transform. The wavelet transform is suited for nonstationary signals (signals with interesting components at different scales) [2].

This makes wavelets interesting and useful.

They therefore do a very poor job in approximating sharp spikes. But with wavelet analysis, we can use approximating functions that are contained nearly in finite domains. Wavelets are well-suited for approximating data with sharp discontinuities [1].

In the present paper a network initialization procedure that takes advantage of the properties of discrete wavelet frames in order to improve the training efficiency of continuous wavelet frames. It will be focus on wavelet frames rather than on orthogonal wavelet bases, this is because the letter must comply with conditions that are seldom feasible.

The present the wavelet networks will be present first in addition to their architectures and the principle of the training. The difference between discrete and continuous wavelet frames will be emphasized, since both approaches will be used at different stages of wavelet network training. After outlining the problem of parameter initialization of a wave net. The genetic algorithm, and subsequently describe the proposed method. Finally, a data base library image takes to experimental the ability of network.

Wavelet theory is a useful tool for function estimation and signal processing. However, wavelets are usually limited to small dimensions because constructing and storing wavelet basis of large dimension is very difficult. In recent years, neural networks, which can handle problems of large dimensions efficiently, are introduced into the wavelet theory.

The combination of wavelet theory and neural networks has led to the development of wavelet networks. Wavelet Networks (WN) are feed forward neural networks using wavelet as activation function. WN has been used in classification and identification [3].

2. Review of Wavelet Network

The origin of wavelet networks can be traced back to the work by Daugman (1988) in which Gabor wavelets were used for image classification. Wavelet networks have become popular after the work by Pati (1991, 1992), Zhang (1992), and Szu (1992). Wavelet networks were introduced as a special feed forward neural network. As mother wavelet, they use the following function [4]:

$$\psi(a,b) = |a|^{-1/2}\psi\left(\frac{(x-b)}{a}\right), \mathrm{a} > 0, \mathrm{b} \in \Re \text{ -----------(1)}$$

Where:

$\{a\}$: The parameter is the dilation (scaling) parameter

$\{b\}$: The translation (shifting) parameter,

$(1/\sqrt{a})$: The constant is used for energy normalization across different scales,

$\Re$: The vector space of real numbers.

Both the dilation and translation parameters ($a \& b$) vary continuously over$\Re$, as mentioned above, with the constant ($a \neq 0$).

The wave-net algorithms consist of two processes: the self-construction of networks and the minimization of error. In the first process, the network structures applied for representation are determined by using wavelet analysis. The network gradually recruits hidden units to effectively and sufficiently cover the time-frequency region occupied by a given target.

Simultaneously, the network parameters are updated to preserve the network

Topology and take advantage of the later process. In the second process, the approximations of instantaneous errors are minimized using an adaptation technique based on the LMS algorithms. The parameter of the initialized network is updated using the steepest gradient-descent method of minimization. Each hidden unit has a square window in the time-frequency plane. The optimization rule is only applied to the hidden units where the selected point falls into their windows. Therefore, the learning cost can be reduced [4], [5].

2.2 Proposed neural Wavelet Network (NWN)

The term "wavelet" as it implies means a little wave. This little wave must have at least a minimum oscillation and a fast decay to zero, in both the positive and negative directions, of its amplitude. This property is analogous to an admissibility condition of a function that is required for the wavelet transform. Fig.1a is an example of a wavelet called "Morlet wavelet" named after Jean Morlet, the inventor [6].

Sets of "wavelets" are employed to approximate a signal and the goal is to find a set of daughter wavelets constructed by a dilated (scaled or compressed) and translated (shifted) original wavelets or mother wavelets that best represent the signal. So, by "travelling" from the large scales toward the fine scales, one "zooms in" and arrives at more and more exact representations of the given signal.

Figs. (1a,b, c and d) display various daughter wavelets where *a* is a dilation and *b* is a translation corresponding to the Morlet mother wavelet[6].

The mother wavelet must satisfy the following admissibility condition and any admissible function can be a mother wavelet.

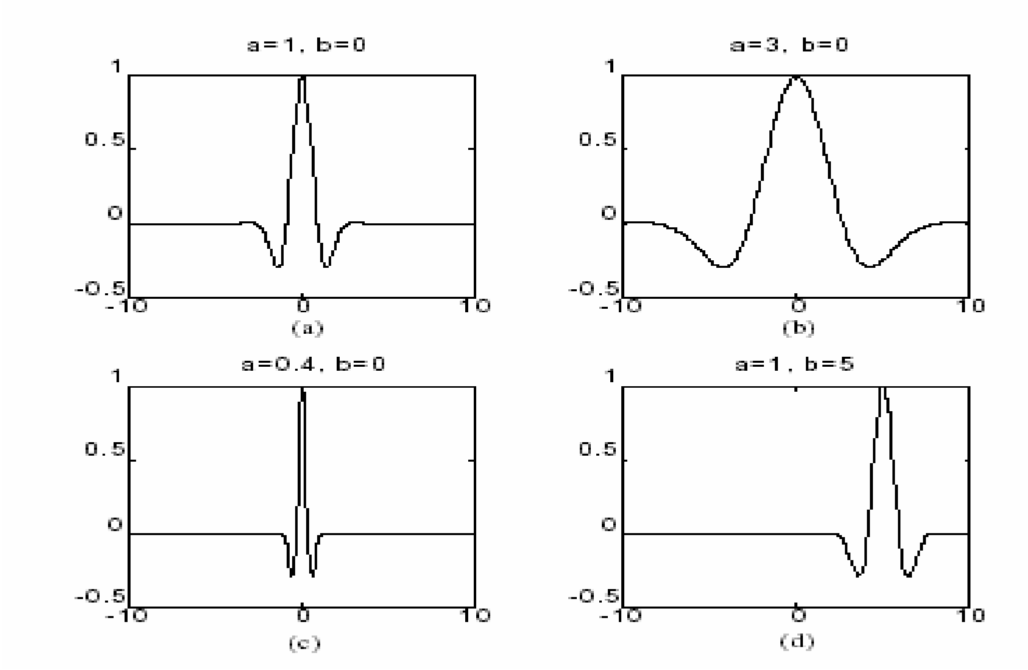


Fig. 1. Dilated and Translated Morlet Mother Wavelets.

$$C_h = \int_{-\infty}^{+\infty} \frac{|H(\omega)|^2}{|\omega|} d\omega < \infty \text{ --------(2)}$$

Where: - ***H* (ω)** is the Fourier transform of *h* (*t*).

The constant ***C_h*** is the admissibility constant of the function *h(t).*

The wavelet transform of a function ***f*** with respect to a given admissible mother Wavelet ***h*(t)** is defined as:

$$wf(a,b) = \int_{-\infty}^{\infty} f(t) h * a,b(t) dt \text{ ---------(3)}$$

Where ***** denotes the complex conjugate. However, most wavelets are real valued. The daughter wavelets are generated from a single mother wavelet ***h(t)*** by dilation and translation:

$$h_{a,b}(t) = \frac{1}{\sqrt{a}} h\left(\frac{(t-b)}{a}\right) \text{ ---------------------(4)}$$

Where **a>0** is the dilation factor and **b** is the translation factor [6].

In 1958, Rosenblatt demonstrated some practical applications using the perceptron. The perceptron is a single level connection of McCulloch-Pitts neurons sometimes called single-layer feed forward networks. The network is

Capable of linearly separating the input vectors into a pattern of classes. In such an application, the network associates an output pattern (vector), and information is stored in the network by virtue of modifications made to the synaptic weights of the network [3]. Fig.2 illustrates perceptron, which is described by:

$$y_i = f\left(\sum_{j=1}^{N} wijxj - vi\right) \text{ ------------(5)}$$

Where **i = 1, 2... M** (output nodes), **j = 1, 2**… (Inputs).

Rosenblatt derived a learning rule based on weights adjusted in proportion to the error between the output neurons and the desired output (target). The weight adaptations are given by

$$\Delta w_{ij}(n) = \mu[y\hat{}_i(n) - y_i(n)] * x_i(n) \text{ ---------(6)}$$

Where ***i* = 1, 2… M** (output nodes), ***j* = 1, 2... N** (inputs), ***$y\hat{}_i$*** is the desired output at node *i* of time *n* and **μ** is a learning rate.

Wavelet transform have proven to be very efficient and effective in analyzing a very wide class of signals because of their attractive feature. Wavelet transform describes signals in terms of their local shifts. Thus, they provide a time-frequency representation, the generation of wavelets and a calculation of all wavelet expansions employing summation, not integrals, that is well matched to be implemented by digital computers [1].

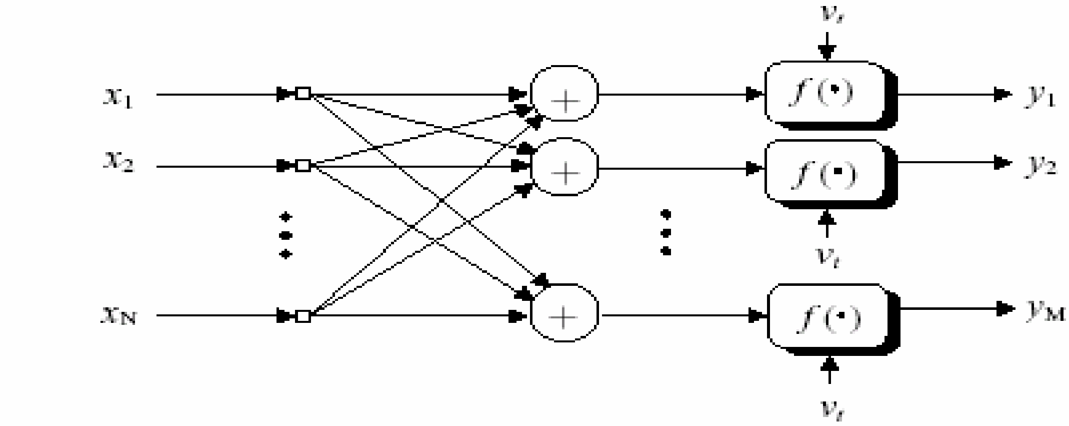


Fig. 2. Single-Layer Perceptron Feedforward

2. 3 Wavenet networks Structure:

ANNs are mathematical constructs that try to mimic biological neural systems. Over the years, ANNs have become recognized as powerful pattern recognition techniques. The networks are capable of recognition spatial, temporal or other relationships and performing tasks like classification, prediction and function estimation. The structure of wavenet networks shown in Fig 3 [4][7].

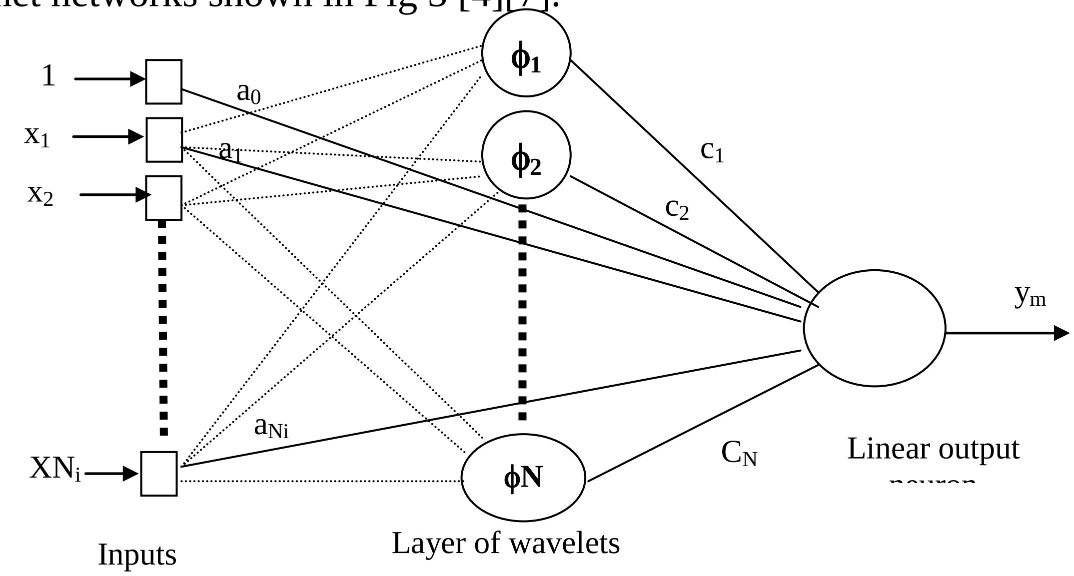


Fig 3: Wave net structure

Considered that, 1 input represent the bias to the network. x_1…………………….. x_{Ni} represent gray level value which taken from read the medical data base library. a_0……………………. a_{Ni} input weight from input layer to output layer. ϕ_1……………………. ϕ_{Ni} activation function as bases mother matrix, the base function considered as Mexican function, which defined:

$$\phi(x) = (1-x)\exp(-0.5 * x^2) ---------(7)$$

Where x (new)=(x old)-translation/dilation

Translation and dilation optimal or best value taken from genetic algorithm generation. c_1……………………. CNw output weight from hidden layer or wavelets layer to output layer. For every input vector there is a target out given by designer (Supervised). The feed process to find actual output as:

$$actualoutput = y(x) = \sum_{j=1}^{N_w} c_j \phi_j(x) + \sum_{k=0}^{N_i} a_k x_k ------(8)$$

Wavelet networks training consists in minimizing the usual total sum Squares error TSSE:

$$TSSE(\theta) = \frac{1}{2} \sum_{n=1}^{N} (t_p^n - y^n)^2 -----------(9)$$

Where θ includes all network parameters to be estimate by genetic algorithm: hidden nodes, Learning rate, translations, and dilations. The connection weight direct from input to output and weight from layer wavelets and output initialize randomly using Gussian distribution curve to give value between –1 to 1 except the zero [8][9].

3. Genetic Algorithm:

Genetic algorithms (GAs) are search procedures based on the mechanics of natural selection and natural genetics. The GAs work as follows (Fig. 4): first code each individual in the search space as a finite-length string (chromosome), which consists of the characters (genes) 1's and 0's. A set of the chromosomes is a population. Second evaluate each chromosome with the fitness by an objective function (fitness function). Third apply basic operations to the population of the chromosomes. The basic operations compose of selection, crossover, and mutation. Selection is an operation which selects the chromosomes according to their fitness values. The higher the chromosome's fitness value is, the higher its probability to produce offsprings into the next generation is. Crossover creates new chromosomes by swapping genes of parent chromosomes for each pair of selected chromosomes. Mutation changes the gene of the chromosomes with a probability (mutation probability), 1 to 0 and 0 to 1. This process is repeated until terminal conditions are satisfied.[10]

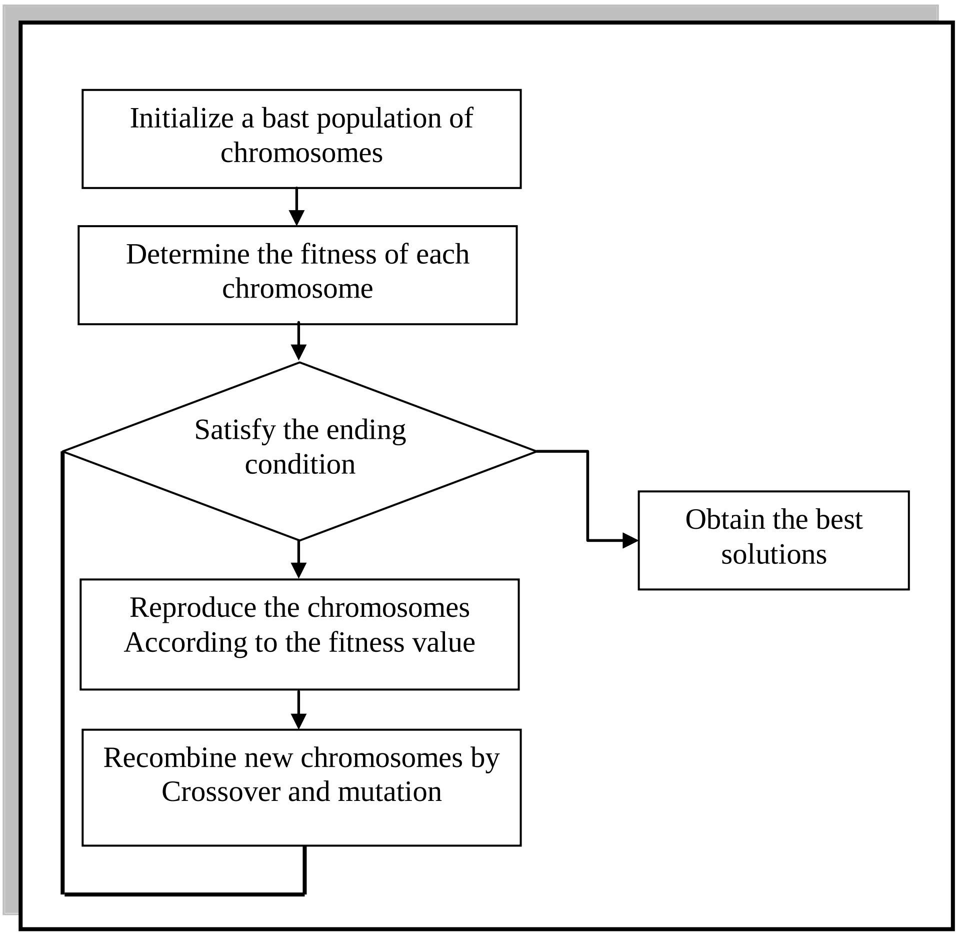


Fig 4: Basic step finding the by solution by GA

3.1 The Chromosome Encoding Structure

Since the *CT* matrix is a signed real-valued matrix, the real value encoding is used for chromosomes representation. Since the *CT* is a 3×3 matrix, the chromosome should contain nine genes. The *CT* matrix and its chromosome encoding that is used in this work is given in Equation (10) and (11).

$$CT = \begin{bmatrix} C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{bmatrix} \quad .... \ (10)$$

$$CT\_Chromosome = [C_{11} C_{21} C_{31} C_{12} C_{22} C_{32} \ C_{13} \ C_{23} C_{33}] \quad .... \ (11)$$

It must be noted that the encoding shown above is performed via a double vector and not a binary string. For more information on how to use double vectors asGA chromosomes the reader is referred to [11][12]

3.2 The Fitness Function

The fitness functions are applied to assess the effectiveness of each GA generated *CT* matrix which is denoted as GA. fitness function is formulated based on the theoretical and mathematical knowledge Root Mean Squared Error (RMSE) Let f be an M×N image and f' is the corresponding reconstructed image after compressing and decompressing of image f, then the RMSE is given by:

$$e = \sqrt{\frac{1}{MN}\sum_{x=1}^{M}\sum_{y=1}^{N}[f(x,y) - f'(x,y)]^2} ------(12)$$

The fitness function assigns to each individual in the population a numeric value that determines its quality as a potential solution. The fitness denotes the individual ability to survive and to produce offspring. In our case, the fitness is the number of regions that can be coded with RMS error less than a fixed value. The RMS is the distance between the region and the domain block determined by its coordinates and the proposed algorithm has been evaluated on various images with different sizes. The following results are obtained for 32x32 image. We present the obtained results for different configurations of error limit, number of iterations and population size. [13]

3.3 Basic components, a GA works as follows.

It starts by using the initialization procedure to generate the first population. The members of the population are usually strings of symbols (chromosomes) that represent possible solutions to the problem to be solved. Each of the members of the population for the given generation is evaluated, and, according with its fitness, it is assigned a probability to be selected for reproduction.Using this probability distribution; the genetic operators select some of the individuals. By applying the operators to them, new individuals are obtained. The mating operator selects two members of the population and combines their respective chromosomes to create offspring. The mutation operator selects a member of the population and changes some part of its chromosome. The elements of the populations with the worst fitness measure are replaced by the new individuals. The algorithm continues until some termination criteria[10][13].

4. Data base Library:

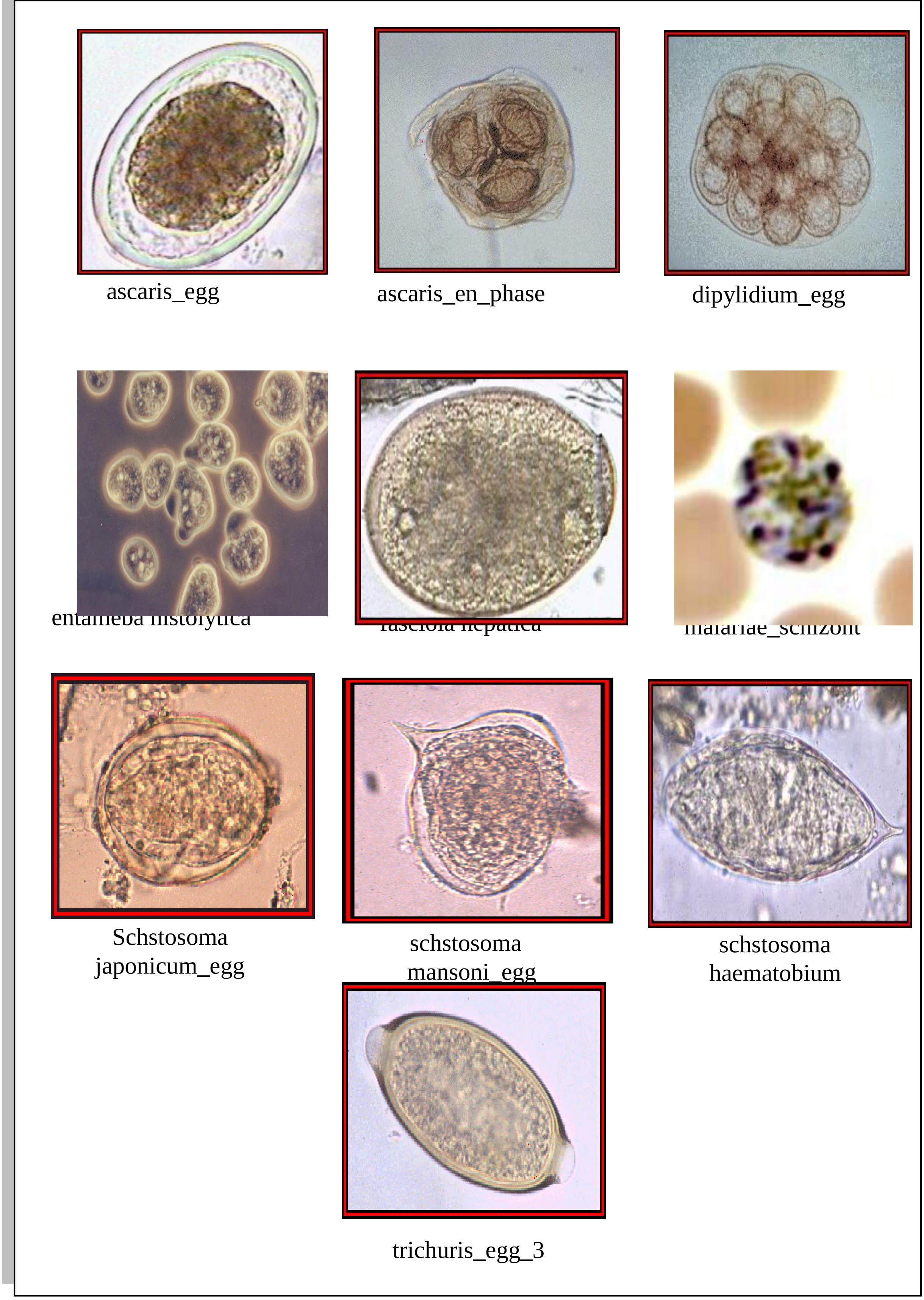


Figure 5. Sample from data base medical images.

5.1 Experiment results:

5.2 Genetic Wavenet algorithm:

The algorithm applied as:

Step 1: Function to be optimized (Minimum error)

TSSE = ½ (target output – actual output) ^{2} (13)

Where Target output given by designer, Actual output calculates using the flowing equation: -

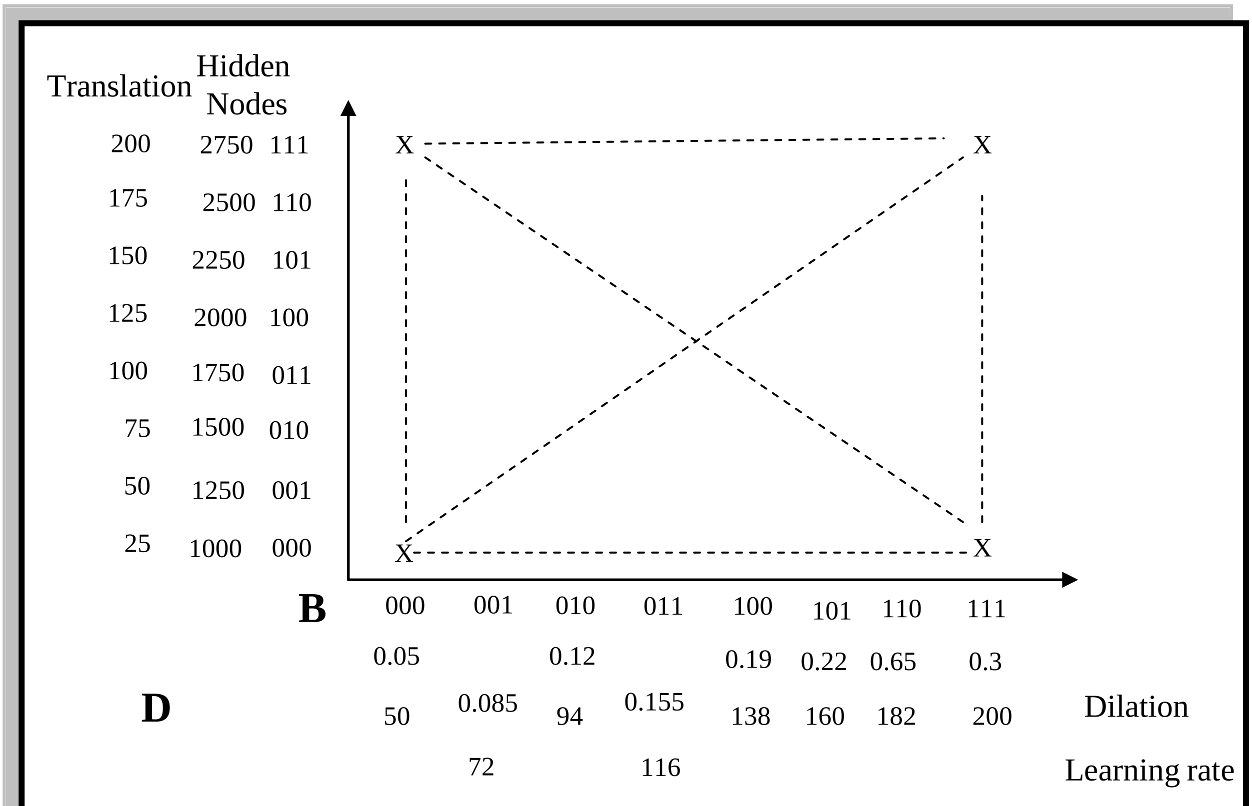
$$actualoutput = y(x) = \sum_{j=1}^{N_w} c_j \phi_j(x) + \sum_{k=0}^{N_i} a_k x_k \tag{14}$$

Weight (new) = Weight (old) + Δ weight (15)

The important parameters in above equation are Learning rate, number of hidden nodes (because we considered momentum rate =zero), Translation, and dilation.

Step 2: Phenotype to Genotype conversations.

In genetic algorithm, the phenotypes (parameter of problem) are usually decoded to genotype (genetic parameters), we used binary genotype as the shown in figure below:



B: Binary, **D:** Decimal

Fig6: Discretization of the search space using binary representations.

Step 3: Chromosome Formulation

Once the genotype is defined, the strings are concatenated to form the chromosome of the function. Length of chromosomes which used = 12 (can be represent as 000 100 101 110).

Step 4: Population Formulation

A set of these chromosomes forms the population. (Considered Population size =8) then:

POP [0] =
000 101 101 111
011 100 010 011
111 011 000 101
010 110 000 111

Step 5: Generations.

Next the optimization operators are applied to the population and based on certain criteria the population is altered. This iteration takes the search to the next generations.

5.3 implementation results:

Considered ten medical color images with size **(100x100)**, cross over rate=**0.167**, mutation rate=**0.05**;popsize = **8**,length of chromosomes **12**, binary genotype , **1000** wave lone, **Morlet function** search for best value of Translation, Dilation, Learning rate, and hidden nodes.

Table 1: Responses of Genetic algorithm.

Population	**Learning rate**	**Hidden**	**Translation**	**Dilation**	**Fitness**
	Generation One				
Pop[1] [Lchrom]	0.079	1276	32	72	0.25
Pop[2] [Lchrom]	0.073	2125	112	124	0.026
Pop[3] [Lchrom]	0.064	1161	78	60	0.004
Pop[4] [Lchrom]	0.165	2500	101	114	0.001
Pop[5] [Lchrom]	0.165	1465	40	76	0.842
Pop[6] [Lchrom]	0.063	1276	32	86	0.09
Pop[7] [Lchrom]	0.063	1748	88	124	0.013
Pop[8] [Lchrom]	0.165	1161	88	76	0.00004
	Generation Two				
Population	Learning rate	Hidden	Translation	Dilation	Fitness
Pop[1]	0.069	1276	40	124	0.0007

Population	Learning rate	Hidden	Translation	Dilation	Fitness
	Generation One				
[Lchrom]					
Pop[2] [Lchrom]	0.073	2125	112	60	0.0066
Pop[3] [Lchrom]	0.064	1161	111	124	0.036
Pop[4] [Lchrom]	0.165	2500	76	76	0.000041
Pop[5] [Lchrom]	0.165	1161	76	76	0.000041
Pop[6] [Lchrom]	0.063	1276	32	124	0.029
Pop[7] [Lchrom]	0.063	1748	111	86	0.0093
Pop[8] [Lchrom]	0.075	1000	25	100	0.046

Table 2: Result obtained best and worst solution.

Best fitness	0.0012	0.0004	0.00004	0.00004	0.00004
Worst fitness	0.842	0.036	0.026	0.032	0.015
generation	0	2	4	6	8

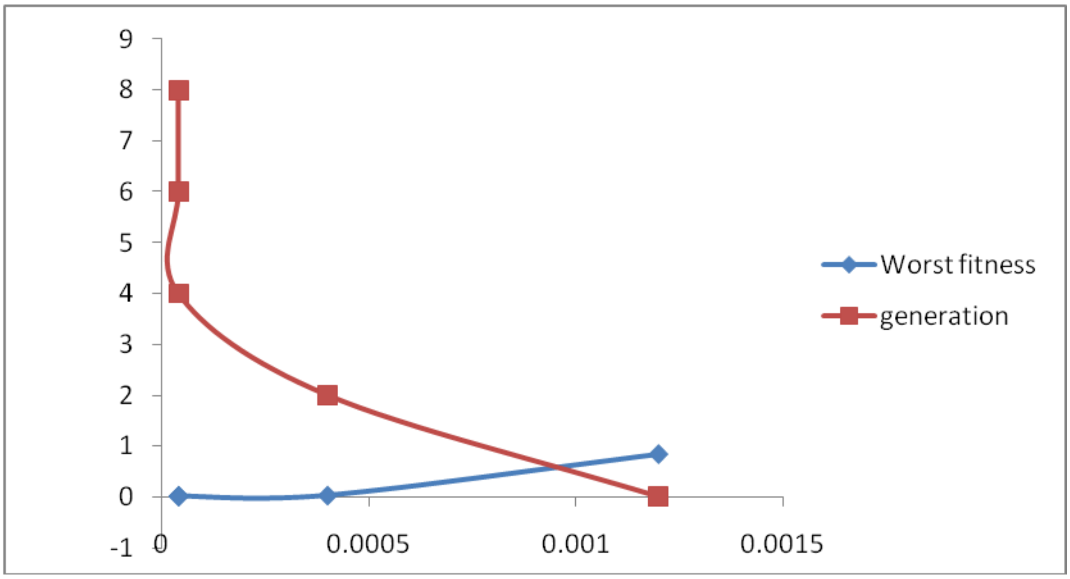


Fig. 7 Relationships between worst fitness with number of generation

Training parameters for data base library for Ten color images (100×100)

Hidden = (output + input)/2=1000**,** Learning rate=0.07, Desired accuracy=90%

Where x (new) = (xold – t)/d, t: translation, d: scaling.

Number of wave-lone = 1000.

Type of wavelet function	Equation	TSSE	Number of iteration
Polog1	**$F(x)=-x \exp(-0.5*x^2)$**	**0.066**	**30**
Morlet	**$F(x)=\cos(w_0x) \exp(-0.5x^2)$**	**0.033**	**15**
Polog4(mexicanhat)	**$F(x)=(1-x^2) \exp(-0.5*x^2)$**	**0.0044**	**60**

The learning epoch will terminate when the desired normalized error of 0.033 is reached. Figure 8 will describe the results of the wave net network performance employing Morlet.

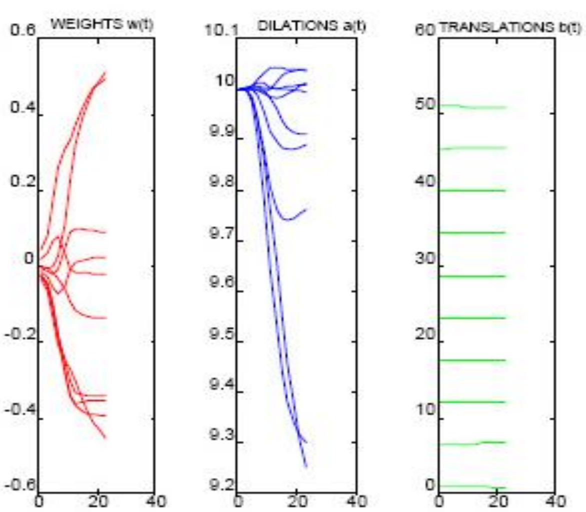


Fig. 8**. a** Wavenet Parameter Updates with 15 Morlet Wavelets

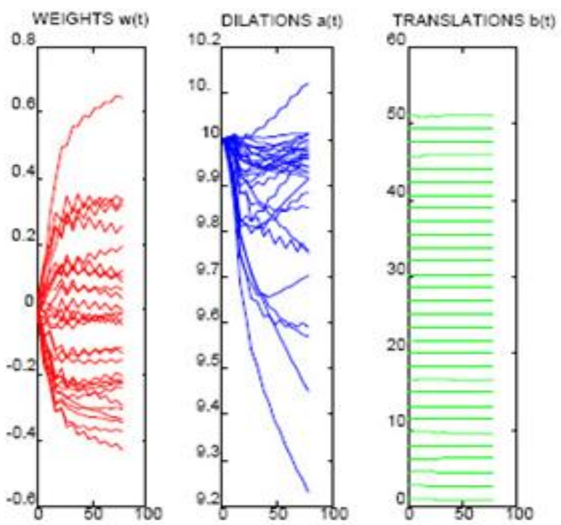


Fig. 8b. Wavenet Parameter Updates with 25 Morlet Wavelets

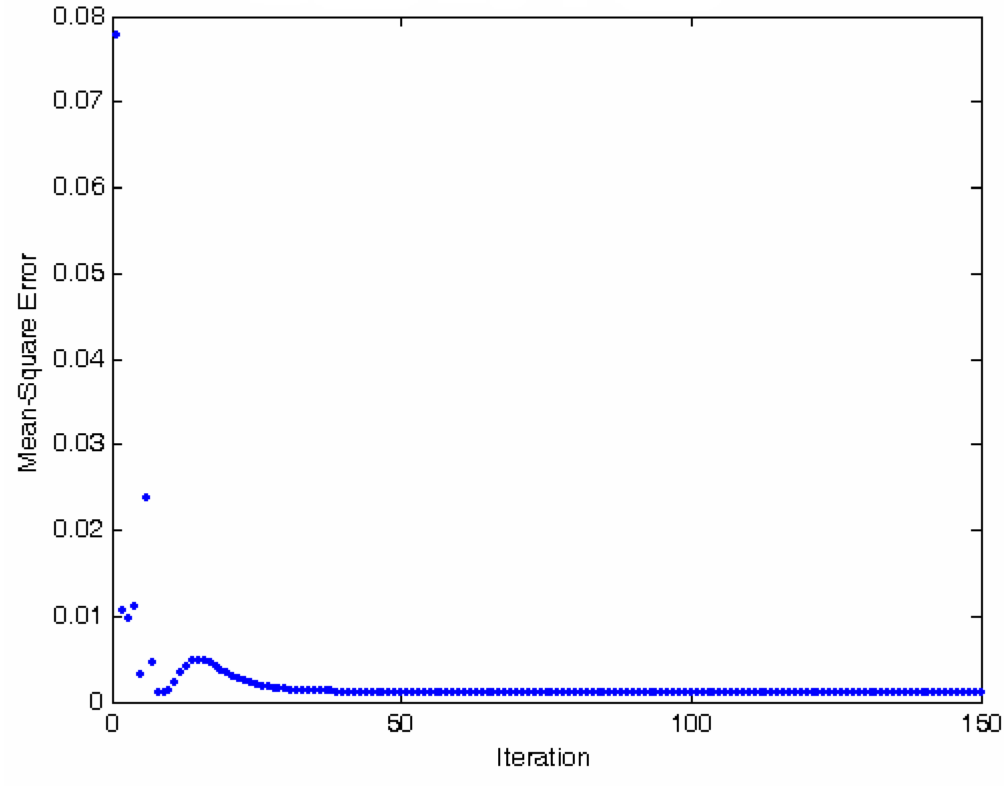


Fig 9 Mean-Square Error per learning iteration

Figs.8a and 8b group the learning performance of the *wavenet* network using 15 and 25 Morlet wavelets, respectively. We can conclude that the *Wave net* network composed of more wavelets can reach initial convergence With reference to the number of iterations very rapidly. However, to reach the desired error goal 0.033, networks with a large number of wavelets cannot converge easily and the error performance starts to oscillate.

6. Conclusion:

In this paper, an advanced wavelet network, called Neural Wavelet Network is presented as an interesting alternative to wavelet networks. This technique absorbs the advantage of high resolution of wavelets and the advantages of learning and feed-forward of neural networks. The algorithm of function identification is designed and implemented using Matlab 7 and Visual c^{++} tool. It can be concluded that this structure achieves an approximation assuming reasonable choice of the number of wavelets and mother wavelet basis functions.

The Neural Wavelet Network (NWN) structure is implemented and an example of a recognize color medical images is carried out to verify this implementation. The Neural Wavelet Network is proved to recognize a color medical image Artificial Neural Network (ANN) has some problem;

such as number of hidden unit, value of learning rate, initial weights. The traditional method (such as trial and error) to find optimal these values is spent in time and the result is not accurate. Wavelet neural network was used the powerful of wavelet transformation to reduce the number of iteration in the learning phase. But need to find the optimal or best value to the translation and dilation parameters, which used in the transformation. Genetic algorithm (GA) can be also used to find the nearest initial weight to the feed forward back propagation algorithm. Note the noise in the color medical images can be acceptance in the level below to 30%.

In general, several generations (successive applications of the operators) are applied. A new population is the result of each generation. If the GA functions the user should notice an improvement in the general fitness of the populations.

REFERENCES:

[1] Parker, T. E., and Matsakis, D.:" **Time and frequency dissemination advances in GPS transfer techniques"**. GPS World Magazine, pp. 32-38, 2004

[2] Mosavi, M. R.: "**Comparing DGPS corrections prediction using neural network,fuzzy neural network, and Kalman filter"**. International Journal of GPS Solutions,10(2), pp. 97- 107, 2006.

[3] A.Judisky, Q. Zhang**,"Wavelets in Identification**: wavelets, splines, neurons, fuzzes: how good for identification?, Rapport INRIA No. 2315, 1994.

[4] Zhang, Q. and Benveniste A. "Wavelet Networks" IEEE Transactions on Neural Networks, Vol. 3, No. 6, November 1992.

[5] Parvez S**.:" Advanced Control Techniques for Motion Control Problem"**. Paper pp. 401-408, at ISA 2003.

[6] Anthony J., and Rolf T**."Nonlinear Adaptive Flight Control using Neural Networks"**,Georgia Institute of Technology School of Aerospace Engineering Atlanta, GA,30332, 1996.

[7] B. Delyon, A. Judisky, and A. Benvensite; **" Accuracy analysis for wave let approximation"**; IEEE Trans. Neural Networks, vol. 6, pp. 332-348,1995.

[8] L. M. Reyneri; **" Unification of neural and wavelet networks and fuzzy systems";** IEEE Trans. Neural Networks, Vol.10, pp.801-815,1999.

[9] T. Kugarajah and Q. Zhang; **" Multidimensional wavelet frame";** IEEE Trans. Neural Networks, Vol. 6, pp. 1552-1556,1995

[10] Frank Y. Shih*, Yi-Ta Wu Computer Vision Laboratory "**Enhancement of image watermark retrieval based on genetic algorithms"**, College of Computing Sciences, New Jersey Institute of Technology, Newark, NJ 07102, United States Received 10 December 2002; accepted 18 May 2004 Available online 11 November 2004.

[11] Eskicioglu A., Fisher P., **"Image Quality Measures And Their Performance",** vol.43, no.12, pp.2959-2965, 1995.

[12] Kountchev R., Milanova M., Todorov V., Kountcheva R., **"Adaptive fuzzy filter for the reduction of blocking artifacts in images compressed with IDP decomposition and JPEG",** WSEAS Transactions on Signal Processing, Vol. 2, No. 7, pp. 941-8, 2006.

[13] Acharya T., Tsai P**.,"JPEG2000 Standard For Iimage Compression Concepts, Algorithms And VLSI Architectures",**John Wiley and sons., 2005.

الخلاصة

الهدف من هذا العمل هو أستخدام الشبكة المويجية والخوارزمية الجينية في التنبؤ المسبق لطريقة التحليل اللاولي للعناصر التغذية الامامية في الشبكة المويجية قبل مرحلة التدرييب باستخدام تقنيات الميول. ان تقنية الخوارزميةِ الجينية هي تحققِ الوصول الى القيم ألمثالية القريبة الى كل من(learning rate, translation, dilation, and hidden nodes). وهذا العمل يُمْكِنُه اختصار وقت التعلّمُ ايَ (يُخفّضُ 50% مِنْ عددِ التكرارِ) وبالدقةِ المطلوبةِ أكثر مِنْ بدون إستعمال خوارزميةِ الجينية.

بأستخدام طريقة الخوارزمية الجينية يمكن أحتساب افضل مجموعة ال wave net التي translation and dilation الى عدد كبير من العناصر ذات القيم المثالية بالنسبه للتعليم، وانّ العُقَدَ المخفيةَ مستعملة كقِيَم أولية لتدريب السلسلةِ الثانويِ لمراحل ما قبل التدريب، اثبتت ان المويجة الجيبيةِ والِخوارزمية الجينية نجحت في التَعْرفُ على الصورَ الطبيةِ الملونة والتي تم تُحوّيلها كمستوى رمادي وكانت النتائج جيدة وعالية الدقة في تصنيف الصور الطبية الملونةوالتي تم تحويلها الى التدرج الرمادي.