Republic of Iraq
Ministry of Higher Education and Scientific Research
Babylon University_Education College
Mathematics Department

Novel Perspective for Combining Fuzzy Logic with Wavelets and Neural Networks

A Thesis submitted
to
the council of the Education College in Babylon University
In partial fulfillment of the requirements
for the degree of
Master of Science in Mathematic

By
Bushra Hussien Aliwi

Supervised
By
Dr. Nabeel Hashem Kaghed

٢٠٠٥
Babylon-Iraq

جمهورية العراق
وزارة التعليم العالي و البحث العلمي
جامعة بابل-كلية التربية
قسـم الرياضيات

منظــور جديد لجمع المنطق المضـــــــبب مع المويجات و الشبكات العصبية

رســـالة مقدمة
إلى
مجلس كلية التربية في جامعة بابل
كجزء من متطلبات نيل
شهادة الماجستير في علوم الرياضيات

من قبل
بشرى حسين عليوي

بأشراف
الأستاذ الدكتور: **نبيل هاشم كاغد**

٢٠٠٥ م -١٤٢٦هـ
بابل- العراق

Supervisor Certification

I certify that this thesis was prepared under my supervision at the Mathematics Department at Education Collage in Babylon University as a partial fulfillment of the requirements needed toward the degree of Master of Science in Mathematics .

Signature:
Name:
Title :
Date:

Republic of Iraq
Ministry of Higher Education and Scientific Research
Babylon University_Education College
Mathematics Department

Novel Perspective for Combining Fuzzy Logic with Wavelets and Neural Networks

A Thesis submitted
to
the council of the Education College in Babylon University
In partial fulfillment of the requirements
for the degree of
Master of Science in Mathematic

By
Bushra Hussien Aliwi

Supervised
By
Dr. Nabeel Hashem Kaghed

٢٠٠٥
Babylon-Iraq

جمهورية العراق
وزارة التعليم العالي و البحث العلمي
جامعة بابل-كلية التربية
قسم الرياضيات

منظـور جديد
لجمع المنطق المضـــــــبب مع
المويجات و الشبكات العصبية

رســــــالة مقدمة
إلى
مجلس كلية التربية في جامعة بابل
كجزء من متطلبات نيل
شهادة الماجستير في علوم الرياضيات

من قبل
بشرى حسين عليوي

بأشراف
الأستاذ الدكتور: **نبيل هاشم كاغد**

٢٠٠٥ م -١٤٢٦هـ
بابل- العراق

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ

وَعَلَّمَ آدَمَ الأَسْمَاءَ كُلَّهَا ثُمَّ عَرَضَهُمْ عَلَى الْمَلاَئِكَ
بِأَسْمَاءِ هَؤُلاَءِ إِن كُنتُمْ صَادِقِ ٣١ قَالُوا
سُبْحَانَكَ لاَ عِلْمَ لَنَا إِلاَّ مَا عَلَّمْتَنَا إِنَّكَ أَنْتَ الْعَلِيمُ الْحَكِيمُ ٣٢

صَدَقَ اللَّهُ الْعَظِيم

من سورة البقرة

Acknowledgements

Praise be to God who the Greatest and Almighty firstly and before any thing
who help me, that with out of his help no one be what it want to be .
I present to my family with high respect for their help and support to be in this position,

I am
introducing by deeply indebted thank and esteem to president of Babylon University
Dr. Nabeel Hashem Kaghed *for his instructions and supports during the research time*
in various ways directly or indirectly in different sides .

I am also owe many thanks to deanery of Education Collage and mathematical department .
I am grateful to computer department in Science Collage and its chief ship for their helps.

The Researcher

The science is not gifting but it is learned to each one need it,

but stay the status that get on for a hard work is a simplest gift with great regard
to
my family

Symbols List

U , V	universal sets
$A, B, \ldots$	fuzzy sets
$\tilde{x}, \tilde{y}, \ldots$	fuzzy variables
x,…	vectors
ε	learning rate
γ	momentum term
ϕ	empty fuzzy set
φ	scaling function
ψ	wavelet function
1_U	universal fuzzy set
?(.)	activation function
μ , $\mu(.)$	membership function ,membership degree
$\mu_A(u)$	Membership degree for member u in a fuzzy set A
R	a set of real numbers
Z	a set of integer numbers
$\mathcal{F}(U)$	family of all fuzzy (sub)set of universal U .
$\mathbf{X}^{T}$	transpose of matrix
$supp(A)$	support of fuzzy set A
Δ	changing (difference)
$\int$	traditional integral
$\bigcup_{\alpha \in [0,1]} A_\alpha$	union of $\alpha-cut$ (*level*) sets
$*$	Cartesian product
sup	smallest upper bound
$\lim_{u \to (.)} \mu_A(u)$	limit of membership function $\mu_A(u)$.

Abbreviations List

AI	Artificial Intelligence
ANN	Artificial Neural Network
BMP	Bit Map Format
BP	Back Propagation
DWNN	Dynamic Wavelet Neural Network
EFuNN	Evolving Fuzzy Neural Network
FL	Fuzzy Logic
FT	Fourier Transform
FuNN	Fuzzy Neural Network
LMS	Least Mean Square error
MF	Membership Function
MLP	Multi layer Perceptron
PDF	Probability density function
PE	Processing element
PR	Pattern Recognition
PSNR	Peak Signal to Noise Ratio
RBFN	Radial basis functions network
SNR	Signal to Noise Ratio
STFT	Short Time Fourier Transform
NN	Neural Network
WSNR	Weighted Signal to Noise Ratio
WNN	Wavelet Neural Network
WT	Wavelet Transform

المــــــلخص

هذه الأيام تتجه التقنيات نحو الأنظمة الهجينة بين التقنيات ,لتعوض النواقص والمشاكل التي قد ترد عند استخدام تقنية واحدة .

محور الحديث في هذا البحث هو استخدام طرق التهجين التي تجمع نظرية المويجة مع الشبكات العصبية والمنطق المضبب ,من خلال اقتراح دالة تنشيط توصف كجمع دالة مويجة و دالة عضوية,وهذه الدوال هي دوال مستمرة.دالة المويجة المختارة هي مويجة(*Mexican hat*) ,بينما دالة العضوية هي دالة العضوية لكاوس (*Gaussian*) .

كذلك تقترح هذه الطريقة اتجاه جديد في اتخاذ القرار من خلال حساب المخرجات (*outputs*) عند فترة مغلقة لقيم المخرجات الفعلية عن الشبكة(*actual output values*) لمدخلات الشبكة(*inputs values*) و التي هي ملفات قيم صفات الصور التي استخدمت كمجموعة تدريب (*training set*), والتي توفر قاعدة واسعة للعمل عليها ولينجز تقارب الأنماط خاصة وان الطريقة المقترحة نفذت على مسالة تمييز الأنماط.

نفذت الطريقة المقترحة على مجموعة من الصور لأشكال هندسية مرسومة يدويا لثمان فئات والتي تتدرج في التعقيد,لإثبات أمكانية الشبكة. في كل فئة ستة أشكال (و في بعض التجارب تم استخدام سبعة أشكال),احدها هو الشكل القياسي للفئة و الذي يميز تلك الفئة, و ستة أشكال أخرى هي من نسـخ مضببة (مشوشـة) عنه .

إن فكرة هذه الطريقة المقترحة هي فكرة حديثة لأنها تعتمد على تمييز الأنماط من خلال اعتمادها على قيم المخرجات الفعلية للأنماط المراد تمييزها كونها يجب على الأقل تنتمي إلى الفترة المغلقة لقيم مخرجات الأنماط في تلك الفئة ,و تحدد تلك الفترات من خلال قيم المخرجات الفعلية لأنماط تلك الفئة ,تمتلك كل فئة فترة مغلقة لقيم مخرجات أنماطها تختلف عن جميع فترات الفئات الأخرى,ويجب أن لا تتقاطع هذه الفترات مع الفترات الأخرى مطلقا ,كذلك يمكن اعتبار هذه الفترات المغلقة كمناطق قرار لاختبار التقارب .بينما تعتبر المناطق بين هذه الفترات أو تلك التي تقع

خارج فضاء الفترات المغلقة كمناطق خطأ , تحاول الشبكة المفترضة قدر ما تستطيع تضييق المناطق بين الفترات المغلقة لتلك الفئات ,و من ثم تقليل احتمالات الخطأ خلال التدريب .

نفذت العديد من التجارب لتوضيح العمل و كفاءة الطريقة المقترحة .عند تدريب الشبكة بواسطة قيم المعلمات المفروضة خلال التجارب يجب أن تكون ثابتة لجميع الفئات في التجربة الواحدة . عملت التجربة الأولى على الأشكال في مجموعة التدريب وكانت النتائج جيدة عند التعامل مع تلك الأشكال و قابلية التعميم للتمييز خلال الفترات .

قدمت التجربة الثانية كمحاولة لزيادة فاعلية الشبكة و لتقليل مناطق الخطأ بين الفترات المغلقة لمخرجات أنماط الفئات ,من خلال تحوير هيكل الشبكة بزيادة عدد العقد المخبأة ,و أيضا لتوضيح كفء هذا النوع من الزيادة .لاحظنا بأن قيم خطأ التدريب قد تضاءلت و بنسبة ليست بسيئة .

في التجربة الثالثة استمرت المحاولات على زيادة الفاعلية ,و حاولت زيادة سعة الفترات المغلقة لمزيد من الأنماط ,من خلال زيادة عدد النسخ المضببة عن كل شكل قياسي في كل فئة .

جمعت التجربة الرابعة أفكار التجربتين السابقتين بتنفيذ الطريقة المقترحة بزيادة عدد العقد المخبأة وعدد الأنماط في كل فئة في مجموعة التدريب .وهذا يثبت بأن عمل الشبكة قد تحسن عند أجراء هذا التحوير .

كذلك نفذت الطريقة المقترحة على الصور الملونة بطرق محتملة ,الرأي المنجز باعتباره تجربة على نفس مجموعة التدريب , و ذلك عند التجربة الخامسة التي أستخدم فيها صور ملونة لنفس الأنماط في مجموعة التدريب ولكنها قد رسمت بخطوط ملونة على خلفيات بيضاء .أظهرت الشبكة قدرة عاليـــــة في التعامل مع هذه الصور .

أنجزت تجارب عمل الطريقة المقترحة بجهاز كمبيوتر من جيل بنتيــوم III (*Pentium III*) , و تمت برمجة البرامج بلغة البرمجة تيــــربو C^{++} .

Abstract

These days the technologies are going toward Hybrid systems between techniques ,that to compensate the imperfects and a problems that may be occurred when using just one technique .

A key word of this research is to use a Hybrid method that combine a wavelet theory with a neural networks and fuzzy logic , through suggest an activation function as a summation of wavelet function and membership function ,these functions are continuous. The chosen wavelet function is a *Mexican hat wavelet* ,while the membership function is a *Gaussian membership function* .

Also this method suggests new way to take decision through compute the output at closed interval of actual output values for inputs values to a network which are a files of features of images used as training set ,which is provide a wide support to work on , and then to accomplish more convergence pattern(s) ,since the suggested method performed on a pattern recognition problem .

The suggested method performed on a set of images for a geometric curves Hand drawing for eights categories that gradual in complexity ,that to proving the network ability .Each category has six (in some experiments used seven) figures ,one a standard figure that characterize the category and five(six) from its fuzzy (vague) versions .

The idea of this suggested method is novel since; that it depends on recognize a pattern(s) through the actual output of a pattern(s) that wanted to be recognized must at least be with in a closed interval for output values for patterns in that category .This interval determined by the actual outputs values for a figures in a category that used to train a network ,each category has a closed interval for its patterns outputs values different on all other categories intervals ,and these intervals must not intersect with other at all ,also these closed intervals can be regard as a decision regions for convergence .While the regions that be between these intervals or out of the space of intervals are regarded as error regions ,the proposed network try to as could as to minimize regions between these closed intervals of that categories ,and then minimize errors probabilities though training .

Many experiment performed to showing the work and efficiency of a suggested method .In training of a network with a supposed parameters values for the experiments it must be fixed for all categories in an one experiment .A first experiment is work on the figures in the training set , the results of training were good for dealing with that figures ,and a generalization ability for recognizing through the intervals .

A second experiment is presented as trail to increase the efficient of a network and to reduce the error regions between the closed intervals for patterns categories outputs ,through modifying the architecture of a network by increase the number of hidden neurons ,also to showing the effect of this type of increasing .We note that error values for training were decreased at not bad rate .

In the third experiment trails on increasing the efficiency continued ,and try to increase the capacity of a closed intervals to more patterns through increase the number of a fuzzy versions for each standard figure at each category .

A fourth experiment is combine the closed ideas of the two previous experiments through performing the suggested method for increasing number of hidden neurons and of number of patterns for each category at training set .This proven that work of a network is improved when do these modifications .

The suggested method also performed on a color images in experimented way ,a view that was accomplished as experiment on the same training set ,that at the fifth experiment the colored images through using the same patterns in the training set but drawn by colored lines on white backgrounds .The network testify high ability in dealings with these images .

Experiments of work the suggested method is accomplished by *Pentium III computer* ,and programmed by *"Turbo C^{++} " language programming .*

٥.٨ *Conclusions*

This scheme introduce a novel method for combining fuzzy logic and wavelets ,that have been exploited to work on solutions for some patterns recognition problems through fixed architecture for a network (except some supposed experiment to improve the work).

Here will view several conclusion drawn from this work ,through seeing a gotten results from all the five performed experiments , we infer success network under suggested method and a supposed parameters values through experimenting in dealings with what that we want to accomplishing from this work ,that led and help us to thing and encourage to work on more and more different problems after this success .

We note through learning this network in first experiment ;a real doing for network in recognize images in training set and even on that in testing set and for all categories normalcy, in ways seems similar ,that is in limits of a closed intervals for each category .

Also note that increasing efficiency of the network with an increasing in number of hidden neurons ,that what introduce through second experiment for work on a same used training set and testing set for first experiment and for all categories, with decreasing in number of iterations for most categories if it compared with previous experiment ,that certify work of a suggested method on patterns under this condition, misrecognition to some categories really occurred ,so we tend to perform a next experiments .

A network efficiency has been increased to good rate for an other way in modification in framework through increasing number of a used images(patterns) in training set ,that what hold through coming third experiment at increasing ability to recognize patterns , increasing not effective in testing set that also improve performance of the suggested method through success in this test.

Fourth experiment recover on a good success in accumulation between ideas of previous experiments ,that add a goodness for work and then results of learning network and clarify effect of increasing in number of hidden and training set patterns on doing of the network on each category .Increasing number of testing set patterns not very active on this work ,that is we used same testing set on all previous experiments .The main thing is increase ability to pattern recognition .

The network proven an excellent ability in dealing with a colored images ,and for all categories in training set that also mean improving ability for this type from images and on testing performance, in spite fail a network in recognize some patterns it may be so noisy(fuzzed) to recognition or not in that category , a largest available error value that used for this job is ٠.٠٠١ even it larger than the used one for uncolored images, that promote to trade with more ways for coloring .Results displayed good training , in addition to testing, in recognize images in testing work .

We have shown that ,when used the colored images leads to results that are much more independent from colors even it different for coloring(in colors) .

Also for all experiments we noted existing a deviation in parameters values for learning ;

- Minimizing iterations and epochs number from first to a second experiment for most categories ,but it increased at third .Also it decreased from third experiment to the fourth ,and then minimized for the colored images at fifth experiment .
- Decreasing mean error tolerance values for categories from first to the second experiment ,in spite using a value some what large for colored images .Also it decreased from third to fourth experiment under same view .
- Increasing ranges of the intervals at first to second experiment for most categories ,also increased at third than two previous experiments which increased at fourth experiment in general .The ranges are increased and decreased at some interval at fifth experiment .

After training ,the network displayed an ability to generalize , that is to recognize images in testing set for all categories in performed experiments ,since it obeyed two techniques ,which represented as universal approximators(fuzzy logic and wavelets) in architecture of the network .

The main important thing at this point is actual output for each pattern not belongs to two (or more) categories intervals in same time(experiment) ,and choice of desired output values for these categories is effective on ordering these categories intervals ,and so these intervals are ordered increasingly ,this since choosing a desired values with increasing order, these intervals are appeared as a sequence of closed interval began with first category interval and end with last category interval(eighth category) for all categories in the performed experiments . Also see that existing a regions between these intervals of categories (as parts from error regions) ,and existing regions out over higher value for all intervals(eighth interval) ,and less than a smallest value for all categories (first interval) .

From previous experiments it be known that a largest available error value that a network able to work under is ٠.٠٠١ for an images drawn with colored lines. While largest available error value is ٠.٠٠٠٨ for uncolored images(black and white) .

Using a single file for weights values of connections for all categories and changed for each category that training is helpful in decrease multiplicity in train each category dividual ,that simplify the testing performance ,and support a static criteria in compare output values .

From points that it important also ; fixing values of translation and dilation on values ٦ ,٢ , respectively for all categories and for all experiments ,this declare the flexibility of a network in working in spite existing fixed sides that offer a constructive side to put a criteria .

It be grateful to know ability of network in dealing with images(for all used categories) ,for any number of image (even it large) .

Finally, this work introduce a way to more works in this subject or that schemes associated with it ,and then gain at less trials to solve some problems .

Contents

Chapter Three; Fuzzy Logic

Appendix I

The reason to choice a ***Mexican hat wavelet*** (it and not other) , since it is ;

- Even: even function

  $\psi(-x) = \psi(x)$ … (١)

  $\forall\ x \in\ domain\ \psi$, and square valued for inputs ,that make it symmetric .
- Real valued: $\psi : E \rightarrow R, \psi(x) \in R$, R is a set of real numbers [٣٠].
- Commonly used [١٤] .
- Rapidly vanishing: satisfies the condition (٢-١) ,and as wave characteristic [٥٥];

$$\int_{-\infty}^{+\infty} \psi(x)dx = 0 \qquad \dots (٢)$$

As it is known the exponential functions[١] such ***Gaussian MF*** is over an infinite support ,that is infinite set of variable values *(x_axis)* and its output values is always real values , singlity of a maxima is really exists since it is a convex function satisfies the condition in a fuzzy logic, also since convex property means having a single maxima (see[٥١]) ,it is satisfies a function to be convex for if x_1, $x_2 \in \text{supp}(A)$ in a support of φ at A ,then;

$$x_1 \leq x_2 \Rightarrow \varphi(x_1) < \varphi(x_2) \qquad \dots (٣)$$

Note that this function is symmetric since it an ***Exponential function*** over squared value of inputs values .The squared in exponential power is refer to a positive value for input values even if it is negative or positive .

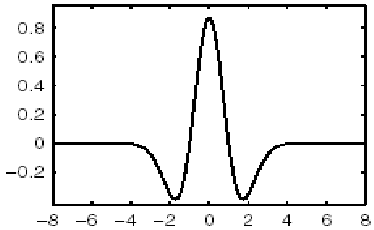
Note: The minus sign in Gaussian function do not led to negative value(as think in first time),but it represent a converse of exponential under one ,this restrict values smaller than ١ and over than ٠ , that is exactly what we want from a MF .

$$e^{-x^2/2} = \frac{1}{e^{x^2/2}} \qquad \dots (٤)$$

---

[١] Exponential function is **real valued function** : for **y=exp(x)** , and **exp: A→R** , such that **y(x)∈R** , and **y(٠)=١** .

Indeed, a symmetric condition is work on an absolute operator for entered values and in depending on shape of Gaussian MF is give positive values not exceed ٠ to a negative values ,the working on this concept is clarify every where positive property ,and symmetric in the same time ,it may be not forgotten that the MFs always only in values between ٠ and ١ ; $\varphi(x) \in [0,1]$.

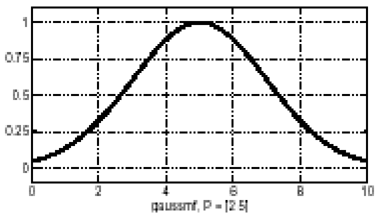


Figure(1) :*Mexican hat wavelet*

For Gaussian MF in fuzzy logic [٣٧] that as;

$$f\ (x,s,t) = e^{\frac{-(x-t)^2}{2s^2}} \qquad \ldots \quad (٥)$$

The symmetric Gaussian function depends on two parameters s and t which are width and center for distribute values, respectively .



Figure(2) :*Gaussian Membership function*

Appendix II

To view the discussion for several commend properties for wavelets and membership functions.

In addition to the properties that a MF satisfies, it is accommodate with wavelets in many sides see [١٨] .

Such that ; input values are scaling with center value of a fuzzy set ,while in wavelets a scaling is to integer values ,but the concept is one .Also inputs to a wavelet are dilute and translate into integer values ,while a MF is operate on a width factor of a fuzzy set ,this simplify a work with wavelets and MFs *.*

From operations that used in NNs and WNNs is a "normalization ".In a fuzzy logic the normalization is not different ,it means applying standard range of [٠,+١] for universe both for inputs and outputs [١٨] .If the standard values do not modified ,actual values will be too large or small ,that is transform a crisp input into a normalized input in order to keep its value within universe ,while an output transform crisp outputs from a normalized universe into an actual output .

Chapter One

General Introduction

١.١ Introduction

Most technological techniques are try to as could as to match the human brain abilities in solving problems in computation or control through using computers and programming to solve these tasks .

Humana are often able to manage complex tasks under significant uncertainty has stimulated the search for alterative modeling and control paradigms ,so called “intelligent modeling”, and control methodologies which employ techniques motivated by biological systems and human intelligence to develop models and controllers for dynamic systems, and from common drawbacks most standard modeling approaches that they cannot make effective use of extra information [١] . Paradigms such as; expert systems ,Neural Networks, Fuzzy logic,…,are used as intelligent techniques, which offer some advantages in real world applications[٢].The ineffective of these methods in dealing with many problems ,it needed to complete its mistakes but not as alternatives on it .Such techniques named “Hybrid Systems” . Hybrid intelligent techniques integrate two or more intelligent techniques ,is such away that each technique enhances capability of the other[٢].The technique that used for fusing these systems called a “Fusion Technology” [٣].

A hybrid methods that combining soft computing methods (intelligent techniques) with wavelet theory have therefore the potential to accommodate two central elements of the brain ; the ability to select an appropriate resolution to the description of a problem and to some what tolerant of imprecision [٤] ,that is computationally ,NNs are performing successfully where other methods do not , in recognizing and matching complicated ,vague ,or incomplete patterns [٥] , and they are ideally suited to pattern recognition ,signal processing ,classification and the predication tasks [٤] ,but in spite it were used in solving a wide variety of problems, it have many drawbacks such as; convergence, local minima, and generalization ,that make it inefficient in many applications .So many ways used as trials to solve these problems such adding a momentum term into adjustment equation to increase the converging speed and to avoid a local minima as in([٦]),or from other researches that attempt to solve this problem and to go past it is by using a wavelet transform (WT) in neural networks as occurred in([٣]) .

Wavelet functions that used as universal approximations give a theoretical basic to their use in the framework of function approximation and process modeling, For wavelet functions this property can be expressed as follows ;

Any function of $L^2(R)$ *can be approximated to any prescribed accuracy with a finite sum of wavelets, the dilation and translation obtained on a chosen mother wavelet [٧] .*

There is so much ways have combined elements of wavelet theory with soft computing .The vast majority of these use wavelet analysis in combination with NNs ,either for feature extraction or to reduce the dimension of the input space [٤].Wavelet have also been combined with NNs in wavelets networks and wavenets . In first one the wavelet part is essentially developed from learning , while in the other a signal is decomposed on some wavelet and the wavelet coefficients into a single method which covers wavelet neural networks WNNs ,wavenets ,and fuzzy wavenets(as will see) Wavelet networks have been successfully implemented to identify and classify rapidly varying signals for example to identify high risk patients in cardiology or for echo cancellation ,and in economic at Forecasting and predication of chaotic signals over two other promising fields [٤].

Wavenets recently extended to biorthogonal constructions , this development allow wavelets to be used to develop fuzzy rules from data in on_line problems .The main advantage of this method is that wavelet theory furnishes simple means to validate the fuzzy rules [٨][٤].

Fuzzy logic has found applications basically in all domains of science from biology to particle physics .The majority of applications are clearly in the domain of control .A linguistic interpretability of fuzzy rules is certainly one of main reasons for success of fuzzy logic[٩] .A major challenge to fuzzy logic is translation of the information contained implicitly in a collection of data points into linguistically interpretable fuzzy rules , what is called a neuro fuzzy methods have been developed for this purpose since a serious difficulty with most neuro fuzzy methods is that they do often furnish rules without a transparent interpretation [٩][٨] .

Fuzzy systems and NNs attempt to determine the transfer function between a feature(s) space and a given class, and they can automatically adapted by computer in an attempt to optimize their classification performance [١٠] ,NNs used before for classification ,but it however can only be initialized in a random state ,thus training of computer to optimize the classifier is usually much faster with fuzzy classifier than a NN classifier .In other hand the drawback of a fuzzy system is difficulty to dealing with multiple features .So there are a ways to combine a fuzzy logic and NNs in variety manners. Generally, hybrid systems of fuzzy logic and NNs are often referred to fuzzy neural networks (which also named neuro fuzzy methods) [١١] classified into ;

- *Fuzzy rule_ based systems with learning ability .*
- *Fuzzy rule_based systems represented by network architectures.*
- *Neural Networks for fuzzy reasoning .*
- *Fuzzified Neural Networks .*

Also there are other approaches .

The classification of fuzzy neural networks into one of the previous categories is not always easy .These systems basically work on *fuzzy if_then rules* that extracted from data ,and adjusted by iterative learning algorithm similar to NNs learning .It relies on a set of fuzzy rules and a fuzzy inference machine ,in connectionist way . A fuzzy neural network is a connection of feed forward layers network architecture ;First layer receives the input information , second layer calculates a fuzzy membership degree to which the input values belong to predefined fuzzy membership functions , e.g. “small”, “medium”, ”large“,…,etc .While third layer represents associations between input and output variables through fuzzy rules .Fourth layer calculates degree to which output membership functions are matched by inputs data,and fifth layer does defuzzification and calculates value(s) for an output variable [١٢].

Also, wavelet analysis has been combined with fuzzy logic in flame detectors for on_line signal processing[٩],and multiresolution fuzzy technique ,that also known as fuzzy wavelet , first concept of a fuzzy wavelet was defined as a normal fuzzy set (its maximum ١), whose membership function meets the admissibility condition for wavelet [١٣] .

١.٢ Literature Review

The previous ideas for the fuzzy wavenets -fuzzy wavelets- are presented through dividual ideas ,and then hybrid systems opened the gate to more and more ideas of combinations[١] .

We can devoted the literature review as its date in declaration ;

- From the ideas for modification the activation function used it as summation of two functions was occurred ,such as a summation of two sigmoidal functions which were presented in publish [١١] in ١٩٩٣.

- Combining a neural network and wavelet network and fuzzy system ; The researches in this direction(From it subject of this thesis) were exploited the similarity among Soft computing paradigms ,as in [١٤] which offer a merged methods for such combination (the date of publishing this, is not available) you can see a last previous reference to know more .

  The beginnings of the fuzzy neural networks back to Kasabov in ١٩٩٦ and its development in ١٩٩٧ ,the architecture ,principles of this network were illustrated through what republished in reference[١٢] and part from [١١] .It is a feedforward network with five layers of neurons similar to a layered structure that given in reference[١] .A (BP) algorithm as training algorithm with fixed or adaptable MFs ,this paper also present several algorithms for rule extraction from FuNN . What can be figured out from this paper is regarded as a corner stone for studying FuNNs, back to a last reference on its location on internet web (see list of references).In [١٥] that was published in ١٩٩٧ also implement wavelet transform and neural networks work with fuzzy max_min inference to construct the fuzzy rules , and employ the orthogonal Daubechies filters of length ٤ to decompose

[١] Here we will present the sets of some researches that work on ideas communicated to introduce the suggested method .

the original signal ,through computing wavelet and scaling coefficients.

This is similar to what discussed for a different connections ways between fuzzy logic and NN that it were occurred in [١١] that was published in ١٩٩٨ as one of FuNN techniques special fuzzy classification ,This book also donates an other direction through introducing a forms for fuzzy arithmetic in NNs. NNs for handling fuzzy inputs ,and NNs with fuzzy weights ,which all in a general form for Fuzzified NNs .Fuzzy NNs awarded in this at about five or more forms, each one represents a beginning to other researches and ideas of work .

From subjects that were be raised ,and it a different for the combination of fuzzy logic and wavelets what is reported in [١٣] at ١٩٩٨ , that introduced a form called a *fuzzy basis functions* and frames ,forms in this publishing were refute showing that a previous forms of fuzzy wavelets are wrong[٢] -this by means what is literarily occurred in main idea for work-, it is also proffer several theorems on this form and some of its properties with some examples ,you can see[١٤] for detailed description.

As manufacturing applications for FuNN what was published in [١٥] ,an other one in ١٩٩٩ in [١٦] through new method of tool wear detection with cutting conditions and detected signals is presented, which includes model of wavelet fuzzy neural network and the model of fuzzy classification with motor current. The results of tool wear estimated by cutting conditions and detected signals (spindle motor current, feed motor current ,…) are fused by fuzzy inference, spindle speed, feed rate and the depth of cut should be taken into account effects of tool wear to estimate the tool wear states by a known spindle current signal, feed current signal and cutting parameters. The structure of fuzzy neural network was established in this paper in addition to the algorithm . Experimental results show that method of tool wear detection is reliable and practical.

[٢] See the reference "***Fuzzy basis functions ,universal approximators ,and orthogonal least square learning***", if got on it because it is not gotten to see for what mean by wrong form .

- In combining fuzzy and wavelets the most papers about this subject were reported firstly in ٢٠٠٠ by *March Thuillard* ,(ideas in these papers[٣] are convergent some what to a general idea of subject of this thesis) ,through suggesting a form for activation function as a summation of a wavelet and membership function and implement a multiresolution analysis through architecture present a series of a networks and then repression a coefficients under validation operations .This work was published in ٢٠٠٠ in [٨] .In a subsequent papers follow it present similar ideas but with some developments in some sides such learning with multiresolution Perceptron_like networks [١٢],(this idea also was discussed in other papers) .

- Also the papers [٨],and [٩] were introduced as developments to a fuzzy wavenet for biorthogonal wavelets in addition to an orthogonal way in previous references, this was presented as a subject of article in the newsletter in[٤] in same last mentioned year .

- In ٢٠٠٠ a paper grant a ways subsidized with numerical examples to initialization a wavelets for wavelet networks with two methods illustrated widely in that paper which is [٧] (from references list at this work) ,and these two ways were a heuristic initialization procedure ,and an initialization procedure using a selection method .

- As published in the scientific Newsletter Synergy of Coil in ٢٠٠٠ that a work shop for practical applications of ideas for Marc Thuillard[٤] through articles for researches . Researchers at the University of the Aegeanand MIT GmbH have developed a system based on wavelet transforms, neural networks and fuzzy logic in order to analyze the noisy time_series data typical of changing price values in equities on the Stock Market ,that in reference [٤] .They aim through this way to construct a prediction and decision tool for investors ,through using a very short-term rates of change to arrive at a short term

[٣] This idea was presented in an other papers but it was most illustrated in this paper -from a researcher view point .

[٤] **Marc Thuillard** (as say the newsletter) the man behind the world first producted fuzzy wavelet device.

buy/sell/hold policy ,and they used a filtering system based on wavelet decomposition ,thresholding ,and reconstruction ,in order to uncover trends in the daily rate of change of closing values of a selected equality .This applies multiresolution analysis to primary inputs ,the reconstructed signal was used to train a multilayer Perceptron in daily_trend prediction .Outputs from the trained NN are fed into a decision system based on a fuzzy set ,which makes use of the predicted trend to arrive at a final buy/sell/hold strategy .The researchers used some neuro_fuzzy techniques (which for our attention[٥]) and genetic algorithm to select the most suitable MFs and rule based for the fuzzy set .This was a part from the article in a previous reference .

From a wavelet based neural network as classifier that exploit an (EFuNN), was presented to University of Otago in ٢٠٠١ in [١٧] , which is a complement for techniques that applied to task of image recognition were published in ٢٠٠٠ (see previous reference).It is used for detection of browning in Braberun apples .A wavelet coefficients are extracted from each image and these features used to train both an EFuNN and (MLP) ,This paper also introduces a comparison between obtained results from previous two networks , which clarify the efficiency of EFuNN ,that combines power of a fuzzy inference system embedded in a neural networks NN with ability for adaptive on-line or real-time learning .The layers of network are represented steps of work of inference as; A first layer neurons represents values as a fuzzy linguistic terms such small, large for variables through a membership functions ,a next layer(s) nodes represent a rule(case) nodes, and a linear functions are used as activation functions of the fuzzy output neurons ,which are output membership functions for the output variable(s) .A new fuzzy input neuron can be created during the adaptation of an EFuNN ,to be aware on a

[٥] Through the literature review we want to focus on a most relevant references ideas with the essential idea of this thesis in all the different sides ,not only the main one .

learning algorithm steps back to reference[١٧] ,which give a minute description for this algorithm[٦] ,see [١٨] also.

- Combining wavelets and NNs ; From the works on images the technique that explore a wavelet and NNs through reported technique in ٢٠٠١ to compress images ,through computing a histogram of wavelet coefficients such variance and Entropy ,but the work here on images from Kodak database in reference [١٩] ,which is a special case of wavelet_based signal processing systems to be conversant with results of set of experiments see[١٩] .

- There is a thesis presented to Babylon university in ٢٠٠٣ see[٣] ,on the mean of activation function as summation of two function ,through suggesting activation function as linear function of three sigmoid functions with form ;

$$?(x) = s(x-2) - 2s(x) + s(x+2) \qquad \dots \quad (١\text{-}١)$$

In other view used activation function as a first derivative of *Gaussian function* .

- Uncertainty data used as input to multilayer Perceptron that it is used in terms of fuzzy rules ,and use a membership functions in these rules such *Gaussian* and *trapezoidal* membership function that exploited equivalence of fuzzy logic with radial base functions ,two situations for single and multidimensional functions were also used for variables in the rules, several examples for implementing this way were presented ,and this in [٢٠] at ٢٠٠٤ see this reference in the list .

- Fuzzy classifiers ; are represent an other direction for search ,which is the ability that led to combination with other methods .In studied established in ٢٠٠٤ presented approximately all known fuzzy classifiers such triangular , Gaussian ,and singleton Fuzzifier ,and in same time it introduced a defuzzification methods such center of gravity ,and others. you can see [٢١] for more details ,and also these methods were occurred in references [١٠] and [١٨] .

[٦] A **Neuro-fuzzy modeling** are techniques exploit the fact that a fuzzy model can be seen as a layered structure (network) ,similar to **ANNs** (the beginning of these back to ١٩٩٣ and ١٩٩٤) [١٨] .

١.٣ *Thesis Aim*

The trails for development are not stopped at point , so we want through this research to;

- Suggest a feedforward network exploit wavelets and fuzzy logic properties in one method .
- Evade the previous problems and errors in each way alone and implement it on a patterns recognition .
- Perform a suggested method through experiments to recognize an images (of geometric curves figures) .
- Perform a suggested method through experiments to improve its efficiency in recognizing colored images (geometric curves) ,through coloring to proving that the recognition is independent on colors .
- Recognize in wide range of patterns space and then more robust in patterns recognition .
- Recognize the patterns even in presence the noise in patterns versions ,at a two latest previous ideas for images .

١.٤ Thesis Framework

The general structure of thesis mainly organized into five chapters to take into account offering a relevant subjects with the essential idea in succession chapters .

Chapter One focuses on a necessary premise information as an introduction ,a literature review on a previous works about a neural network ,wavelets and fuzzy logic all with(or not) different combination ways .

In ***Chapter Two***, we present an impractical sides for the neural networks and wavelets .Wavelet transforms specially a continuous wavelet transform ,and wavelet networks in their special case wavenets .

Chapter Three is contained fuzzy logic and concepts of fuzzy set theory subsidized with illustrations examples which refer to the key features of this chapter, they were carefully designed and included to help guide the reader in development through mathematical ways . In addition to membership function and some of its properties .

In ***Chapter Four*** ,we suggest a method through devote a mixing of the previous three subjects ,that works on a geometric curves and extract features and then works through structure of the network. All sides of this method were reported form structure to the algorithm steps and images that works on .

Chapter Five consists of the results of implementing the suggested method through set of experiments and declare results of training ,testing , conclusions and some future works that were reported in this chapter .

Last part presented with two appendixes; first one to recover some properties of the chosen wavelet and membership function ,while the other discuss some commend properties between manner of wok with wavelets and fuzzy logic .

Chapter Two

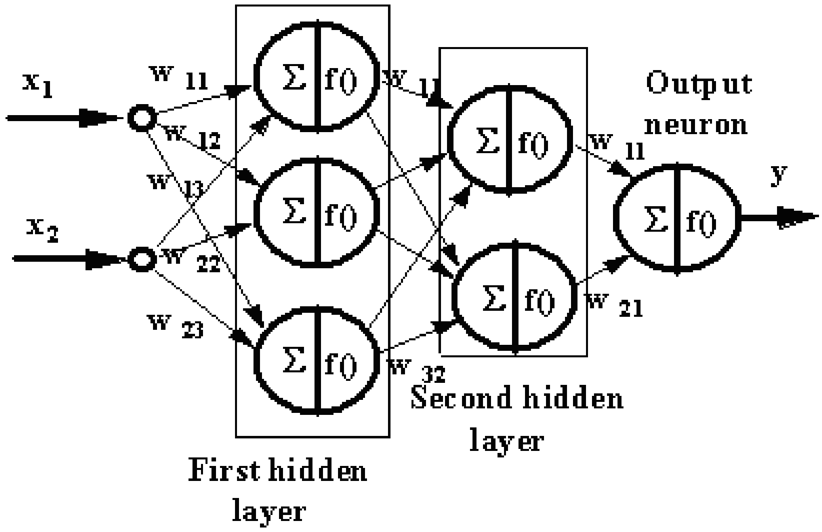
Neural Networks and Wavelets

٢.١ Neural Networks

The attention for the NNs is increased since ١٩٨٠'s ,as one of *"Artificial Intelligent"* applications [٣][٦] ,since the wanted to knowing how to store the information, and learning ability at biological human brain, from these networks abilities a generalization ability and learning not programming which were exploited in most applications of NNs [٦].

NNs emerged from psychology as a learning paradigm, which mimic how the brain learns [٢٢]. Mathematically ,this basic idea can be defined as mathematical models composed of large number of processing elements organized into layers, so this clarifies that the description of NNs is conjugated with a description of a biological neuron and considering a human brain as a network of such neurons [١٨].PEs(are an artificial neurons) handles with several basic functions. It evaluate input signals ,and determining the strength of each one ,then calculate a total for a combined input signals and compare that total against some threshold level, and so depending on the results, that determine what the output should be[١٨][٣].

These PEs(neurons) are connected with (each) others by a weighted connections .The work of each neuron represents an activation of that neuron and basic function that this neuron deals with which change this activation is called *activation function* that suitable for input and lead to get on an output [٣] .Figure(٢-١) clarify the sample of architecture.



Figure(٢-١):The sample for Neural network architecture

٢.١.١ Learning in Neural Networks;

The brain learning is by experiments ,teaching ,training ,and testing .Artificially ,the networks also learn by training, but through changing of network connections weights (to learn solution of a problem) [٥],but the learning which represents training NNs is an external process through external data [٢٣].

Learning ability of the NNs is determined by its architecture and by the algorithm that is chosen for training ,so there exists many laws for learning ,The learning laws are mathematical algorithms used to update the connection weights between layers [٢٢][٥] . These laws are variety for one which in common use ,or the oldest law such Hebb's Rule ,Delta Rule, and competitive learning ,see the references [٦] and [٥] .

The learning in NNs is divided into two basic learning modes [١٧][٣][٢٣][٢٤];

- *Supervised Learning ;*In which ,inputs apply to the network along with an expected response ,also it be used with external criteria .At each input applied to a network a desired output(which is a response) of the system is provided by the user (or teacher) , through computing the distance between actual output and desired output ,it serves as an error measure used to correct the network parameters extremely .

- *Unsupervised Learning ;*or called *Learning by doing* as in [٥] ,is training without teaching, the desired output (response) is not known ,explicit error information cannot be used to improve network behavior ,since no information is available as to correctness or incorrectness of response ,so the learning some how be accomplished based on inputs and outputs ,to adjust the weights . Automatically ,a network learns it self by self .Also it is termed as self_organization such Hopfield network .

Some references are regard an other kind of learning which is a Reinforcement Learning ,it is also known as a supervised learning, or as middle situation between supervised and unsupervised learning ,supervised because it requires a teaching, either as training set of data or an observer who grades performance of the network results. In this way a connections among the neurons in hidden layer are randomly arranged ,and then reshuffled as network in telling how close it is to solve the problem [٣][٥].

٢. ١. ٢ Kinds of Neural Networks

The NNs depending on their varied components are varied to more kinds ,as since neurons in various layers with types of connections among neurons for different layers ,as well as among the neurons within one layer ,deciding way that neuron receives inputs and produces output, learning way ,and a purpose that used was for [٥].

Neurons are connected via network of paths carrying the output of one neuron as input to other neuron, these paths are normally unidirectional ,because there may be an other path in reverse direction .A neuron receives input from many neurons,but produces single output which communicated to other neurons [٦] . Connections between the neurons in different layers called inter layer connections ,while connections between neurons in a single layer called intra layer [٥] .

NNs can classify into many kinds depending on its architecture as [٢٢] ;

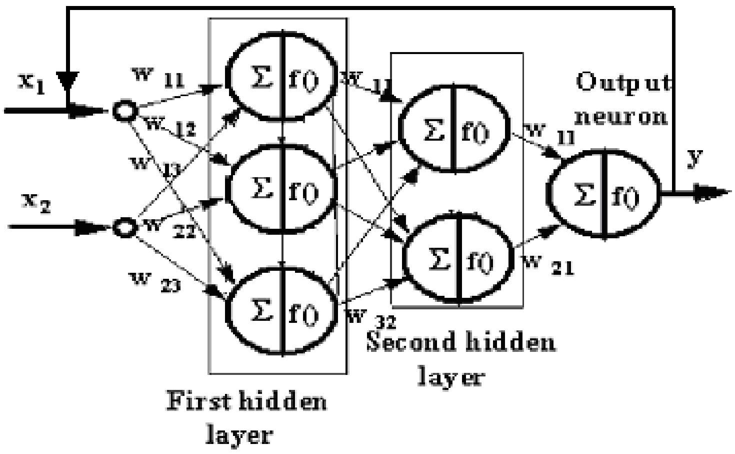
- *Feedforward Neural Networks.*
- *Feedback Neural Networks.*

٢.١.٢.١ *Feedforward Neural Networks* [٦][٣][٢٢][٢٤]

This network consists of a single or multi_layers from neurons with nonlinear activation functions ,a neuron in a layer connected to the neurons in second layer(follow it in the order) .In such a network the outputs of layers are inputs to the following layer through this set of connections.Learning in this network is under teaching ,that is must be supervised learning needs to know the desired output for a network ,such this network; error back propagation network .

٢.١.٢.٢ *Feedback Neural Networks* [٣][٦][٢٣]

The neurons in this kind are connected with each other, that is in addition to the connections between neurons in layers, the neurons also connected with themselves ,that connecting neuron's outputs to their inputs .Network begins with state transitions as one state or sequence, and its goes through states until find an equilibrium state at minimum energy, such network a Hopfield network .



Figure(٢- ٢):The sample for Feedback Neural network architecture

٢.١.٣ *Multilayer Feed Forward Networks*

This type is the most commonly type of NNs **,**because its multiple purposes. This called in many times a *Layered networks* [٦] ,it consists of a set of inputs neurons connected with a set of output neurons, through one or more layer(hidden) modifiable inter mediate connections*,* and for most functional maps at least one hidden layer of neurons or sometimes two hidden of neurons are required [٢٢] .It can implement arbitrary complex input/output mappings or decisions surfaces separating pattern classes ,and the most important attribute of a multilayer feed forward networks is that it can learn a mapping of any complexity ,while the network learning is based on repeated presentations of training samples [٣].

From applications of this type are ,function approximations [٢٥][٢٦][٧], hand written recognition [٢٧] ,expert systems [١١] ,and from other developed applications in personal computers [٦].

From types of this layered network ; a multilayer Perceptron network ,the algorithm of error back propagation training is used to train a multilayer feed forward networks as in[٨][٩][٢٨].

٢.١.٣.١ *Problems of Multi layer Neural Networks*

In spite of the efficiency of NNs in many fields through different applications ,it accomplished with some problems in many sides ,in performing at most applications ,from these problems ;

١. *Design problem* ;

There is no fixed rule in designing a NN at suitable design to solve all problems .The design is consists of determining number of layers , number of neurons in these layers ,deciding the types of connections among neurons for different layers ,as well as for neurons in the same layer ,or determine the strength of connection in the network ,through learning for appropriate values of connections weights by using training set [٥][٢٣] ,and what activation function use .The designing is almost through a period of trail and error in the design decision before coming up with a satisfactory design that make design issues in NNs are complex,or by using wavelet coefficients for determining connections [٣][٧] (as it used in reference [٣]) .

٢. *Generalization* ; [٦][٢٨]

One of the distinct strength of NNs ,is the ability of a network in knowing(or recognizing) a new patterns not in training set ,viz new to network ,that extends to patterns a network not trained on it before at all ,this problem occurred in a classification problems for patterns ,in spite of it a main ability of networks for such problems .

٣. *Convergence* ;

The learning of NNs is to minimize error value ,and get a minimum error value at end learning is not easy since the domain of cost function has features avoid the algorithm to converge ,such existing many local minima in the cost function curve .Learning rate is effected on convergence speed ,if a learning rate value chosen large out of a determined value at ((٠,١)as in [٦]) is main reason to be in this trouble ,or lead to overshooting in convergence .So it must be accuracy in choice a learning rate ,this trouble is occurred in approximation problems of NNs,since they are universal approximators [٧][٣] .

٤. *Premature Saturation* ;

This problem is result for choice initial weights with large values that leads to unlearning ,the weights modify continuously to be large values ,so neurons outputs will be converge to maximum or minimum boundary(٠ or ١ at sigmoidal unipolar function),and the derivative will converge to zero, so the updated value in weights will be zero and stay a weight at its values ,this all make error values large while a neurons is in saturation case [٣].

٢.٢ Wavelets

The term "wavelet" as it simplest means a little wave [٢٩][٣٠][٣١],as linguistically. This little wave(looks like wave) must have at least a minimum oscillation and a fast decay to zero in both the positive and negative directions of its amplitude [٣٠]. This feature help us to introduce wavelets in different ways ,as wave _like functions but clipped to a finite domain,which is why they so named[٣٢],i.e. localized(wave_like)functions (graphically) of mean zero .

Mathematically, wavelets occurred in family of functions[٣٣] as a member of an infinite set of functions with finite support(support refers to the region where the function is nonzero ,it can exhibit a support that is compact or rapidly vanishing ,in this light there are many possible sets of wavelets which one can choose [٣٢] ; one trade off between different wavelet set ,is between their compactness and smoothness ,that is type of support(domain) ,and values of frequency(oscillation)fast decay to zero [٣٠] ,this since wavelets are localized in frequency as well as space, that is their rate of variation is restricted .Wavelet functions of compact support are used in wavelet transforms (that will discussed later),which is computationally efficient ,since it defined on a limited and finite domain .also compactness has implications for computational complexity [٣٢] ,while wavelets that are rapidly vanishing has the admissibility condition as;

$$\int_{-\infty}^{+\infty} \psi(x)dx = 0 \qquad \ldots \quad (٢\text{-}١)$$

this admissibility condition with the wavelets centered at origin , which called analyzing wavelet .

Each wavelet form a family of wavelets, each one is defined by dilation ,which controls the position of a single function which is named mother wavelet [٣٤] .

A wavelet level is defined by a wavelet of constant dilation that is translated to cover the region of interest ,a typical way to construct a wavelet family is to use analyzing wavelet located at the origin with dilated and translated ,it can be write any wavelets as function for input variable x,and its variables by shift b,and scale a (dilation) factor [٣٥] ;

$$\psi_{a,b} = \psi(\frac{x-b}{a}) \qquad \ldots \quad (٢-٢)$$

since a mother wavelet centered at origin ,these factors has a values shift $b=0$,and the scale $a=1$,So ;

$$\psi_{a,b} = \psi(x) \qquad \ldots \quad (٢-٣)$$

Each version from a mother wavelet that a daughter wavelet , which has all values for shift and scale factor ,except the values that mother wavelet has, that is $b \neq 0$,and $a \neq 1$so get on the form (٢-٢).

The mother wavelet can take many forms subject to some admissibility constrain to a best choice ,that is different functions and the choice of a mother wavelet for a particular application is not given a priori[٣٢] ,or that a wavelet appropriate for all applications, the choice is depended on trial and error in many works as efficient values of a wavelet ,and the goal is to find a set of daughter wavelets ,so by traveling from the large scales toward the fine scales one" zooms in" and arrive at more and more exact representations of the given signal [٣٠] .

A sets of wavelets are employed to approximate functions[٣٠] , analysis functions and syntheses [٢٥] ,and signal(images) processing [٣٤] .Indeed wavelets ,in their brief history with in the signal processing field, have already proven that themselves to be an indispensable addition to the analysis collection of tools and continue to enjoy a burgeoning popularity today [٣٦] .

٢. ٢. ١ *Scaling and Shifting* [٣٦][٣٧]

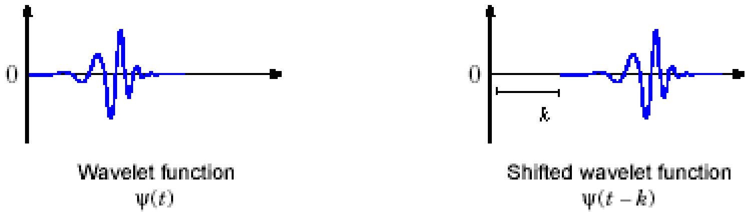
The fact that wavelets features through the window produce a time_scale view of a signal.Scaling and shifting wavelets are from a properties of functions ,the meaning of this terms mathematically known through dealing with its meaning nume-rically. So it must be know meaning and effect of these two terms in work of wavelets with signals .

Scaling a wavelet simply means stretching(or compressing)it, this illustrate underlying descriptions such as" stretching," through *scale factor*, often denoted by the letter a .The smaller scale factor, then more "compressed" wavelet.

$$f(t)\ ;\ a=1$$
$$f(2t)\ ;\ a=1/2 \qquad \ldots \quad (٢-٤)$$
$$f(4t)\ ;\ a=1/4$$

this satisfies for a functions for example a sine function .

Shifting a wavelet simply means delaying (or hastening) its onset. Mathematically, delaying a function $f(t)$ by k is represented by $f(t-k)$ .For wavelets see Figure(٢-٣) .



Figure(٢-٣):The shifting

٢.٢.٢ The Window of Wavelet [٢٥][٢٦]

From features of a wavelet is localizing in a scale/frequency plane, or some times in time/frequency plane when dealing with time varying signals ,so one of the most important concepts is window of a wavelet which is not fixed as on Fourier Transform FT or as in STFT which has a slide window change to time only for the signal .The needed for a window of a wavelet is mainly in approximation problems ,so when talking on a window of wavelet refers to a rectangular region in the scale(time)/frequency plane defined for a mother wavelet ,that can see such region and no other region ,this property results in the identification of localization ,for this function .

For a mother wavelet ψ ,the support of this function in scale(time) domain , $\mathrm{supp}_{\xi}(\psi)$ is defined as ;

$$\mathrm{supp}_{\xi}(\psi) = [x_{\min}\psi, x_{\max}\psi] \qquad \ldots \quad (٢-٥)$$

where $x_{\min}\psi$ and $x_{\max}\psi$ satisfies the following inequality ;

$$1-\xi < \frac{\int_{x_{\min}}^{x_{\max}} /\psi(x)/^2\,dx}{\int_R /\psi(x)/^2\,dx} \qquad \ldots \quad (٢-٦)$$

where ξ is small real number .

It is implies that energy of ψ in $[x_{\min}\psi, x_{\max}\psi]$ is at least concentrated in the time domain at rate $1-\xi$ .Similarly for Fourier transform of ψ , $\Psi(\lambda)$,the support in frequency domain is defined as;

$$\operatorname{supp}_{\xi}(\Psi) = [\lambda_{\min}\psi, \lambda_{\max}\psi] \quad \dots \quad (٢\text{-}٧)$$

, where

$$1-\xi < \frac{\int_{\lambda_{\min}}^{\lambda_{\max}} /\Psi(\lambda)/^2 \, d\lambda}{\int_R /\Psi(\lambda)/^2 \, d\lambda} \quad \dots \quad (٢\text{-}٨)$$

therefore ψ has a time_frequency window

$$[x_{\min}\psi, x_{\max}\psi] \times [\lambda_{\min}\psi, \lambda_{\max}\psi] \quad \dots \quad (٢\text{-}٩)$$

In general, the daughter $\psi_{a,b}(x)$ has a window

$$[\mathrm{b} + \mathrm{a}x_{\min}\psi, \mathrm{b} + \mathrm{a}x_{\max}\psi] \times [\lambda_{\min}\psi, \lambda_{\max}\psi] \quad \dots \quad (٢\text{-}١٠)$$

The size of window can be almost freely variable by two parameters ,thus wavelet can identify the localization of unknown signals at any level .

٢.٢.٣ Wavelet Transform

The WT is an operation that transform a function by integrating it with modified versions (daughters wavelets)of some kernel function (mother wavelet)[٣٠],that is change in its shape to be suitable to process .The WT maps the input data into a new space ,the basis functions of which are quite localized in space .They are usually of compact support (as accrued in[٣٢]),also it is allow more freedom in the choice of basis functions ,so that the basis functions can be better match to the shape of the signal [٣٨].

Wavelet transform perform through out analysis the incomings signals or images or data.The so called wavelet analysis (aim to analysis of both time and frequency).Wavelets(and so wavelet analysis) are localized in frequency as well as space .i.e. their rate of variation is restricted unlike Fourier analysis ,which is not local in space ,but it local in frequency and is unique but wavelet analysis is not since there are many possible sets of wavelets [٣٢].

If the property of integrating a wavelet to zero(their frequency components are zero),so it say to satisfies the *Grossmann_Morlat admissibility condition* [٣٠] ,so a function is admissible if [٣٨][٣٩][٣٠] ;

$$C_h = \int_{-\infty}^{+\infty} \frac{/\Psi(w)/^2}{/w/} dw < \infty \qquad \ldots \quad (٢\text{-}١١)$$

and $\Psi(0) = 0$ … (٢-١٢)

where C_h is the admissibility constant , $\Psi(w)$ is a *Fourier transform* of $\psi(w)$,and w is frequency .

WT is often efficient in computational complexity ,also wavelet transform incorporate a pyramidal representation of the result, the set of successively smoother versions of an image are not down sampled [١٩][٣٢] .

Another point of view on wavelet transform is by means of filter banks[٤][٤٠](the filtering of the input signal is some transformation of it).The first filter acts as a low pass filter means convolution with a smoothing function, that is smoothing the signal, while second high pass _filter furnishes details of the signal (wavelet coefficient) , their complementary use provides signal analysis. With digital signals(images),the transform is accomplish using a filter bank with several levels in resolution[٤١],and these filters are based on the mother wavelet selected for the transform. Details of the signal at a lower level of resolution are obtained by applying iteratively a same two filters to smoothed signal. Overall picture that results could be called a stylisation of the original signal -a caricature-, which highlights features of interest such as spikes, discontinuities and periodic components[٤].Many applications which use this transform specially signals .The features of WT are from wavelets and the window of function in time(scale)/frequency plane .

From all in all above ,WT provides a decomposition of the original data ,allowing operations to be performed on the wavelet coefficients and then data reconstituted .In analysis of multivariate data a wavelet transform integrate with a methods such the neural networks ,and (supervised and unsupervised)pattern recognition [٣٢] .

There are two kinds of WT*(depending on a type of functions in each one)* :

- *Continuous wavelet transform .*
- *Discrete wavelet transform .*

Here the Continuous and Discrete wavelet transform will discuss ,to know more information on these types of transforms , see any reference on WT .

٢. ٢. ٣. ١ Continuous Wavelet Transform

Mathematically, the Fourier transform is represented through the sum over all time of the signal multiplied by complex exponential .Similarly ,the *continuous wavelet transform* CWT ,is defined as the sum over all time of the signal multiplied by scaled , shifted versions of the wavelet function ψ as[٣٦] ;

$$C(scale, shift(position)) = \int_{-\infty}^{+\infty} f(t).\psi(scale, shift, t)dt \qquad \ldots \quad (٢\text{-}١٣)$$

Morlet_Grossmann define the CWT for a *one*-dimensional signal, $f(t) \in L^2(R)$ (space of all square integrable functions [٣٢]),or as finite energy function[٣٠] ,as [٣٩][٣٨] ;

$$W(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x).\psi^*(\frac{x-b}{a})dx \qquad \ldots \quad (٢\text{-}١٤)$$

where $W(a,b)$ a wavelet coefficient of function $f(x)$ as result of CWT,
$\psi(x)$ is(real valued) analyzing wavelet (mother wavelet),
$(a > ٠)$ is scale parameter,
b is shift (position, translation)parameter,
ψ^* is conjugate function of ψ , and (*)means a complex conjugate ,
the constant $\frac{1}{\sqrt{a}}$ is term for energy normalization .

In CWT , the input signal(in one-dimension or more) combining with an analyzing continuous(mother) wavelet ,that time varying at various scales and shifts , (if used with it) through choosing parameter set of these versions of a mother wavelet ,are basis functions [٢٥],so the original signal can be reconstituted from its wavelet transform [٣٢] ,through wavelet coefficients that are with mean zero.

٢. ٢. ٣. ١. ٢ Continuous Wavelet Transform Properties:

Since WT is a function deals with functions, then if the signal function have some properties this may be have influence of CWT function this properties were characterized as [٣٢];

١. CWT *is a linear transformation :*

- If $f(x)$ is a linear signal function as

$$f(x) = f_1(x) + f_2(x) \qquad \ldots \quad (٢\text{-}١٥)$$

then , $W_f(a,b) = W_{f_1}(a,b) + W_{f_2}(a,b) \qquad \ldots \quad (٢\text{-}١٦)$

Proof; $W_f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x).\psi^*(\frac{x-b}{a}).dx$

Since , $f(x) = f_1(x) + f_2(x)$

$$= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} (f_1(x) + f_2(x)).\psi^*(\frac{x-b}{a}).dx$$

$$= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} (f_1(x).\psi^*(\frac{x-b}{a}) + f_2(x)).\psi^*(\frac{x-b}{a})).dx$$

From properties of integral;

$$= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f_1(x).\psi^*(\frac{x-b}{a})dx + \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f_2(x).\psi^*(\frac{x-b}{a}).dx$$

$$= W_{f_1}(a,b) + W_{f_2}(a,b)$$

- If $f(x) = kf_1(x) \qquad \ldots \quad (٢\text{-}١٧)$

then , $W_f(a,b) = kW_{f_1}(a,b) \qquad \ldots \quad (٢\text{-}١٨)$

Proof; Similarly ,can prove this;

$$W_f(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x).\psi^*(\frac{x-b}{a}).dx$$

Since , $f(x) = kf_1(x)$

$$= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} (kf_1(x)).\psi^*(\frac{x-b}{a}).dx$$

From properties of integral;

$$= \frac{k}{\sqrt{a}} \int_{-\infty}^{+\infty} f_1(x).\psi^*(\frac{x-b}{a}).dx$$

$$= kW_{f_1}(a,b)$$

٢. CWT *is covariant under translation :*

If $f_0(x) = f(x - x_0)$ … (٢-١٩)

then , $W_{f_0}(a,b) = W_f(a, b - x_0)$ … (٢-٢٠)

Proof; $W_{f_0}(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f_0(x).\psi^*(\frac{x-b}{a})dx$

Since $f_0(x) = f(x - x_0)$, then;

$$= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x - x_0).\psi^*(\frac{x - b + x_0 - x_0}{a})dx$$

$$= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x - x_0).\psi^*(\frac{(x - x_0) - (b - x_0)}{a})dx$$

Since x regarded as $x \to x - x_0$ for signal

$$= \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f(x).\psi^*(\frac{x - (b - x_0)}{a})dx$$

$$= W_f(a, b - x_0)$$

٣. CWT *is covariant under dilation as :*

If $f_s(x) = f(sx)$ … (٢-٢١)

then , $W_{f_s}(a,b) = \sqrt{s}\; W_f(sa, sb)$ … (٢-٢٢)

Proof; $W_{f_s}(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{+\infty} f_s(x).\psi^*(\frac{x-b}{a})dx$

$$= (\frac{\sqrt{s}}{\sqrt{s}})(\frac{1}{\sqrt{a}}) \int_{-\infty}^{+\infty} f(sx).\psi^*((\frac{x-b}{a})(\frac{s}{s}))dx$$

$$= \frac{\sqrt{s}}{\sqrt{sa}} \int_{-\infty}^{+\infty} f(sx).\psi^*(\frac{sx - sb}{as})dx$$

$$= \sqrt{s}.W_f(as, sb)$$

this property makes WT very suitable for analyzing ,it is like a mathematical microscope with properties that do not depend on the magnification .

٢. ٢. ٣. ١. ٢ *Inverse of Continuous Wavelet Transform* [٣٢][٣٠]

The inversion formula ,can restore a signal(function) $f(x)$ by inverse transform ,if considering WT $W(a,b)$ of a given function ;

$$f(x) = \frac{1}{C_\chi} \int_0^{+\infty} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{a}} W(a,b).\chi(\frac{x-b}{a}).\frac{dadb}{a^2} \quad \ldots \quad (٢\text{-}٢٣)$$

Generally; $\chi(x) = \psi(x)$ … (٢-٢٤)

,but other choices can enhance certain features for some applications .The reconstruction is only possible if C_χ define admissibility condition in the case at form(٢-١١) , this condition implies $\hat{\psi}(0)=0$,that is the mean of wavelet function is zero ,where ;

$$C_\chi = \int_{-\infty}^{0} \frac{\psi^*(v)\hat{\chi}(v)}{v} dv = \int_{0}^{+\infty} \frac{\psi^*(v)\hat{\chi}(v)}{v} dv \qquad \ldots \quad (٢\text{-}٢٥)$$

v is a variable .

In practice ,some discrete versions of this continuous transform may be used ,which gives a beginning to the other type WT named *discrete wavelet transform* DWT .

٢.٢.٣.٢ *Discrete Wavelet Transform*:

In most practical applications ,the dealing with signals is often through take samples for these signals ,that is a signals discretely sampled and this amounts to considering an input vector of values[٣٢] ,and then implement WT. Calculation wavelet coefficients at every possible scale is a fair amount of work and it generates an awful lot of data and choose only a subset of scales and positions at which to make calculations .If choose scales and positions based on power of two (so-called dyadic scales and positions) then analysis will be much more efficient and just as accurate[٣٧]. In the framework of the discrete WT ,a family of wavelets can be defined as ;

$$\Omega_d = \{a^{m/2}\, \psi(a^m x - nb), (m,n) \in \mathbf{Z}^٢\} \qquad \ldots \quad (٢\text{-}٢٦)$$

where a and b are constants that fully determined together with mother wavelet ψ in family Ω_d ,where

$$\begin{cases} m_j = n a^{-m} b \\ d_j = a^{-m} \end{cases} \qquad \ldots \quad (٢\text{-}٢٧)$$

these relations show that ,unlike the continuous approach, wavelet parameters cannot be varied continuously[٧] .

CWT is useful for many applications ,since it can deals with continuous signals at time (or non stationary signals) so it used in applications of signals analysis such; signals of radar(Doppler signals)[٣٩],and images video[٣٨],it is unlike DWT ,is not restricted to time scales and there for offers more flexibility in the analysis [٣٩],then can be note that difference between them ,is that reconstitute a signal in CWT through integral to get high accuracy information,in contract to DWT that do this through summation .

٢.٢.٣.٢.١ *Inverse of Discrete Wavelet Transform* [٣٥]

The inverse discrete wavelet transform(IDWT) is ;

$$f(x) = \sum_a \sum_b \frac{1}{\sqrt{a}} W(a,b).\psi(\frac{x-b}{a}) \qquad \ldots \quad (٢\text{-}٢٨)$$

٢.٢.٤ *Multidimensional Wavelets*

A functions can be regard as multidimensional functions or not depending on dimensions of input domain and then input variable(s) defined on that function either with only one variable or number of variables (vector of variables).Wavelet as functions also be acted upon these situations.

Wavelets were used for one_dimensional data ,which encompassed most ordinary signals [٣٦] .However ,wavelet analysis can be applied to two_dimensional data (images)and in principle ,to higher dimensional data .

For modeling multi_variable processes the multidimensional wavelets must be defined ,multidimensional wavelets constructed as the product of v_j scalar wavelets(v_j being number of variables) as[٧] ;

$$\Psi_j = \prod_{k=1}^{N_i} \psi(z_{jk}) \qquad \ldots \quad (٢\text{-}٢٩)$$

with

$$z_{jk} = \frac{x - m_{jk}}{d_{jk}} \qquad \ldots \quad (٢\text{-}٣٠)$$

where m_{jk} and d_{jk} are translation and dilation vectors ,respectively.

Families of multidimensional wavelets generated according to this scheme have been frames of $L^2(R^{v_j})$.

The wavelets have been initialized from a mother wavelet in $L^2(R^{v_j})$ space in two ways ,either by using a single scaling(which were used in a suggested method for this work) it use a scalar dilation parameter command for all dimensions through this relation [٣];

$$\Psi_{a,b} = \{\psi_{l,k}(\mathrm{x}) = a^{\frac{1}{2}l} \psi(a^l \mathrm{x} - bk) : l \in Z, k \in Z^n\} \qquad \ldots \quad (٢\text{-}٣١)$$

wavelets constructed as it produced,

l index of dilation , a is dilation factor ,

k index of translation , b is translation factor,

x is multidimensional inputs vector , and Z^n integer numbers with n_dimensions .

Or using a multi scaling ,there is a scalar dilation parameter independent for each dimension ,that there exists two matrices , one for scaling and the other for translation as ;

$$\Psi_{a,b} = \{\psi_{l,k}(\mathrm{x}) = D_l^{\frac{1}{2}}\psi(D_l\mathrm{x} - Tk) : l\ , \ k \in Z^n\} \qquad \ldots \quad (٢-٣٢)$$

D is scaling matrix , T is translation matrix ,

$l = (l_1, \ldots\ldots, l_n)$, $k \in Z^n$ vectors ,

Z^n integer numbers with n_dimensions .

The operations for computing a multidimensional_wavelets be as frames ,from the ways that compute frames of multidimensional wavelets is by (*Tensor product*) [٣],for one dimensional wavelets through eq(٢-٣١) as ;

$$\Psi(\mathrm{x}) = \psi(x_1).\psi(x_2).......\psi(x_n) \qquad \ldots \quad (٢-٣٣)$$

where

$$\mathrm{x}=(\ x_1,\ x_2, \ldots., \ x_n) \qquad \ldots \quad (٢-٣٤)$$

since Tensor product of one dimensional wavelets frames lead to gotten on a frames ,too [٣] .

٢.٢.٥ *Wavelet Neural Network*

From ways that exploit WT the *Wavelet Neural Networks* (WNNs). A WNNs mean combining wavelet transform theory with the basic concept of neural networks ,it is introduce as a class of feed forward networks (or a ٣_layers feed forward neural networks from structure or learning algorithms) using wavelets as activation functions [٨][٣٠][٩],and it have recently attracted great interest because of their advantage over (RBFNs),as they are universal approximations ,but achieve in addition to NNs faster convergence[٣٣] .

A (WNNs) constructed in both the static and dynamic modeling versions, but it is first presented in the framework of static modeling ,and it is amenable to accommodating learning routines (on_line and off_line) so that the algorithm can be improved with time [٧]. (WNN) is static model in the sense that it establishes a static relation (such linear function)between its inputs and outputs that is all signals flow in a forward direction ,while a dynamic as DWNN (or recurrent NNs)on other hand ,are required to model the time evolution of dynamic systems [٣١] .

Signals(which change over time ,such video images [٣٤]) in such a network configuration can flow not only in forward direction , but also can propagate backwards ,in a feedback sense ,from output to input nodes it have recently been proposed to address the prediction /classification issues [٣١].

WNNs have been also ,used as a multilayer (such three layers[٨][٤٢][٩]) or single layer Perceptron [٤٣] ,and nodes of this network are wavelet coefficients of the function expansion that have a significant value in spite similarity in a structure with Perceptron network , that is each node has a window in a space _frequency (some times if used in a time _frequency)plane ,and so it can reflect time_frequency properties of functions (more accuracy than RBFN),and so composed of local nature of data patterns ,thus they are efficient tools for classification and identification problems[٨] (with some success) ,and approximation problems, since WNN can be regard as alternatives universal approximators for standard NNs .In standard NNs the non_linearity's are approximated by supper position of sigmoidal functions, these networks are universal approximators [٣٤] .

From practical point of view, determination number of wavelets and their initialization represent a major problem with WNs ,a good initialization of WNN is extremely important to obtain a fast convergence of the algorithm .Number of methods have been implemented [٧].(To know what these methods see reference[٧]).

The reason why wavelets are use instead of other transfer functions in the networks is due to features of wavelets; firstly wavelets have a high compression abilities(so used in image compression[٤١][١٩]) ,because wavelets are localized in space (time[٤٣]) since it joint input_space/frequency localization , computing value at single point involves small subset of coefficients, robust against coefficient errors and updating the function estimate from a new local measure (so used for approximate functions),that is it have a universal approximation properties .Since the self_similar multiple resolution nature of wavelets offers a natural framework for analysis physical signals and images[٣١][٢٩] .On other ANNs which constitute a powerful class for nonlinear function approximation. A common ground between these two technologies may be coherently exploited by introducing a WNN .In addition to the salient features of approximating any nonlinear function , WNN outperforms MLP and RBFN due its capability in dealing with the so called "*curse of dimensionality*" ,and non stationary signals and in faster convergence speed [٣٣] .

٢.٢.٥.١ Wavelet Neural Network Structure and Algorithm

The structure of WNN is different for different purposes used in it, here will only discuss WNN with three layers (single hidden layer, and with single output neuron) which is similar to a ٣_*layers* Perceptron network .

Figure(٢-٥) shows the WNN architecture ,this network comprised of three layers of artificial neurons :Input layer ,a middle or hidden layer, and output layer .The structure used to accommodate a large number of inputs ,inputs(signals) flow forward through the network that is from input layer to hidden layer, and then to output layer ,such an architecture is known as a multi layer feed forward network or multi layer Perceptron[٣٩] .Input neuron(s) layer perform no processing ,it merely provides means for coupling input vector to hidden layer .A process is in hidden layer (which some times called layer of wavelets in[٧]) ,the neurons in hidden layer sum weighted network inputs as an inputs to it after that it normalized beforehand in input layer ,along with an internal basis for each neuron ,then apply wavelets (daughter wavelet functions)as activation functions $\psi_{a,b}$,through implement any kind of a wavelets functions ,through the square window of each neuron at time(scale)/frequency plane[٣٠], this from nature of wavelets functions in analysis functions(transform)since each neuron has a scaled and shifted version of the chosen mother wavelet (analysis wavelet) [٣٩].Outcomes of neurons at hidden layer are wavelet coefficients (of function expansion or signal)that gives significant values[٣٥] .

The(single neuron, used at universal function approximator [٣٨]) output neurons compute a weighted sum of outputs at hidden neurons (wavelet coefficients) along with its internal bias(simple linear summation) without applying the sigmoidal functions or wavelet function or any nonlinear function ,given sufficient number of hidden neurons N_h in hidden layer .

Response of a network to input signal as output [٣٩]:

$$y_j(n) = \sum_{i=1}^{N_h} w_i \psi_{ai,bi} \qquad \ldots \quad (٢\text{-}٣٥)$$

where y_j actual output of neuron j ,

N_h number of hidden neurons ,

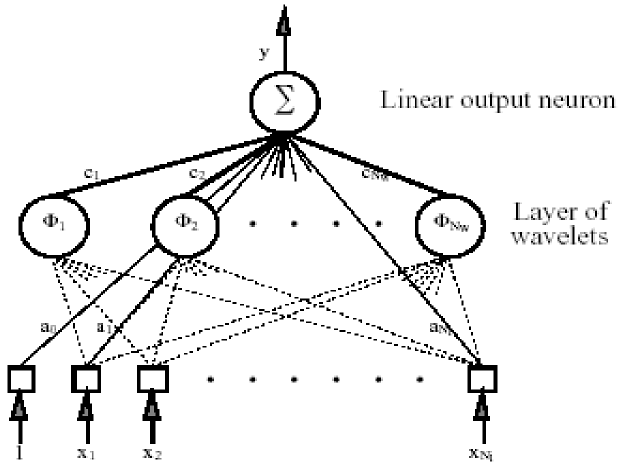
w_i weights between hidden and output layer,

$\psi_{ai,bi}$ wavelet function at hidden neuron i .

But when using a direct connections between input and output layers neurons ,the output form will be in [٧],to minimize error ;

$$y = y(x) = \sum_{i=1}^{N_h} c_i \psi_i(x) + \sum_{k=0}^{No} a_k x_k \qquad \ldots \quad (٢\text{-}٣٦)$$

Also this measure reflects the approximation quality performed by the network over parameters space [٣٣] since it's a summation of wavelet coefficients to restore a signal function, which prove efficiency of network .The parameters of translation ,dilation and weights are optimized through learning [٨][٣٤] ,also the task of training WNN involves estimating parameters in the network by minimizing some cost function [٣٨] .



Figure(٢- ٥): Architecture of a three layer feed forward WNN

Various learning algorithms exist for training this network(WNN) , through computing the network weights and biases for a given problem and parameters network .Most popular learning algorithm is back ward error propagation ,the concept of this method is minimize a squared error of the network (cost function)in weight space[٣٩][٣٨].

An error signal for neuron j is ;

$$e_j(n) = d_j(n) - y_j(n) \qquad \ldots \quad (٢\text{-}٣٧)$$

where n indexes training vector,

$d_j(n)$ is a desired(output) response of neuron j,

$y_j(n)$ actual output of neuron j .

The (instantaneous) value of summation of (half) squared errors;

$$\frac{1}{2}\left(e_j^2(n) \right) \qquad \ldots \quad (٢\text{-}٣٨)$$

Over all neurons in output layer of the network ,a form(٢-٣٨) can be written as[٣٩];

$$E(n)=\frac{1}{2}\sum_{j\in C}e_j^2(n) \qquad \ldots \quad (٢\text{-}٣٩)$$

where the set C includes all neurons in output layer ,and N is number of vectors in training set indexes by n.

Note in the case under studying output layer has only one neuron, C has single member, this leads to as expected view (suggested) ;

$$E(n)=\frac{1}{2}e^2(n) \qquad \ldots \quad (٢\text{-}٤٠)$$

The (squared) error averaged over all training vectors is then ;

$$E_{av}=\frac{1}{N}\sum_{n=1}^{N}E(n) \qquad \ldots \quad (٢\text{-}٤١)$$

N number of vectors in the training set,

E_{av} average (squared)error .

E_{av} constitutes a cost function (that needed to minimize it)

By applying correction delta rule $\Delta w_{ji}(n)$ to update weights $w_{ji}(n)$;

$$\Delta w_{ji}(n)=-\varepsilon\frac{dE(n)}{dw_{ji}(n)} \qquad \ldots \quad (٢\text{-}٤٢)$$

where ε is a parameter that determines rate of learning ,minus sign in eq(٢-٤٢) in gradient descent in weight space means that weights values are moved in apposite direction to error gradient .

After ending training of the network ,it tested with transforms of (denoised)versions of image (signal ,function) .

To minimize $E(n)$ it may use a steepest descent[٣٠] ,which require the gradients $\frac{\partial E}{\partial w_i}$, $\frac{\partial E}{\partial a_i}$,and $\frac{\partial E}{\partial b_i}$,which are refer to a parameters; weighs ,dilation ,and translation that wanted to updated .The incremental changes to each particular parameter w_i , a_i , and b_i ,respectively .

So a gradient of E as ;

let $\tau=\frac{x-b_i}{a_i}$ $\qquad \ldots \quad (٢\text{-}٤٣)$

$$\frac{\partial E}{\partial w_i}=-\sum e(n).\psi(\tau) \qquad \ldots \quad (٢\text{-}٤٤)$$

$$\frac{\partial E}{\partial b_i}=-\sum e(n).w_i.\frac{\partial\psi(\tau)}{\partial b_i} \qquad \ldots \quad (٢\text{-}٤٥)$$

$$\frac{\partial E}{\partial a_i}=-\sum e(n).w_i.\tau.\frac{\partial\psi(\tau)}{\partial b_i}=\tau\frac{\partial E}{\partial b_i} \qquad \ldots \quad (٢\text{-}٤٦)$$

The delta of change as ;

$$\Delta w = \frac{-\partial E}{\partial w} \qquad \ldots (٢-٤٧)$$

$$\Delta b = \frac{-\partial E}{\partial b} \qquad \ldots (٢-٤٨)$$

$$\Delta a = \frac{-\partial E}{\partial a} \qquad \ldots (٢-٤٩)$$

Also it must concerned with modeling problem, that is with the fitting of a data set by a finite sum of wavelets for this purpose

- From orthogonal wavelet decomposition theory ,what is known that with a suitable choice of ψ ,and if m_j and d_j are integers satisfying some conditions the family $\{\psi_j\}_{j=1,\ldots,N}$ forms an orthogonal wavelet basis. A weighted sum of such functions with appropriately chosen m_j and d_j can thus be used; in this way, only the weights have to be computed [٣٤][٨].
- Another way to design a wavelet network is to determine m_j and d_j according to space_frequency analysis of data; this leads to a set of wavelets which are not necessarily orthogonal [٣٤][٢٥] .
- Alternatively, in other way one can consider a weighted sum of wavelets functions whose parameters m_j and d_j are adjustable real numbers, which are trained together with weights[٣٤] .(such architecture discussed beforehand).

In the latter approach, wavelets are considered as a family of parameterized nonlinear functions.

٢.٢.٦ Wavenets :

An interesting alternative to a wavelet networks consists of using a dictionary of dyadic wavelets and to optimize only the weights [٩] .This approach is generally refereed to as *wave nets* or *wavenets* [٨][٤٢] ,which is an other term to describe WNN , or as *adaptive wavelets* (by [٣٠]) .It was first proposed by *Bakshi et al* in ١٩٩٤ [٨] . In its simplistic version a wavenet corresponds to a feed forward neural network using wavelets as activation functions .

Originally, wavenets did referred to NNs using dyadic wavelets (form from wavelets)as activation function ,and in it translation and dilation of wavelets are fixed ,that is not updated ,and only the weights are optimized by the network [٩], So dyadic activation functions will be with form [٨] ;

$$\psi_{m,n} = \psi(2^{-m}(x-n)) \qquad \dots \qquad (٢\text{-}٥٠)$$

with m,n are integers .
since this form is of dyadic wavelets[١].

Wavenets architecture is same for WNN[٣٠] (this is not strange in light of the wavenets definition),while the algorithms consists of two processes ; self construction of network and the minimization of error [٣٠][٤٤] .In first process ,network structure applied for representation are determined by using wavelet analysis (this is exactly work of WNNs) ,the network gradually through hidden neurons cover the scale(time)/frequency plane .In second process, approximations of instantaneous errors are minimized using an adaptation techniques based on the Least mean square error LMS algorithm[٣٠] ,this is clear in WNN .Selected point falls into square windows of hidden layer neurons in scale(time)/frequency plane . So the optimization is applied to hidden neurons ,and then reduce learning cost function .

Form this and previous one can be define a wavenets as a special case of WNN its train depends on optimize only weights at the network .So the algorithm will be same for WNNs but it is stopping at equations *(٢-٤٤)*and *(٢-٤٧)* for updating the weights than a previous discussed forms .

[١] There is another form ,which is **trydic wavelets** with (3^{-m}) ,this out of study here to have more information see the reference [٤٤] .

Chapter Three

Fuzzy Logic

٣.١ Fuzzy Logic

In crisp logic, that is deals with propositions that may or may not be true a proposition can be true on one occasion (has truth value ١), and false on other (has failed value ٠)[٤٥] ,also the approaches base on comentional logic and classical probability theory do not provide an appropriate conceptual framework for dealing with the representation of common sense knowledge [٤٦] ,since such knowledge by it's nature is both lexically imprecise and non categorical .Really ,only two values not enough ,some situations require more than true or false(١ or ٠) , such as applications of AI [٤٧] .This is as much as to use more values to represent all possible cases through combine these propositions to generate others by means of logical operations ,and then get many values(situations) which are graded between these values [٢٣] .

The truth value of a proposition that were combined to generate it (this use widely control problems ,through if_then rules in subject to fuzzy rules ,and applications of fuzzy logic).Since most of its applications involves construction and processing of fuzzy rules [١٨] .

Here the keynote[١] is existing logic which is called a "*Fuzzy Logic*" , which is a generalization of a classical (crisp, conventional) logic ,and it (fuzzy logic)is proposed with new operation on logic [٤٦] .

---

[١] We saw to extend the work on fuzzy logic that is from a mathematical view point to robust the mathematical sides and since it is most connect with the specials of the researcher .

A FL is basically a multi_valued logic [٤٨][٣٧] ,that allows intermediate values to be defined between conventional evaluations like; true/false ,black/white ,yes/no ,high/ low [٤٥] ,and it is translate the fuzziness of linguistic concepts(variables) and approach it mathematically into form that computers (machines)can understand and manipulate it [١٩] . So linguistic notions like ; rather warm ,or pretty cold ,rather tall or very fast ,can be formulated mathematically and processed by computers, and capture the uncertainties associated with human cognitive process thinking and reasoning [٤٦] depending on a mathematical strength of fuzzy logic .

There are two different meanings for FL; a narrow and a broad(wide) sense. In its narrow sense FL is a system of logical operators ,which is an extension of multi_valued logic [٣٧] ,defined by a calculus of interaction that reasoning which are approximate rather than exact [١٨] .

But in a wider (broad) sense ,which is in predominant used today, FL is almost(synonymous refer to a general class of fuzzy set theory) , describing a family of classes with unsharp boundaries in which membership is a matter of degree [٣٧] .

Even (FL)in its narrow sense ,the agenda of fuzzy logic(framework) is very different from the agenda of traditional multi_valued logical systems both in spirit and substance ,since the basic concepts underlying in it are that of a linguistic variable and fuzzy (*if_then*) rules [٣٧]. That is it operate with a mathematical formulas ,words ,and terms of natural language beside the numeric values in (linguistic variables) , words are inherently less precise than numbers ,or use hedges to restrict some variables on limited values [١٨],such terms their use is close to human intuition .

Systems of FL, special rule_based system ,which basically depends on these rules ,which (rules) play a central role in most of FL applications [٣٧] .But on other hand ,other systems such as control system depends on consequents and antecedents part of rules .See now this example ;

Example ٣.١ :

Nakamora saw that his friend working hard on the report ,he expect that his friend will get a good mark .This expect is either true or false that is the truth take value ١ and false take value ٠ ;this is crisp concept.

In addition to these ,Nakamora can say his expect with degree of certainty by use fuzzy logic ,that his friend may be get the good mark , the degree of his expected here about ٠.٦ rather than ١ .

Example ٣.٢ *:* [٢٣]

You say it will rain today ,you are making statements with certainty , of course your statements in this can be either true or false(either ١ or ٠) ,your statements then can be said to crisp .On other hand there are statements you can’t make with such certainty .You can saying that you think it will rain today ,your level of certainty is about ٠.٨ rather than ١ ,in this case you are using fuzzy logic .

FL provide a different way to approach a central or classification problems [٤٥] ,so we can view to it as a methodology for computing (with words rather than numbers), and a control system [١٨],this type of controller includes a knowledge_based system consisting of fuzzy rule ,and a special mechanism of their processing ,Inference engine as its(mechanism) main parts ,that enable approximate human reasoning capabilities to be applied to knowledge_based systems [٤٦] .

It is be clear now that with any system we can use fuzzy values or grade of degrees to get wider logic for working of that system rather than the logical one .

From essential characteristics of fuzzy logic which recognize it on an other logical systems ;

١. Any logical system can be Fuzzified .
٢. In FL ,every thing is a matter of degree .
٣. In FL, exact reasoning is viewed as a limiting case of approximate reasoning .
٤. In FL ,knowledge is interpreted collection of elastic or,equivalently , fuzzy constraint on a collection of variables .
٥. Inference is viewed as a process of propagation of elastic constraints [٤٦] .
٦. A major challenge to FL is the translation of the information contained implicitly in a collection of data points to linguistically interpretable fuzzy rules [٨] .
٧. FL provides a bridge between the continuous world of our perceptions and the digital world of computers . Since it can transform linguistic variables(through rules) into numerical variables (correspond it) with out jettisoning partial truth along the way it allows us to construct vastly improved models of human reasoning and expert knowledge [٤] .

٣.٢ *Fuzzy Set Theory*

Fuzzy set theory operate with mathematical models as any other mathematical theory does ,it is replace one sort of mathematical model with one [١٨] .Also allow us to model terms of natural language with the help of linguistic variables, in addition to the numbers sets and operations on these sets .So a fuzzy set theory which is an extension of a classical set theory [١١],deals with two kinds of variables ;linguistic variables and numerical variables .

The fuzzy set is basic idea in fuzzy set theory ,and a multi_valued logic based on it called fuzzy logic [٤٦][٤٩] .

Mathematics of fuzzy set theory is that what described by Zadeh , this theory proposed making the membership function(or values false and True) operate over range of real numbers [٠,١] [١٨] .

Operations on a classical set theory are satisfies for this theory , and the variety of the available operators for conjunction (intersection), union (disjunction) ,and complement of set ,and other operators .We can consider these operations and other principles the flexibility of using fuzzy set theory [٥٠] .

We say that this theory operate on natural language and linguistic variables ,so it has a variety applications in real life .From the applications of fuzzy set theory ;the fuzzy controller [١٨],the fuzzy classifier [٤٥]

٣.٢.١ *Fuzzy Sets*

In classical mathematic ,we are familiar with what called crisp sets[٤٥][٤٨] .In fuzzy logic this different and we deals with fuzzy sets. In crisp sets an element is either a member in a set or not (True or False) (١ or ٠)[١٠].Fuzzy sets on the other hand allow members (elements) to be partially in a set (in some degree) .

Definition ٣.١; [٤٦][١١]

Let U (a non empty set) the universal set ,and let a function $\mu_A(u):U \to [0,1]$, for $u \in U$,called a membership function , its value which represent the degree of belonging u to a set U . Then the fuzzy set A is defined on U to be a set of ordered pairs of a member(element)in the universal set U and the value of a membership function at that member(element) which is called a membership degree and denoted by $\mu_A(u)$,for a member $u \in U$.

So A can written as[٢٦] ;

$$A = \{(u, \mu_A(u)); u \in U\} \qquad \ldots \quad (٣\text{-}١)$$

that is, it is characterized by its membership function [٤٦] .

Note that elements of a fuzzy set are ordered pairs in form $(u, \mu_A(u))$ [٢٣] ,first side is member included in the set A(satisfies its conditions) while second part refer to degree of this inclusion(value between ٠ and ١) .

The idea of fuzzy sets proposed to represent data information that possess non statistical uncertainty [٥٠] ,vague [١١] values .

A family of all fuzzy sets on U is denoted by $\mathcal{F}(U)$[٤٦],A fuzzy concept is a generalization of the binary concept (of crisp set)by allowing more values between ٠ and ١(true and false ,black and white). In fact infinitely (finite in some cases) many alternatives (more values) can be allowed between these two boundaries [٤٥] (interpretation of numbers now assigned to all members in much more difficult) since values refer to gradual membership in the set .So if $U = \{u_1, u_2, ..., u_n\}$ is a finite (universal)set and A is a fuzzy set in U then we use notation[٢] [٤٦];

$$A = \mu_1 / u_1 + ... + \mu_n / u_n \qquad \ldots \quad (٣\text{-}٢)$$

where term μ_i / u_i , $i = 1, ..., n$,means that μ_i is a grade of membership of u_i in A ,and plus sign represents the union .

Remark; *The crisp set is a subset of fuzzy set .*

Proof:

For

if $x \in U$*, the (membership degree) or range of function;*

$\mu_A(x) \in \{0,1\}$ *in crisp set ,*

but $\mu_A(x) \in [0,1]$ *in fuzzy set .*

and the mapping as ;

$\mu_A(x) : U \to \{0,1\}$ *in crisp set ,*

but $\mu_A(x) : U \to [0,1]$ *in fuzzy set .*

and so $\{0,1\} \subset [0,1]$*, led to the subsets .*

٢ Or as a set form in[١٨]; $A = \{u_1 / \mu_1 + ... + u_n / \mu_n\}$

Note: *A fuzzy set can have a finite or an infinite number of elements depending on the universal set (if it is a set of real , integers, or other situations),and a membership function , these reflect on the support of a fuzzy set see (٣.٢.١٠) ,also it depends on the property(s) that fuzzy set has in its members .*

It is clear that if one only allowed the extreme membership values of ٠ and ١ [١٠] ,that is would actually be equivalent to crisp set .This is wide in discrete membership function on one point .

Example ٣.٣*:* Let R be a universal set ,the fuzzy set A on R is set of all and only real numbers which near zero and that between -1 and 1 .

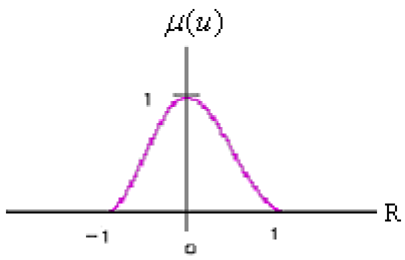


Figure (3-1); A fuzzy set "near zero "

Note: *This fuzzy set will be infinite if we take it on real numbers without giving determined domain ,i.e. the domain will be* R *.*

Example ٣.٤ :[١٨] Consider a set of five pencils found in box a universal set U ,determine a fuzzy sets of U ;

A is a set of "short pencils" ,

and B is a set of "long pencils" ,

as;

$$A = \{(1st,0.2),(2nd,0.5),(3rd,1.0),(4th,1.0),(5th,0.9)\}$$
$$B = \{(1st,0.8),(2nd,0.5),(5th,0.1)\}$$

Note that in A a ٣rd and ٤th pencils are exactly short ,while the ٥th is almost short ,٢nd pencil more or less short ,and ١st pencil is almost exactly not short .

In B , ١st pencil is almost long (belongs to B) , ٢nd pencil is more or less long ,and ٥th pencil is almost exactly not long ,while the ٣rd and ٤th pencils not long (not belong to B) .

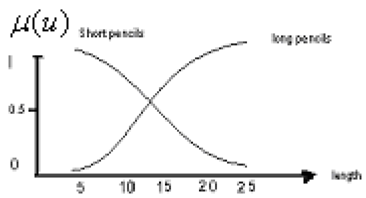


Figure (3-2); A fuzzy set "short and long pencils "

٣.٢.٢ Empty Fuzzy Set :

The empty fuzzy (sub)set of universal (non empty) set U ,is defined as the fuzzy subset ϕ of U ,such that ;

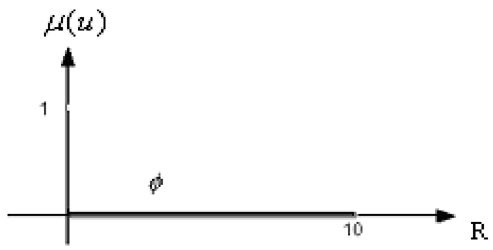
$$\mu(u) = 0 \qquad \dots \quad (٣-٣)$$

for each $u \in U$,
or some times denoted as $\phi(u)$ [٤٦] , see figure(٣-٣) .

Notes:

١. *It is easy to see that* $\phi \subset A$, *holds for any fuzzy subset* A *of* U [٤٦] *,this concept also holds as known for conventional sets .*
٢. *Empty fuzzy subset is also in a family of all fuzzy sets on universal set* U *(any set) , i.e.* $\phi \in \mathcal{F}(U)$.
٣. *The operations on a fuzzy sets also holds on the empty fuzzy sets as any other fuzzy set .*

The graph of the empty fuzzy set ,as example on universal set $U = \mathrm{R}$,define empty fuzzy set on interval $[0,10]$,then ϕ can be written as just ϕ since membership degrees of its members (on domain which defined on it)are zero.



Figure(3-3) :Empty fuzzy set in u=[0,10]

٣.٢.٣ Universal Fuzzy Set

The larges fuzzy set in universal (non empty) set U ,is called "*universal fuzzy set*" ,and denoted by 1_U ,and defined as [٤٦][١١] ;

$$\mu_{1_U}(u) = 1 \qquad \dots \quad (٣-٤)$$

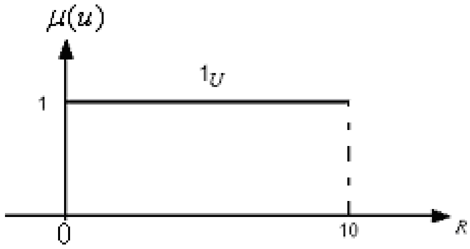
for each $u \in U$.

Notes:

١. *It is easy to see that* $A \subset 1_U$ *holds for any fuzzy subset* A *of* U [٤٦].
٢. *Universal fuzzy set* 1_U *is also in family of all fuzzy sets* $\mathcal{F}(U)$*,for* U *(any set), i.e.* $1_U \in \mathcal{F}(U)$.

٣. *The operations on a fuzzy sets also holds on 1_U ,as any other fuzzy set .*

The graph of the universal fuzzy set ,as example on universal set $U = \mathbb{R}$,define universal fuzzy set on interval $[0,10]$,then 1_U can written with membership degree equal to ١ for each member in $[0,10]$ see figure(٣-٤).



Figure(3-4) :Universal fuzzy set in u=[0,10]

٣.٢.٤ *Membership Functions ;*

A (MF)is basic idea in fuzzy set theory [٥٠] , its value(s) measure the degree(s) which called a *membership degree* for the fuzzy set objects satisfy imprecisely defined properties ,to be member in determined set of members has these determined properties ,that is if the object(member) belongs to a fuzzy set or not and in what degree ,viz to be either member or not in a (fuzzy) set. This contradiction to conventional(crisp) sets which contain objects (members)that satisfy precisely properties require for membership in a
set [٥٠] ,since it is a relationship between values of an elements and its degrees of membership [١٠] .

The membership degree and MF are complete a fuzzy set ,since MF gives the value of membership degree of a member in set ,while membership degree gives value of MF at that member .

Definition ٣.٢ *;* [٥٠][١٨][٤٦]

A function which characterize a fuzzy set A on a universal set U is said to be a membership function ,and denoted by $\mu_A(u)$,for $u \in U$, and its can be written as [١١][٢٣] ;

$$\mu_A(u): U \to [0,1] \quad \text{,for } u \in U \quad .$$

That is ,a mapping from universal set into determined set (a closed interval $I = [0,1]$).This operation called the fuzzification ,when it is mapping a crisp set from universal set .

For any subset in universal set ,a MF is not unique [٥٠].We can choose many types for same set(range) ,in spite of that they different in its values of degree for same member(element) ,(see this example).

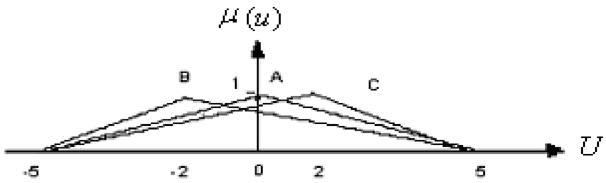
These MFs can be chose from a wide set, which is called a *“dictionary of* MFs*”* [٨][٤٢] .

Example ٣.٥*:* See[٥٠].

Let a universal set U is a set of real numbers between -5 and 5 (or in a closed interval $[-5,5]$),define a fuzzy sets on U as ;

A is a set of numbers close to zero ,
B is a set of numbers close to -2 ,
and C is a set of numbers close to 2 .

We can use a same MF to describe these different fuzzy sets,these shapes are shown in figure(٣-٥) .



Figure(3-5) : *Different fuzzy sets on same range*

These are same MFs on a same domain at different fuzzy sets, for the same case we can use different MFs on each fuzzy set, since it is determine a shape of classes(term) [٥١] .

In universal set which has different classes(fuzzy sets),we can use same MF to describe each class, or we can use different MFs for different classes(fuzzy sets) in universal set on different ranges,see example in[١٠] .

Note that, we don’t forget that MFs describe different classes should overlaps in different formats and wide [١٨] ,and this not necessary means that the classes(fuzzy sets) must overlaps .

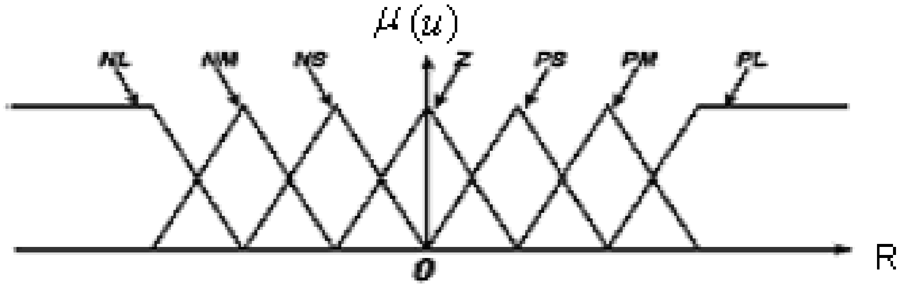
Example ٣.٦ :[١٠]

A set of real numbers is a universal set , fuzzy sets

NL = numbers that are negative large,
NM = numbers that are negative medium,
NS = numbers that are negative small,
Z = numbers that are near zero,
PS = numbers that are positive small,
PM = numbers that are positive medium,
and PL = numbers that are positive large .

The shape of these classes is in figure(٣-٦) .Note the MF is same for all classes in the middle ,and different in classes PL and NL .

Note: *Example (٣.٦) can be take on the set of integer numbers .*



Figure(3-6) : *Membership Function for the set of all numbers (R or Z)(N=Negative ,P=Positive ,L=Large ,M=Medium ,S=Small)*

From a properties of a fuzzy set is that every member in a universal set has some positive membership in fuzzy set [٥٠] . But don't except it as a member in fuzzy set ,so approximate its membership degree to zero, that isn't belong to a fuzzy set ,also membership function may be defined on just one point(or interval) in the universal set. In this case the membership graphic is as point(s) corresponds to (single)point(s) which defined on it ,which(previous case) is a special case of discrete MF,described graphically in figure(٣-١) . Now depending on the influence of fuzzy MF in fuzzy sets ,it must given more attention in choose and in graphic and the properties which make it appropriate for some range on another one ,from sides that must be focus on ;is choosing of general parameters which determine a MF ,a number of classes to describe all values of (linguistic)variable on universal set [١٠][١٨] , position of different MFs on universal set , width ,and concrete parameters ,such as shape of particular membership functions.

The width is refer to range of fuzzy set that it is has on a universal set .If we choice the absolute value for width it is not appropriate [١٨] .Since it do not compare width of separate MF with number of classes and universal set . A classes has different ranges of values in spite its is same universal set or same MF,or it is overlaps.

٣.٢.٥ *Some of Membership Functions Properties ;*

Like other functions MFs upon on their potential applications and properties desired for U [٥٠],and since they refer to a graphical representation of magnitude of participation to each value in U (a set) [١٤] ,it may have some properties as any function these properties are such ;

٣.٢.٥.١ *Monotonic* on each side of the center of range(fuzzy set)[١٤][٥٠],i.e. closer value to center .

٣.٢.٥.٢ *Symmetric* around a center value [١٤][٥٠].That is number equally for left and right of center value .These can be express as center of fuzzy set A [١٤].

$$\mu_A(u) = H(\frac{/u - dA/}{sA}) \qquad \ldots \quad (٣\text{-}٥)$$

$H(z)$ gives the shape of MF μ_A ,
$dA > 0$ center factor of a fuzzy set A ,
and $sA > 0$ width factor of A (step of moving) .

Note: In the case when MF *defined on one point ,this means universal set has just one point .*

Example ٣.٧ *:* The set of moons around the earth .

٣.٢.٥.٣ *Continuity ;*

The concept of continuity is same as in other functions, that say a function f is continuous[٣] at some number c if ;

$$\lim_{x \to c} f(x) = f(c) \qquad \ldots \quad (٣\text{-}٦)$$

for all x in range of f .
That require existing; $f(c)$,and $\lim_{x \to c} f(x)$

In fuzzy set theory the condition will be;

٣ Back to this definition in the mathematic books .

$$\lim_{x \to c} \mu_A(x) = \mu_A(c) \qquad \ldots \quad (٣\text{-}٧)$$

such that x and $c \in A$.

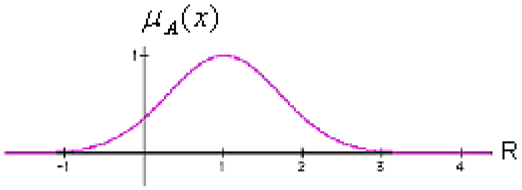
A graphic of MF which is continuous, seen that shape is as continuous curve(or line if MF f is linear) ,see this example ;

Example ٣.٨ :[٤٦]

Let A be a fuzzy set of real numbers "close to 1" , a MF of A can be defined as ;

$$\mu_A(x) = \exp(-y(x-1)^2) \qquad \ldots \quad (٣\text{-}٨)$$

,for $x \in A$, y is a positive real number .See figure(٣-٧).

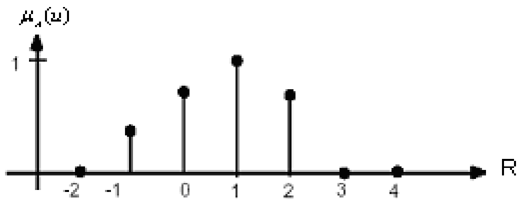


Figure(3-7) : A continuous membership function for "x is close to 1"

But the continuity is not always satisfies for each functions (MFs) . So here refer to a discrete functions MFs ,and the graphic of these functions is drawn as points in the interval [٠,١] ,each point corresponds to membership degree of a member in fuzzy set ,see example(٣.٩) .

Example ٣.٩:[٤٦]

Let A be a fuzzy set of real numbers "close to 1" .The graphic of discrete MF of A is shown in figure (٣-٨) .



Figure(3-8) : A discrete membership function for "x is close to 1"

Hint :We will intensify our attention on the properties ;normality and convex, since we will need it later in other definitions and purposes.

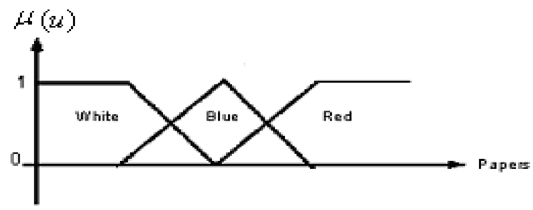
٣.٢.٥.٣ *Normality:*

The shapes of fuzzy sets are generated by its certain accepted MFs [٥١],and than in values of these functions on members of sets. These shapes have adopted several conventions influence on fuzzy sets ,the normality is one of these .

Definition ٣.٣ *:* [٥١][٥٠][١٨] A fuzzy set A in a universe of discourse U ,is said to be normal fuzzy set if $\exists u' \ni, u' \in U$ and $\mu_A(u') = 1$ that is $\max \mu_A(u) = 1$ ... (٣-٩) ,for $u \in U$.

This normalization of fuzzy set means that it have a maximum membership value(degree) of 1 [٥١].It might be thought that the normality of the fuzzy sets in linguistic variable(s) is self_evident [٤٧],as it is in numerical variables. For more understanding see example(٣.١٠) .

Example ٣.١٠ *:*Let the universal set is a set of color papers in a weekly magazine ,and suppose color a linguistic variable with classes "red", "blue" ,and "white" .See figure(٣-٩).It seems that each of these terms(classes) reach value 1 at some point in fuzzy sets.



Figure(3-9) : *Linguistic variable Color with normal classes(MFs.)*

٣.٢.٥.٥ *Sub_normal* [٥١]

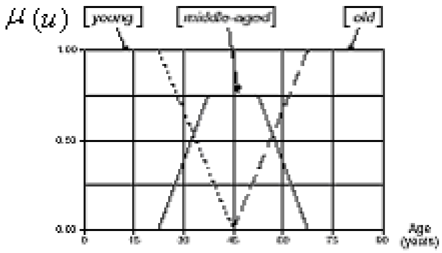
Definition ٣.٤ *:* A fuzzy set A ,is sub_normal if it is not normal , so $\exists$ no $u' \ni, u' \in U$,and $\mu_A(u') = 1$, that is,

$$\max \mu_A(u) \neq 1 \quad \ldots \quad (٣\text{-}١٠)$$

Example ٣.١١ :

Consider the linguistic variable (age) with classes young, middle aged and old ,on a set of people as universal set ,see figure(٣-١٠), while it would seem commonsense that young and old would both reach normality (although exactly where is debatable) ,when it comes to class"middle aged" the situation is more difficult,since no one ever definitely middle-aged[٤].This situation particularly in the context of specific application conceived by no .

[٤] One may be died before he has a middle-aged .



Figure(3-10) :*Linguistic variable age with subnormal MF*

We note that a class middle-aged is expressed quite naturally ,by a subnormal MF. Also note that there is an obstruct notion of a set of middle-aged people, but no specific person(one) under consider will ever be considered to be fully compatible with membership of that set[٤٨], or at least not as definitely as a person who is less than one year old is considered to be "young" ,it is just a baby .

Note: In some situations ,we may find out that all classes of some linguistic variable are subnormal(this is strange some thing),but it really exists .Such situations does not occur frequently.We can imagine a plausible example for this note .

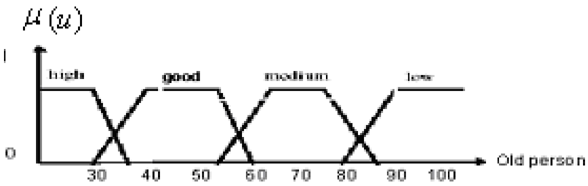
Example ٣.١٢:

Let the set of old persons as a universal set ,consider a linguistic variable "health", and a set of classes of variables at health as; low , medium ,good ,and high .In spite of these classes not exactly represent the categories of classes of this variable(dividual classes).

No person(one) will satisfy these expressions(criteria's),so all will be sub_normal classes ,that is ;

$$\max \mu(x) < 1 \qquad \ldots \quad (٣\text{-}١١)$$

and $\forall \; x \equiv$ *person* .



Figure(2-11) :*Linguistic variable Health with subnormal classes*

٣.٢.٥.٦ *Convex*:

Definition ٣.٥: [٥١] A fuzzy set A , is said to be *convex fuzzy set* ,if and only if all of its $\alpha -$*cuts (levels)* are convex in classical sense, that is for each $\alpha -$*cut (level)* A_α ,for any $r, s \in A_\alpha$,and any $\lambda \in [0,1]$, then ;

$$\lambda r + (1-\lambda)s \in A_\alpha \quad \ldots \quad (٣\text{-}١٢)$$

Convex means that any $\alpha - cut$ *(level)* which parallel to the horizontal axis through interval(for example) [١٨] .

$$A_\alpha = [a_1(\alpha), a_3(\alpha)] \quad \ldots \quad (٣\text{-}١٣)$$

, yield nesting property ,that is for α_1, α_2 _ *cuts* ,

$$\alpha_1 < \alpha_2 \quad \text{but} \quad a_1(\alpha_2) \le a_1(\alpha_1) \quad \ldots \quad (٣\text{-}١٤)$$

$$, \text{ and } \quad a_3(\alpha_1) \le a_3(\alpha_2) \quad \ldots \quad (٣\text{-}١٥)$$

Where $\alpha_1 - cut$ set is;

$$A_{\alpha_1} = [a_1(\alpha_1), a_3(\alpha_1)] \quad \ldots \quad (٣\text{-}١٦)$$

,and $\alpha_2 - cut$ set is ;

$$A_{\alpha_2} = [a_1(\alpha_2), a_3(\alpha_2)] \quad \ldots \quad (٣\text{-}١٧)$$

See figure(٣-١٢) ,then the condition of convexity implies that if ;

$$\alpha_2 < \alpha_1 \Rightarrow A_{\alpha_1} \subset A_{\alpha_2} \quad \text{or} \quad A_{\alpha_2} \supset A_{\alpha_1} \quad \ldots \quad (٣\text{-}١٨)$$

That is, for

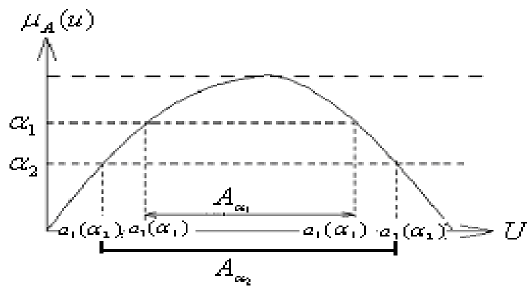
$$\alpha_2 < \alpha_1 \Rightarrow \mu(a_1(\alpha_2)) \le \mu(a_1(\alpha_1)) \quad \ldots \quad (٣\text{-}١٩)$$

,where

$$a_1(\alpha_1) \ge a_1(\alpha_2) \quad \ldots \quad (٣\text{-}٢٠)$$

and

$$a_3(\alpha_1) \ge a_3(\alpha_2) \quad \ldots \quad (٣\text{-}٢١)$$



Figure(3-12) :Convex set

The intervals° on real axis which represent $\alpha - cut$ sets are nested if fuzzy set is convex and normal [١٨] ,so as will see fuzzy set can be considered as a series of $\alpha - cut$ sets which are represented by interval

° These intervals called **intervals of confidence** [١٨] .

determined on real axis ,that is a fuzzy set(fuzzy number) is a union of $\alpha - cut$ sets .

٣.٢.٦ Membership Degree ;

Definition ٣.٦ *:* [١٨] A real number between 0 and 1 , which characterize the degree to which a member in universal set belongs to fuzzy set.

This degree is complement with MF ,and it is necessary and basic in fuzzy set ,since each member in fuzzy set is as order pair from a member and its membership degree [٢٣][١١] ,and it is a value of MF on member of fuzzy set. So a membership degree of a member u by a MF μ_A denoted by;

$$\mu_A(u) \in [0,1] \qquad \ldots \quad (٣\text{-}٢٢)$$

If $\mu_A(u) = 0$ the member is not belongs to A ,that it has membership degree zero ,so has no membership in a fuzzy set , while if $\mu_A(u) = 1$ the member is exactly in A .

Example ٣.١٣ *:*

Consider $U = \{-3,-2,-1,0,1,2,3\}$ a universal set, and let $A = \{(-2,0.6),(-1,0.8),(0,1.0),(1,0.8),(2,0.6)\}$ be a fuzzy set on U .

Note that membership degree of -2 is 0.6 belongs to A ,while membership degrees of -3 and 3 are zero ,i.e. not belong to A .

Note: *The values of membership degrees are help us to determine a shape of MF,and if it is finite or not .*

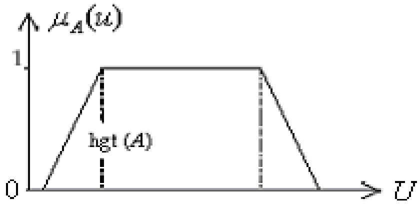
Note: *If the support of a fuzzy set is infinite ,then the degrees of members which approximate to infinite ,are approximate to zero. See (quasi fuzzy number in (٣.٤.٢)) .*

٣.٢.٧ Height of a Fuzzy Set ;

Definition ٣.٧*:* [١٨] The height of a fuzzy set A on universal set U , denoted by $hgt(A)$ is given by a supremmum of MF over all $u \in U$ and written as;

$$hgt(A) = \sup \mu_A(u) \qquad \ldots \quad (٣\text{-}٢٣)$$

which is refer to highest possible membership degree see figure(٣-١٣);



Figure(3-13) :Height of a fuzzy set

Note: In many situations not necessary the $\sup \mu_A(u)$ *equal to* 1 *as will see in sub_normal MFs .*

٣.٢.٨ Crossover Point;

Definition ٣.٨ *:* [١٨] The member(element)of universal set ,for which the MF of a fuzzy set has value 0.5 is called a *crossover point* .

Crossover element marks point where the possibility of belonging becomes higher or lower than the possibility of not belonging . Although an exact position of a crossover point(s) is usually not very important ,it is (are) characterizes a shape of MF .

Note : In some situations the MF has more than one crossover point .

Example ٣.١٤*:*

Suppose set of five balls in different size is a universal set U .
Define a fuzzy sets on U ;

A is a set of "*small balls*" ,
and B is a set of "*large balls*" ,as ;

$A = \{(1st, 0.3), (2nd, 0.5), (3rd, 1.0), (4th, 0.9), (5th, 0.8)\}$,

and $B = \{(1st, 0.7), (2nd, 0.5), (3rd, 0.0), (4th, 0.1), (5th, 0.2)\}$.

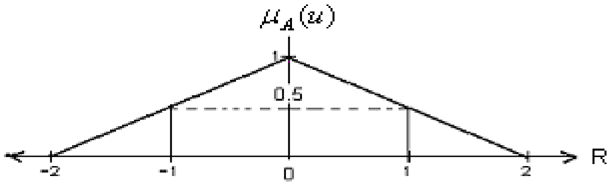
Note that in A a ٣rd ball and ٤th are almost small ,also ٥th ball is nearly almost small ,while the ٢nd ball is more or less small ,and ١st ball is almost exactly not small .

The members(balls) in B ,which are almost large is a ١st ball, ٢nd ball is more or less large ,while ٤th and ٥th ball are almost not large ,and ٣rd ball is exactly not large . Note that the crossover point is same for two fuzzy sets ,on two membership functions .

In example(٣.١٥), we will explain that fuzzy set can be with two crossover points .

Example ٣.١٥ *:*

Consider a set of real numbers R as a universal set ,define a fuzzy set A on R as a real numbers "*close to zero*" ,the graphic of this set is shown in figure (٣-١٤) . It is clear a two crossover points which are -1 and 1 ,fuzzy set taken for positive and negative real numbers, and value of MF at -1 and 1 is 0.5 .



Figure(3-14) :Two Crossover Points for a fuzzy set

٣.٢.٩ *level Set (Cut Set)*

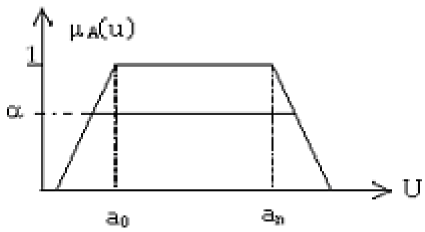
Definition ٣.٩*:* The set of elements that belongs to fuzzy set A on universal set U ,at least to the (not lower than) degree α is called $\alpha-$*level* or $\alpha-$*cut* set [١٨] ,

and it can be defined as[١١] ;

$$A_\alpha = \{u \in U; \mu_A(u) \geq \alpha\} \qquad \ldots \qquad (٣\text{-}٢٤)$$

for $\alpha \in [0,1]$.

This set of elements determined in a closed interval in universal set , if R (real numbers) is a universal set also for other situations . See figure(٣-١٥) [١١] .So A_α will be $[a_0, a_n]$,



Figure(3-15) : *Level Set*

Example ٣.١٦ *:*

Assume $X = \{-2,-1,0,1,2,3,4\}$ a universal set , and

$$A = \{(-2,0.0),(-1,0.3),(0,0.6),(1,1.0),(2,0.6),(3,0.3),(4,0.0)\},$$

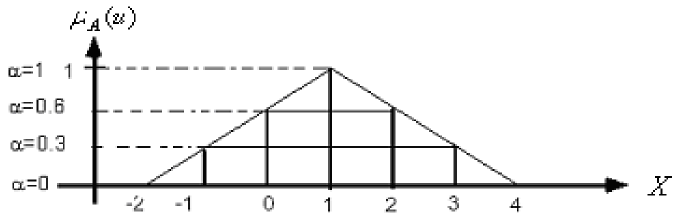
as fuzzy set ,

The level set $\alpha = 0.6$ is $A_{0.6} = \{0,1,2\}$,

at $\alpha = 0.3$ is $A_{0.3} = \{-1,3,0,1,2\}$,

and for $\alpha = 1$ is $A_1 = \{1\}$.

Note that $A_1 \subset A_{0.6} \subset A_{0.3}$,while $1 > 0.6 > 0.3$. See figure(٣-١٦).



Figure(3-16) : *Level Sets ;α= ٠.٣ ,α= ٠.٦ ,and α= ١*

Note :

١. *If $\alpha = 0$ (in some situations),the $\alpha - level$ set will equal to the support of a fuzzy set .*

٢. *The union of $\alpha - level$ sets , will equal to support if we take $\alpha - level$ set for each value between $[0,1]$, see subject (fuzzy numbers),and finally get the fuzzy set for each element with its membership degree .*

٣. *In example (٣-١٦), the $\alpha - level$ set in one subset from an other this true in general situations for $\alpha_1 < \alpha_2$,we get $A_{\alpha_2} \subset A_{\alpha_1}$, which clear the convex principle ,through the nesting property .*

٣.٢.١٠ *Support of a Fuzzy Set ;*

Definition ٣.١٠ *:* [١٨][٤٦][١١] Let A be a fuzzy (sub)set on universal set U ,then *support of* A ,denoted by $supp(A)$ or A_{sup} ,is crisp set of all members of the universal set U ,for which a members have non zero membership grades(degrees) in A, and written as ;

$$supp(A) = \{u \in U; \mu_A(u) > 0\} \qquad \ldots \quad (٣\text{-}٢٥)$$

The support can be a finite set[١٨] ,that is just a finite number of members of the universal set have a non_zero membership degree .

Example ٣.١٧ *:*

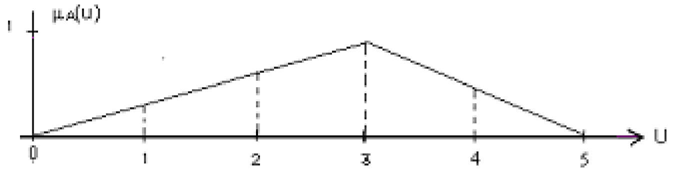
Consider set of natural numbers in $[0,5]$ as a universal set ,define fuzzy set A a set of numbers "*about* 3" ,that written as ;

$$A = \{(0,0),(1,0.4),(2,0.8),(3,1.0),(4,0.6),(5,0)\}$$

The support of A will be ;

$$supp(A) = \{1,2,3,4\}$$

See here support is finite set, see figure(٣-١٧).



Figure(3-17) :Finite Support

Also the support can be an infinite set, we can consider a membership functions depending on a fuzzy set of universal set if it have a finite or an infinite number of elements ,this mean determine the shape of membership function ,then we will know that support is finite or not .

Example ٣.١٨ *:*

Consider universal set is a set of real numbers R (infinite set) ,and define fuzzy set A as a set of numbers "*about* 3" ,it can written as;

$$A = \{(u, \mu_A(u)) : \text{either } u \geq 3 \text{ or } u \leq 3, u \in R\} \qquad \ldots \quad (٣\text{-}٢٦)$$

This set is infinite .Note support of A will be infinite also, that is ;

$$supp(A) = \{u \in R : \mu_A(u) > 0\} \qquad \ldots \quad (٣\text{-}٢٧)$$

Example ٣.١٩*:* Consider set of positive real numbers as universal set,

determine a fuzzy sets ;

A a set of "*small numbers*" ,

and B a set of "*large numbers*" .

Note that both these sets are infinite , so supports of A and B are also infinite(not defined on set) ,this gave us the infinite supports .

٣.٢.١١ *Singleton*

*Definition*٣.١١ *:* [١١]

A fuzzy set A whose support $supp(A)$ contains a single point u in U with;

$$\mu_A(u) = 1 \qquad \ldots \quad (٣\text{-}٢٨)$$

is referred to as a *fuzzy (set)singleton* .

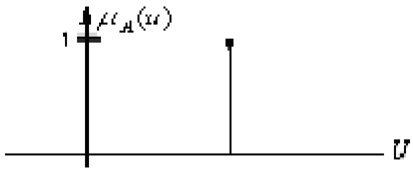
Some times called a *fuzzy point* by [٤٦],which is a single real number [١]

That is say ;

$$A_\alpha = [u, u] = \{u\} \qquad \ldots \quad (٣\text{-}٢٩)$$

, $\forall \alpha \in [0,1]$

A function which define this point is with a unity value at this one particular point and zero every where else called a singleton function [١٨].See figure(٣-١٨) ,which is a special case of MF as we noted in (MF properties in(٣.٢.٥)) .From all a fuzzy singleton means a fuzzy set has only one point .



Figure(3-18) : *Singleton*

٣.٣ *Extension Principle*

In systems that deal with control problems or systems of making decision ,…,etc ,interconnect the different variables with each other with algebraic formulas (relationships)and equations [١٨] .These relationships may reflect physical laws or a system structure ,or an influence of one variable on other variables about it's value or it's degree in a fuzzy set as will see ,is one of these systems by use this variables described linguistically as fuzzy sets .

The base of this systems (such control system) is mathematic through relations between variables for inputs to systems and variables of outputs for systems. Output is not necessary be in same kind with input variables (not membership) ,that is get a fuzzy model for a variable if we know the fuzzy model for an other variable and the functional relationship between them .This is a start for extension principle which is one of fundamentals of fuzzy set theory , is also a basic for rules of fuzzy logic algebra [١٨] ,this extension is under knowledge of structure of the system, by means of extension principle a crisp function can be evaluated for a fuzzy argument [١] ,i.e. a system description is crisp ,but with a fuzzy inputs and outputs .

Definition ٣.١٢ *:* [١٨]

If A is fuzzy set in universal U ,and f is mapping from U to universal set Y ,then extension principle allow us to define a fuzzy set B in Y as ;

$f : A \to B$,

$$B = f(A) = \{(y, \mu_B(y)) : y = f(u), u \in U, y \in Y\} \qquad \ldots \quad (٣\text{-}٣٠)$$

where $$\mu_B(y) = \begin{cases} \sup \mu_A(u) & ,if \quad f^{-1}(y) \neq 0 \ ,u \in f^{-1}(Y) \\ 0 & ,if \quad f^{-1}(y) = 0 \end{cases} \qquad \ldots \quad (٣\text{-}٣١)$$

$f^{-1} : Y \to U$, $f^{-1}(y) \in U$, $y \in Y$, $u \in U$,and $f^{-1}(Y) \subset U$

Note supremmum of operation is applied because the function f may maps different elements of universal set U into one element of universal set Y ,and so this element may inherit a few membership degrees [١٨] ,see example(٣.٢٠) .

The extension principle is a mathematical framework for system that is use input data as fuzzy and the resulting output data also fuzzy ,and a system description is as crisp [١] .Through rules which connect these information in fuzzy system and this outcomes give us an imaginary for behavior of the system .

Example ٣.٢٠ *:* Let $A = \{(-2,0.3),(-1,0.4),(0,0.8),(1,1),(2,0.7)\}$ be a fuzzy set on universal set U (a set of real numbers),define function f on U as;

$$f : U \to U \quad , \quad y = f(u) = u^2 \quad ,$$

for $u, y, \in U$.

So by extension principle $B = f(A)$,and then ;

$$B = \{(4,0.3),(1,0.4),(0,0.8),(1,1),(4,0.7)\} \ .$$

So $B = \{(0,0.8),(1,1),(4,0.7)\}$

Note that the element 4 in B has two values for membership degrees ,this means it is an image of two different elements in B (for -2 and 2 in A), so we use the definition in choose value of membership degree for element 4 in B ,by supremmum value for these two values then it is 0.7 .A same thing for the element 1 in B is image of two different members(elements) in A (-1 and 1) with two different membership degrees ,and supremmum membership value is 1 in B .

Extension principle by means the rule which defining fuzzy set(or more ,descriptive variables) on universal set which has functional relationship(s) on the universal set(s) [١٨] ,the variables of input universal set(s) may be one or more ,and also the variables of output universal set(s) may be one or more ,this add to structure of system(s) how to process these variable(s) .

Also we can describe the system by number of inputs and outputs variables as for ;Single input_Single output system ,Multi input _Multi output system , Multi input_Single output system [١].

Depending on this discussion the extension principle will be more general for inputs and outputs through use more universal sets for variables ,so the relation(function) will generalize for many variables .

Definition ٣.١٣ *:* [١٨]

Let X be a Cartesian product of universal sets $U_1, U_2, .. U_n$

$X = U_1 * U_2 * ... * U_n$,and $A_1, A_2, ..., A_n$ are fuzzy sets in these sets as , A_1 on U_1 , A_2 on $U_2, \ldots, A_n$ on U_n ,and define f as mapping from X to universal set Y ; $f : X \to Y$,where

$$Y = f(U_1, U_2, ..., U_n) \qquad \ldots \quad (٣-٣٢)$$

,then fuzzy set B in Y will be ;

$$B = \{(y, \mu_B(y)) : y = f(u_1, u_2, ..., u_n) \in Y, (u_1, u_2, ..., u_n) \in X\} \quad \dots \quad (٣\text{-}٣٣)$$

where

$$\mu_B(y) = \begin{cases} \sup\min\{\mu_{A_1}(u_1), ..., \mu_{An}(u_n)\}, if \quad f^{-1}(y) \neq 0, (u_1, u_2, ..., u_n) \in X \\ 0 \qquad , if \quad f^{-1}(y) = 0 \end{cases} \quad \dots \quad (٣\text{-}٣٤)$$

That is means we choose minimum value of membership degree for values of input variables (to output variables),and then choose the supremmum of all these minimums .

Note : The number of elements of each set is not necessary be equal.

Example ٣.٢١ *:* Consider a fuzzy sets defined on universal sets as ;

$A_1 = \{(-2,0.3), (-1,0.4)\}$ on U_1 ,a set of negative real numbers,

$A_2 = \{(-1,0.5)\}$ on U_2 ,a set of negative integers ,

and $A_3 = \{(2,0), (1,0.4), (0,0.8)\}$ on U_3 ,a set of real numbers .

so $X = U_1 * U_2 * U_3$ … (٣-٣٥)

Let $A = A_1 * A_2 * A_3$ … (٣-٣٦)

, $u \in A$, $u = (u_1, u_2, u_3) \in X$, $\forall \; u_1 \in A_1, u_2 \in A_2, u_3 \in A_3$,

define $f : X \to Y$, for a fuzzy set B on Y ,

$y = f(u) = u_1 + u_2 + u_3$ … (٣-٣٧)

then elements of A as ;

$$A = \{(-2,-1,2), (-2,-1,1), (-2,-1,0), (-1,-1,2), (-1,-1,1), (-1,-1,0)\}$$

and then B will be ;

$$B = \{(y, \mu_B(y)) : y = f(u) , u = (u_1, u_2, u_3) \in X\} \quad \dots \quad (٣\text{-}٣٨)$$

So

$$B = \{(-1,0), (-2,0.3), (-3,0.3), (0,0), (-1,0.4), (-2,0.4)\}$$
$$= \{(-3,0.3), (0,0), (-1,0.4), (-2,0.4)\}$$

Note that the element -1 has two values of membership degrees in B (0 and 0.4) ,so we used extension principle in chose the value of this membership degree by supremmum after we choose it as minimum from membership degrees of components of elements in A.Also note the element -2 has two values of membership degrees in B (0.3 and 0.4) a same manner used to choose membership of -2 .

٣.٤ *Fuzzy Numbers*

The use of linguistic terms to describe variables which has descriptive features ,like the variables which refer to human features such as ;intelligent ,health ,behavior ,tall ,age ,…etc ,(and more many features others),is simple and more favor in dealing with these features than numeric one .

In many situations, people are only able to characterize numeric information imprecisely [٤٦] ,for example ,people use terms such as; ; *about* ١٠٠٠٠٠٠ ,*near zero* ,or *essentially bigger than* ٥٠٠٠ , these are examples of what are called *fuzzy numbers*. That is use linguistic terms(words) to describe numerical values ,spite of it is vague and not exactly give values(measurements) of the variables, i.e. not one but interval or set of numeric values.

Using theory of fuzzy (sub)sets ,we can represent these fuzzy numbers as fuzzy (sub)sets of set of real numbers [١٨].

Definition ٣.١٤ *:* [٤٦][١٨]

A fuzzy set A on real line R (a set of real numbers),as a universal set with a normal,(fuzzy) convex ,and continuous membership function $\mu_A(.)$ of bounded support(finite set) ,is said to be a *fuzzy number*.

The family of fuzzy numbers will be denoted by $\mathcal{F}$ [٤٦] .In some references we note that fuzzy numbers are denoted by uppercase letters with tildes such as ; $\tilde{A}$, $\tilde{B}$, …,etc[٦],or use numeric symbols to refer to values of some variables . For example [١] ;

The equation :

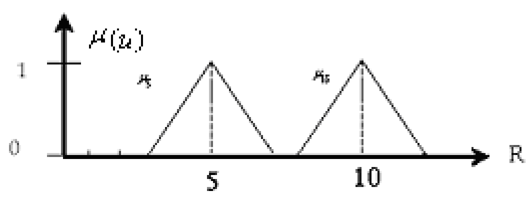
$$y = \tilde{5}\,x_1 + \tilde{2}\,x_2 \qquad \ldots \quad (٣-٣٩)$$

,where $\tilde{5}$ and $\tilde{2}$ are fuzzy numbers *"about five*" and "*about two*", respectively defined by MFs instead of use exact values .

Let us take example on fuzzy numbers .

Example ٣.٢٢ *:* [١١] Consider a fuzzy sets " *about* 5 " ,and " *about* 10" ,

which can denoted as ; $\tilde{5}$ and $1\tilde{0}$, respectively ,defined on real axis R with triangular MFs ,see figure(٣-١٩) .



Figure(3-19) :*Examples of fuzzy number*

By depending on the fact ,that a fuzzy set A can be considered as a union of $\alpha-cut$ sets [١٨][١١] as ;

٦ * The uppercase letters is refer to uncertainty values for variable see[٤١][١٨] .
** This is wide use in systems which defined by an algebraic differential equations see [١] .

$$A = \bigcup_{\alpha \in [0,1]} A_\alpha \qquad \ldots \quad (٣\text{-}٤٠)$$

Fuzzy numbers can approximately represented by a collection of it's $\alpha-cut$ (*level*) sets for various values of α [١١] .Here we will explain this situation ,the $\alpha-cut$ (*level*) set of a fuzzy number(set) A is defined as [١٨][١١];

$$A_\alpha = \{(u, \mu_A(u)) : \mu_A(u) \geq \alpha \ , u \in R\} \qquad \ldots \quad (٣\text{-}٤١)$$

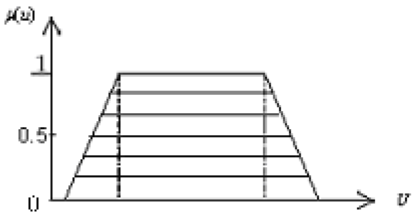
for $\alpha \in R$, A_α the $\alpha-level$ set

If we generalize this concept for each $\alpha \in [0,1]$,we get a series of $\alpha-level$ sets ,which are represented by intervals determined on real axis ,i.e. $\alpha-level$ set for A will be (for example) ;

$$A_\alpha = [a_1(\alpha), a_{10}(\alpha)] \qquad \ldots \quad (٣\text{-}٤٢)$$

So we have interval $\forall \ \alpha \in [0,1]$,here upon union of these $\alpha-level$ sets(intervals) we will approximate fuzzy number A,for more details see figure(٣-٢٠) .

Note that level sets of fuzzy numbers are closed intervals [١١] ,the called out of these intervals was occurred previously .



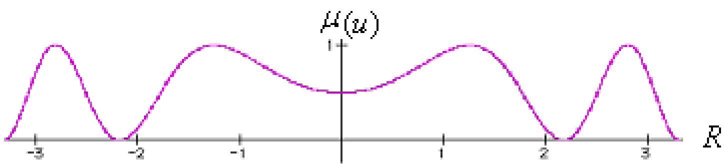
Figure(3-20) :Approximate representation of a fuzzy number by its level sets

Note : Using examples with piecewise linear membership functions is to get more precise of numeric to explain the idea only ,but that is not mean the other functions are not efficient .

٣.٤.١ Not Fuzzy Number

Definition ٣.١٥:[٤٦] A set A is a not fuzzy number if there exists β , such that $\beta \in [0,1]$, that A_β (*level* set) is not convex subset of R.

That is not satisfies condition for a set to be convex,see figure(٣-٢١), in spite of it is continuous MF (on a set) ,and normality satisfied in this situation .



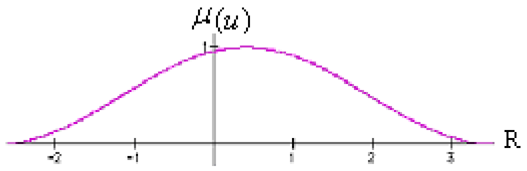
Figure(3-21) :Not fuzzy number

٣.٤.٢ *Quasi Fuzzy Number*

Definition ٣.١٦ *:* A fuzzy set A on real line R with a normal ,fuzzy convex, and continuous membership function satisfies the limiting conditions as [٤٦];

$$\lim_{u \to \infty} \mu_A(u) = 0 \quad , \quad \lim_{u \to -\infty} \mu_A(u) = 0 \qquad \ldots \qquad (٣\text{-}٤٢)$$

called a *quasi fuzzy number* .See figure(٣-٢٢) .



Figure(3-22) : *A quasi fuzzy number*

Now we must intensify on the point that a classical(typical) arithmetic and algebra deal with crisp (typical) numbers ,while fuzzy arithmetic deals with fuzzy numbers .

In problems of control ,physics ,and other scientific and industrial fields ,fuzzy numbers used to model real system variable(s) [١٨] ,it deal with variables with real values(at minimum value zero),not imaginary or negative values ,which has no mean in real world applications .

Chapter Four

The Suggested Method

٤.١ The Suggested Method Idea

The suggested method is a new perspective on fuzzy wavenets .It is combining wavelet theory with NN and FL in a single method as it applied in one of a wavelet_based techniques[١] ,and a subset of wavelet networks ,specially wavenets that in which the dilation and translation factors are fixed (not optimized) and only weights are optimized through learning (see [٨] for more) .In contradiction to other types of wavelets that a weights ,dilation factor and translation factor are optimized through learning .This not all thing, the suggested method as a novel idea not occurred in previous known applications of NNs ,and different in many structures sides .

It is represent a modification in many sides proposed on traditional methods. These sides are;

- Firstly in the training set structure .
- The main concept which is an important side is activation function nature .
- Network architecture ;which represented in using a single output layer neuron ,and in chose the desired output .
- In last; the output values structure and a taken decision depending on it .

[١] Form a wavelet_based techniques a wavelet analysis and wavelet networks , The structure of training set is chosen as a set of categories ;

$$S=\{ ١^{st} \ category , ٢^{nd} \ category , ٣^{rd} \ category ,\ldots, i^{th} \ category,\ldots, Nc^{th} \ category \} \quad \ldots \quad (٤\text{-}١)$$

Where i is index of categories ,Nc is number of categories.

Each category is consist of a set of patterns as versions of figure characterize that category ;

$$i^{th} \ category = \{ ١^{st} \ pattern , ٢^{nd} \ pattern , \ldots , j^{th} \ pattern , \ldots, Np^{th} \ pattern \} \quad \ldots \quad (٤\text{-}٢)$$

Where j is index of patterns , Np is number of patterns

The structure of activation function is depends on fact that the wavelets (as activation functions) can be decomposed into a sum of a scaling functions[٨][٩] ,but this wavelets with some special properties that a scaling functions associated to these wavelets must be ;

_ *Symmetric ,*
_ Every where positive ,
_ and with a Single Maxima .

So under these conditions ,a scaling functions can be interpreted as MFs (this idea represent a basic connection to fuzzy logic in this method ,since most of MFs have these properties) .

Essentially these MFs are chosen from a dictionary pre_defined of MFs *(see*[٨][٩][٤]*)* ,among the families of scaling functions that have properties to be symmetric ,every where positive ,and with a single maxima (these properties were discussed for MFs in *Chapter* ٣) , and among other splines and some radial functions .

The main advantage of using a dictionary of MFs (which were used in fuzzy wavenets after extended for biorthogonal wavelets and multiresolution analysis)(*see previous references*),is to associate a linguistic interpretation to each MF,and each term such as “small” ,”large“ ,…,etc ,is well defined beforehand ,through the general properties of MFs in chosen dictionary ,as in [٨] .

From that all and a previous discussion ,the activation function is expressed as a summation of two continuous functions which give output real values in outputs space ,a wavelet function and MF which from previous will be substitute the scaling functions for wavelet ,this is part from a continuous wavelet transform .

The wavelets as in wavenets is with form ;

$$\psi_{m,n} = \psi(2^{-m}(x-n)) \qquad \dots \qquad (٤-٣)$$

where $m , n \in \boldsymbol{Z}$.

A membership function[٢] is dilated and translated versions of scaling functions with steps of wavelets ,that is works step by step with wavelets $\psi_{m,n}$,as;

$$\mathrm{f}_{\mathrm{m,n}} = \mathrm{f}(2^{-m}(x-n)) \qquad \dots \qquad (٤-٤)$$

So the activation function which is a summation of these two forms(functions);

$$?_{m,n} = \psi_{m,n} + \mathrm{f}_{\mathrm{m,n}} \qquad \dots \qquad (٤-٥)$$

or

$$?_{m,n} = \psi_{m,n} + \varphi_{\mathrm{m,n}} \qquad \dots \qquad (٤-٦)$$

Note : There is ability to implement any functions from wavelets or MFs .

The suggested method is used for pattern recognition problem , steps of this method start from structure of chosen training set into ending by output structure and take decision ,these steps will illustrate in figure(٤-٣) ,but through that we will clarify some necessary sides .

Since the network is performed on a patterns recognition problem ,and a patterns with many features ,multidimensional wavelets are used ,and used a single scaling for wavelets that is a single dilation and translation parameters in all dimensions ,so it be a single_scaling multidimensional wavelet to construct a translated and dilated versions of a mother wavelet .

Once again ,since a membership function is work as a scaling function and it is work at steps with wavelet (since a random of a proposed activation function) ,single scaling _multidimensional is performed for a membership function ,too .

A network trained through a back propagation algorithm ,which is a successful algorithm in training of a multilayered networks ,in addition that it is simple and elegant .

---

[٢] Since **f** as a scaling function ,we will denote it by φ instead of **f** in later .

A weights values were chosen randomly for connections in a network ,and a weights values file is single for all patterns categories in a work ,to put a fixed criteria in compare a values for categories and then get on an excellent results .

This not all ,the structure of inputs to a network for some pattern is as a vector of features values for that pattern ,and as matrix of feature values for a patterns category .The aim through such structure of inputs is to improve a work of a network in dealing with recognition problem that implemented on .

The inputs file structure as a matrix for category patterns and its features ,patterns number represent rows number ,and the features number of that pattern is represented by number of column .So the file of inputs as matrix with dimensions as $Np \times Nf$

$$\begin{bmatrix} p_1 f_1 & p_1 f_2 & \cdots & p_1 f_f \\ p_2 f_1 & p_2 f_2 & \cdots & p_2 f_f \\ . & . & \cdots & . \\ . & . & \cdots & . \\ p_i f_1 & p_i f_2 & \cdots & p_i f_f \\ . & . & \cdots & . \\ p_{Np} f_1 & p_{Np} f_2 & \cdots & p_{Np} f_{Nf} \end{bmatrix} \quad \ldots \quad (٤-٧)$$

while vector of features values for a pattern p_i with dimension ١ × Nf ,

$$[p_i f_1 \quad p_i f_2 \quad \cdots\cdots \quad p_i f_f] \quad \ldots \quad (٤-٨)$$

where i=١,٢,...,Np index of patterns, and f=١,٢,...,Nf ,index of features number

Also the structure of outputs ,which as a closed intervals to determine a size that a values can change on it for some category.

When a traditional NNs were depending on one output value for pattern output ,that be work on to recognize it through compare outputs of the testing patterns to this single value .This using is syndrome with errors such generality or unlearning ,additionally it confine work and make it exclusive on that value and consequently decrease a number of patterns that converge to pattern .

The suggested method offer an interval of values to work on , available for output values instead that the single one previously used, this through introduce a set of fuzzifies copies different in the noise rate or fuzzy rate .From outputs of this set we can determine a closed interval characterize a category ,to know high value of outputs in that interval which means that the pattern of this output is very noisy , but network can recognize it at any what .

Also ,smallest value of outputs in the same interval ,means its pattern is very noisy but in direction that decrease output value ,but the network can recognize it .

Values in these intervals are outputs of fuzzifies copies around standard value for output ,the work of interval is in two directions into a decreasing direction and increasing direction ,some times (as will see) the standard value is a least value in interval or may be largest one .

This work was generalized from one pattern’s category to a set of categories of patterns ,which represent the modification in training set structure on other problems in NNs ,see training set at (٤.١.٤) .

All in all ,the aim through all modifications in structure is to design a network that able to recognize a pattern(s) category (converge)even in the presence of noise ,in addition ,the network must also be capable of handling with repeated pattern(s) within the testing set in future .

٤.١.١ *Conditions Survey*

The choice of a three properties of functions in a dictionary pre_defined is with regard to some facts and essential functions properties .

When a function property (function values) to be with single maxima this led to dealing with a convex functions that mean there is a standard value characterize this variable values(feature value), which be a criteria to other values around it in two directions left and right ,that is this value is a center led to take one and only one decision .For if two or more function values for one variable value , choose a maximum value .

The values from left and right represent a different degrees on this standard values until reach a minimal value in two directions remote on the center value ,which represent unbelonging to this variable (feature) value .

Also these values from left to right are in same remote on the center through the width (index)that chosen for variable values ,and then these values have an equal output values ,that in the same level (α_level set)

For x , $x_{،}$

$$|x-x_{،}| < \epsilon \qquad \ldots \quad (٤\text{-}٩)$$

$\forall$ x , $x_{،}$ in some set ,and x is a center of values , ϵ is constant

A symmetric property is avoid the function to be in one of its directions tend to converge to infinity than other side direction

For $x_١ < x < x_٢$

$/ x-x_١ / < \epsilon$ ... (٤-١٠)

and $/ x- x_٢/ < \ni$ ... (٤-١١)

that is $/ x-x_١ / \neq / x- x_٢/$ ... (٤-١٢)

$\forall\ x\ ,\ x_١\ , x_٢$ in some set , and x is a center of values ,
ϵ and $\ni$ are constants , such that $\epsilon \neq \ni$, this reflected through MFs .

This led to variation data ,and then an error in a practical applications ,since in practical work there is needs to work on a directions *on /off* (*switch on , switch off*) .

The last property is that the functions are every where positive

$\forall\ x \in\ E\ \Rightarrow f(x) > ٠$ ... (٤-١٣)

, E is an any set .

This property is necessary ,since the negative values mean nothing in the practical applications .Importance of these properties is in two sides ,in practical and even in numerical applications(mathematical) ,since the most applications of NNs are in a practical sides ,this encouraged us to use a MFs for this job .

٤.١.٢ *Satisfaction of the Conditions for Membership Functions*

As we saw in (٣.٢.٥) ,these properties of MFs can extracted from serious facts and analysis of work of a fuzzy logic through MFs , in spite than other types of functions even it satisfies these properties (some of it or all) .

Most MFs satisfies previous properties exactly and it is traditionally have these properties .

- The thing that we do not differ on it is that values of MFs are in [٠,١],which is in fact a positive values ,that is it wholly have a property to be every where positive;

$\forall\ u \in\ A\ ,\ \mu_A(u) \in [٠,١]$ ... (٤-١٤)

that is $\mu_A(u) \geq ٠$.

- The rest properties ,symmetric and a single maxima are satisfies for most MFs ,but it depend on a distribution of values in support and a membership degrees around the center value of support for that fuzzy set ,that if it satisfies a condition in symmetry definition(٣.٢.٥.٢) ,and on an absolute operator for input values .

- Also the singlity of maxima is occurred for many MFs ,or that existing many maxima for one fuzzy set (in spite it is symmetric) ,

$$\forall\ (u,\mu_A(u)) \in A\ ,\exists\ (u',\mu_A(u')) \in A\ ,\ \ni$$
$$\mu_A(u') \geq \mu_A(u) \qquad \ldots \quad (٤\text{-}١٥)$$

That if there is many values in support its membership degree is equal to maxima which in most cases equal to ١ .

Using of such MFs means that there are many values have same output and then this practically means an equal in patterns in category (if used for pattern recognition) ;

$$\forall\ (u,\mu_A(u)) \in A\ ,\exists\ (u',\mu_A(u')),\ (u'',\mu_A(u'')) \in A\ ,\ \ni$$
$$(u',\mu_A(u')) \neq (u'',\mu_A(u'')) \qquad \ldots \quad (٤\text{-}١٦)$$
$$,\ \mu_A(u') \geq \mu_A(u) \quad \& \quad \mu_A(u'') \geq \mu_A(u) \qquad \ldots \quad (٤\text{-}١٧)$$

Even when

$$\mu_A(u'') = \mu_A(u') \qquad \ldots \quad (٤\text{-}١٨)$$

that is if supposed ;

$$(u',\mu_A(u')) = (u'',\mu_A(u'')) \qquad \ldots \quad (٤\text{-}١٩)$$

we can write it as

$$\mu_A(u') \geq \mu_A(u) \qquad \ldots \quad (٤\text{-}٢٠)$$

then $\mu_A(u')$ is unique .

Using a single maxima means also ,that there is one value (here patterns in category) satisfies a standard value (maximum value) or exactly satisfies a property ,and the symmetric values around it is a fuzzifies (noisy)copies from it .

A membership functions that convex are having a single maxima[٤٨].In concept of fuzzy logic the final result will be obtained after applying all the results from rules ,then one will have a fuzzy set of results .Once again, if any element (pattern) of the universe has two or more different membership degrees as a result of different rules processing ,one can choose a maximum value [١٨] .

Note: Not necessary that the maxima is ١ [٣] *,less than one is also true .*

Note :In fuzzy logic the problem of choice a MF had not yet been solved since did not existing a significant dependence of the performance (see[١٨] ,for more details) *.*

The suggested method through a dictionary and a proposed properties not has this problem ,exactly can choose a MF has the requests .

Note :Choice of a MF for this work is lifted for user ,in condition; satisfies previous properties .

Since a MF in a suggested method, is work with wavelet as one activation function ,the scaling of these functions (a MF and wavelet) is the same ,but through an experiments performed. Also a normalization for inputs(features) values as noted is same ,since the inputs to a wavelet are the same for MF because a structure of a single function for modification these two functions .

٤.١.٣ Network Structure

Figure(٤-١) shows the structure of a fuzzy_wavenet for suggested method .It is a multilayered feedforward neural network, that is the layers are fully connected .

It is consist of three layers (input layer ,single hidden layer ,and output layer) ,this structure is similar to a ٣_layer Perceptron and general WNNs in references (*[٨][٩][٣٠][٧][٣٥]*) ,and use a dyadic wavelets specially as in(reference *[٨]*).

The input layer includes many neurons due to that patterns in a training set are with many features studied . Also since number of neurons in the input layer N_i determines dimensions of the inputs , and number of neurons in output layer N_o determines dimensions of the outputs (see *[٢٣]*) .

So that we can determine number of neurons that it wanted in input and output layer .Input neurons N_i receive inputs as numerical files for images features ,reads as values of features vector for each image in some category . The number of input neurons that used is five neurons, N_i =٥ .

[٣] When **maximum value =١** ,that means **normalized** .

The hidden layer consists of many neurons N_h,too ,due to multidimensionality of inputs which are features values of patterns in categories ,and this number was determined before start learning .The process is chiefly in these neurons ,and activation function ,that proposed as a sum of wavelet function and membership function ,the causes were declared previously so it work instead of the scaling functions for wavelets .Activation function will be a keyword for explain the novel suggested idea in this method .

In this work number of hidden neurons that were used as; five neurons in first experiment and eight for second one(N_h =٥ and N_h =٨),to exceed the computations ability and then get a good results .

The output layer as shown in figure(٤-١) consists of a single neuron .A single output layer neuron compute on a weighted sum of outputs for hidden neurons by applying a non linear function , this different from WNN which implement linear function .

Continuous functions were used .This choice of functions is to complement a work for a suggested idea ,instead using other linear one ,to avoid a conflict ,and to get a real valued which offer a continuous and wide support to determine the outputs(actual outputs) values in categories intervals ,and more accuracy , instead a linear which are in many times with integer values for outcomes and discrete values ,and then can not get on intervals .

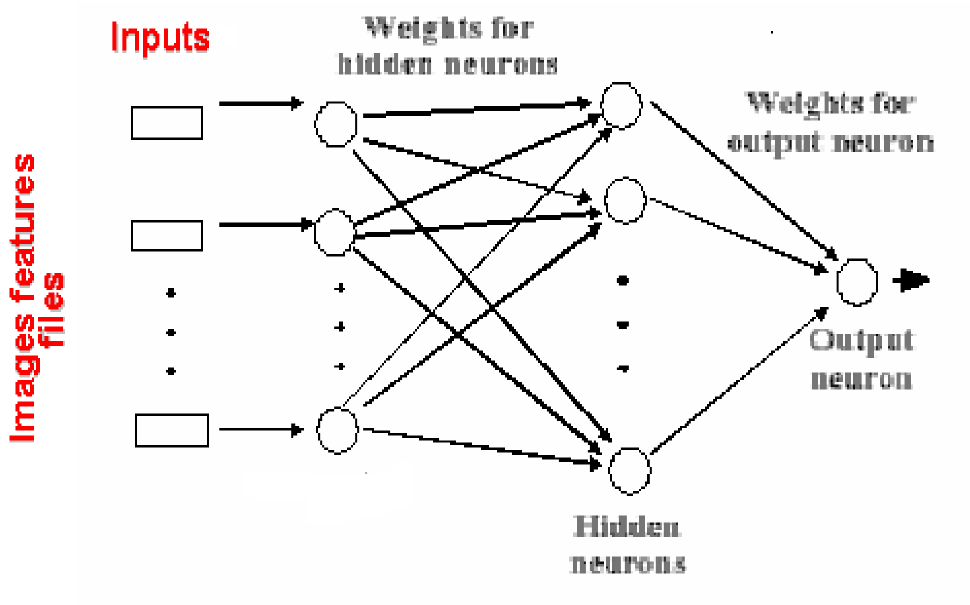
In many structures of the networks that used in recognition problem the output layer has at least two neurons(in a WNN even if it is consist of a single neuron it will be with a linear function) ,since in many networks there are a ٣ neurons or more but it were must be with a finite number of output layer neurons .

It was a need to put values of network outputs ,that it is corresponding desired outputs to be in a binary values ٠ or ١ . As in each output layer neuron the actual output is corresponds to desired value either ٠ or ١ (black or white ,true or false) .For example if we want to recognize three patterns at network with output layer has three nodes ,then the desired values for outputs will be for ;

First pattern	٠	٠	١
Second pattern	٠	١	٠
Third pattern	١	٠	٠

or any order consists of ٠ and ١ code[٤]. This manner may led to error in recognition or with the most famous problems a generalization .

Through the chosen network structure with this single neuron is led to the following things; firstly minimization number of elements of weights matrix between hidden layer and output layer(at output layer),that is minimization in number of computational operations , also minimize the outputs values ,in addition to that all decreasing in the effort .



Figure(4-1) : *Wavenet structure in the suggested method* .

The second thing is that desired values for outputs were chosen (not ٠ or ١) such previous situations .But it is chosen as a gradual values between ٠ and ١ ,(in [٠,١]) ,with different from one category than all others ,with increasing order stared with a first category at a least value ٠ ,and increased to the order of categories until end at the last category at highest value ١ .

٤ This code called a **binary coded** [٣٩] ,

٤.١.٤ Training Set

The suggested method is trained through a set of experiments on a set of images for eight categories of geometric curves (figures),not very complex figures with ٢D_dimensions ,that is not complex mean not has different details ,it is just curves can drawn with hands(if possible) in black and white colors only, do not use more colors ,this added an complexity in computations and this led to more errors ,since it is need to capture on a colors tights for different parts of image and to different images ,but in experiments of performing suggested method on colored images used the same images with colored in lines as will illustrate in experiments ,then compute more features .Consequently ,we need to have an information on work with color image processing , this need wide work ,in addition that we favor to dealing with geometric curves than other images , see figure(٤-٢) ,for patterns in a training set .

Images are with format " *Bit Map format* " BMP ,(will discussed widely in read image file) ,which is the most commonly used than other formats and it is corresponds with window systems by[٥٣][٣]. Image is entered to be processed by scanner from its recourse , and use "*Paint Program*" to capture on some sides of images .

A categories of images that used in training set are varies in complexity ,the first one rather way simple ,so on the second is seems complex one ,while the third and the fourth categories are in first look tend to be reflected ,that is the third category represent the reflection of fourth category figures(curves) in direction from left to right or vice versa in other look angle .This give an testing for network ability in recognize images may be similar ,but in inverse direction ,that is exam ability of network for generalization ,which is a most drawbacks of networks in recognitions problems .

Fifth category is a more complex than figures in two previous categories, variance values of features are large .

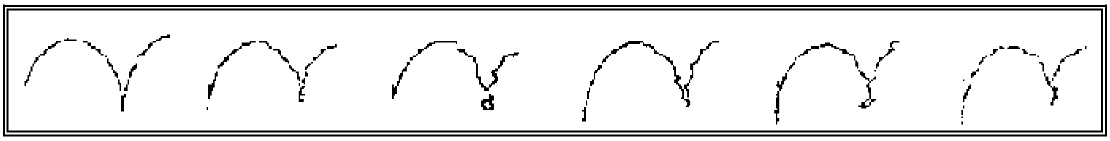
The sixth category is a more complex than fifth one ,to increase a network fortitude in identification more complex curves ,this category is complex since it have many details due to it intricacy . Seventh and eighth categories seems simple ,and looks likes ,but it sure difference .

Each figure from the eight categories is taken with six gradual fuzzifies versions(copies) ,then number of images will be six images for each category ,one standard (original) image and its fuzzifies versions .

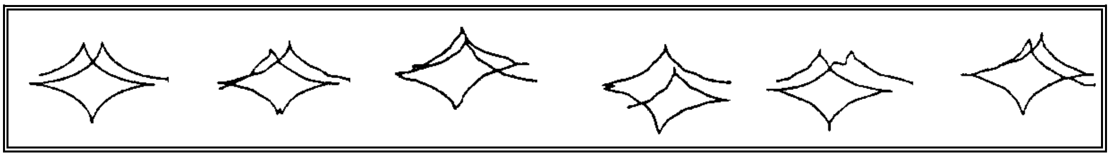
So with eight categories the number of images will be ٤٨ *images ,* that is means a number of members of training set .

Later ,will show outputs of network for the categories figures determine the closed interval of output for each category .
Then network can recognize any figure looks like any category and its output value be in this interval .

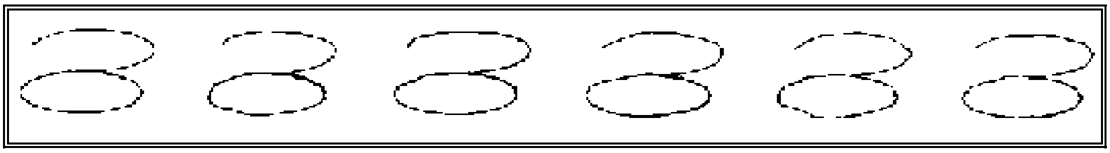
Since the work basically ,deals with mathematical side and the problems of images processing is out of limits of this research , then we will focus on the features extracted for images, what these features, and their values .



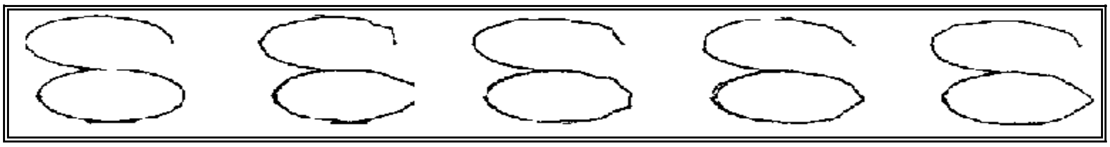
first category patterns



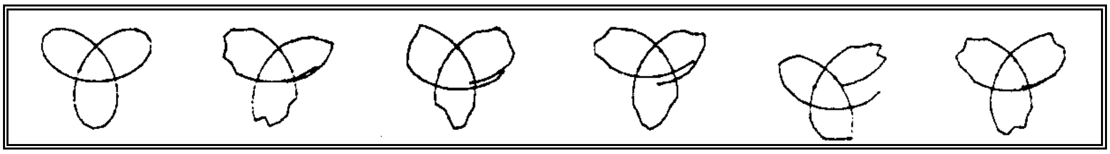
second category patterns



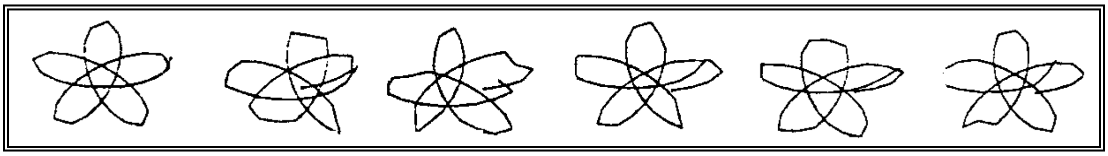
third category patterns



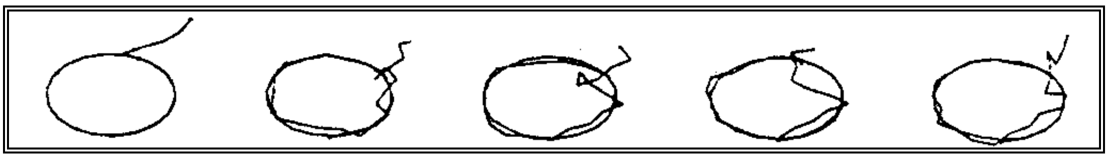
fourth category patterns



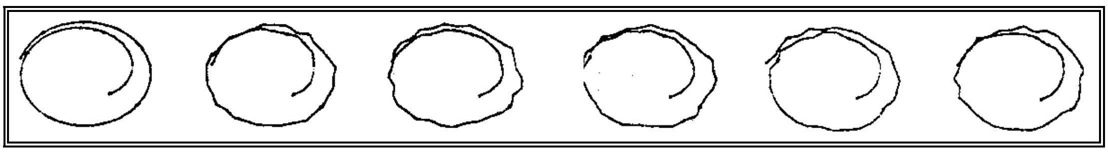
fifth category patterns



sixth category patterns



seventh category patterns



eighth category patterns

Figure(4-2): Training Set

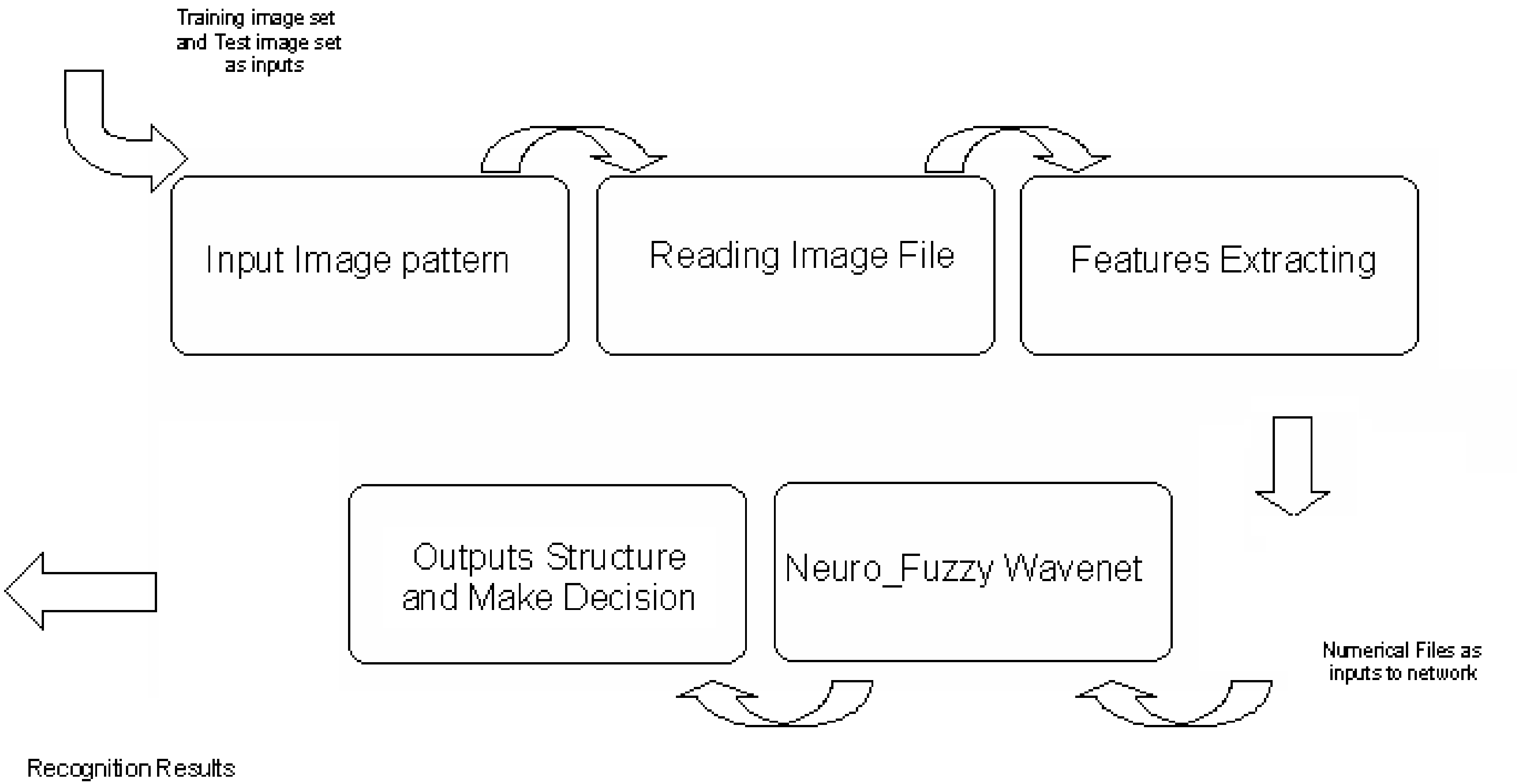
The variability between any one from a noisy versions and a standard image for some category ,can be computed in accuracy by calculations criteria called " *Fidelity Criteria* " [٥٢] ,which compute error rate between image and its noisy versions ,such "*SNR"* ,but the commonly used is a "*PSNR "* , since we need the tight accuracy if dealing with versions of a standard image of some category and the images that would be entered to be a testing set .

In same time, we can use an other type of criteria is a "*Subjective Fidelity Criteria* " in [٤٠] ,which in this situation disadvantageous ,since it is evaluate the difference through "*Human Visual System* " (see[٤٠]) ,which is in many problem is helpful but not in geometric curves ,which are mainly depend on numbers language .

Also as known that Human visual system is capable of recognize mistakes in dark parts for image very well ,than that ones in light parts ,and it is affected by errors in edges larger than an image background .That is ,if we used this criteria on images under study (in which the curves were drawn by black line on a white background) ,it is may be able to recognize the errors (differences) between images ,but still the errors in the white background or edges less to classify if not be impossible .

٤.٢ *Suggested Method Stepping*

To introduce to steps of a suggested method ,firstly see figure(٤-٣) which clarify the steps of working through this method at the diagram ,these steps start when prepare a set of images (or set of images categories) that to work on ,and then implement the suggested method through its steps as would explain exactly ;



Figure(4-3): Suggested Method Diagram

٤.٢.١ *Input Image Pattern*

To put the images under study through computer device it must be extracted from the resource(s) utilize from some ways , either by employment programs were installed in computer to initialize images such ; " *Paint Brush*" or "*Micro Soft Power Point Paint"* ,or extracted from external resource(s) by "*Scanner* " from resource papers (as example) .The images in this type either drawn by hand or using the previous discussed programs on computer.

The images that used in this work are extracted by last way from reference(*[٢٨]*) from a references list .These images were hand drawn and extracted from resource papers by Scanner .Each image separated alone to processed ,a *Paint Program* is appropriate for this purpose .

We used a closed size for all images (original and copies) by using efficient program for this objection called "*Corel Draw* ١٠" to capture on images limits(edges) ,these limits dimensions is

١٠٠ *pixels* for width ,
and ١٠٠ *pixels* for length ,

as maximum limits ,that to control on at less on one side that causes errors in measurements, or not efficiency network ,and to get an accurate calculations ,since difference in dimensions for images in one category led to increasing or decreasing in measurements values(features values) ,which led to change in outputs values for images in outputs space .That is ,it may be out of a closed interval that determined for each category ,and then fail a network in recognize through static criteria .

Also if we chose different dimensions for limits for each category than other categories ,this also led to either intersection of outputs interval of some category with other categories intervals ,this intersection put a network in generalization problem ,so not efficient recognition ,or since the values of weights that were fixed for network and for all categories ,in addition to other training parameters with fixed previous values for all categories and increasing or decreasing values for images features caused exit a network error out of the tolerance in previous situation that is fail a network in compare with fixed criteria between different categories.

Images that extracted are stored in format BMP to be prepared to applied in second stage of work .

٤.٢.٢ Reading Image File

The first process stage on the image(s) ,(that were extracted in before stage) before any modification ,is reading the image file [٤٠] .This stage depends on the image file format ,that is must know what is the image file format . There exists many formats for images files define images on computer devise .from this formats as occurred in [٥٣] ; BMP ,JPG ,TIF ,PCX ,PNG ,…,etc .

We say before that the format for image file is BMP°, , which is the most commonly used than other formats,
The BMP file consists of three parts ;

- A first and main part is the "*Header*", which include ٥٤ *bytes* and it has the necessary information (documents) that command with the rest parts of image file .These information describe the image through ;image width and high (depth) in pixels ,number of bit per pixel for each image pixel ,…,etc. [٥٣][٥٢] ,such file modification data [٥٣] .

- *The Color Pallet* represent second part for image file, which refers to color tight for (RGB) true colors (Red , Green, and Blue) gray [٥٢][٤٠] . The size that this part is depend on is effected by number of bit per pixel that computed in first part from image file .

- The third part of image file is *Image Data* [٥٢] .

Note: There are many stages can be implemented on image(s) , dependence on image Segmentation ,it is out of our specialty in this research .

° It is some times reads as ***windows Bit Map Format*** [٥٣] .

٤.٢.٣ Features Extracting

It is means decreasing image dimensions through extract some numerical measurements from image data [٣] ,from it's benefits , get high compression ratio [٦] to decrease calculation time .

In this stage of work ,"*Histogram Features*" [٥٢] ,were exploited for gray level values into a number of image pixels at each value .

A *Histogram* is a graph used in image analysis ,that shows the distribution of intensities in an image .The information in a histogram can be used to choose an appropriate enhancement operation, for example ,if an image histogram shows that the range of intensity values is small ,we can use an intensity adjustment function to spread the values a cross a wider range [٥٣] .

So it is use a probability distribution ,and then extract a statistical features from this distribution .Statistical features represent a histogram features [٣].

The probability density function(PDF) is computed at each gray level by [٥٢];

$$P(g) = \frac{N(g)}{M} \qquad \dots \quad (٤\text{-}٢١)$$

where $P(g)$ is a (color) probability density function for level g ,
g is index gray levels,
$N(g)$ is number of image pixels at level g ,
and M is number of image pixels(multiply columns number by rows number),since image as numerical matrix .

Hint:The order of presenting these features is as it computed in a program and it is the same order also for features of images in features files for the categories .

A number of features in the experiments of suggested method are five features ,which are ; the mean , standard deviation , skewness ,energy ,and entropy .While number of images for patterns category are six images; original image and it's copies, so we get on a matrix dimensions .

[٦] When used in image compression applications .

Probability density function exploited to calculate or extract the following features ;

١. *Mean* :

The mean gray levels for image ,is computed as[٤٠];

$$\bar{g} = \sum_{g=0}^{L-1} g \; .P(g) \qquad \ldots \quad (٤\text{-}٢٢)$$

$\bar{g}$ is mean gray levels , g is index of gray levels ,
L is a gray levels total number .

So the index g will be from ٠ to L-1, if bits number is ٨ then L =٢^٨ =٢٥٦ ,since L =٢^n , such that n is number of levels bits .

٢. *Standard Deviation* :

A standard deviation between levels is computed as;

$$\sigma_g = \sqrt{\sum_{g=0}^{L-1} (g - \bar{g})^2 \, .P(g)} \qquad \ldots \quad (٤\text{-}٢٣)$$

where σ_g is a standard deviation at level g , [٣].

٣. *Skewness* :

The skewness for image is computed as[٥٢] ;

$$Skew = \frac{1}{\sigma_g^3} \sum_{g=0}^{L-1} (g - \bar{g})^3 \, .P(g) \qquad \ldots \quad (٤\text{-}٢٤)$$

٤. *Energy* :

It is computes for image levels as [٣][٥٢];

$$Energy = \sum_{g=0}^{L-1} [\, P(g)]^2 \qquad \ldots \quad (٤\text{-}٢٥)$$

٥. *Entropy* :

Entropy measurement for bits number required for “Coding Operation ”[٧] [٤٠] ,and it is computed as ;

$$Entropy = \sum_{g=0}^{L-1} P(g).\log_2[P(g)] \qquad \ldots \quad (٤\text{-}٢٦)$$

Note : These features that extracted from image(s) are saved within a numerical files for each pattern. It is a measurements of features values for standard image ,and its versions for some pattern saved in file different from other patterns files.

Note : In ending the last step of previous stages , it start the process in network ,that shown in figure(٤-٣).

The extracted features values are in numerical form ,for more accuracy ,not described by using linguistic variables values ,we saw such this description in fuzzy logic .The numerical files for features values are represented for each images(patterns) category ,that is each features file for each patterns category ,as a matrix of image features number(represent rows number), and number of images in category (the column number represented by images number for that determined patterns category) .

Features values in this numerical files computed through program for feature extraction(see a program that communicated with a reference [٣]), performed by programming language “ *Turbo C++* ” ,this program exploit the equations compute these features through algorithm for extract these features from each entered image(s) to be process by this program .

[٧] It depends on number of image pixels nearby that have same color value ,that is the frequency of these values for pixels in levels ,it is also reflect the efficiency of coding if the image is simple less number of colored levels .

The algorithm and steps of extract features can be implemented by " *Math works*(*MATLAB*) " with any version .

Also we can get on these features values by using "*Coral Draw* ١٠ " ,which compute these features in addition to more features such a median ,number of pixels in each image ,…,etc ,indirectly with out implementing any algorithm just entire image(s) to be processed by program .

In a numerical files ,a matrix elements are a features values for a patterns in this category .Each row is a vector of a five features values for that image .Each value is a feature value from these five features for some image in that category ,that is a data of each category are in matrix determined for this category ,and each row is a point in a features space .

The patterns that belong to same category have approximately the same values for features .

Note : *There is a problem with a fuzzy system is that it is difficult to deals with two many features* [١٠] , *but the use of MFs here as scaling functions ,while the wavelet represent the expansion of data is in* $L^2(R^n)$ [٤٣][٣] .

٤.٢.٤ Neuro_Fuzzy Wavenet

The steps of working in this stage are gradual as follows;

٤.٢.٤.١ Inputs Network Structure

After the past steps implemented on images under study for training or testing set ,and extract features values from all these images we get numerical files for features values of images for each pattern . In other word these numerical files are represent the inputs to network in the suggested method to be studied or processed .

Consider a set of N_f features (that be extracted from each image),detectors as;

$$(f_١ ,\dots ,f_{Nf})$$

these values represent a vector of values for some pattern .

The set of patterns(entered to network) for *p_patterns* where p =١,…, N_p , N_p is number of patterns in that file .

Each file represent a data of some category from the eight categories used .

That if consider the set S_j of inputs of patterns at *j_th category* ,and N_c number of categories under study ;

$$S_j = \{ (x_١ ,y_١(f_f)) , (x_٢ ,y_٢(f_f)) ,.., (x_p,y_p(f_f)) ,\dots, (x_{Np},y_{Np}(f_f)) \}$$
$$= \{ (x_١ ,y_١(f_١ ,\dots ,f_{Nf})) , (x_٢ ,y_٢(f_١ ,\dots ,f_{Nf})) , ..., (x_p,y_p(f_١ ,\dots ,f_{Nf}))$$
$$, \dots, (x_{Np},y_{Np}(f_١ ,\dots ,f_{Nf})) \} \qquad \dots \qquad (٤\text{-}٢٧)$$

where f=١,…, N_f index of features numbers ,
N_p number of patterns in category ,
p index of patterns numbers
$y_p(f_f)$ is a point in a features space, that is with ٥_dimensions , while $(x_{Np},y_{Np}(f_f))$ if it expressed by values of $y_p(f_f)$ for f=١,…, N_f ,it will be as a vector or pattern features values vector x_p , the input numerical file is consist of these vectors .

Each file for category can be expressed in previous form .If we index for inputs of the eights categories S_j , j=١,.., N_c, and N_p for training set; N_p, that is number of members in S_j . While in testing set for each category N_p is different .

From that all each category has a set of its patterns information .

With multidimensionality of features space ,we so have multi neurons in input layer ,so that the input layer can assimilate the entire values to it .

٤.٢.٤.٢ *Weights Initialization*

To apply the back_propagation algorithm the network weights must first be initialized [٥٤] .These weights are initialized to small random values ,since it has much significance in working of NN , and characterization of a network .It would optimized through learning ,but chose initial weights too large will make a network untrainable since it reach a " *Premature Saturation* " .

The reason to choice a weights randomly to give a wide and general basis for weights values and appropriate for all patterns in a training set .

There is method to compute weights depending on inputs values to a network[٨] ,this method means that weights values change for each input ,so the parameters values and results will change also ,and then success for pattern or/and fail for other . So each pattern will has weight file different from other patterns ,but this in a suggested method not give a constant criteria to compare between patterns ,so we used just one weights file for all patterns ,change for each trained patterns category to test it ,that give network high criteria in compared between outputs of the network .

Choice of a random weights and basis is for all weights between input and hidden layer ,also between the hidden and output one chosen randomly since weights on hidden and output layer is less effective on a network work ,than a first one .

In other hand a normalization operation is accomplished to this random weights that were initialized in this work and for all experiments ,which performed the random weights normalized to values in [٠ ,٠.٢],(*see program of performing*) ,to avoid a high values in random values .Weights values between an input and hidden layer are in a matrix ,since there are many neurons in input layer and many neurons in hidden layer ,number of hidden neurons is refers to columns number ,while number of input neurons is refers to rows number as in form fellow ;

[٨] This method was occurred in reference [٣] .

$$\begin{bmatrix} w_{11} & w_{12} & .... & w_{1j} \\ w_{21} & w_{22} & .... & w_{2j} \\ . & . & .... & . \\ . & . & .... & . \\ w_{i1} & w_{i2} & .... & w_{ij} \\ . & . & .... & . \\ w_{N_i1} & w_{N_i2} & .... & w_{N_iN_h} \end{bmatrix} \quad ... \quad (٤\text{-}٢٨)$$

A weights values between hidden and output neurons as vector values its rows number is a number of hidden neurons ,while the column number is output neurons number ,but there is a single output neuron so these values as vector ;

$$[w_{11} \quad w_{12} \quad .... \quad w_{jNo}]^T \quad ... \quad (٤\text{-}٢٩)$$

where T is a vector transform .

٤.٢.٤.٣ *Fuzzy Wavelets Coefficients with a Dilation and Translation Parameters*

A wavelets in this work were chosen as a single scaling wavelet that is contain a single scaling parameter and translation in all dimensions ,since the work is with multidimensional wavelets due to a many features of patterns and a many patterns in a category (multidimensionality is necessary) .A single scaling multidimensional wavelets to construct a translated and dilated versions of mother wavelet .

From this wavelet which dilated and translated through single scaling ,and since a MFs work as a scaling functions ,and it work communicate with wavelets a single scaling_ multidimensional is implemented on it ,too, as in form(٤-٤) .

If a multi_scaling wavelets were used ,a wavelets would have independent dilation and translation parameters for each dimension .Since for the form[٧];

$$\psi_j = \psi(\frac{x - m_j}{dj}) \quad ... \quad (٤\text{-}٣٠)$$

and m_j, d_j are the index of translation and dilation factors , respectively .

So what happen ,multidimensional(many dimension) case from a computational point of view to multi_scaling wavelets result in a large complexity . Also for MFs associated to these wavelets would effected by this multiplicity which led to complexity in a fuzzy logic part framework and then complicate a problem more and more .

So would suppose a values for dilation and translation by trying, with a constant values for all wavelets in NN ,with integer values by dilation as in the form(٤-٣) with;

$$m = integer\ number\ ,$$

and translation as ;

$$n = integer\ number\ .$$

so the form(٤-٣)(٤-٤) will be;

$$\mathrm{f}_{m,n} = \mathrm{f}(2^{-m}(x_i - n)) \quad \ldots \quad (٤\text{-}٣١\ a)$$

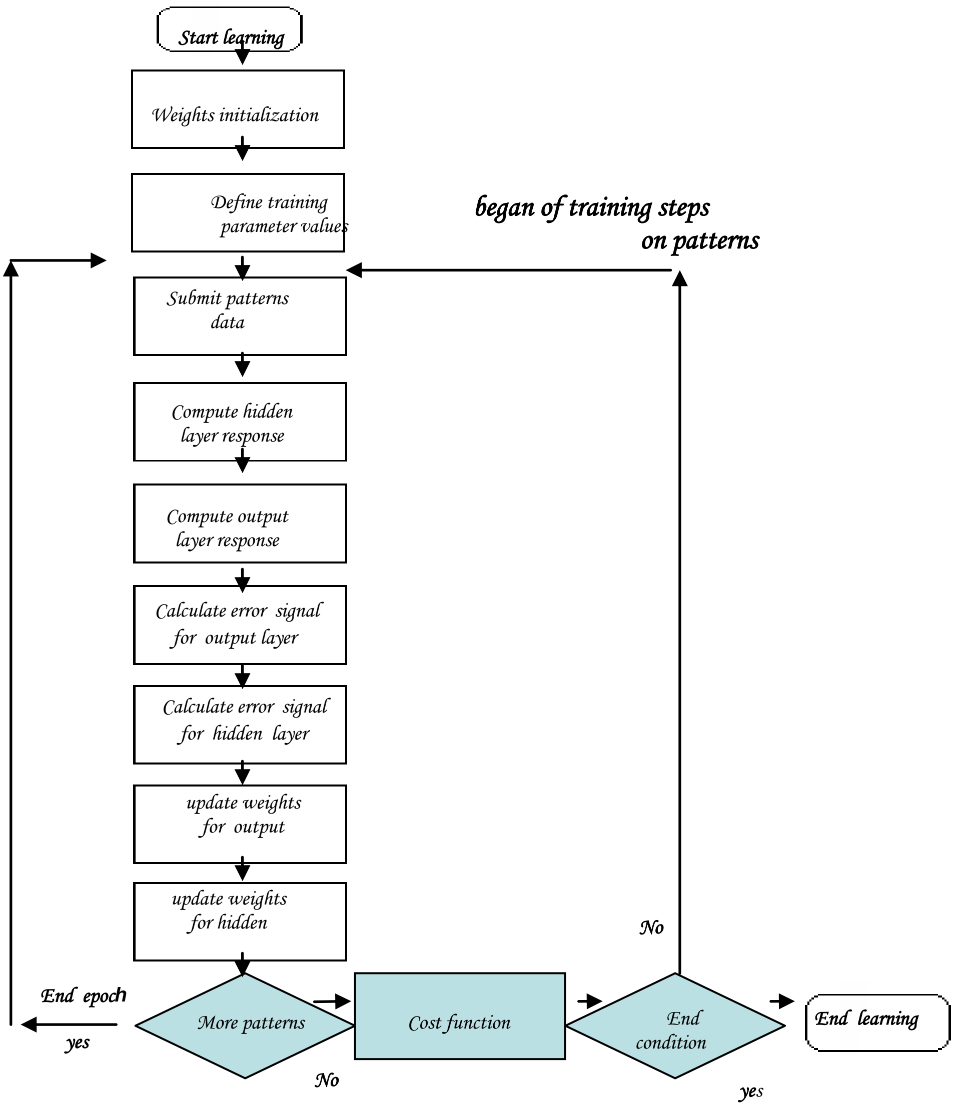
$$\psi_{m,n} = \psi(2^{-m}(x_i - n)) \quad \ldots \quad (٤\text{-}٣١\ b)$$

they were chosen as $m >١\ ,\ n > ٠$.

٤.٢.٤.٤ Back Propagation Algorithm

Over all process of back_propagation learning including both forward and back ward bass [٢٤] .Forward of the input training patterns ,and a backward of the associated error and adjustment of weights ,figure(٤-٥) illustrate a flow chat of this algorithm with its two directions .

Back propagation is used here due to its simplicity , since it is suitable for any continuous function . A necessary weights of a network were initialized previously in (٤.٢.٤.٢) ,which were randomly, then know can began a working ,see the references ([٦][٢٤][٣]) that demonstrate this traditional algorithm .



Figure(4-5) : Back propagation flow chat .

Step ١ :

Given a primary values for a network training parameters which were chosen by the NN user ,This parameters are; a *learning rate* (ε)is proposed at most experiments with value (ε=٠.١), *momentum* (γ) is also proposed the same for all experiments with (γ=٠.٠٠١) ,a *maximum iterations* to do before stopping (*MAXITER*)[٩] is given a number ٤٠٠٠٠٠ for all experiments in this work ,and a *minimum error tolerance* that required for convergence(*ERRTOL*), its value is change for a supposed experiment to work .

Step ٢ :

Submit patterns data is represented by introduce information about entered values for each pattern in training set (which represented in a next steps) .These data are a desired output value for patterns categories in training set ,which were proposed beforehand ,since we are dealing with supervised training ,that is the desired output value is known to us (or supposed randomly) for patterns .In implementing a suggested method on training set through experiments a desired output values were chosen within [٠,١] .Each category has a desired output value different on other categories ,a patterns that in same category has a same desired output value ,each category then has a desired outputs file as a vector of number of patterns in a category and its desired outputs values ,which were equals (these files as vectors since there is a single neuron in output layer) .

From other parameters information a total patterns in training set (*NumPatterns*) ,order to present patterns (*PatPresOrder)* , the index of current pattern used in training (*CurrPat*=٠) ,which will increase during learning ,and begin the index of iteration number with ٠ .

Define a primary values for ; maximum error over a full epoch Worst error(*WrstErr*) ,average error over a full epoch(*AvgErr*) ,mean squared error(*MSE*) ,all with values=٠ ,also index of epochs which started with ٠ ,and increased for more patterns .

Additionally, calculate a learning rate for current epoch updated on primary learning rate value and depending on maximum iteration puts to training ;

---

[٩] The symbols words with lighting line such *MAXITER* , *ERRTOL* ,…etc , have been explained clearly through the communicated program for this work .

$$\varepsilon(n+1)=\varepsilon(n).(1-n/N) \quad \dots \quad (٤\text{-}٣٢)$$

N is max iteration ,
n current epoch ,
and $\varepsilon(n)$ is a previous learning rate value .

This step is a beginning of forward pass side from algorithm , which starts at input layer .

Step ٣ :

Submit patterns and compute a hidden layer response ,that is load training set , training step starts here .Patterns features values are presented and entered to input layer ,as known no action on input layer ,but the Normalization[١٠] ,must be accomplish on inputs values .Input patterns values in this work are normalized to a maximum value for patterns values in a features values vector , as if a feature (f_f) value is a maximum value then the input x_p will normalized with respect to ;

Where $(x_{Np}, y_{Np}(f_f))$ input point

$$(x_p / (f_f)) \quad \dots \quad (٤\text{-}٣٣)$$

for f=١ ,…,N_f ,
and p=١,…, N_p ,for pattern p

In other hand the normalization operation accomplished to a random weights that initialized in previous step ,that is random weights normalized to [٠ ,٠.٢] .

The process is in the hidden layer on a weighted sum of inputs , a number of hidden neurons is determined before start learning(as we say). Act of each hidden neuron is through multi_ dimensional activation function[١١] ,since here dealing with patterns data with many features , and continuous values .This activation function is chosen as a sum of multi_dimensional wavelet function and membership function .as in form (٤-٥) .

[١٠] The need for such procedure ,that the data may be in text form in a disk file(this mean order patterns to submitted into network) ,or may be in integer form but the network input array requires a real type ,also the data to be tearned (input pattern values)may be outside the allowed range of network , so the input data must be adjusted to accommodate the needs of the network .
[١١] the form of activation function here is part of wavelet transform (analysis) .

Activation function used in this work as a sum of ***Mexican hat wavelet*** ,which is a second derivative Gaussian function as wavelet side ,and ***Gaussian membership function*** ,since it is satisfies a previous discussed properties ,then the form is;

$$\Phi_j = (1-x^2).e^{-x^2/2} + e^{-x^2/2} \qquad \dots \quad (٤-٣٤)$$

where j =١,…, N_h is index hidden neurons,

N_h number of hidden neurons .

Here with ,the input value x to activation function is translated and dilated as in form of dyadic wavelet (٤-٣),

Since

$$x_i = \sum_{i=1}^{N_i} x_i \,.w_{ij} \qquad \dots \quad (٤-٣٥)$$

N_i is number of input neurons ,

i is index of input neurons .

So the input to hidden layer is in form ;

$$z_{ij} = [2^{-m}(x_i - n)] \qquad \dots \quad (٤-٣٦)$$

$m, n \in Z$

Then the outputs values for hidden layer neurons computed as;

$$h_j = \Psi_j(\mathrm{x}) + \Phi_j(\mathrm{x}) \qquad \dots \quad (٤-٣٧)$$

where Ψ , Φ are from multi_dimensionality ,and it is computed as;

$$\Psi_j(\mathrm{x}) = \prod_{i=1}^{N_i} \psi(z_{ij}) \qquad \dots \quad (٤-٣٨)$$

$$= \psi(z_{1j}) \cdot \psi(z_{2j}) \cdot \dots \cdot \psi(z_{3j}) \cdot \dots \cdot \psi(z_{N_i j}) \qquad \dots \quad (٤-٣٩)$$

Also

$$\Phi_j(\mathrm{x}) = \prod_{i=1}^{N_i} \varphi(z_{ij}) \qquad \dots \quad (٤-٤٠)$$

$$= \varphi(z_{1j}) \cdot \varphi(z_{2j}) \cdot \dots \cdot \varphi(z_{3j}) \cdot \dots \cdot \varphi(z_{N_i j}) \qquad \dots \quad (٤-٤١)$$

so the output which expressed in equation (٤-٣٧) ,will be as ;

$$h_j = \prod_{i=1}^{N_i} \psi(z_{ij}) + \prod_{i=1}^{N_i} \varphi(z_{ij}) \qquad \dots \quad (٤-٤٢)$$

$$= \psi(z_{1j})\cdot\psi(z_{2j})\cdot\ldots\cdot\psi(z_{3j})\cdot\ldots\cdot\psi(z_{N_i j}) + \varphi(z_{1j})\cdot\varphi(z_{2j})\cdot\ldots\cdot\varphi(z_{3j})\cdot\ldots\cdot\varphi(z_{N_i j}) \quad \ldots \quad (٤-٤٣)$$

For a chosen functions this formula will be ;

$$h_j = \prod_{i=1}^{N_i}[(1-z_{ij})^2 . e^{-z_{ij}^2/2}] + \prod_{i=1}^{N_i}[e^{-z_{ij}^2/2}] \quad \ldots \quad (٤-٤٤)$$

$$= \left[\prod_{i=1}^{N_i}\left[(1-z_{ij})^2\right]\right].e^{\sum_{i=1}^{N_i} -z_{ij}^2/2} + e^{\sum_{i=1}^{N_i} -z_{ij}^2/2}$$

Note : The w_{ij} is weights values matrix between input and hidden layer ,since input and hidden layer are with many neurons ,so the dimensions of this matrix is $N_i \times N_h$,where N_i is number of input neurons ,while N_h is number of hidden neurons .

Step ٤ :

Outputs of the hidden layer are entered to output layer as inputs .A received outputs (of hidden layer) of neurons h_j is multiplicated with random weights w_{jk} between hidden and output layer .The output response by implementing continuous function on ;

$$z_{ij} = \sum_{j=1}^{N_h} w_{jk} \, . h_j \quad \ldots \quad (٤-٤٥)$$

k index output neurons , $k = 1,\ldots,N_o$.

The output is

$$O_{p_k} = C(\sum_{j=1}^{N_h} w_{jk} \, . h_j) \quad \ldots \quad (٤-٤٦)$$

p is index of patterns N_p .

But we used in architecture of network just one neuron in output layer ,that is $k = 1 = N_o$.

We can implement a continuous function that in a hidden neurons as in output neuron ,so;

$$O_{p_k} = \Psi(z_{jk}) + \Phi(z_{jk}) \quad \ldots \quad (٤-٤٧)$$

since $k = 1$.

Note :The weights w_{jk} in this situation is weights values vector, not a matrix .

Note : As an other direction , we can implement , any continuous function ,also we can use wavelet(in hidden layer neurons) or membership function, or even sigmoid .

The ending of these steps represent ending of forward pass part from this learning algorithm .

Step ٥ :

Error values vector at output layer neurons for pattern p ,is represented by vector since it is computed at one layer ,but in the light of architecture of network ,error values here is not as vector put it is only single element for pattern p under training .

$$\delta_k \text{ or } \delta_{N_o=1} = (d_p - O_{p_k}).(O'_{p_k}) \quad \ldots \quad (٤\text{-}٤٨)$$

$$O'_{p_k} = \Psi'_k + \Phi'_k \quad \ldots \quad (٤\text{-}٤٩)$$

Step ٦ :

Error values vector at hidden layer computed through error values at hidden neurons ,which represented as;

$$\delta_j = (\sum_{k=1}^{N_o} \delta_k . w_{jk}).h'_j \quad \ldots \quad (٤\text{-}٥٠)$$

So by ;

$$h'_j = \Psi'_j + \Phi'_j \quad \ldots \quad (٤\text{-}٥١)$$

when implement this for numbers it will be as[٣] ;

$$= \varphi(\overset{p}{z_{1j}}) .\varphi(\overset{p}{z_{2j}}). \;\ldots\ldots \varphi'(\overset{p}{z_{ij}}). \ldots\ldots \varphi'(\overset{p}{z_{N_i j}})$$

$$+ \psi(\overset{p}{z_{1j}}) .\psi(\overset{p}{z_{2j}}). \;\ldots\ldots \psi'(\overset{p}{z_{ij}}). \ldots\ldots \psi'(\overset{p}{z_{N_i j}}) \quad \ldots \quad (٤\text{-}٥٢)$$

where $\varphi'(\overset{p}{z_{ij}})$ is a wavelet derivative at z_{ij} point value for pattern p , and $\psi'(\overset{p}{z_{ij}})$ is a membership function derivative at z_{ij} point value for pattern p to hidden layer .

Step ٧ :

Update weights values between hidden and output layer at epoch ($n+1$) depending on epoch (n) as follows [٢٤][٦]; Firstly calculate a weight correction term which used to update a weights values later ;

$$\Delta w_{jk} = \varepsilon(\delta_k . O_{p_k}) \qquad \ldots \quad (٤\text{-}٥٣)$$

and for bias ;

$$\Delta b_k = \varepsilon . \delta_k \qquad \ldots \quad (٤\text{-}٥٤)$$

In practice a momentum term γ is frequently added to equation as an aid to more rapid convergence in certain problem domains . It has effect of smoothing the error surface in weights space by filtering out high_frequence variations .

Weights are adjusted in the presence of momentum by ;

$$\Delta w_{jk}(n+1) = \varepsilon(\delta_{p_k} . O_{p_k}) + \gamma \Delta w_{jk}(n) \qquad \ldots \quad (٤\text{-}٥٥)$$

Δw_{jk} is a change in weights between hidden and output layer,
δ_{p_k} error signal at output neuron at pattern p ,
n index of epochs , γ is smoothing parameter (momentum) ,
ε is a learning rate .

Since ;

$$\Delta w_{jk}(n+1) = w_{jk}(n+1) - w_{jk}(n) \qquad \ldots \quad (٤\text{-}٥٦)$$

so current weight value;

$$w_{jk}(n+1) = w_{jk}(n) + \Delta w_{jk}(n+1) \qquad \ldots \quad (٤\text{-}٥٧)$$

The same steps are implemented for adjust bias values at output layer ;

$$\Delta b_k(n+1) = \varepsilon . \delta_{p_k} + \gamma \Delta b_k(n) \qquad \ldots \quad (٤\text{-}٥٨)$$

also since

$$\Delta b_k(n+1) = b_k(n+1) - b_k(n) \qquad \ldots \quad (٤\text{-}٥٩)$$

then it can be written as ;

$$b_k(n+1) = b_k(n) + \Delta b_k(n+1) \qquad \ldots \quad (٤-٦٠)$$

Step ٨ :

Update the weights values between hidden layer and input layer as[٣][٦];

$$\Delta w_{ij} = \varepsilon.\delta_j.h_j \qquad \ldots \quad (٤-٦١)$$

and,

$$\Delta w_{ij}(n+1) = \varepsilon(\delta_j.h_j) + \gamma \Delta w_{ij}(n) \qquad \ldots \quad (٤-٦٢)$$

Once again since ;

$$\Delta w_{ij}(n+1) = w_{ij}(n+1) - w_{ij}(n) \qquad \ldots \quad (٤-٦٣)$$

then ;

$$w_{ij}(n+1) = w_{ij}(n) + \Delta w_{ij}(n+1) \qquad \ldots \quad (٤-٦٤)$$

where $j = 1,\ldots,N_h$,and $i = 1,\ldots,N_i$

and for bias ;

$$\Delta b_j(n+1) = \varepsilon.\delta_j \qquad \ldots \quad (٤-٦٥)$$

since ;

$$\Delta b_j(n+1) = b_j(n+1) - b_j(n) \qquad \ldots \quad (٤-٦٦)$$

then ;

$$b_j(n+1) = b_j(n) + \Delta b_j(n+1) \qquad \ldots \quad (٤-٦٧)$$

Note : We do not adjust the translation and dilation factors as it previously supposed in WNNs ,but it is stay fixed at training and for all inputs values .

Step ٩ :

This step is a condition step .It is work at complete loading the training patterns in training set .It is conditioned if the current pattern is less than total patterns in training set ,so must back to (*Step* ٢) to get a next training pattern and so on to other steps .

If contradiction ,that is all patterns have been loaded into the network algorithm ,then go on to next step from algorithm .

Step ١٠ :

Compute cost function in aim to minimize the usual error measurement which is measured errors (MSE) [٣][٢٤].

Mean square error (E_p) for pattern p is computed as ;

$$E_p = \frac{1}{2}\sum_{k=1}^{N_o}(d_p - O_{p_k})^2 \qquad \ldots \quad (٤-٦٨)$$

but $k=1$,so the equation (٤-٦٩) will be;

$$E_p = \frac{1}{2}(d_p - O_{p_k})^2 \qquad \ldots \quad (٤-٦٩)$$

For generalization ,this formula to all patterns in training set N_p (total number of patterns) ;

$$MSE = \frac{1}{2}\sum_{p=1}^{N_p}\sum_{k=1}^{N_o}(d_p - O_{p_k})^2 \qquad \ldots \quad (٤-٧٠)$$

which is a sum of N_p errors computed for singled patterns using equation(٤-٧٠),since $k = N_o = 1$ so ;

$$MSE = \frac{1}{2}\sum_{p=1}^{N_p}(d_p - O_{p_k})^2 \qquad \ldots \quad (٤-٧١)$$

In light of this measured values we could compute *AvgErr,*and *Worst Err* values for training on a patterns categories ,(in *Step* ٢) .

Step ١١ :

To ending a learning steps ,it is must be examined .This end is occurred through two dependence conditions ;

- If error is within tolerance then this situation need take decision ,also exit error value than tolerance ,and that the current epoch is exceed on a maximum iteration is led to take decision .

- If epochs are still within tolerance maximum iterations the algorithm pass will going to a *Step* ٢ .

٤.٢.٤.٥ Outputs Structure and Make Decision

٤.٢.٤.٥.١ Network Outputs Structure

Output of neural network is represent a response of a network for entered patterns(values) ,that is the result of implementing functions and network parameters on input values through possible value for error (available error) for work of this net ,and then outputs will be points (or elements) in outputs space available for network .

An outputs structure in the suggested method is the essential view led to suppose this method ,it is a novel idea in recognize patterns through real valued in closed intervals in outputs space which is a part of real numbers field R ,that offers a wide basis in classify a pattern(s) to its category(s) through its actual output(s) from network in a closed interval(s) contain all real valued that converge to that patterns ,this real values represent the output values for other patterns in the same category ,or may be other patterns converge to that pattern of category .

This basis is offered by using the proposed activation function in structure of a network as a summation of two continuous functions a wavelet(***Mexican hat***) and membership function (***Gaussian*** MF) which with infinite support ,since choose a linear or function with a finite ,this restrict the network through just output values in many situations .

Choice of this MF offers a necessary factors for satisfies and success the suggested idea .Main idea is clarified by implementing a suggested method on pattern recognition problem led to that actual output is real value in R,and for recognition a pattern(s) (that the network was trained on) by determine an output for each pattern from some category in closed real interval in outputs space , then a range in R.For each category(in training set) there is a closed interval in range of the network ,determine an available values for pattern(s) in each category, these values characterize a pattern(s) to be in this interval ,and then a category ,or not also for pattern(s) that to be tested if it in some category depending on its output(s) that converge to values in some closed interval .

In traditional NNs ,the available range for network is just one point (for each pattern)in outputs space ,this point is represent an actual output for pattern ,and the output of a pattern(s) that wanted to be compare with (or to exam) if it in the same category of the first pattern ,must be at less equaled .

Since the output layer in many times contain many neurons . The output is computed through output of these neurons ,these led a network to be with generalization problem for define pattern(s) not belong to the same category with patterns in training set ,that is network generalize it's converge for some pattern into pattern(s) not converge to it at all ,which lose accuracy of network in recognize patterns ,or fail a network in converge to any pattern in training set, and consequently, fail in find result for recognition problem .

In suggested method ,the response of a network is through actual output for entire pattern(s) in network's range .The range is divided into a set of closed intervals (as say represent an outcomes of a patterns of category).actual output for some pattern is a one value in category interval ,which is a set of all values of patterns for that category in training set ,since from the structure of network there is one neuron in output layer for this purpose .

Input pattern(s) if entered to be recognized to its category its output value(s) must at less lie within closed interval of patterns outputs values for this category ,so pattern(s) output(s) (to be tested) must be in the interval(s) of category(s) in training set .A network must find a convergence between outputs into one interval of range ,that is able to identify a category of pattern(s) under test .

Note: *When the network fail in converge to any one from these intervals ,then it is fail in recognize that pattern(s) .*

Note: *The output(s) of pattern(s) must belongs to just one closed interval for some category, and it is must not belong to two (or more) closed intervals of categories at all.*

A basic ,necessary and important thing that must heeded to it is that when we divide the range of network into a set of real intervals this intervals represent a decision regions ,that is a network can make decision or response for inputs in these regions. A values in outputs space that lie out these closed intervals (range) or between these intervals or out of range ,represent a non decision regions or neutral regions on other decision regions (intervals).

These regions are also an intervals between categories intervals .Result of network with such situation is failed convergence in this regions ,since there are no pattern(s) output(s) (from training set) in it ,which mean no thing to be recognize ,there is no patterns of categories that these values converge to it ,but inversely it is an error region that the network can not make or even take decision in it .

Now ,let us explain this all mathematically and in light of patterns outputs used in training set .The range as a set of closed intervals N_j in R ,where j is index of $\mathcal{N}_c$ categories ,for $j=1,\dots,\mathcal{N}_c$, $\mathcal{N}_c$ is number of categories .

We can put it in form;

$$N_j = [s_j, t_j] \qquad \dots \quad (٤-٧٢)$$

s_j is minimum boundary in interval N_j ,
t_j is maximum boundary in interval N_j ,
and $s_j, t_j \in R$.

So the range T will be a set of the intervals N_j ,and $T \subset R$ as;

$$T=\{ N_1, N_2, \dots, N_j, \dots, N_{\mathcal{N}_c} \} \qquad \dots \quad (٤-٧٣)$$

$$=\{ [s_1, t_1], [s_2, t_2], \dots, [s_j, t_j], \dots, [s_{\mathcal{N}_c}, t_{\mathcal{N}_c}] \} \qquad \dots \quad (٤-٧٤)$$

or $$T=\{ N_j \} = \{ [s_j, t_j] \} \qquad \dots \quad (٤-٧٥)$$

for $j=1,\dots,\mathcal{N}_c$

The difference c between these intervals ,which refer to the errors between outputs for two consequent intervals are computed as different between maximum boundary for one category and minimum boundary for consequence category as form ;

$$c = s_{j+1} - t_j \qquad \dots \quad (٤-٧٦)$$

where $j= ١,..., N_c$.

Regions between decision regions is computed as difference between two consequent interval, using the law of difference between interval operation ;

$$[s_j ,t_j]- [s_{j-١} ,t_{j-١}]= [s_j - t_{j-١} , t_j - s_{j-١}] \qquad \dots \quad (٤-٧٧)$$

A range of category interval r ,which reflect the intervals capacity to output(s) value(s) for pattern(s) may be within a maximum boundary and a minimum boundary ,then for the interval $[s_j ,t_j]$, as ;

$$r = t_j - s_j \qquad \dots \quad (٤-٧٨)$$

for $j= ١,..., N_c$

٤.٢.٤.٥.٢ *Make Decision*

A networks that used for categorization as in [٥] ,and pattern recognition as in [٢٨],is need to be as a decision tool in such applications .

The need to take decision is for satisfying (or not) a previous conditions .

- If error is within tolerance and the current epoch is less than maximum iteration ,then declares success learning which led to saving a trained weights to be fixed values used in testing stage later .
  Also save archive parameters values .

- The declaring failure learning is decision ,that is occurred if error value is exit over than tolerance ,and the current epoch is exceed on a maximum iteration ,that a network is did not converge to pattern(s) that need to be recognized .

- While the converge is occurred if the network is able to recognize pattern(s) submitted to it to its category .

Note : The take decision is depending on outputs space which is part from R (real valued field) .

At testing step the making decision is difference .The patterns output values compared with the determined closed interval values of any category(to work on at training set) .

- If the pattern output is with in the interval(for some category) ,the declaration here is as : This pattern is within that category (for example in second category).

- If it do not the declaration as :This pattern is not in this category ,try an other category interval .

- If the pattern output do not belong to any one from previous interval, the decoration is same for a previous decoration at any category interval worked on .

٥.٦ Fourth Experiment :

A fourth experiment idea is proposed after showing the effective of increasing in number of hidden neurons in a network structure at a second experiment ,and increasing the number of patterns in each category from a training set at a third experiment. In both ,the aim was to decrease error value computed through training . So number of hidden neurons increased from ٥ *neurons* into ٨ *neurons* ,which is a same used in second experiment . Also increasing number of patterns in each category at training set to ٧ patterns for each category .So number of patterns in a training set at the fourth experiment be ٥٦ *patterns* .These patterns (as said before)were extracted in a same way ,and features also were extracted in a same file for a category (in a third experiment) .

Desired output values for each category patterns outputs are same for other experiments ,so no needed to recall it here again . The inputs (as in a previous experiments) were normalized to enter a network .Once again ,same steps would accomplish on inputs and then outputs values such in a first experiment, which means more computational operations .

Output values would be ordered increasingly (that we depend it on) to determine a boundaries of a closed interval of outputs values for each category in the light of output values for training patterns .Then compute a range (which reflect a capacity of each closed interval) for each category to be with knowledge how much can this interval convergence output value(s) to a standard output value (actual output for each pattern) for that category .The ranges(scopes) of error regions were also computed ,and compared with others in a previous experiments .

Aim of this experiment is associate the aims from a second and third experiments .In generally It is aim to decreasing errors values that computed at training on each category . Also try to eliminate ranges(capacities) of error regions between the decision regions (closed intervals) ,that positively reflected on an efficiency of a suggested method that generally improve the work and then implementing in more complex applications .

Before clarify a gotten results for training ,firstly will show the table values that proposed for a training parameters and other essential structure parameters in a network for this experiment .

A training parameters which are; a learning rate is proposed at ٠.١ ,it will be modified during learning ,momentum value is equal to ٠.٠٠١ . Also the training is under maximum number of iterations for work at ٤٠٠٠٠٠ .

Learning rate	٠.١
Momentum	٠.٠٠١
Max. iterations	٤٠٠٠٠٠
Translation	٦
Dilation	٢
Available error	٠.٠٠٠٧
Hidden neurons	٨
Number of patterns in category	٧

Table (5-6):Proposed training parameters values at fourth experiment

First Category :

Results of training on this category were illustrate in figure (٥-٣٧) ,which clarify the changing in error value over an iterations number within a proposed limited number for iterations .

Mean square error value is equal to ٠.٠٠٠٦٨, which a maximum error value during training ,and is a largest error can a network reaches at this experiment ,after ٤٤٠٣ iterations through ٦٢٩ epochs , also it a maximum number of epochs and iterations for training a network on all categories in a depended training set at this experiment .

Worst error value over a full epochs is equal to ٠.٠٣٦٩٥ , which is a worst error value that a network reaches during a training on all categories at this experiment ,and the average error value over a full epochs is equal to ٠.٠٠٥٢٨ ,which is also a largest average error value from error values through training on all categories at the fourth experiment .

Learning rate is traditionally updated through training ,and its value after end the training is ٠.٠٦٠٩٢ ,which a minimum rate for learning through training on this category at this experiment .

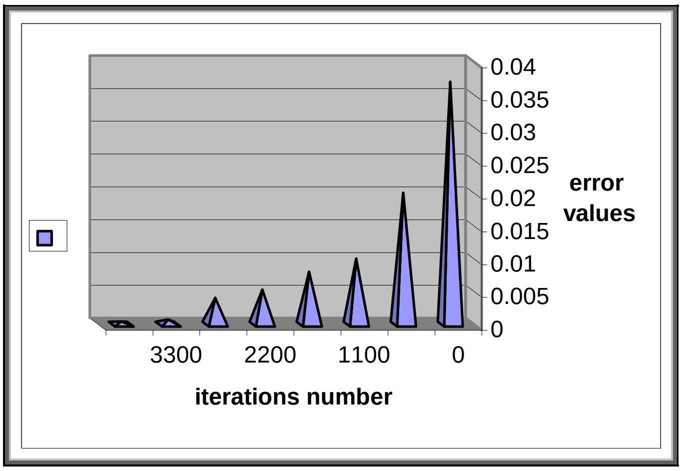


Figure (٥-٣٧):Changing error values over iterations number for a first category at the fourth experiment

Output values for each pattern in this category under training are;

output[٧]= ٠.٠٣٨٣١
output[١]= ٠.٠١٩٧٨
output[٢]= ٠.٠١٢٤٣
output[٣]= ٠.٠٤٤٩٥
output[٤]= ٠.٠٢٠٠٠
output[٥]= ٠.٠٤٠٢٦
output[٦]= ٠.٠٣٦٩٢

This order for patterns output values is same for a third experiment output for first category ,which appears here as a patterns ordered in this category after adding the seventh pattern. Also note the outputs values changed for patterns than a first experiment .

Standard output value is;
output[١]= ٠.٠١٩٧٨

which traditionally ,different from that values in previous experiment .

After reordering output values ,under a depended order in this work ,which is the increasing order will be;

output[٢]= ٠.٠١٢٤٣
output[١]= ٠.٠١٩٧٨
output[٤]= ٠.٠٢٠٠٠
output[٦]= ٠.٠٣٦٩٢
output[٧]= ٠.٠٣٨٣١
output[٥]= ٠.٠٤٠٢٦

output[٣]= ٠.٠٤٤٩٥

Differences values for patterns outputs values around a standard output value through depended order be;

٠.٠٠٧٣٥

٠.٠٠٠٢٢

٠.٠١٧١٤

٠.٠١٨٥٣

٠.٠٢٠٤٨

٠.٠٢٥١٧

The real closed interval for a first category ,at this experiment ,that determined by a patterns outputs values for this category ;

[٠.٠١٢٤٣, ٠.٠٤٤٩٥]

which is -as said- a decision region for pattern(s) recognition in first category at this experiment (if used) ,through it a network can recognize a pattern(s) if it in a patterns of this category or not .

The range of this closed interval for a first category is ٠.٠٣٢٥٢ which is larger than that one in previous experiments .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٩ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Second Category

Results of training on this category were illustrate in figure (٥-٣٨) ,which clarify the changing in error value over an iterations number within a proposed limited number of iterations .

A mean square error value is equal to ٠.٠٠٠٥٥ , which is less than an available proposed error value through training on this category after ٧٠٠ iterations through ١٠٠ epochs .Worst error value over a full epochs is equal to ٠.٠٣٣٤ ,and the average error value over a full epochs is equal to ٠.٠٠٤٧٦ .Learning rate after update is ٠.٠٩٨٧٤ .

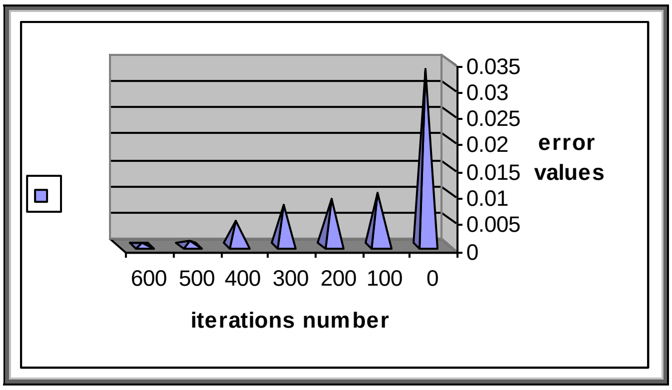


Figure (٥-٣٨):Changing error values over iterations number for a second category at the fourth experiment

Output values for each pattern in this category under training;

output[٧]=٠.١٩١٢٩
output[١]=٠.١٨٩٥٩
output[٢]=٠.١٩٢٢٤
output[٣]=٠.١٨٦٦٩
output[٤]=٠.٢٠٩٩٢
output[٥]=٠.١٩٩٦٤
output[٦]=٠.١٨٣٠٦

This order for patterns output values is appear here as a patterns ordered in this category after adding the seventh pattern . Also note outputs values changed for patterns than a first experiment ,once again changing a standard output value at this experiment .

Standard output value is;

output[١]=٠.١٨٩٥٩

After reordering output values ,under a depended order in this work will be;

output[٦]=٠.١٨٣٠٦
output[٣]=٠.١٨٦٦٩
output[١]=٠.١٨٩٥٩
output[٧]=٠.١٩١٢٩
output[٢]=٠.١٩٢٢٤
output[٥]=٠.١٩٩٦٤
output[٤]=٠.٢٠٩٩٢

Differences values for patterns outputs values around a standard output value through previous order be;

٠.٠٠٦٥٣
٠.٠٠٢٩٠
٠.٠٠١٧٠
٠.٠٠٢٦٥
٠.٠١٠٠٥
٠.٠٢٠٣٣

The real closed interval for a second category ,at a fourth experiment ,that it determined by a patterns outputs values for this category ,and after restrict the limits of a values ,will be;

[٠.١٨٣٠٦ , ٠.٢٠٩٩٢]

which is a decision region for pattern(s) recognition for a second category at this experiment (if used) ,through it a network can recognize a pattern(s) if it in a patterns of this category or not .

The range of this closed interval for a second category is ٠.٠٢٦٨٦ ,that is different from a previous values .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٥ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Third Category

Results of training on this category were illustrate in figure (٥-٣٩) ,which clarify the changing in error value over an iterations number within a proposed limited number for iterations .

Mean square error value is equal to ٠.٠٠٠٣٦, is less than an available proposed error value ,and it is a least error value for training on all categories ,that gotten on after ٩١ iterations through ١٣ epochs ,this number of epochs and iterations is a minimum number through training on all categories in a depended training set at this experiment .Worst error value over a full epochs is equal to ٠.٠٢٦٩٨ , which is a least worst error value for training on all categories ,that is the error be less at this category training, and the average error value over a full epochs is equal to ٠.٠٠٣٨٥ .

Learning rate value after updating through training is equal to ٠.٠٩٩٩٧ ,which a largest updated value during the training at this category at fourth experiment ,that means a proposed value for training is suitable for this category to learning on .

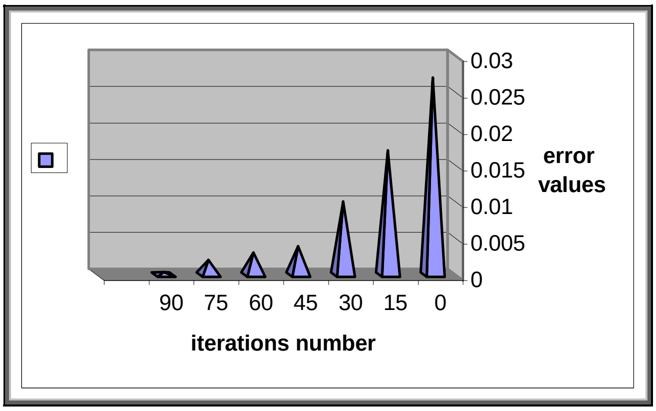


Figure (٥-٣٩):Changing error values over iterations number for a third category at the fourth experiment

Output values for each pattern in this category under training;

output[٧]=٠.٢٦٨٧٢

output[١]=٠.٢٧٦٥٨

output[٢]=٠.٢٧٨٨٥
output[٣]=٠.٢٧٨٠٣
output[٤]=٠.٢٧٩٢٩
output[٥]=٠.٢٧٧٠١
output[٦]=٠.٢٧٦٦١

This order for patterns output values appears here as a patterns ordered in this category after adding the seventh pattern (that were talking on beforehand).

Standard output value is;

output[١]=٠.٢٧٦٥٨

After reordering output values ,under a depended order in this work ,which is the increasing order will be;

output[٧]=٠.٢٦٨٧٢
output[١]=٠.٢٧٦٥٨
output[٦]=٠.٢٧٦٦١
output[٥]=٠.٢٧٧٠١
output[٣]=٠.٢٧٨٠٣
output[٢]=٠.٢٧٨٨٥
output[٤]=٠.٢٧٩٢٩

Differences values for patterns outputs values around a standard output value through depended order be;

٠.٠٠٧٨٦
٠.٠٠٠٠٣
٠.٠٠٠٤٣
٠.٠٠١٤٥
٠.٠٠٢٢٧
٠.٠٠٢٧١

The real closed interval for a third category ,at a fourth experiment ,that it determined by a patterns outputs values for this category ,and after restrict limits of values will be;

[٠.٢٦٨٧٢, ٠.٢٧٩٢٩]

which is a decision region for pattern(s) recognition for a third category at this experiment (if used) ,through it a network can recognize a pattern(s) if it in a patterns of this category or not .

The range of this closed interval for a third category is ٠.٠١٠٥٧ that different from a previous values .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category,

the network was able to recognize ١٣ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Fourth Category

Results of training on this category were illustrate in figure (٥-٤٠) ,which clarify changing in error value over an iterations number within a proposed limited number for iterations .

Mean square error value is equal to ٠.٠٠٠٤٦ ,after ٣٤٣ iterations through ٤٩ epochs .

Worst error value over a full epochs is equal to ٠.٠٣٠٤٣ , and the average error value over a full epochs is equal to ٠.٠٠٤٣٤ .

Learning rate after update through training is ٠.٠٩٩٦٩ .

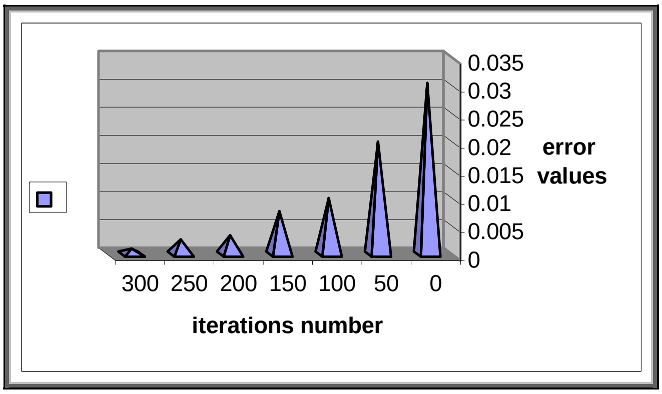


Figure (٥-٤٠):Changing error values over iterations number for a fourth category at the fourth experiment

Output values for each pattern in this category under training are;

output[٧]=٠.٣٠٧٤٠
output[١]=٠.٣١٦٧٠
output[٢]=٠.٣١٨١٢
output[٣]=٠.٣١٩٧٦
output[٤]=٠.٣٢٠٧٩
output[٥]=٠.٣١٦٢٣
output[٦]=٠.٣٢٠٠٤

This order for patterns output values is appear here as a patterns ordered in this category after adding the seventh pattern .

Also note outputs values changed for patterns than first experiment ,once again changing a standard output value at this experiment .

Standard output value is;

output[١]=٠.٣١٦٧٠

After reordering output values ,under a depended order in this work ,which is the increasing order will be;

output[٧]=٠.٣٠٧٤٠
output[٥]=٠.٣١٦٢٣
output[١]=٠.٣١٦٧٠
output[٢]=٠.٣١٨١٢
output[٣]=٠.٣١٩٧٦
output[٦]=٠.٣٢٠٠٤
output[٤]=٠.٣٢٠٧٩

Differences values for patterns outputs values around a standard output value through a depended order be;

٠.٠٠٩٣٠
٠.٠٠٠٤٧
٠.٠٠١٤٢
٠.٠٠٣٠٦
٠.٠٠٣٣٤
٠.٠٠٤٠٩

The real closed interval for a fourth category at fourth experiment ,that it determined by a patterns outputs values for this category ,and after restrict the limits of a values ,will be;

[٠.٣٠٧٤٠ , ٠.٣٢٠٧٩]

,which is a decision region for pattern(s) recognition for a fourth category at this experiment (if used) ,through it a network can recognize a pattern(s) if it in a patterns of this category or not .

The range of this closed interval for the fourth category is ٠.٠١٣٣٩ that is different from a previous values .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٤ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Fifth Category

Results of training on this category were illustrate in figure (٥-٤١) ,which clarify the changing in error value over an iterations number within a proposed limited number for iterations .

Mean square error value is equal to ٠.٠٠٠٠٣٩ , that is gotten on after ٨١٩ iterations through ١١٧ epochs .

Worst error value over a full epochs is equal to ٠.٠٢٨١٨ ,and the average error value over a full epochs is equal to ٠.٠٠٤٠٢ .

Learning rate after update through training is ٠.٠٩٨٢٨ .

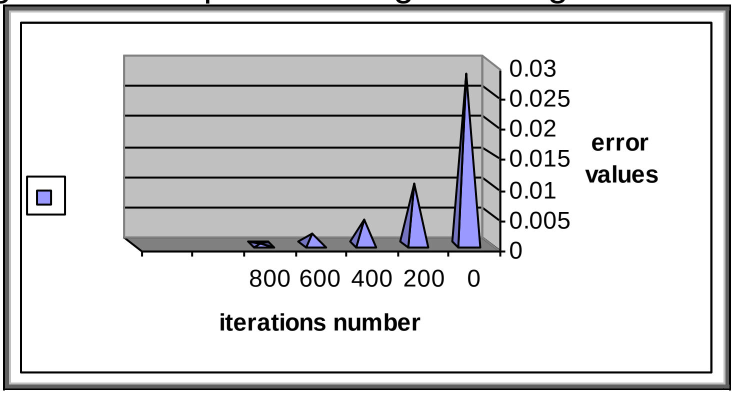


Figure (٥-٤١):Changing error values over iterations number for a fifth category at the fourth experiment

Output values for each pattern in this category under training ;

output[٧]=٠.٤٦١٢٦
output[١]=٠.٤٦٨١٣
output[٢]=٠.٤٧١١٥
output[٣]=٠.٤٧٤١٠
output[٤]=٠.٤٦٩٧٩
output[٥]=٠.٤٨٣٥٣
output[٦]=٠.٤٧٢٣٨

This order for patterns output values is appear here after adding the seventh pattern .

Standard output value is;

output[١]=٠.٤٦٨١٣

it traditionally ,different from that values in previous experiment .

After reordering output values ,under a depended order in this work ,which is the increasing order will be;

output[٧]=٠.٤٦١٢٦
output[١]=٠.٤٦٨١٣
output[٤]=٠.٤٦٩٧٩
output[٢]=٠.٤٧١١٥
output[٦]=٠.٤٧٢٣٨

output[٣]=٠.٤٧٤١٠
output[٥]=٠.٤٨٣٥٣

Differences values for patterns outputs values around a standard output value by depending order be;

٠.٠٠٦٨٧
٠.٠٠١٦٦
٠.٠٠٣٠٢
٠.٠٠٤٢٥
٠.٠٠٥٩٧
٠.٠١٥٤٠

The real closed interval for a fifth category, at fourth experiment ,that it determined by a patterns outputs values for this category ,and after restrict the limits of a values ,will be;

[٠.٤٦١٢٦, ٠.٤٨٣٥٣]

which is decision region for pattern(s) recognition for a fifth category at this experiment ,through it a network can recognize a pattern(s) if it in a patterns of this category or not .

The range of this closed interval for the fifth category is ٠.٠٢٢٢٧ that is different from a previous values .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٨ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Sixth Category

Results of training on this category were illustrate in figure (٥-٤٢) ,which clarify the changing in error value over an iterations number within a proposed limited number for iterations .

Mean square error value is equal to ٠.٠٠٠٣٧, that is gotten on after ٨١٩ iterations through ١١٧ epochs .Worst error value over a full epochs is equal to ٠.٠٢٧٣٧ ,and average error value over a full epochs is equal to ٠.٠٠٣٩١ .

Learning rate after update through training is ٠.٠٩٨٢٨ .

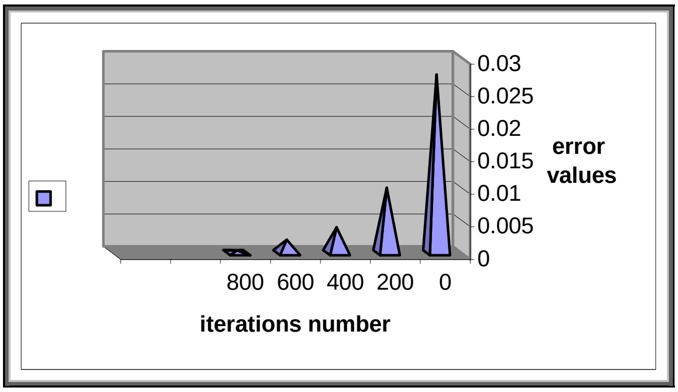


Figure (٥-٤٢):Changing error values over iterations number for a sixth category at the fourth experiment

Output values for each pattern in this category under training are;

output[٧]=٠.٦٠٨٧٦
output[١]=٠.٦٢٤٨٣
output[٢]=٠.٦٢٧٣٥
output[٣]=٠.٦٢٨٧٦
output[٤]=٠.٦٢٢٠٧
output[٥]=٠.٦٢٣٨٩
output[٦]=٠.٦٢٣٢٥

This order for patterns output values is same for a third experiment output for sixth category .which appear here as a patterns ordered in this category after adding the seventh pattern. Also note the outputs values changed for patterns than a first experiment ,once again changing a standard output value at this experiment .

Standard output value is;

output[١]=٠.٦٢٤٨٣

After reordering output values ,under depending order in this work ,which is the increasing order will be;

output[٧]=٠.٦٠٨٧٦
output[٤]=٠.٦٢٢٠٧
output[٦]=٠.٦٢٣٢٥
output[٥]=٠.٦٢٣٨٩
output[١]=٠.٦٢٤٨٣
output[٢]=٠.٦٢٧٣٥

output[٣]=٠.٦٢٨٧٦

Differences values for patterns outputs values around a standard output value by depending order be;

٠.٠١٦٠٧
٠.٠٠٢٧٦
٠.٠٠١٥٨
٠.٠٠٠٩٤
——————
٠.٠٠٢٥٢
٠.٠٠٣٩٣

The real closed interval for a sixth category ,at a fourth experiment ,that it determined by a patterns outputs values for this category ,and after restrict the limits of a values ,will be;

[٠.٦٠٨٧٦, ٠.٦٢٨٧٦]

which is a decision region for pattern(s) recognition for a sixth category at this experiment,through it a network can (some what) recognize a pattern(s) if it in a patterns of this category or not .

The range of this closed interval for sixth category is ٠.٠٢٠٠٠ that is different from a previous values .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١١ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Seventh Category

Results of training on this category were illustrate in figure (٥-٤٣) ,which clarify changing in error value over an iterations number within a proposed limited number for iterations .

Mean square error value is equal to ٠.٠٠٠٣٩, that is gotten on after ٨٩٦ iterations through ١٢٨ epochs .Worst error value over a full epochs is equal to ٠.٠٢٨٢٣,and average error value over a full epochs is equal to ٠.٠٠٤٠٣ .

Learning rate after update through training is ٠.٠٩٧٩٥ .

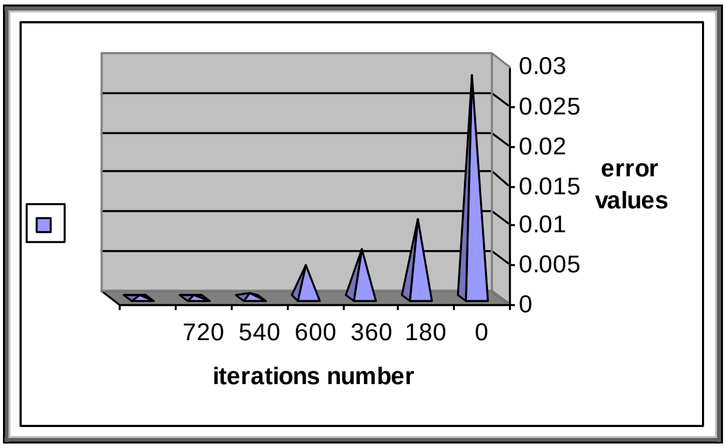


Figure (٨-٤٣):Changing error values over iterations number for a seventh category at the fourth experiment

Output values for each pattern in this category under training;

output[٧]=٠.٧٠٥٢٠
output[١]=٠.٧١٤٢٤
output[٢]=٠.٧٢٣٩٠
output[٣]=٠.٧٢٣٤٤
output[٤]=٠.٧٢٤١١
output[٥]=٠.٧٢٢٦٩
output[٦]=٠.٧٢٢٣٥

This order for patterns output values is appear here as a patterns ordered in this category after adding the seventh pattern. Also note the outputs values changed for patterns than a first experiment ,once again changing a standard output value .

Standard output value is;

output[١]=٠.٧١٤٢٤

After reordering output values ,under a depended order in this work ,which the increasing order will be;

output[٧]=٠.٧٠٥٢٠
output[١]=٠.٧١٤٢٤
output[٦]=٠.٧٢٢٣٥
output[٥]=٠.٧٢٢٦٩
output[٣]=٠.٧٢٣٤٤
output[٢]=٠.٧٢٣٩٠
output[٤]=٠.٧٢٤١١

Differences values for patterns outputs values around a standard output value through depending order be;

$$\frac{٠.٠٠٩٠٤}{٠.٠٠٨١١}$$

٠.٠٠٨٤٥
٠.٠٠٩٢٠
٠.٠٠٩٦٦
٠.٠٠٩٨٧

The real closed interval for a seventh category ,at a fourth experiment ,that it determined by a patterns outputs values for this category ,and after restrict the limits of a values ,will be;

[٠.٧٠٥٢٠, ٠.٧٢٤١١]

which is a decision region for pattern(s) recognition for a seventh category at this experiment ,through it a network can (some what) recognize a pattern(s) if it in a patterns of this category or not .

The range of this closed interval for the seventh category is ٠.٠١٨٩١ that is different from a previous values .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٦ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Eighth Category

Results of training on this category were illustrate in figure (٥-٤٤) ,which clarify the changing in error value over an iterations number within a proposed limited number for iterations .

A mean square error value is equal to ٠.٠٠٠٦٣, that is gotten on after ١٩٧٤ iterations through ٢٨٢ epochs .Worst error value over a full epochs is equal to ٠.٠٣٥٥٧ ,and average error value over a full epochs is equal to ٠.٠٠٥٠٨ .

Learning rate after update through training is ٠.٠٩٠٥٠ .

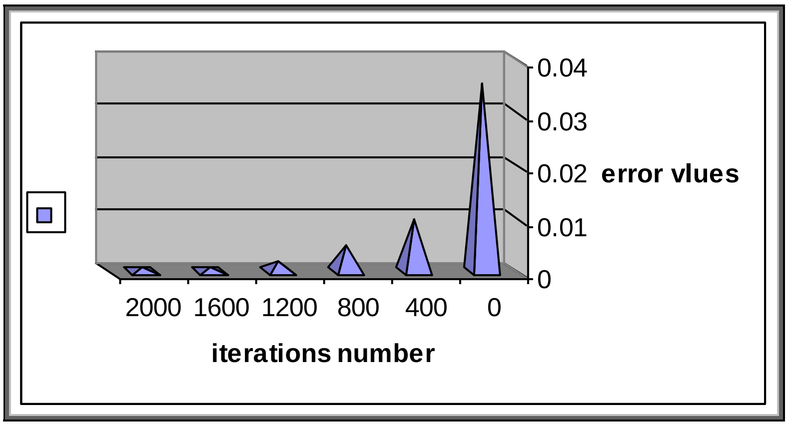


Figure (٥-٤٤):Changing error values over iterations number for the eighth category at the fourth experiment

Output values for each pattern in this category under training are;

output[٧]=٠.٩٤١٠٤
output[١]=٠.٩٦٧١١
output[٢]=٠.٩٦٨٧٩
output[٣]=٠.٩٦٧٨٠
output[٤]=٠.٩٦٧٤٠
output[٥]=٠.٩٦٣٥٨
output[٦]=٠.٩٦٤٥٦

This order for patterns output values is appear here as a patterns ordered in this category after adding the seventh pattern

Standard output value is;

output[١]=٠.٩٦٧١١

After reordering output values ,under a depended order in this;

output[٧]=٠.٩٤١٠٤
output[٥]=٠.٩٦٣٥٨
output[٦]=٠.٩٦٤٥٦
output[١]=٠.٩٦٧١١
output[٤]=٠.٩٦٧٤٠
output[٣]=٠.٩٦٧٨٠
output[٢]=٠.٩٦٨٧٩

Differences values for patterns outputs values around a standard output value through depended order be;

٠.٠٢٦٠٧
٠.٠٠٣٥٣
٠.٠٠٢٥٥
٠.٠٠٠٢٩
٠.٠٠٠٦٩
٠.٠٠١٦٨

The real closed interval for a sixth category ,at a fourth experiment ,that it determined by a patterns outputs values for this category ,and after restrict the limits of a values ,will be;

[٠.٩٤١٠٤, ٠.٩٦٨٧٩]

which is a decision region for pattern(s) recognition for a eighth category at this experiment ,through it a network can recognize a pattern(s) , if it in a patterns of this category or not .

The range of this closed interval for the eighth category is ٠.٠٢٧٧٥ that is traditionally different from a previous values .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٤ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Discussion of the Fourth Experiment :

The intervals of a training set categories at a fourth experiment are ;

First category:	[٠.٠١٢٤٣ ,٠.٠٤٤٩٥]
Second category:	[٠.١٨٣٠٦ ,٠.٢٠٩٩٢]
Third category:	[٠.٢٦٨٧٢ ,٠.٢٧٩٢٩]
Fourth category :	[٠.٣٠٧٤٠ ,٠.٣٢٠٧٩]
Fifth category:	[٠.٤٦١٢٦ ,٠.٤٨٣٥٣]
Sixth category:	[٠.٦٠٨٧٦ ,٠.٦٢٨٧٦]
Seventh category:	[٠.٧٠٥٢٠ ,٠.٧٢٤١١]
Eighth category:	[٠.٩٤١٠٤ ,٠.٩٦٨٧٩]

By compute ranges of closed intervals for the eights categories as difference between maximum boundary and minimum boundary for each closed interval ordinary as follows ;

٠.٠٣٢٥٢
٠.٠٢٦٨٦
٠.٠١٠٥٧
٠.٠١٣٣٩
٠.٠٢٢٢٧
٠.٠٢٠٠٠
٠.٠١٨٩١

٠.٠٢٧٧٥

Not that a least range is for third category interval (which is also least range at third experiment) at it equal to ٠.٠١٠٥٧ .While largest range is at first category interval (which is largest one in first experiment) that has big capacity for outputs values(actual outputs) for patterns in this category equal to ٠.٠٣٢٥٢ .

Note :

١.Outputs values for patterns(actual outputs) do not belong to two(or more) categories interval in the same time at a fourth experiment also .

٢.The outputs values at categories intervals(for actual output values of patterns)are ordered increasingly,this may because choice of desired outputs values with increasing order ,then it appear as sequential intervals in numbers line(if it is ordered), this sequence began with first category interval and then the second interval and the third and the fourth ,and so on until end of all eights categories intervals .

٣.In addition to regions between these closed intervals of the categories ,there are regions out over the maximum value for all intervals(a maximum boundary of the eighth category interval in this experiment) ,and a less than a minimum value for all intervals categories(minimum boundary of a first category interval in this experiment).

It are also an error regions but it be larger ,since it expended to infinity with the two directions ,a minus and plus infinity ,to $-\infty$ *and* ∞ *,respectively .*

Differences regions between these intervals which a parts of error regions determined between each two successive intervals .

Ranges that lies between these intervals are error regions computed at each two successive intervals as difference between maximum boundary for first one with minimum boundary of second one ordinary as follows :

٠.١٣٨١١

٠.٠٥٨٨٠

٠.٠٢٨١١

٠.١٤٠٤٧

٠.١٢٥٢٣

٠.٠٧٦٤٤

٠.٢١٦٩٣

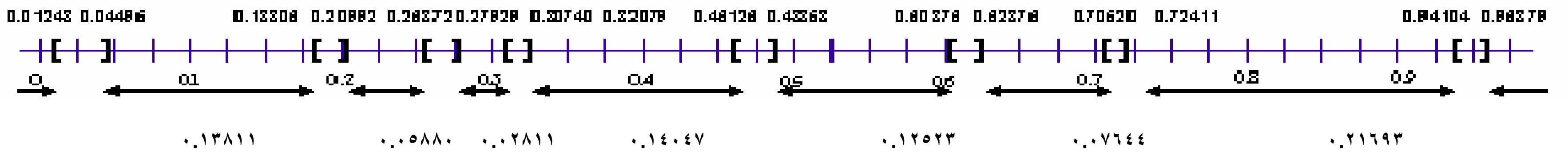
A least difference range occurred between third and fourth category this also reflect that error region between these two categories intervals is small equal to ٠.٠٢٨١١, while the largest difference is between seventh and eighth category and it equal to ٠.٢١٦٩٣ ,which also a largest range for error regions in third experiment .

The resultant of testing through this experiment, was at recognize ١٠٠ *patterns* for all categories from the number of patterns ١٦٠ *patterns* that were used .

Note the improving in performance at this experiment through increasing number of patterens that were recognized .

Categories	Epoch	It .No.	MSR	Average err.	Worst err.	Eta. updated	Min. boundary	Max. boundary	Range
١st category	٦٢٩	٤٤٠٣	٠.٠٠٠٠٦٨	٠.٠٠٠٥٢٨	٠.٠٠٣٦٩٥	٠.٠٠٦٠٩٢	٠.٠٠١٢٤٣	٠.٠٠٤٤٩٥	٠.٠٠٣٢٥٢
٢^{n} category	١٠٠	٧٠٠	٠.٠٠٠٠٥٥	٠.٠٠٠٤٧٦	٠.٠٠٣٣٣٤	٠.٠٠٩٨٧٤	٠.١٨٣٠٦	٠.٢٠٩٩٢	٠.٠٠٢٦٨٦
٣rd category	١٣	٩١	٠.٠٠٠٠٣٦	٠.٠٠٠٣٨٥	٠.٠٠٢٦٩٨	٠.٠٠٩٩٩٧	٠.٢٦٨٧٢	٠.٢٧٩٢٩	٠.٠٠١٠٥٧
٤th category	٤٩	٣٤٣	٠.٠٠٠٠٤٦	٠.٠٠٠٤٣٤	٠.٠٠٣٠٤٣	٠.٠٠٩٩٦٩	٠.٣٠٧٤٠	٠.٣٢٠٧٩	٠.٠٠١٣٣٩
٥th category	١١٧	٨١٩	٠.٠٠٠٠٣٩	٠.٠٠٠٤٠٢	٠.٠٠٢٨١٨	٠.٠٠٩٨٢٨	٠.٤٦١٢٦	٠.٤٨٣٥٣	٠.٠٠٢٢٢٧
٦th category	١١٧	٨١٩	٠.٠٠٠٠٣٧	٠.٠٠٠٣٩١	٠.٠٠٢٧٣٧	٠.٠٠٩٨٢٨	٠.٦٠٨٧٦	٠.٦٢٨٧٦	٠.٠٠٢٠٠٠
٧th category	١٢٨	٨٩٦	٠.٠٠٠٠٣٩	٠.٠٠٠٤٠٣	٠.٠٠٢٨٢٣	٠.٠٠٩٧٩٥	٠.٧٠٥٢٠	٠.٧٢٤١١	٠.٠٠١٨٩١
٨th category	٢٨٢	١٩٧٤	٠.٠٠٠٠٦٣	٠.٠٠٠٥٠٨	٠.٠٠٣٥٥٧	٠.٠٠٩٠٥٠	٠.٩٤١٠٤	٠.٩٦٨٧٩	٠.٠٠٢٧٧٥

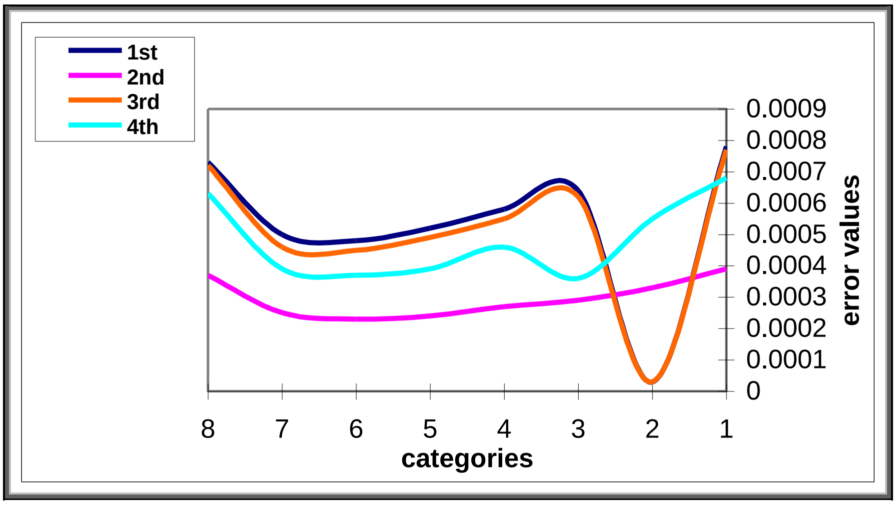
Table(5-7): Conclusions values for the eight categories from the fourth experiment



Figure(5-45): Decision Line for the fourth experiment

Categories	*First experiment*	*Second experiment*	*Third experiment*	*Fourth experiment*
١st	٠٫٠٢٤٧١	٠٫٠٣١٥٧	٠٫٠٢٤٧٠	٠٫٠٣٢٥٢
٢nd	٠٫٠١٨٧١	٠٫٠٢٦٤٥	٠٫٠١٨٧١	٠٫٠٢٦٨٦
٣rd	٠٫٠٠١٩٨	٠٫٠٠٢٧٣	٠٫٠٠٧٣٠	٠٫٠١٠٥٧
٤th	٠٫٠٠٣٩٣	٠٫٠٠٤٦٢	٠٫٠١٠٦٥	٠٫٠١٣٣٩
٥th	٠٫٠١٢٥٥	٠٫٠١٥٥٥	٠٫٠١٨٠٨	٠٫٠٢٢٢٧
٦th	٠٫٠٠٧٤٠	٠٫٠٠٦٧٠	٠٫٠١٨٠١	٠٫٠٢٠٠٠
٧th	٠٫٠٠٩٤٣	٠٫٠٠٩٩٠	٠٫٠١٦٠٤	٠٫٠١٨٩١
٨th	٠٫٠٠٤٨٠	٠٫٠٠٥٣٠	٠٫٠٢٣٧٤	٠٫٠٢٧٧٥

Table(5-8): Conclusions ranges values for the eight categories at the first fourth experiment



Figure(5-46) Error distribution curve at the first fourth experiments

٥.٧ Fifth Experiment :

After sawing performance of a suggested method on images that were occurred in training set drawn with only white and black line (drawn in black line over white background)which some what simple .It be a main challenge for suggested method to be performed on colored images for the same patterns in training set .

So the initialized idea of this experiment is ; perform a suggested method on a same patterns in previous experiments but by colored lines on white background for each image ,different colors were used for different images in same category .

Images data which are a features that extracted in same way for images in previous experiments ,also these data are normalized ,and feed to network .The parameters that proposed for work must be known, and desired output values were also proposed same for that one at previous experiments.

In this experiment parameters values were proposed in way to match the change in images data ,so value of learning rate is supposed different than previous .Also an available error value will change or in more details it increased to assimilate the changes in data values .Do not forget that the structure of a network is the same for one in first experiment(original structure) for all categories (from number of layers ,and number of neurons in these layers).

The aim from this experiment in chiefly ,to prove that the work efficiency is independent on colors ,that same figure or its versions , which is good property can added to work of a proposed network for suggested method .Parameters values that we talked on are illustrate in table(٥-١٠)

Learning rate	٠.١
Momentum	٠.٠٠١
Max. iterations	٤٠٠٠٠٠
Translation	٦
Dilation	٢
Available error	٠.٠٠١
Hidden neurons	٥
Number of patterns in category	٦

Table (5-9):Proposed training parameters values at fifth experiment

Note : Patterns order is the same one used in first experiment.

First Category

This experiment is done on a first category through take a patterns in a first category in training set and each figure taken with color different on all other figures in the category ,for example the first figure drawn by red line ,a second with yellow ,and third with green ,…,so on .Parameters values for training on this category are same for other categories as in table(٥-١٠) .

The Results:

Result of training a network on a first category after the discussed changes in colors is illustrated in figure(٥-٤٧) for changing error over an iterations number during the training .

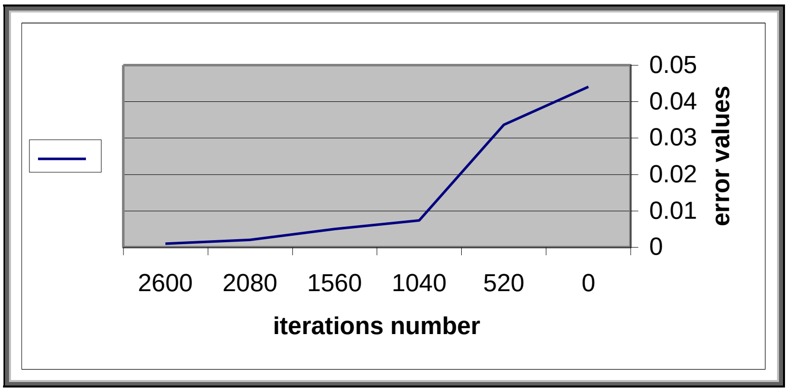


Figure (٥-٤٧):Changing error values over iterations number for the first category at the fifth experiment

Mean square error value is ٠.٠٠٠٩٦ which is largest value gotten for error through training than all other categories ,the training is completed after ٢٦١٦ iterations ,through ٤٣٦ epochs these numbers are a maximum through training on all categories .

A worst error value over this number of epochs is ٠.٠٤٤٠٢ ,and the average error value is equal to ٠.٠٠٧٣٣ which is a largest average for errors for training on all other categories at this experiment, a learning rate is updated through training than a proposed one to be equal to ٠.٠٧٨٨٠ , this is a least updated value for learning rate through training .

Output results of patterns in this category at this experiment ;

output[١]=٠.٠٤٤٩٥
output[٢]=٠.٠٣٠١٨
output[٣]=٠.٠٤٩٥٧
output[٤]=٠.٠٤٤٠٣
output[٥]=٠.٠٤٧٧٣
output[٦]=٠.٠٤٣٩٦

Standard output value is ;

output[١]=٠.٠٤٤٩٥

After reordering output values ,under the depended order;

output[٢]=٠.٠٣٠١٨
output[٦]=٠.٠٤٣٩٦
output[٤]=٠.٠٤٤٠٣
output[١]=٠.٠٤٤٩٥
output[٥]=٠.٠٤٧٧٢
output[٣]=٠.٠٤٩٥٧

The differences values for patterns outputs values around a standard output value through this order be;

٠.٠١٤٧٧
٠.٠٠٠٩٩
٠.٠٠٠٩٢
٠.٠٠٢٧٧
٠.٠٠٤٦٢

Depending on reordering a patterns output values for this category with the general used manner for ordering in this work as increasingly, to determine a closed interval for this category at this experiment ;

[٠.٠٣٠١٨, ٠.٠٤٩٥٧]

The range of this interval for output values that converge graphically to a figures in this experiment is equal to ٠.٠١٩٣٩ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٧ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Second Category

The experiment done on a second category through take a patterns in a second category in training set and each figure taken with color different on all other figures in the category with similar way for a first category but somewhat different in colors .

The Results:

Result of training a network on a second category after the discussed changes in colors is illustrate in figure(٥-٤٨) for changing error over an iterations number during the training .

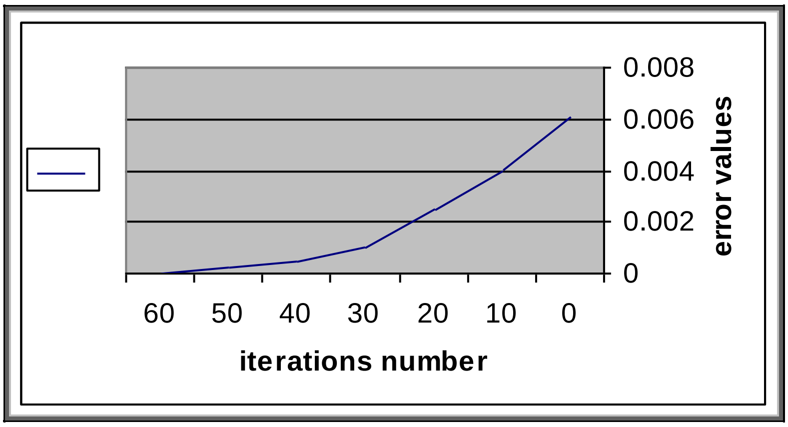


Figure (٥-٤٨):Changing error values over iterations number for a second category at the fifth experiment

Mean square error value is ٠.٠٠٠٠١ which is a smallest value gotten for error through training on all the other categories ,the training is completed after ٦٦ iterations through ١١ epochs , these numbers are less than all other categories .

Worst error value is equal to ٠.٠٠٦٠٩ ,and average error value is ٠.٠٠١٠١ , which are a least worst and average for errors at training than all rest categories at this experiment, this because a small value for error over small number of epochs .

Learning rate is updated through training to be at value ٠.٠٩٩٩٨, which is a largest value for learning rate through training, that means the chosen value is suitable for this category .

Output results of patterns in this category at this experiment are;

output[١]=٠.١٤٥١٨
output[٢]=٠.١٤٤٠٣
output[٣]=٠.١٠٦٤٣
output[٤]=٠.١٥٤٤٩
output[٥]=٠.١٤٤٤٠
output[٦]=٠.١٤٣٩٤

Standard output value is ;
output[١]=٠.١٤٥١٨

After reordering output values ,under the depended order;

output[٣]=٠.١٠٦٤٣
output[٦]=٠.١٤٣٩٤
output[٢]=٠.١٤٤٠٣
output[٥]=٠.١٤٤٤٠
output[١]=٠.١٤٥١٨

output[٤]=٠.١٥٤٤٩

The differences values for patterns outputs values around a standard output value through this order be;

٠.٠٣٨٧٥
٠.٠٠١٢٤
٠.٠٠١١٥
٠.٠٠٠٧٨
٠.٠٠٩٣١

Depending on reordering a patterns output values for this category increasingly to determine a closed interval for this category at this experiment;

[٠.١٠٦٤٣, ٠.١٥٤٤٩]

The range of this interval for output values converge graphically to a figures in this experiment is equal to ٠.٠٤٨٠٦ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٥ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Third Category :

The experiment done on a third category through take a patterns in a third category in training set and each figure taken with color different on all other figures in the category with similar way for a first category but somewhat different in colors .

The Results:

Result of training a network on third category after the discussed changes in colors is illustrated in figure(٥-٤٩) for changing error over iterations number during the training .

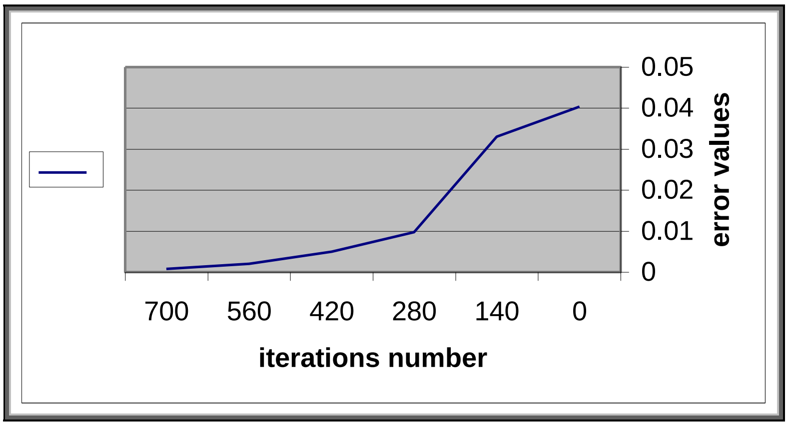


Figure (٥-٤٩):Changing error values over iterations number for the third category at the fifth experiment

Mean square error value is ٠.٠٠٨١٤ that through training , the training is completed after ٦٩٠ iterations ,through ١١٥epochs .

A worst error value over all epochs is ٠.٠٤٠٣٥,and the average error value over a full epochs is ٠.٠٠٦٧٢.Learning rate is updated through training than a proposed one to be equal to ٠.٠٩٨٣٤ .

Output results patterns in this category in this experiment are ;

output[١]=٠.٢١٧١٤
output[٢]=٠.٢١٥١٠
output[٣]=٠.٢١٦٦٥
output[٤]=٠.٢٠٤٤٣
output[٥]=٠.٢٠٦٦٦
output[٦]=٠.٢١٠٠٠

Standard output value is ;
output[١]=٠.٢١٧١٤

After reordering output values ,under the depended order;

output[٤]=٠.٢٠٤٤٣
output[٥]=٠.٢٠٦٦٦
output[٦]=٠.٢١٠٠٠
output[٢]=٠.٢١٥١٠
output[٣]=٠.٢١٦٦٥
output[١]=٠.٢١٧١٤

Differences values for patterns outputs values around a standard output value through this order be;

٠.٠١٢٧١
٠.٠١٠٤٨
٠.٠٠٧١٤
٠.٠٠٢٠٤
٠.٠٠٠٤٩

Depending on reordering a patterns output values for this category with the general used manner for ordering in this work as increasingly to determine a closed interval for this category at this experiment :

[٠.٢٠٤٤٣, ٠.٢١٧١٤]

The range of this interval for output values converge graphically to a figures in this experiment is equal to ٠.٠١٢٧١ .

Note :*The standard output is a largest value than all other colored* *versions ,it as maximum boundary for a closed interval for this category .*

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٣ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Fourth Category :

The experiment done on a fourth category through take a patterns in a fourth category in training set and each figure taken with color different on all other figures in the category with similar way for a previous categories but somewhat different in colors .

The Results:

Result of training a network on fourth category after the discussed changes in colors is illustrated in figure(٥-٥٠) for changing error over an iterations number during the training .

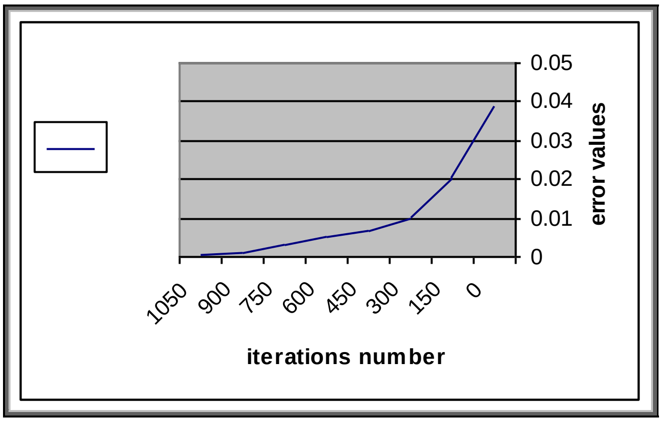


Figure (٥-٥٠):Changing error values over iterations number for a fourth category at the fifth experiment

The mean square error value is ٠.٠٠٠٧٥ that through training, the training is completed after ١٠٢٦ iterations ,through ١٧١ epochs .

A worst error value over all epochs is ٠.٠٣٨٩٠ ,and the average error value over a full epochs is ٠.٠٠٦٤٨ .Learning rate is updated through training than a proposed one to be equal to ٠.٠٩٦٣٩ .

Output results patterns in this category in this experiment are as;

output[١]=٠.٣٣١٧٤

output[٢]=٠.٣٣٣٥٨

output[٣]=٠.٣٣٣٦٨

output[٤]=٠.٣٣١٠٩
output[٥]=٠.٣٢٣٦٦
output[٦]=٠.٣١١٥٧

Standard output value is ;
output[١]=٠.٣٣١٧٤

After reordering output values ,under the depended order;
output[٦]=٠.٣١١٥٧
output[٥]=٠.٣٢٣٦٦
output[٤]=٠.٣٣١٠٩
output[١]=٠.٣٣١٧٤
output[٢]=٠.٣٣٣٥٨
output[٣]=٠.٣٣٣٦٨

Differences values for patterns outputs values around a standard output value through this order be;

٠.٠٢٠١٧
٠.٠٠٨٠٨
٠.٠٠٠٦٥
٠.٠٠١٨٤
٠.٠٠١٩٤

Depending on reordering a patterns output values with the general used manner for ordering in this work as increasingly, to determine a closed interval for this category at this experiment;

[٠.٣١١٥٧, ٠.٣٣٣٦٨]

The range of this interval for output values converge graphically to a figures in this experiment is equal to ٠.٠٢٢١١ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٥ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Fifth Category :

This experiment was done on a fifth category through take a patterns in a fifth category in training set and each figure taken with color different on all other figures in the category with similar way for a previous categories but somewhat different in colors .

The Results:

Result of training a network on fifth category after the discussed changes in colors is illustrated in figure(٥-٥١) for changing error over an iterations number during the training .

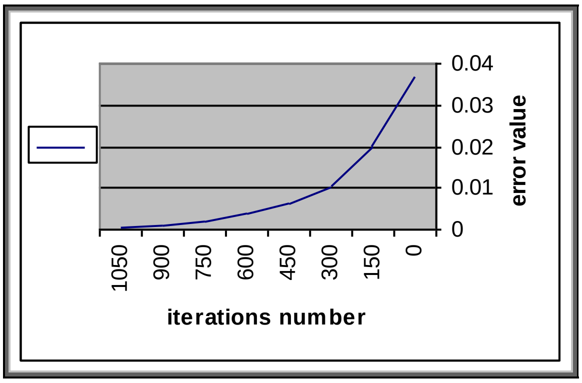


Figure (٥-٥١):Changing error values over iterations number for a fifth category at the fifth experiment

The mean square error value is ٠.٠٠٠٦٦ that through training , the training is completed after ١٠١٤ iterations ,through ١٦٩ epochs .

A worst error value over all epochs is ٠.٠٣٦٤٠ ,and the average error value over a full epochs is ٠.٠٠٦٠٦ .Learning rate is updated through training than a proposed one to be equal to ٠.٠٩٦٤٧ .

Output results patterns in this category in this experiment are ;

output[١]=٠.٤٦٦٥٢
output[٢]=٠.٤٧٧٦٧
output[٣]=٠.٤٧٧٧١
output[٤]=٠.٤٨١٧٨
output[٥]=٠.٤٨٨٠١
output[٦]=٠.٤٦٤٢٣

Standard output value is ;

output[١]=٠.٤٦٦٥٢

After reordering output values ,under the depended order;

output[٦]=٠.٤٦٤٢٣
output[١]=٠.٤٦٦٥٢
output[٢]=٠.٤٧٧٦٧
output[٣]=٠.٤٧٧٧١
output[٤]=٠.٤٨١٧٨
output[٥]=٠.٤٨٨٠١

Differences values for patterns outputs values around a standard output value through this order be;

٠.٠٠٢٣٠

٠.٠١١١٥
٠.٠١١١٩
٠.٠١٥٢٦
٠.٠٢١٤٩

Depending on reordering a patterns output values increasingly, to determine a closed interval for this category at this experiment
[٠.٤٦٤٢٢, ٠.٤٨٨٠١]

The range of this interval for output values converge graphically to a figures in this experiment is equal to ٠.٠٢٣٧٩ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٩ *patterns* from a testing set with ٢٠ *patterns*, for this category that within a determined training interval.

Sixth Category :

This experiment done on the sixth category through take a patterns in a sixth category in training set and each figure taken with color different on all other figures in category with similar way for a previous categories but somewhat different in colors .

The Results:

Result of training a network on sixth category after the discussed changes in colors is illustrated in figure(٥-٥٢) for changing error over an iterations number during the training .

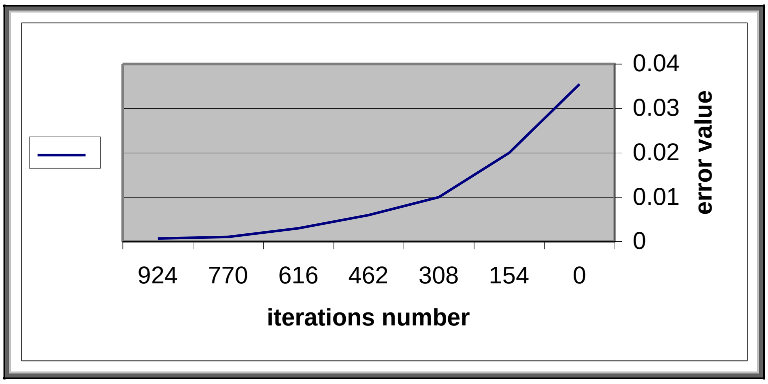


Figure (٥-٥٢):Changing error values over iterations number for a sixth category at the fifth experiment

Mean square error value is ٠.٠٠٠٦٢ that through training ,the training is completed after ٩٢٤ iterations ,through ١٥٤ epochs .

A worst error value over all epochs is ٠.٠٣٥٤٤ ,and average error value over a full epochs is ٠.٠٠٥٩٠ .Learning rate is updated through training than a proposed one to be equal to ٠.٠٩٧٠٦.

Output results patterns in this category in this experiment are ;

output[١]=٠.٦٠٩٩٧
output[٢]=٠.٦٣٠٥٦
output[٣]=٠.٦١٧٥٥
output[٤]=٠.٦١٢٨٤
output[٥]=٠.٦٢٠٧١
output[٦]=٠.٦١٥٢٦

Standard output value is ;

output[١]=٠.٦٠٩٩٧

After reordering output values ,under the depended order;

output[١]=٠.٦٠٩٩٧
output[٤]=٠.٦١٢٨٤
output[٦]=٠.٦١٥٢٦
output[٣]=٠.٦١٧٥٥
output[٥]=٠.٦٢٠٧١
output[٢]=٠.٦٣٠٥٦

Differences values for patterns outputs values around a standard output value through this order be;

٠.٠٠٢٨٧
٠.٠٠٥٢٩
٠.٠٠٧٥٨
٠.٠١٠٧٤
٠.٠٢٠٥٩

Depending on reordering of a patterns output values with the general used manner for ordering in this work as increasingly, to determine a closed interval for this category at this experiment;

[٠.٦٠٩٩٧ , ٠.٦٣٠٥٦]

The range of this interval for output values converge graphically to a figures in this experiment is equal to ٠.٠٢٠٥٩ .

Note : A standard output is a smallest value than all other colored versions ,it as minimum boundary for a closed interval for this category ,this mean ability of a network in recognize pattern(s) its output(s)larger than a standard output .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٢ *patterns* from a testing set of ٢٠ *patterns* ,for this category within a determined training interval

Seventh Category :

The experiment done on a seventh category through take a patterns in a seventh category in training set and each figure taken with color different on all other figures in the category with similar way for a previous categories but somewhat different in colors .

The Results:

Result of training a network on seventh category after the discussed changes in colors is illustrated in figure(٥-٥٣) for changing error over an iterations number during the training .

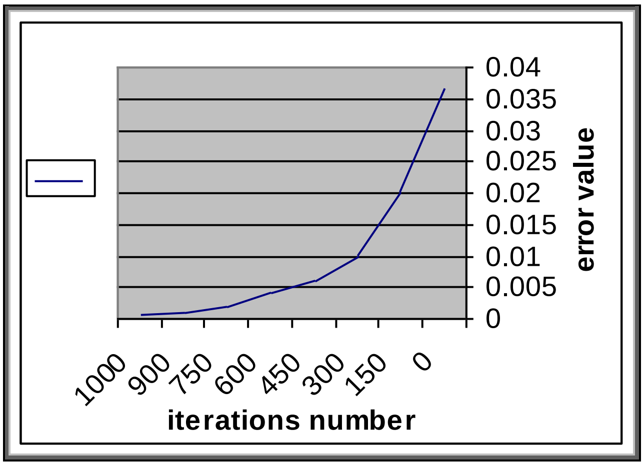


Figure (٥-٥٣):Changing error values over iterations number for a seventh category at the fifth experiment

Mean square error value is ٠.٠٠٦٦١ that through training ,the training is completed after ٩٩٠ iterations ,through ١٦٥ epochs .
A worst error value over all epochs is ٠.٠٣٦٣٦ ,and average error value over a full epochs is ٠.٠٠٦٠٦ .Learning rate is updated through training than a proposed one to be equal to ٠.٠٩٦٦٣ .

Output results patterns in this category in this experiment are ;

output[١]=٠.٧٣٩٤٠
output[٢]=٠.٧٤٩٩٨
output[٣]=٠.٧٥١٤١
output[٤]=٠.٧٣٦٣٠
output[٥]=٠.٧٤٥١٣

output[٦]=٠.٧١٤٢٩

Standard output value is ;
output[١]= ٠.٧٣٩٤٠

After reordering output values ,under the depended order;

output[٦]=٠.٧١٤٢٩
output[٤]=٠.٧٣٦٣٠
output[١]=٠.٧٣٩٤٠
output[٥]=٠.٧٤٥١٣
output[٢]=٠.٧٤٩٩٨
output[٣]=٠.٧٥١٤١

Differences values for patterns outputs values around a standard output value through this order be;

٠.٠٢٥١١
٠.٠٠٣١٠
٠.٠٠٥٧٣
٠.٠١٠٥٨
٠.٠١٢٠١

Depending on reordering a patterns output values increasingly, to determine a closed interval for this category at this experiment;

[٠.٧١٤٢٩, ٠.٧٥١٤١]

The range of this interval for output values converge graphically to a figures in this experiment is equal to ٠.٠٣٧١٢ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١١ *patterns* from a testing set of ٢٠ *patterns* ,for this category within a determined training interval.

Eighth Category :

The experiment done on eighth category through take a patterns that in training set and each figure taken with color different on all other figures in the category with similar way for a previous categories but somewhat different in colors .

The Results:

Result of training a network on the eighth category after discussed changes in colors is illustrated in figure(٥-٥٤) for changing error over an iterations number during the training .

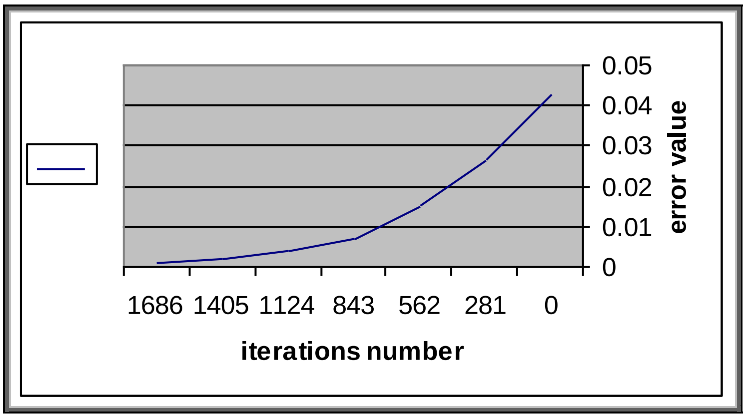


Figure (٥-٥٤):Changing error values over iterations number for the eighth category at the fifth experiment

Mean square error value is ٠.٠٠٠٩٠ that through training ,the training is completed after ١٦٨٦ iterations ,through ٢٨١ epochs .
A worst error value over all epochs is ٠.٠٤٢٥٢ ,and average error value over a full epochs is ٠.٠٠٧٠٨ .Learning rate is updated through training than a proposed one to be equal to ٠.٠٩٠٥٦ .
Output results patterns in this category in this experiment are as;

output[١]=٠.٩٥٩٠٤
output[٢]=٠.٩٦٠٤٩
output[٣]=٠.٩٦٠٣٩
output[٤]=٠.٩٦٠٣٤
output[٥]=٠.٩٤٥٢٨
output[٦]=٠.٩٥٧٦٤

Standard output value is ;
output[١]= ٠.٩٥٩٠٤

After reordering output values ,under the depended order;

output[٥]=٠.٩٤٥٢٨
output[٦]=٠.٩٥٧٦٤
output[١]=٠.٩٥٩٠٤
output[٤]=٠.٩٦٠٣٤
output[٣]=٠.٩٦٠٣٩
output[٢]=٠.٩٦٠٤٩

The differences values for patterns outputs values around a standard output value through this order be;

٠.٠١٣٧٦

٠.٠٠١٤٠
٠.٠٠١٢٠
٠.٠٠١٢٥
٠.٠٠١٤٥

Depending on reordering a patterns output values increasingly, to determine a closed interval for this category at this experiment

[٠.٩٤٥٢٨, ٠.٩٦٠٤٩]

The range of this interval for output values converge graphically to a figures in this experiment is equal to ٠.٠١٥٢١ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٦ *patterns* from a testing set of ٢٠ *patterns* ,for this category within a determined training interval.

Discussion of the Fifth Experiment :

The intervals of a training set categories at a fifth experiment are;

First category :	[٠.٠٣٠١٨ , ٠.٠٤٩٥٧]
Second category :	[٠.١٠٦٤٣ , ٠.١٥٤٤٩]
Third category :	[٠.٢٠٤٤٣ , ٠.٢١٧١٤]
Fourth category :	[٠.٣١١٥٧ , ٠.٣٣٣٦٨]
Fifth category :	[٠.٤٦٤٢٢ , ٠.٤٨٨٠١]
Sixth category :	[٠.٦٠٩٩٧ , ٠.٦٣٠٥٦]
Seventh category :	[٠.٧١٤٢٩ , ٠.٧٥١٤١]
Eighth category :	[٠.٩٤٥٢٨ , ٠.٩٦٠٤٩]

By compute ranges for a closed intervals of the eights categories as a difference between maximum boundary and minimum boundary for each closed interval ordinary ,as follows ;

٠.٠١٩٣٩
٠.٠٤٨٠٦
٠.٠١٢٧١
٠.٠٢٢١١

٠.٠٢٣٧٩
٠.٠٢٠٥٩
٠.٠٣٧١٢
٠.٠١٥٢١

Not that a least range is for third category interval (which is also a least range for fourth experiment) at it equal to ٠.٠١٢٧١ , while largest range is at second category interval that has biggest capacity for outputs values(actual outputs)for patterns in this category that equal to ٠.٠٤٨٠٦ .

Note :

١.The outputs values for patterns(actual outputs)do not belong to two(or more) categories interval in the same time at this experiment also .

٢. The outputs values at categories intervals(for actual output values of patterns)are ordered increasingly ,this may because choice of desired outputs values with increasing order ,then it appear as sequential intervals in numbers line (if it is ordered) , this sequence began with first category interval and then the second interval and the third and the fourth ,and so on until end of all the eights categories intervals.

٣. In addition to regions between these closed intervals of categories ,there are regions out over the maximum value for all intervals(a maximum boundary of the eighth category interval in this experiment) ,and a less than a minimum value for all intervals categories(minimum boundary of a first category interval in this experiment).

It are also an error regions but it be larger ,since it expended to infinity with the two directions ,a minus and plus infinity ,to $-\infty$ *and* ∞ *,respectively .*

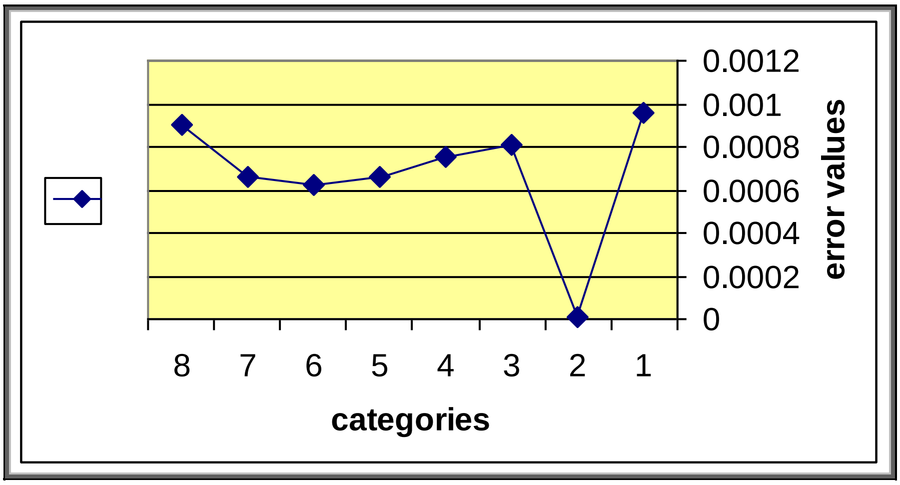
The differences regions between these intervals which is part of error regions determined between each two successive intervals ranges that lies between these intervals which are the error regions computed for each two successive intervals one ordinary , as follows :

٠.٠٥٦٨٦
٠.٠٤٩٩٤
٠.٠٩٤٤٣

٠.١٣٠٥٤
٠.١٢١٩٦
٠.٠٨٣٧٣
٠.١٩٣٨٧

Note that largest error region range is occurred at difference range between seventh and eighth category intervals and it equal to ٠.١٩٣٨٧,it is also a largest range for the error regions in fourth experiment ,while least range is occurred between second and third category interval and it is equal to ٠.٠٤٩٩٤ .

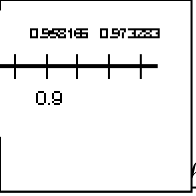
The resultant of testing through this experiment ,was at recognize ١٠٨ *patterns* for all categories from the number of patterns ١٦٠ patterns that were used .

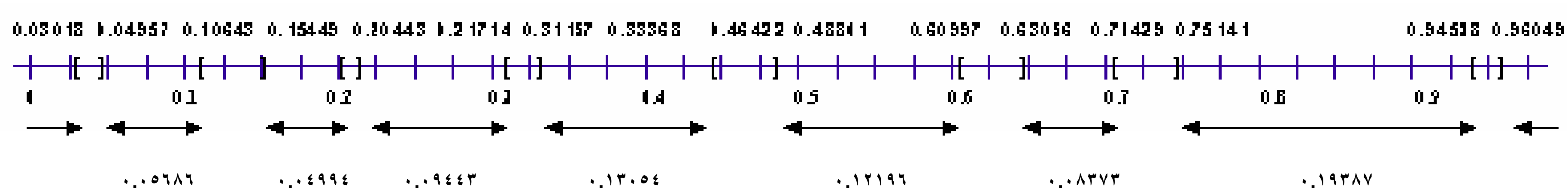


Figure(5-55): Error values curve for the eights categories at the fifth experiment

Categories	Epochs	It .No.	MSR	Average err.	Worst err.	Eta. updated	Min. boundary	Max. boundary	Range
١st category	٤٣٦	٢٦١٦	٠.٠٠٠٩٦	٠.٠٠٧٣٣	٠.٠٤٤٠٢	٠.٠٧٨٨٠	٠.٠٣٠١٨	٠.٠٤٩٥٧	٠.٠١٩٣٩
٢n category	١١	٦٦	٠.٠٠٠٠١	٠.٠٠١٠١	٠.٠٠٦٠٩	٠.٠٩٩٩٨	٠.١٠٦٤٣	٠.١٥٤٤٩	٠.٠٤٨٠٦
٣rd category	١١٥	٦٩٠	٠.٠٠٠٨١	٠.٠٠٦٧٢	٠.٠٤٠٣٥	٠.٠٩٨٣٤	٠.٢٠٤٤٣	٠.٢١٧١٤	٠.٠١٢٧١
٤th category	١٧١	١٠٢٦	٠.٠٠٠٧٥	٠.٠٠٦٤٨	٠.٠٣٨٩٠	٠.٠٩٦٣٩	٠.٣١١٥٧	٠.٣٣٣٦٨	٠.٠٢٢١١
٥th category	١٦٩	١٠١٤	٠.٠٠٠٦٦	٠.٠٠٦٠٦	٠.٠٣٦٤٠	٠.٠٩٦٤٧	٠.٤٦٤٢٢	٠.٤٨٨٠١	٠.٠٢٣٧٩
٦th category	١٥٤	٩٢٤	٠.٠٠٠٦٢	٠.٠٠٥٩٠	٠.٠٣٥٤٤	٠.٠٩٧٠٦	٠.٦٠٩٩٧	٠.٦٣٠٥٦	٠.٠٢٠٥٩
٧th category	١٦٥	٩٩٠	٠.٠٠٠٦٦	٠.٠٠٦٠٦	٠.٠٣٦٣٦	٠.٠٩٦٦٣	٠.٧١٤٢٩	٠.٧٥١٤١	٠.٠٣٧١٢
٨th category	٢٨١	١٦٨٦	٠.٠٠٠٩٠	٠.٠٠٧٠٨	٠.٠٤٢٥٢	٠.٠٩٠٥٦	٠.٩٤٥٢٨	٠.٩٦٠٤٩	٠.٠١٥٢١

Table(5-10): Conclusions values for the eight categories from the fifth experiment that for colors





Figure(5-56): Decision Line for the fifth experiment

Chapter Five

The Experiments

5.1 Presentation Idea

In this side of work ,we are accomplish the suggested method which was discussed beforehand on a set of categories of patterns, each category has a standard figure(image) characterize this category ,while all other patterns in the category are a fuzzifies copies for that standard figure ,then testing ability of the network in identification and then recognize the patterns under work in this experiments ,and (if developed) on more complex problems .

The patterns that we will work on are a sets of images for geometric curves Hand drawn with its fuzzifies copies divided into eight categories ,that were discussed in training set(٤.١.٤).

To show the efficiency of the discussed suggested method ,we perform it with different ways at a set of experiments[١] with different ideas on a suggested activation function as a summation of wavelet and MF .

We are perform these experiments with activation function as summation of ***Mexican hat wavelet*** with ***Gaussian membership function*** **,**as ;

$$?(x) = (1 - x^2)\ .e^{-x^2/2} + e^{-x^2/2} \qquad \ldots \quad (٥\text{-}١)$$

where x in R .

---

[١] The outline character of these experiments in seems alike ,since through this to clarify documents of each experiment and declared in way some what identical to other experiments through same measurements criterias ,but the values really different .

We were chose these two continuous functions to be complemented each to other and for requests were discussed in chapter four .

The considered set of $\mathcal{N}_f$ features (that were extracted from each image) ,we used five features ,detectors as;

$\mathcal{N}_f$=٥ ,and ($f_١ , f_٢, f_٣, f_٤, f_٥$) ,

these values represent a vector of values for some pattern .

A set of patterns(that entered to network) for *p_patterns* where *p* =١,…, $\mathcal{N}_p$, $\mathcal{N}_p$,is number of patterns in that file .The number of patterns supposed equal to ٦ ; $\mathcal{N}_p$=٦ , that is number of members in S_j at *j_th category,* in form (٤-٣٠) . While in testing set ,for each category $\mathcal{N}_p$ is supposed equal to ٢٠ *pattern* (*image*) .

Then the file of inputs as matrix with dimensions as ٦×٥ for each category ,but as will see in some experiments use ٧ versions for the standard figure ,so it will be ٧×٥ ,and for a testing set we will use for all experiments number of patterns equal to ٢٠ .

The form (٤-٧) will be in light of previous notations

$$\begin{bmatrix} p_1 f_1 & p_1 f_2 & .... & p_1 f_5 \\ p_2 f_1 & p_2 f_2 & .... & p_2 f_5 \\ . & . & .... & . \\ . & . & .... & . \\ p_6 f_1 & p_6 f_2 & .... & p_6 f_5 \end{bmatrix} \qquad \ldots \quad (٥\text{-}٢)$$

and vector of features values for a pattern p_i in form(٤-٨) with dimensions ١×٥ ;

$$[p_i f_1 \quad p_i f_2 \quad ...... \quad p_i f_5] \qquad \ldots \quad (٥\text{-}٣)$$

Each file represent a data of a some category from the eight categories used ,where $\mathcal{N}_c$ number of categories under study , that consider $\mathcal{N}_c$=٨ , *j*=١,..,٨ .

Note:*The choice of a desired outputs values for these categories is effort on order of these categories intervals . If we chose desired values randomly ,the position of these intervals is also be randomly.*

We suppose here the desired values increasingly ,so we noted that intervals are ordered increasingly in line decision in figures that communicated with experiments .

The desired outputs values for categories were proposed with increasing order with a categories(by previous note) ,that is the desired output value for a first category is a minimum value ,and so on increasingly into last category with maximum desired output value,

$$d_{١} = ٠ \; ,\ldots, \; d_{\mathcal{N}c} = ١$$
$$, \quad d_{١} < d_{٢} < \ldots < d_{\mathcal{N}c} \qquad \ldots \quad (٥\text{-}٤)$$

that is
for $\mathcal{N}_c = ٨$;

$$d_{١} < d_{٢} < d_{٣} < d_{٤} < d_{٥} < d_{٦} < d_{٧} < d_{٨}$$

$\mathcal{N}_c$ is number of categories

So a desired output values for a training set through its eights categories ,were considered as ;

First category	$d_{١} = ٠$
Second category	$d_{٢} = ٠.١٥$
Third category	$d_{٣} = ٠.٢٥$
Fourth category	$d_{٤} = ٠.٣٥$
Fifth category	$d_{٥} = ٠.٥$
Sixth category	$d_{٦} = ٠.٦٥$
Seventh category	$d_{٧} = ٠.٧٥$
Eighth category	$d_{٨} = ١$

Which all belong to a closed interval [٠,١] as;

$d_j \in [٠,١]$ or can be written as $٠ \le d_j \le ١$

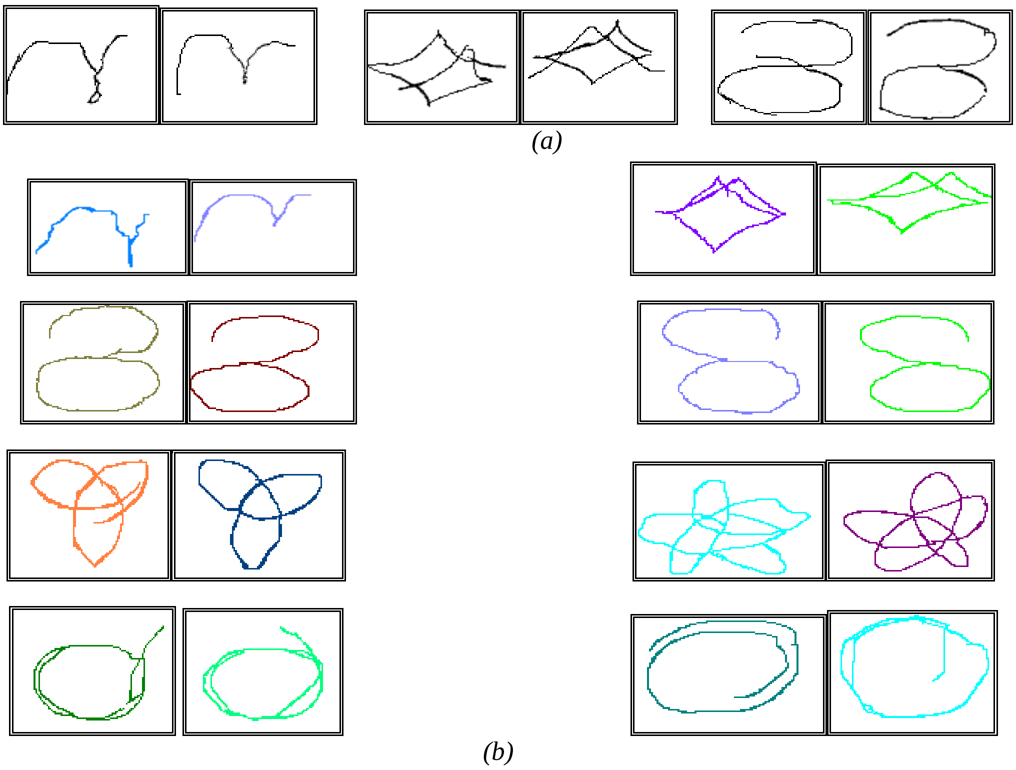
This order for values will effect on order of a closed intervals for outputs values of patterns in categories.

A dilation and translation values were supposed with a constant values for all functions in the network ,as in the form(٤-٣) ,and with integer values fixed at n=٦ ,and m=٢ for all experiments on the suggested method .

٥.٢ Testing Set

For testing the abilities of a network under conditions of a suggested method ,the network has been tested by using a set of images with number of images as two–folds number of a training set that for the experiments ; a first to the fourth experiment ,and fifth experiment . Since the second, third and fourth are similar to a first experiment in nature and type of images in the testing set, and its accomplishing and even in its idea .

So the number of testing set images is ٢٠ *images* for each category, so for eights categories will be ١٦٠ *images* ,for each experiment ,but for a two different views experiments will be ٣٢٠ *images*[٢] different from other .Even if some colored were reoccurred , it different in that drawn figure with colored ,or even recurred some figures that mean it drawn by different colors ,for different categories (no way to view all testing set images since its number is very large to occurring) .



Figure(٥- ١):Some samples form the testing set for;(a) the first experiment (b)and the colored images at fifth experiment

٥.٣ First experiment :

[٢] The number of images may be (seems for some ones) large but it good for work .

The method was performed with its essential idea by implementing the activation function in equation(٥-١) . Structure of a network was discussed previously . When do this experiment it consists of ;

٥ (input layer) *neurons*, ٥ (hidden layer)*neurons*, and with a single (output layer)*neuron*.

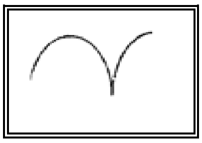
A proposed values for training parameters and other essential parameters in network for first experiment were not changed for all categories .The number of patterns in each category is ٦ *patterns* ,with available error value appropriate for training chosen through trials–error way, Momentum value is ٠.٠٠١ ,this value will not changes after training[٣] .

Learning rate	٠.١
Momentum	٠.٠٠١
Max. iterations	٤٠٠٠٠٠
Translation	٦
Dilation	٣
Available error	٠.٠٠٠٨
Hidden neurons	٥
Number of patterns in category	٦

Table (5-1):Proposed training parameters values at a first experiment

First category :

A standard figure that characterize this category is illustrated in figure(٥-٢) .The file of features values of patterns at this category be an input to a supposed network ,architecture of network -as said- is not change for this category on than other categories .



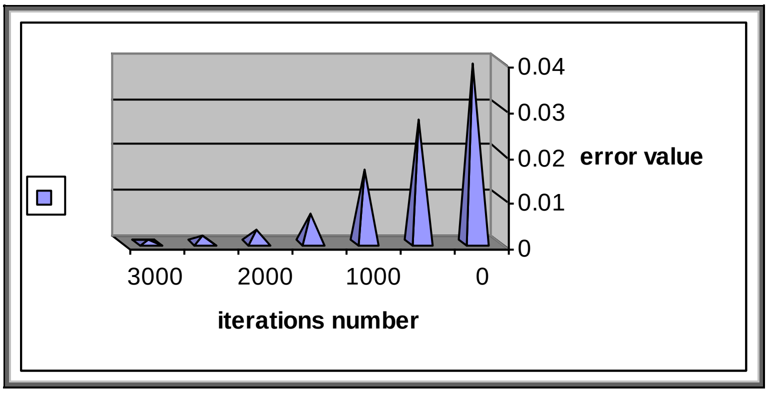
Figure(5-2) : *First Category*

The learning rate is ٠.١ ,we can choose less than this value put it make a category irregular than other categories even if we get good results ,but this is go out of a general criteria of measurement for work.

The Results :

[٣] The momentum value will be supposed at value ٠.٠٠١ for all experiments and for all categories ,(not adjusted after training),this will depend for the coming experiments .

The results of a network training on a first category is represented through quadratic distribution in figure(٥-٣),with respect to errors for each pattern in the category within limits of maximum iterations for training .



Figure(5-3): Change error value over iterations number for a first category at the first experiment

A proposed network with the proposed form was able to training on patterns in the first category with mean square error value equal to ٠.٠٠٠٧٨ this error value is a maximum value for error through training among the eights categories after ٣١٦٢ iterations through ٥٢٧ epochs, this number of iterations is a largest number of iterations in training network on training set . Average error value over these number of iterations is ٠.٠٠٦٥٨ ,while a worst error value that a network reached for training on this category was ٠.٠٣٩٥٢, also these values for error are a maximum for training at this experiment on all categories .

Learning rate value after training over all patterns in this category and epochs is updated to be equal to ٠.٠٧٠٦١ ,which is a least learning rate computed with learning on training set .

Hint : *The outcomes for performing any experiment on any category in training will use to determine a closed interval that for outputs values that we talked on .*

Hint : *The first value for these outputs, and also outputs for later categories, in all experiments represents the output value of a standard image (pattern) for each category .*

Outputs values for each version of original image under work as ;

output[١]=٠.٠٢٦٥٦
output[٢]=٠.٠٢٠٦٩
output[٣]=٠.٠٤٥٤٠
output[٤]=٠.٠٢٦٦٤
output[٥]=٠.٠٤١٩٧
output[٦]=٠.٠٣٩٤٧

It is clear ,that the range of changes between values is increase once and decrease in other ,this exactly because a variance in noise rate between images .

The standard output value of original figure is :
output[١]=٠.٠٢٦٥٦

After reordering outputs ,with increasing order then ;

output[٢]=٠.٠٢٠٦٩
output[١]=٠.٠٢٦٥٦
output[٤]=٠.٠٢٦٦٤
output[٦]=٠.٠٣٩٤٧
output[٥]=٠.٠٤١٩٧
output[٣]=٠.٠٤٥٤٠

Note : *We can reordering outputs as decreasing order, this not problem .*

An average values for outputs around the standard value which represent a different between outputs for patterns in same category through order as :

٠.٠٠٥٨٧
٠.٠٠٠٠٨
٠.٠١٢٩١
٠.٠١٥٤١
٠.٠١٨٨٤

The closed interval as;

[٠.٠٢٠٦٩, ٠.٠٤٥٤٠]

which represent an available region for patterns outputs values in this category that led to take decision for pattern(s) recognition , that if it belongs to this category or not .Then the network is able to identify any pattern(s) output(s) value(s) within this interval to be from a first category patterns .

Desired output value that was chosen for this category as a least value .

The range of a closed interval for this category at this experiment is equal to ٠.٠٢٤٧١ .

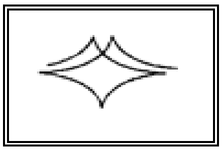
Hint :The line among the values is refer to that the difference values between outputs values is for the outputs values less than the standard output .

Note : While output value for standard pattern(image) that it characterize some category is exactly in interval for that category .In some situations a standard pattern output is represents the left(lower, minimum) boundary for interval and outputs of versions is larger than it ,or inversely .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٩ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Second category :

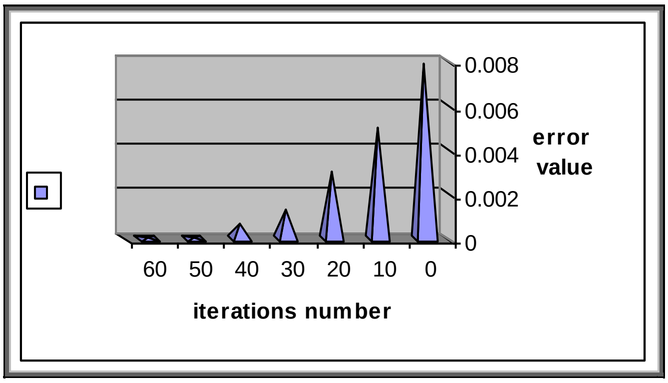
The standard figure of pattern that characterize the second category is illustrate in figure(٥-٤) .The network was trained on patterns of this category which are noisy versions from pattern in figure(٥-٤) , initial values for training parameters and other parameters are same for values used in other categories which were proposed in table(٥-١) .



Figure(5-4) :Second Category

The Results :

The results of network training on a patterns in second category are represented through quadratic distribution in figure(٥-٥) ,with respect to errors values of network for each pattern in this category within limits of maximum iterations for training .



Figure(5-5): Change error value over iterations number for second category at the first experiment

A proposed network was able to training on patterns in second category with mean square error value equal to ٠.٠٠٠٠٣ this error value is a least error value at network training on the eights categories ,after ٦٦ iterations through ١١ epochs, it is a smallest number of iterations and epochs than all other categories. Average error value over these number of iterations is ٠.٠٠١٣٠ , while worst error value that this network reached through training on this category was ٠.٠٠٧٨٥ .

Learning rate value after training over all patterns in this category and epochs is ٠.٠٩٩٩٨ ,note that value of learning rate is a largest value after training .

Outputs results for each version of original image under work ,as

output[١]=٠.١٤٦٩١
output[٢]=٠.١٤٨٥٦
output[٣]=٠.١٤٤٧١
output[٤]=٠.١٦٠٩٠
output[٥]=٠.١٥٣٧١
output[٦]=٠.١٤٢١٩

The scope of change between values is increasing once and decreasing in other such previous category for same reason .

Standard output of original figure is :
output[١]=٠.١٤٦٩١

After reordering outputs ,with increasing order then ;

output[٦]=٠.١٤٢١٩
output[٣]=٠.١٤٤٧١
output[١]=٠.١٤٦٩١
output[٢]=٠.١٤٨٥٦
output[٥]=٠.١٥٣٧١

output[٤]=٠.١٦٠٩٠

Difference between these outputs is small and difference values around a standard output value are ;

٠.٠٠٤٧٢
٠.٠٠٢٢٠
٠.٠٠١٦٥
٠.٠٠٦٨٠
٠.٠١٣٩٩

The outputs values differ from standard one ,must be within limits of this differences .

Then the real closed interval for a second category patterns ,in light of these outputs values is;

[٠.١٤٢١٩, ٠.١٦٠٩٠]

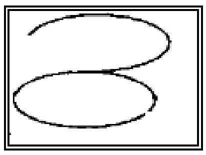
,which represents an available region for patterns output values in this category ,that led to take decision for a pattern(s) recognition that if it belongs to this category or not , then a network be able to identify any pattern(s) it output value(s)within this interval to be from second category patterns .

The range of a closed interval for this category at this experiment is equal to ٠.٠١٨٧١ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٣ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Third category :

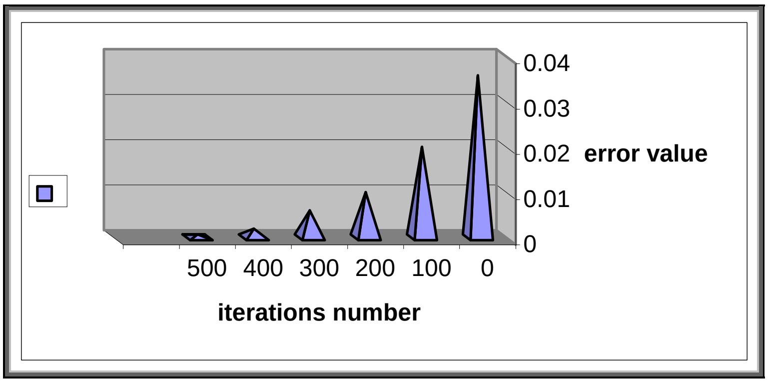
The standard figure of pattern that characterize this category is illustrated in figure(٥-٦) .The network was trained on patterns of this category which are noisy versions from pattern in figure(٥-٦) , initial values for training parameters and other parameters are same for values used in other categories which were proposed in table(٥-١).



Figure(5-6) :Third Category

The Results :

The results of a network training on a patterns in third category are represented through quadratic distribution in figure(٥-٧) ,with respect to errors values of network for each pattern in this category within limited number of maximum iterations for training .



Figure(5-7): Change error value over iterations number for third category at the first experiment

Proposed network was able to training on the patterns in third category with mean square error value equal to ٠.٠٠٠٦٤ ,which traditionally less than available error tolerance for training a network, after ٥٢٨ iterations through ٨٨ epochs .Average error value over this number of iterations was ٠.٠٠٥٩٦, while worst error value that a network reached through training on this category of patterns was ٠.٠٣٥٧٨.

Learning rate value after training over all patterns and epochs is adjusted to be equal to ٠.٠٩٩٠٢ ,it is a nearer value to a proposed value .

The outputs values for patterns in this category ,were;

output[١]=٠.٢١٤٥٣
output[٢]=٠.٢١٦١٩
output[٣]=٠.٢١٥٦٢
output[٤]=٠.٢١٦٥١
output[٥]=٠.٢١٤٨٣
output[٦]=٠.٢١٤٥٦

A variance between outputs values is occurred but is more less from that shown in previous two categories .

Standard output value for third category is :

output[١]=٠.٢١٤٥٣

Also by ordering outputs ,in increasing order have been;

output[١]=٠.٢١٤٥٣
output[٦]=٠.٢١٤٥٦
output[٥]=٠.٢١٤٨٣
output[٣]=٠.٢١٥٦٢
output[٢]=٠.٢١٦١٩
output[٤]=٠.٢١٦٥١

Note that a standard output is a minimum value than all other outputs .

Differences between versions outputs around standard output;

٠.٠٠٠٠٣
٠.٠٠٠٣٠
٠.٠٠١٠٩
٠.٠٠١٦٦
٠.٠٠١٩٨

This differences are seems small, and in higher direction. So the network can recognize outputs values(patterns output) which larger than standard value, and its different from a standard must be at less converge or equal to these differences values, that is not increase larger than a largest one .Also note that differences are only for higher than standard value.

A closed interval for third category patterns ,depending on patterns outputs values in this category is ;

[٠.٢١٤٥٣, ٠.٢١٦٥١]

,which is also represent an available region that led to take decision for pattern(s) recognition that if it belong to this category or not ,that is the network be able to identify any pattern(s) it output(s) within this interval to be from patterns in third category .

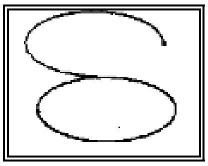
The range of a closed interval for this category at this experiment is equal to ٠.٠٠١٩٨ .

Note : The standard value is a minimum boundary for a closed interval this category .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٤ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Fourth category :

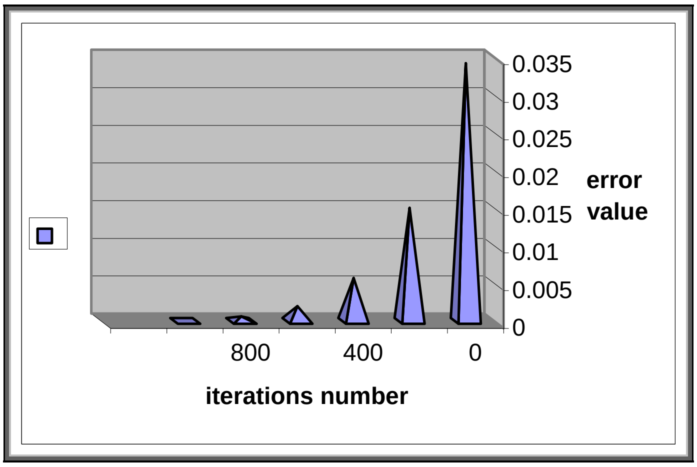
The standard figure of pattern that characterize this category is illustrated in figure(٥-٨) .The network was trained on patterns of this category which are noisy versions from a characterizer pattern , initial values for training parameters and other parameters are same for values used in other categories which proposed in table (٥-١) .



Figure(5-8) :Fourth Category

The Results :

Results for network training on a patterns in the fourth category is represented through quadratic distribution in figure(٥-٩) with respect to errors values of network for each pattern in this category within limited number of maximum iterations for training , which is ٤٠٠٠٠٠ iterations .



Figure(5-9)Change error value over iterations number for the fourth category at the first experiment

Proposed network was able to training on a patterns in the fourth category with mean square error value equal to ٠.٠٠٠٥٨ , which traditionally less than available error tolerance for training a network after ٨٤٦ iterations through ١٤١ epochs . Average error value over this number of iterations was ٠.٠٠٥٧٠ , while worst error value that a network reached through training on this category of patterns was equal to ٠.٠٣٤٢١ .

Learning rate value after training over all patterns and epochs is adjusted to be equal to ٠.٠٩٧٥٢ .

Outputs values for patterns in thus category ,were;

output[١]=٠.٣١٣٣٣
output[٢]=٠.٣١٤٥١
output[٣]=٠.٣١٥٩٨
output[٤]=٠.٣١٦٩٠
output[٥]=٠.٣١٢٩٧
output[٦]=٠.٣١٦٢٧

Standard output value for a fourth category is :

output[١]=٠.٣١٣٣٣

After an increasing order have been;

output[٥]=٠.٣١٢٩٧
output[١]=٠.٣١٣٣٣
output[٢]=٠.٣١٤٥١
output[٣]=٠.٣١٥٩٨
output[٦]=٠.٣١٦٢٧
output[٤]=٠.٣١٦٩٠

Also see that standard output is a minimum value than other outputs .

Differences between versions outputs around standard output ;

٠.٠٠٠٣٦
_____
٠.٠٠١١٨
٠.٠٠٢٦٥
٠.٠٠٢٩٤
٠.٠٠٣٥٧

The range of differences is only for higher values than standard. Pattern(s) that tested must to be different from standard value at less in a value(s) lesser than or equal to these differences to be in a fourth category interval .

Closed interval for this category patterns ,depending on patterns outputs values in this category is ;

[٠.٣١٢٩٧, ٠.٣١٦٩٠]

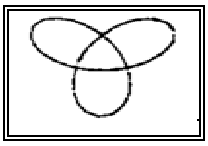
which represent an available region for outputs values in this category that led to take decision for pattern(s) recognition if it belong to this category or not .So a network be able to identify any pattern(s) it output(s) within this interval to be from its patterns .

The range of a closed interval for this category at this experiment is equal to ٠.٠٠٣٩٣.

When a testing step performed through a testing set that contain also a fuzzy versions for a standard figure for this category, the network was unable to recognize ٧ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval

Fifth category :

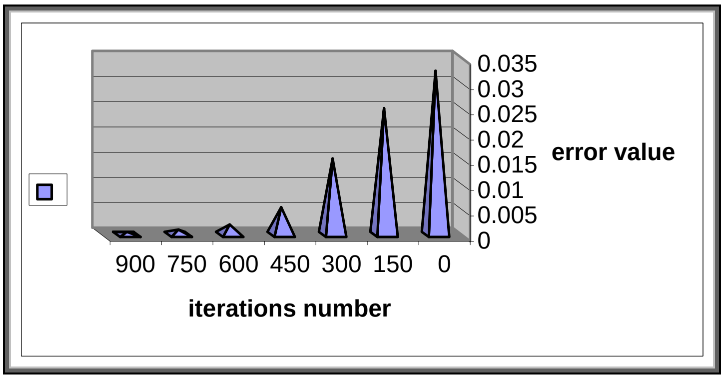
The standard figure of pattern that characterize this category is illustrated in figure(٥-١٠) .The network was trained on patterns of this category which are noisy versions from a characterizer pattern ,initial values for training parameters and other parameters are same for values used in other categories which proposed in table (٥-١) .



Figure(5-10) :Fifth Category

The Results :

The results for network training on a patterns in the fifth category are represented through quadratic distribution in figure(٥-١١),with respect to errors values of network for each pattern in this category within limited number of maximum iterations for training .



Figure(5-11):Change error value over iterations number for the fifth category at the first experiment

Proposed network was able to training on a patterns in the fifth category with mean square error value equal to ٠.٠٠٠٥٢ ,which traditionally less than available error tolerance for training a network after ٩٧٢ iterations through ١٦٢ epochs . Average error value over this number of iterations was ٠.٠٠٥٣٩ , while worst error value that a network reaches through training on this category of patterns was equal to ٠.٠٣٢٣٧ .Learning rate value after training over all patterns and epochs is adjusted to be equal to ٠.٠٩٦٧٥ .

Outputs values for patterns in this category ,were;

output[١]=٠.٤٦٥٠٦
output[٢]=٠.٤٦٧٢٥
output[٣]=٠.٤٦٩٦٣
output[٤]=٠.٤٦٦١١
output[٥]=٠.٤٧٧٦١
output[٦]=٠.٤٦٨٢١

Outputs values are variances ,increased once and decrease in other due to the variance in noise ratio between versions ,as said before .

Standard output value patterns for fifth category is :

output[١]=٠.٤٦٥٠٦

After ordering in same way for previous categories with increasing order have been;

output[١]=٠.٤٦٥٠٦
output[٤]=٠.٤٦٦١١
output[٢]=٠.٤٦٧٢٥
output[٦]=٠.٤٦٨٢١
output[٣]=٠.٤٦٩٦٣
output[٥]=٠.٤٧٧٦١

Standard value is a minimum value from outputs ,and then represent a minimum boundary of interval that characterize a fifth category .

Differences between versions outputs around standard output , were ;

٠.٠٠١٠٥
٠.٠٠١١٤
٠.٠٠٣١٥
٠.٠٠٤٥٧
٠.٠١٢٥٥

The differences are also only for higher values than standard value .

A closed interval for the fifth category patterns ,depending on patterns outputs values in this category is ;

[٠.٤٦٥٠٦, ٠.٤٧٧٦١]

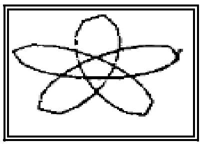
which represent an available region for outputs values in this category that led to take decision for pattern(s) recognition that if it belong to this category or not ,that is a network be able to identify any pattern(s) it output(s) within this interval to be from its patterns

The range of a closed interval for this category at this experiment is equal to ٠.٠١٢٥٥ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٣ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval .

Sixth category :

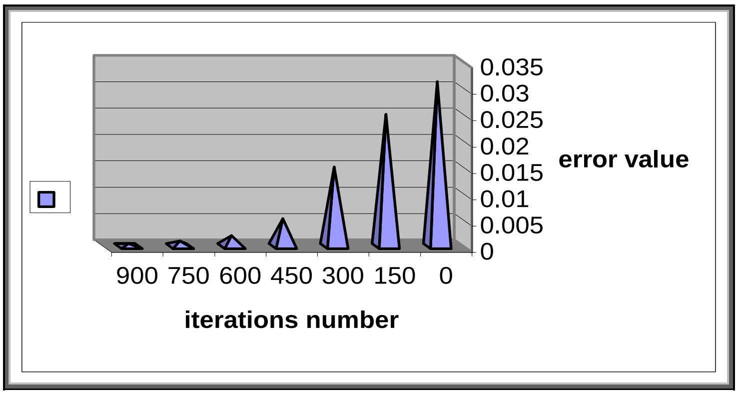
The standard figure of pattern that characterize this category is illustrated in figure(٥-١٢).The network was trained on patterns of this category which are noisy versions from a characterizer pattern initial values for training parameters and other parameters are same for values used in other categories which proposed in table (٥-١) .



Figure(5-12) :Sixth Category

The Results :

Results of a network training on a patterns in sixth category are represented through quadratic distribution in figure(٥-١٣) ,with respect to errors values of network for each pattern in this category within limited number of maximum iterations for training .



Figure(5-13): Change error value over iterations number for the sixth category at the first experiment

Proposed network was able to training on a patterns in the sixth category with mean square error value equal to ٠.٠٠٠٤٨ ,which is traditionally less than available error tolerance for training a network ,after ٨٩٤ iterations through ١٤٩ epochs .Average error value over this number of iterations was ٠.٠٠٥١٩ , while worst error value that a network reaches through training on this category of patterns was equal to ٠.٠٣١١٧. Learning rate value after training over all patterns and epochs is adjusted to be equal to ٠.٠٩٧٢٤ .

Outputs values for each pattern in this category ,were;

output[١]=٠.٦٢٠٧٣
output[٢]=٠.٦٢٣٦٠
output[٣]=٠.٦٢٥٦٤
output[٤]=٠.٦١٨٢٤
output[٥]=٠.٦١٩٩٦
output[٦]=٠.٦١٩٤٨

Standard output value for patterns of sixth category is :
output[١]=٠.٦٢٠٧٣

After ordering in same way for previous categories as increasing order have been;

output[٤]=٠.٦١٨٢٤
output[٦]=٠.٦١٩٤٨
output[٥]=٠.٦١٩٩٦
output[١]=٠.٦٢٠٧٣
output[٢]=٠.٦٢٣٦٠
output[٣]=٠.٦٢٥٦٤

Note that a standard output is not in boundaries of interval but it is a point in this category interval .

Differences between versions outputs around standard output ;

٠.٠٠٢٤٩
٠.٠٠١٢٥
٠.٠٠٠٧٧

٠.٠٠٢٨٧
٠.٠٠٤٩١

Closed interval for the sixth category patterns ,in light of patterns outputs values in this category is;

[٠.٦١٨٢٤, ٠.٦٢٥٦٤]

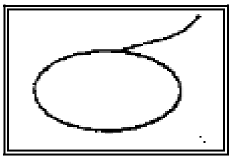
,which represent an available region for outputs values in this category that led to take decision for pattern(s) recognition that if it belong to this category or not ,that is a network be able to identify any pattern(s) it output(s) within this interval to be from patterns in the sixth category .

The range of a closed interval for this category at this experiment is equal to ٠.٠٠٧٤٠.

When a testing step done(performed) through a testing set that contain also a fuzzy versions for a standard figure for this category, the network was able to recognize ٦ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval .

Seventh category :

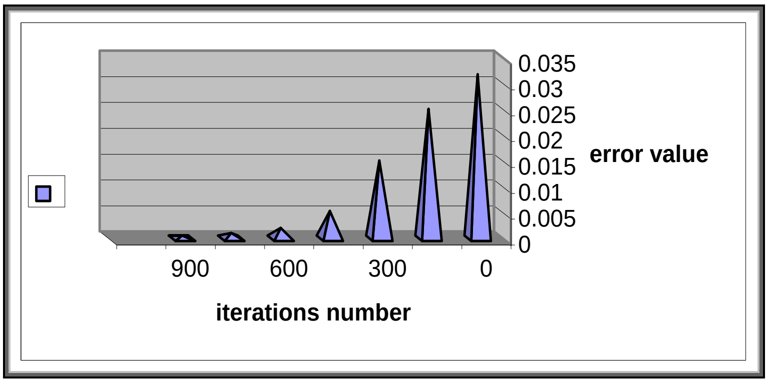
The standard figure of pattern that characterize this category is illustrate in figure(٥-١٤).The network were trained on patterns of this category which are noisy versions from a characterizer pattern, initial values for training parameters and other parameters is the same for values used in other categories which proposed in table (٥-١) .



Figure(5-14) :Seventh Category

The Results :

The results for network training on a patterns in the seventh category are represented through quadratic distribution in figure (٥-١٥) ,with respect to errors values of network for each pattern in this category within limited number of maximum iterations for training .



Figure(5-15): Change error value over iterations number for the seventh category at the first experiment

Proposed network was able to training on a patterns in the seventh category with mean square error value equal to ٠.٠٠٠٥٠ , which traditionally less than available error tolerance for training a network, after ٩٣٦ iterations through ١٥٦ epochs . Average error value over this number of iterations was ٠.٠٠٥٢٨, while worst error value that a network reaches through training on this category of patterns was equal to ٠.٠٣١٦٩ .

Learning rate value after training over all patterns and epochs is adjusted to be equal to ٠.٠٩٦٩٨ .

Outputs values for each pattern in this category were;

output[١]=٠.٧١١١٠
output[٢]=٠.٧٢٠٣٨
output[٣]=٠.٧١٩٨٦
output[٤]=٠.٧٢٠٥٣
output[٥]=٠.٧١٩١٤
output[٦]=٠.٧١٨٩١

Standard output value for patterns of seventh category is :

output[١]=٠.٧١١١٠

Ordering the outputs with increasingly order then;

output[١]=٠.٧١١١٠
output[٦]=٠.٧١٨٩١
output[٥]=٠.٧١٩١٤
output[٣]=٠.٧١٩٨٦
output[٢]=٠.٧٢٠٣٨
output[٤]=٠.٧٢٠٥٣

Note that a standard output is a minimum value, and though it will be minimum boundary for the seventh category interval .

Differences between versions outputs around standard output;

٠.٠٠٧٨١
٠.٠٠٨٠٤
٠.٠٠٨٧٦
٠.٠٠٩٢٨
٠.٠٠٩٤٣

The differences are also only for higher values than standard value .

Closed interval for the seventh category patterns ,in light of patterns outputs values in this category is;

[٠.٧١١١٠, ٠.٧٢٠٥٣]

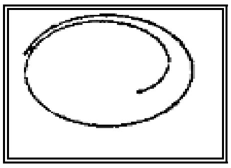
,which represent an available region for outputs values in this category that led to take decision for pattern(s) recognition that if it belong to this category or not ,that is a network be able to identify any pattern(s) it output(s) within this interval to be from patterns in the seventh category .

The range of a closed interval for this category at this experiment is equal to ٠.٠٠٩٤٣ .

When a testing step performed through a testing set that contain also a fuzzy versions for a standard figure for this category, the network was able to recognize ٤ *patterns* from a set of ٢٠ *patterns* ,for this category within a determined training interval.

Eighth category :

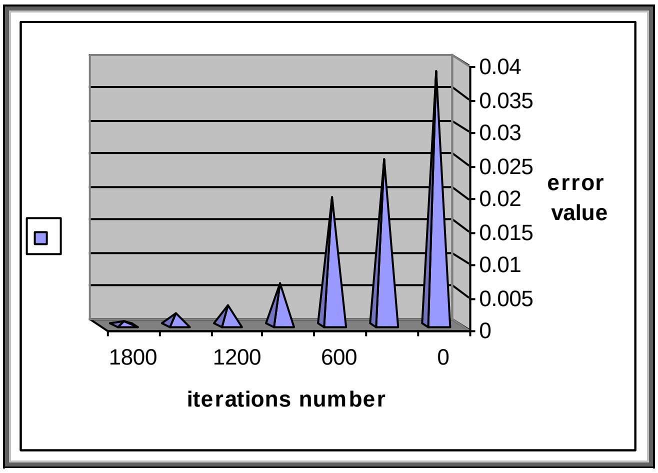
The standard figure of pattern that characterize this category is illustrated in figure (٥-١٦) .The network was trained on patterns of this category which are noisy versions from a characterizer pattern, initial values for training parameters and other parameters are same for values used in other categories which proposed in table(٥-١) .



Figure(5-16) :Eighth Category

The Results :

Results of a network training on a patterns in eighth category are represented through quadratic distribution in figure(٥-١٧) ,with respect to errors values of network for each pattern in this category within limited number of maximum iterations for training .



Figure(5-17): Change error value over iterations number for the eighth category at the first experiment

Proposed network was able to training on a patterns in eighth category with mean square error value equal to ٠.٠٠٠٧٣ ,which traditionally less than available error tolerance for training a network after ١٧٩٤ iterations through ٢٩٩ epochs . Average error value over this number of iterations was ٠.٠٠٦٣٨ , while worst error value that a network reaches through training on this category of patterns was equal to ٠.٠٣٨٣٢ .

Learning rate value after training over all patterns and epochs is adjusted to be equal to ٠.٠٨٩٣١ .

Outputs values for each pattern in this category ,were;

output[١]=٠.٩٦٤١٤
output[٢]=٠.٩٦٥٧٨
output[٣]=٠.٩٦٤٨٣
output[٤]=٠.٩٦٤٣٦
output[٥]=٠.٩٦٠٩٨
output[٦]=٠.٩٦١٨٢

Outputs values are traditionally variant ,since change in noise ratio .in spite it seem converged each to other .

Standard output value for patterns of eighth category is :

output[١]=٠.٩٦٤١٤

Ordering the outputs with increasingly order then;

output[٥]=٠.٩٦٠٩٨
output[٦]=٠.٩٦١٨٢
output[١]=٠.٩٦٤١٤
output[٤]=٠.٩٦٤٣٦
output[٣]=٠.٩٦٤٨٣
output[٢]=٠.٩٦٥٧٨

Computed differences around a standard value were;

٠.٠٠٣١٦
٠.٠٠٢٣٢
٠.٠٠٠٢٢
٠.٠٠٠٦٩
٠.٠٠١٦٤

Closed interval for the eighth category patterns ,in light of patterns outputs values in this category is;

[٠.٩٦٠٩٨, ٠.٩٦٥٧٨]

which represent an available region for outputs values in this category that led to take decision for pattern(s) recognition that if it belong to this category or not ,that is a network be able to identify any pattern(s) it output(s) within this interval to be from patterns in the eighth category .

The range of a closed interval for this category at this experiment is equal to ٠.٠٠٤٨٠ .

When a testing step done through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٤ *patterns* from a set of ٢٠ *patterns* , for this category .

Discussion of the First Experiment :

The intervals of a training set categories at first experiment are;

First category [٠.٠٢٠٦٩, ٠.٠٤٥٤٠]
Second category [٠.١٤٢١٩, ٠.١٦٠٩٠]
Third category [٠.٢١٤٥٣, ٠.٢١٦٥١]
Fourth category [٠.٣١٢٩٧, ٠.٣١٦٩٠]
Fifth category [٠.٤٦٥٠٦, ٠.٤٧٧٦١]
Sixth category [٠.٦١٨٢٤, ٠.٦٢٥٦٤]
Seventh category [٠.٧١١١ ٠,٠.٧٢٠٥٣]
Eighth category [٠.٩٦٠٩٨, ٠.٩٦٥٧٨]

By compute the ranges for a closed intervals of the eights categories as a difference between maximum boundary and minimum boundary for each closed interval ordinary , as follows ;

٠.٠٢٤٧١
٠.٠١٨٧١
٠.٠٠١٩٨
٠.٠٠٣٩٣
٠.٠١٢٥٥
٠.٠٠٧٤٠
٠.٠٠٩٤٣
٠.٠٠٤٨٠

The largest range is at first category interval that has a biggest capacity for actual outputs for patterns in this category equal to ٠.٠٢٤٧١ , while smallest range is occurred at third category interval it is equal to ٠.٠٠١٩٨ ,which is seems narrower than other intervals .

Note :

١. *Actual output for each pattern not belong to two(or more) categories intervals in the same time ,this error or -in more* *accuracy- a generalization problem and then fail to take a right decision .*

٢. Intervals for actual outputs values of patterns in categories are ordered increasingly ,this may because choice of desired values with increasing order ,then it appear as sequential intervals in numbers line(if it is ordered),this sequence began with first category interval and then the second interval and the third and the fourth ,and so on until end of all eights categories intervals.

٣. In addition to regions between these closed intervals of categories at this experiment ,there are regions out over the maximum value for all intervals(a maximum boundary of the eighth category interval in this experiment) ,and a less than a minimum value for all intervals categories(minimum boundary of a first category interval in this experiment).

It are also an error regions but it be larger ,since it expended to infinity with the two directions ,a minus and plus infinity ,to $-\infty$ *and* ∞ *,respectively .*

The differences regions between these intervals which is a part of error regions determined between each two successive intervals.

Ranges of these error regions computed for each two successive intervals as difference between upper(maximum) boundary for first one with lower(minimum) boundary of the second one ordinary , as follows :

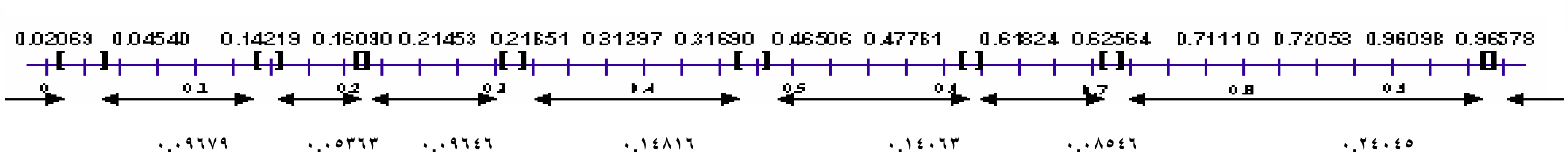
٠.٠٩٦٧٩
٠.٠٥٣٦٣
٠.٠٩٦٤٦
٠.١٤٨١٦
٠.١٤٠٦٣
٠.٠٨٥٤٦
٠.٢٤٠٤٥

We can see that a least range occurred between second interval and third interval ,equal to ٠.٠٥٣٦٣ ,this may reflect that figures (images) in these categories rather way alike or may be similar in some features but a network was capable on recognize it .The same view can be for other small values .Largest range occurred between seventh and eighth category equal to ٠.٢٤٠٤٥ ,see decision line in figure (٥-١٨) .

The resultant of testing through this experiment was at recognizing ٧٠ *patterns* for all categories from the total number of patterns ١٦٠ *patterns* that were used .

Categories	Epoch	It .No.	MSR	Average err.	Worst err.	Eta. updated	Min. boundary	Max. boundary	Range
١st category	٥٢٧	٣١٦٢	٠.٠٠٠٧٨	٠.٠٠٦٥٨	٠.٠٣٩٥٢	٠.٠٧٠٦١	٠.٠٢٠٦٩	٠.٠٤٥٤٠	٠.٠٢٤٧١
٢n category	١١	٦٦	٠.٠٠٠٠٣	٠.٠٠١٣٠	٠.٠٠٧٨٥	٠.٠٩٩٩٨	٠.١٤٢١٩	٠.١٦٠٩٠	٠.٠١٨٧١
٣rd category	٨٨	٥٢٨	٠.٠٠٠٦٤	٠.٠٠٥٩٦	٠.٠٣٥٧٨	٠.٠٩٩٠٢	٠.٢١٤٥٣	٠.٢١٦٥١	٠.٠٠١٩٨
٤th category	١٤١	٨٤٦	٠.٠٠٠٥٨	٠.٠٠٥٧٠	٠.٠٣٤٢١	٠.٠٩٧٥٢	٠.٣١٢٩٧	٠.٣١٦٩٠	٠.٠٠٣٩٣
٥th category	١٦٢	٩٧٢	٠.٠٠٠٥٢	٠.٠٠٥٣٩	٠.٠٣٢٣٧	٠.٠٩٦٧٥	٠.٤٦٥٠٦	٠.٤٧٧٦١	٠.٠١٢٥٥
٦th category	١٤٩	٨٩٤	٠.٠٠٠٤٨	٠.٠٠٥١٩	٠.٠٣١١٧	٠.٠٩٧٢٤	٠.٦١٨٢٤	٠.٦٢٥٦٤	٠.٠٠٧٤٠
٧th category	١٥٦	٩٣٦	٠.٠٠٠٥٠	٠.٠٠٥٢٨	٠.٠٣١٦٩	٠.٠٩٦٩٨	٠.٧١١١٠	٠.٧٢٠٥٣	٠.٠٠٩٤٣
٨th category	٢٩٩	١٧٩٤	٠.٠٠٠٧٣	٠.٠٠٦٣٨	٠.٠٣٨٣٢	٠.٠٨٩٣١	٠.٩٦٠٩٨	٠.٩٦٥٧٨	٠.٠٠٤٨٠

Table(5-2) : Conclusions values for the eight categories from the first experiment



Figure(5-18): Decision Line[2] for the first experiment

[٢] * The symbol ( ) is refer to that the differences is from this point into the infinity in positive direction .
** The symbol ( ) is refer to that the differences is from this point into the infinity in negative direction .

٥.٤ Second Experiment

In this experiment the suggested method performed with its essential idea by implementing an activation function used in first experiment as in equation(٥-١) ,and other essential structures for inputs and outputs ,on an extracted features values for patters of a categories in training set ,through a same steps from normalization to be suitable for a network capacity .

The aim from this experiment is to minimize the error regions that may be occurred between a closed real intervals of categories as could as ,and to minimize an error values for a network during training on all categories under training .

To satisfy this aim ,the suggested method is accomplished through increasing number of hidden neurons ,which effect on a numerical computations and then on final results in all directions . Input neurons number is fixed at ٥ *neurons* ,while number of output neurons is stay ١ *neuron* ,since it fixed in a network structure (that we cannot change it through all experiments) .The number of hidden neurons increased from ٥ *neurons* in first experiment into ٨ *neurons* in this experiment . A proposed available error value is modified from ٠.٠٠٠٨ in first experiment into ٠.٠٠٠٤ at training in this experiment, which is less than previous proposed value .

This experiment is also performed on a same training set in first experiment ,on a same proposed training parameters values .

The proposed values for training parameters and other essential structure parameters in a network for a second experiment used as in table(٥-٣) .Which is constant for all categories in training set through this experiment .

Learning rate	٠.١
Momentum	٠.٠٠١
Max. iterations	٤٠٠٠٠٠
Available error	٠.٠٠٠٤
Translation	٦
Dilation	٢
Hidden neurons	٨
Number of patterns in category	٦

Table(5-3):Proposed training parameters values at second experiment

First Category :

The network trained on this category with a new proposed values and a modification in number of hidden neurons as in table(٥-٣) .

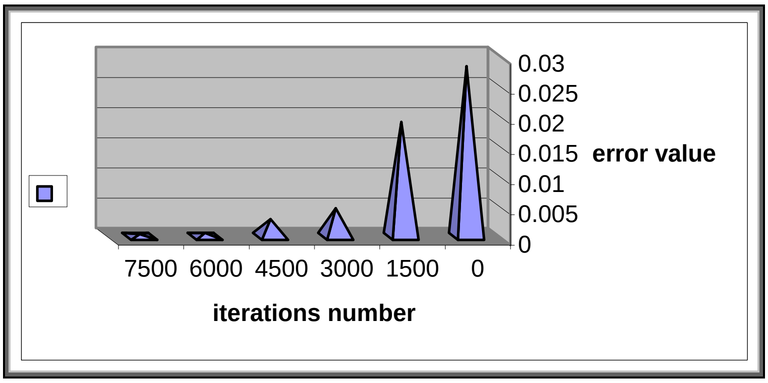
The Results

Results of training of a network on this category through a second experiment were ;

Mean squared error that a network reached at this category is equal to ٠.٠٠٠٣٩ ,which is less than that value get on in a first experiment ,also it is a maximum value for error in training a network on a training set in this experiment ,it is converged to a proposed value for error after ٧١٧٠ iterations ,through ١١٩٥ epochs. These numbers are increased over than numbers in a first experiment ,and they were a largest values for work of a network on all categories in training set.

Worst error value that a network reached is equal to ٠.٠٢٨٢٢ , and average error value is equal to ٠.٠٠٤٧٠ these values are small if it compared with that values in a first experiment .

Leaning rate value is updated on a proposed values until be ٠.٠١٦٧٢ at end the training ,this value is a least one for training in a second experiment than all other categories ,it is also less than that value for learning rate in the first experiment .



Figure(5-19): Change error value over iterations number for the first category at the second experiment

An outputs results for patterns in the first category at second experiment are as ;

output[١] = ٠.٠١١٥٨
output[٢] = ٠.٠٠٤٤٤
output[٣] = ٠.٠٣٦٠١
output[٤] = ٠.٠١١٧٩
output[٥] = ٠.٠٣١٤٦

output[٦] = ٠.٠٢٨٢١

The outputs values for patterns are appear as its ordered in a first category(patterns) in a training set .

A negative values don't mean a different or variance , it as any got on real values ,this because a variance in a noise rate between images (patterns) .

Standard output value first category is ;

output[١] = ٠.٠١١٥٨

To determine a real closed interval for outputs of patterns in this category ,that is to get a minimum and ,maximum boundaries , after reordering the outputs values with increasing order ,so ;

output[٢] = ٠.٠٠٤٤٤
output[١] = ٠.٠١١٥٨
output[٤] = ٠.٠١١٧٩
output[٦] = ٠.٠٢٨٢١
output[٥] = ٠.٠٣١٤٦
output[٣] = ٠.٠٣٦٠١

Hint: Similarly to a first experiment outputs values could ordered with decreasing order ,is true .

Differences values for patterns around the standard value through previous order for patterns ;

٠.٠٠٧١٤
٠.٠٠٠٢١
٠.٠١٦٦٣
٠.٠١٩٨٨
٠.٠٢٤٤٣

Pattern(s) to be in this category must at less not exceed than this range ,if it done it will be out of a determined closed interval for this category .The network can recognize with in these requisites .

Then a patterns outputs values in this category determine a real closed interval for this category as ;

[٠.٠٠٤٤٤, ٠.٠٣٦٠١]

which is a decision region to take a decision for pattern(s) recognition to be in first category or not .Then the network able to recognize any pattern(s) its output(s) value(s) within this interval to be from its patterns .

The range of a closed interval for a first category at a second experiment is ٠.٠٣١٥٧ ,which is larger than a range of this category interval at first experiment ,this range represent the capacity of this category to pattern(s) converge to it .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٩ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Second Category :

The network trained on this category with a new proposed values and a modification in number of hidden neurons as in table(٥-٣) .

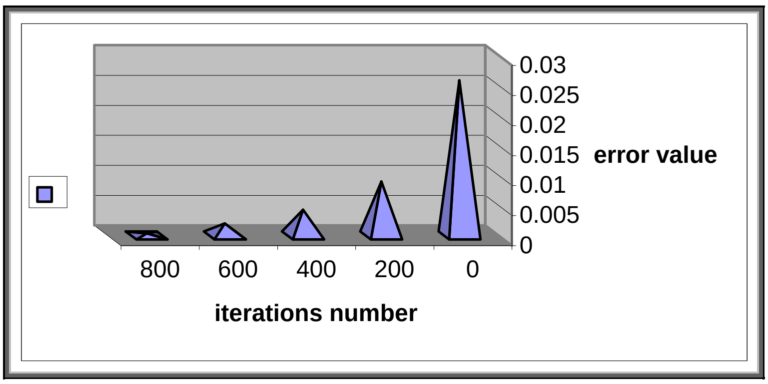
The Results

Results of training of a network on a second category depending on a proposed parameters values through a second experiment increase operations on inputs patterns values and then traditionally change the results than in second experiment, were ;

Mean squared error for a network at second category is equal to ٠.٠٠٠٣٣ in limits of a proposed tolerance ,which is higher than that value that got on in a first experiment ,after ٧٩٢ iterations, through ١٣٢ epochs ,which are increased over than numbers in a first experiment .

Worst error value that a network reached is equal to ٠.٠٢٥٨٣ is a worst error reached a network through training in this experiment on all categories ,while the average error value is equal to ٠.٠٠٤٣٠ .Note that this average value is larger than that average value in a first experiment ,while the worst error value is also larger than a value in a first experiment (this situation need attention) .

Leaning rate value is updated through training in this experiment until be ٠.٠٩٧٨٢ ,it is less than a value in a first experiment .



Figure(5-20): Change error value over iterations number for second category at the second experiment

Output results for patterns in a second category at this experiment are as ;

output[١] = ٠.١٨٢٠٥
output[٢] = ٠.١٨٤٦٦
output[٣] = ٠.١٧٩٢٠
output[٤] = ٠.٢٠٢٠٨
output[٥] = ٠.١٩١٩٥
output[٦] = ٠.١٧٥٦٣

The outputs values are appear for patterns as it were ordered in a second category(patterns) in a training set .

Standard output value is;

output[١] = ٠.١٨٢٠٥

To determine a limits of the real closed interval for outputs of patterns in this category by ordering outputs increasingly will be ;

output[٦] = ٠.١٧٥٦٣
output[٣] = ٠.١٧٩٢٠
output[١] = ٠.١٨٢٠٥
output[٢] = ٠.١٨٤٦٦
output[٥] = ٠.١٩١٩٥
output[٤] = ٠.٢٠٢٠٨

Differences values for patterns around the standard value through a previous order for patterns are;

٠.٠٠٦٤٢
٠.٠٠٢٨٥
―――――
٠.٠٠٢٦١
٠.٠٠٩٩٠
٠.٠٢٠٠٣

Then a real closed interval for a second category ,which is determined by the patterns outputs values in this category as ;

[٠.١٧٥٦٣, ٠.٢٠٢٠٨]

,which is a decision region to take a decision for pattern(s) recognition ,that determine if a pattern(s) in this category or not . Then a network be able to recognize any pattern(s) its output(s) value(s) within this interval to be from its patterns .

The range of a closed interval for a second category at a second experiment is ٠.٠٢٦٤٥ ,which is larger than a range of this category interval at first experiment ,this range represent that capacity of this category to pattern(s) converge to it is increased in this experiment .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٤ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Third Category :

The network trained on this category with a new proposed values ,and a modification in number of hidden neurons as in table(٥-٣) .

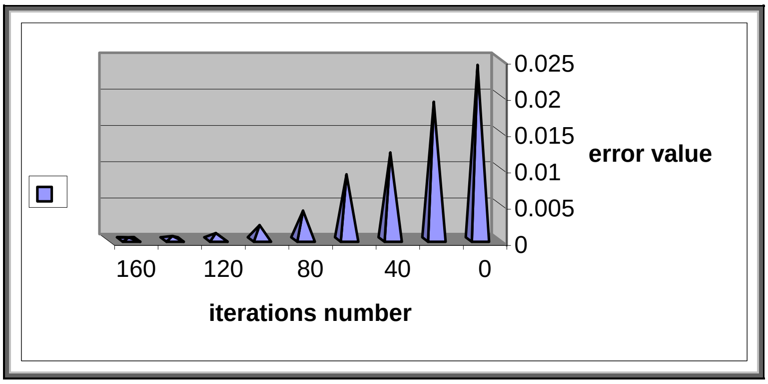
The Results :

Results of a network training on a third category depending on a proposed parameters values in a second experiment are changed than in a first experiment .

Mean squared error for training is equal to ٠.٠٠٠٢٩ ,which is less than error value at first experiment ,after ١٦٨ iterations through ٢٨ epochs .Note that number of iterations epochs are decreased . These numbers of iterations and epochs are a minimum numbers on training on all categories in a second experiment .

Worst error value that a network reached at training on this category is equal to ٠.٠٢٤٠٨ ,which is less than a value in a first experiment .An average error value is ٠.٠٠٤٠١ , which is also less than a previous value in a previous experiment .

Leaning rate value is updated through training in this experiment until be ٠.٠٩٩٨٩ ,which is maximum value for training than other categories ,and nearer value to a proposed value for learning rate. This means a proposed value is suitable for this category and it larger than a value in first experiment .



Figure(5-21): Change error value over iterations number for the third category *at the second experiment*

Outputs values of patterns in third category this experiment are;

output[١] = ٠.٢٧٣٧٢
output[٢] = ٠.٢٧٦٠١
output[٣] = ٠.٢٧٥١٩
output[٤] = ٠.٢٧٦٤٥
output[٥] = ٠.٢٧٤١٥
output[٦] = ٠.٢٧٣٧٥

This outputs values are appear for patterns as ordered in a third category in a training set .

Standard output value is;

output[١] = ٠.٢٧٣٧٢

To determine a limits of a real closed interval for outputs of patterns in this category by ordering outputs with increasing order ;

output[١] = ٠.٢٧٣٧٢
output[٦] = ٠.٢٧٣٧٥
output[٥] = ٠.٢٧٤١٥
output[٣] = ٠.٢٧٥١٩
output[٢] = ٠.٢٧٦٠١
output[٤] = ٠.٢٧٦٤٥

Differences values for patterns around the standard value through a previous order of patterns are;

٠.٠٠٠٠٣
٠.٠٠٠٤٣
٠.٠٠١٤٧
٠.٠٠٢٢٩
٠.٠٠٢٧٣

A real closed interval for a third category ,which is determined by the patterns outputs values in this category as ;

[٠.٢٧٣٧٢, ٠.٢٧٦٤٥]

,which is a decision region to take a decision for pattern(s) recognition ,that determine if a pattern(s) in this category or not .

Then the network is able to recognize any pattern(s) its output(s) value(s) within this interval to be from its patterns .

The range of a closed interval for a third category at a second experiment is ٠.٠٠٢٧٣ ,which is larger than a range of this category interval at first experiment .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٤ *patterns* from a testing set with *patterns* , for this category within a determined training interval .

Fourth Category :

Once again ,the network trained on this category with a new proposed values , and a modification in number of hidden neurons as in table(٥-٣) .

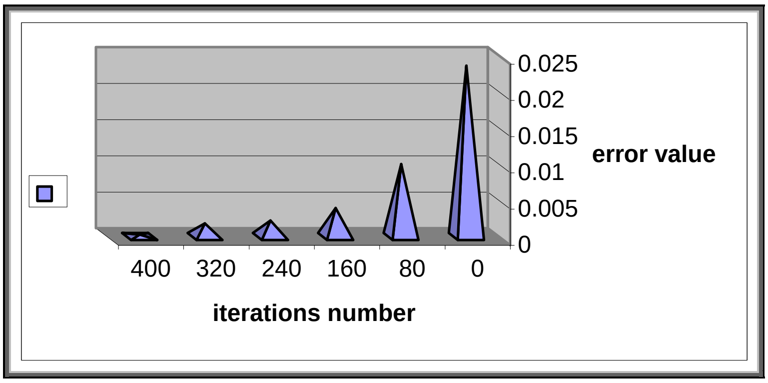
The Results :

Results of a network training on a fourth category in light of a proposed parameters values in a second experiment are changed than in a first experiment .

Mean squared error value is equal to ٠.٠٠٠٢٧ ,which is less than error value at first experiment ,after ٣٧٨ iterations and through ٦٣ epochs ,which are less than that number of iterations and epochs on training to this category at a first experiment . Note that the decreasing in error value is occurred in spite decreasing a number of iterations and epochs .

A worst error value that a network reached on training on this category is equal to ٠.٠٢٣٦٤ ,while the average error value is equal to ٠.٠٠٣٩٤ .These worst and average error values are also less than that values in a first experiment .

Leaning rate value is updated through training on this category until be ٠.٠٩٩٤٩ ,but it is larger than value that got on in the previous experiment on the same category .



Figure(5-22): Change error value over iterations number for the fourth category at the second experiment

Outputs values for patterns in a fourth category are ;

output[١] = ٠.٣٢٣٣٥

output[٢] = ٠.٣٢٤٧٨

output[٣] = ٠.٣٢٦٤٥

output[٤] = ٠.٣٢٧٤٩
output[٥] = ٠.٣٢٢٨٧
output[٦] = ٠.٣٢٦٧٤

This outputs values are appear for patterns as its ordered in a fourth category .

Standard output value is;
output[١] = ٠.٣٢٣٣٥

By ordering outputs values of patterns with increasing order as ;
output[٥] = ٠.٣٢٢٨٧
output[١] = ٠.٣٢٣٣٥
output[٢] = ٠.٣٢٤٧٨
output[٣] = ٠.٣٢٦٤٥
output[٦] = ٠.٣٢٦٧٤
output[٤] = ٠.٣٢٧٤٩

Differences values for patterns around the standard value through a previous order for patterns are;

٠.٠٠٠٤٨
٠.٠٠١٤٣
٠.٠٠٣١٠
٠.٠٠٣٣٩
٠.٠٠٤١٤

Pattern(s) output(s) value(s) must not exceed on this range , other wise it be out of the determined closed interval for this category .

A real closed interval for a fourth category ,that determined by the patterns outputs values in this category as ;

[٠.٣٢٢٨٧, ٠.٣٢٧٤٩]

,which is a decision region for take a decision for pattern(s) recognition ,that determine if a pattern(s) in this category or not .

Then the network is able to recognize any pattern(s) its output(s) value(s) within this interval to be in its patterns .

Range of a closed interval for fourth category at a second experiment is ٠.٠٠٤٦٢ ,which is larger than a range of this category interval at a first experiment .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category,

the network was able to recognize ٦ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Fifth Category :

Once again ,the network trained on this category with a new proposed values , and a modification in number of hidden neurons as in table(٥-٣) .

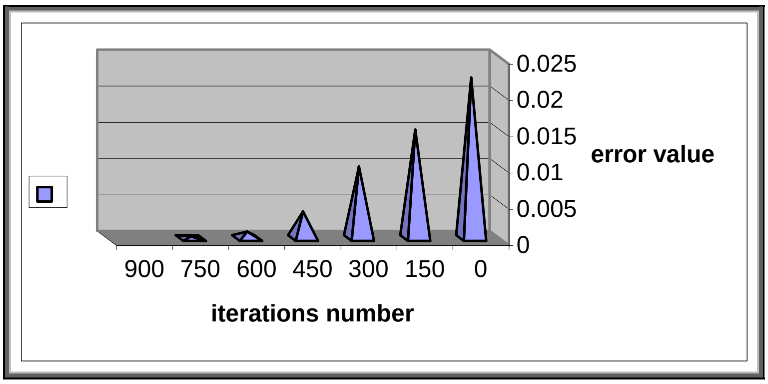
The Results :

Results of a network training on a fifth category depending on a proposed parameters values at a second experiment are changed than a first experiment .

Mean squared error value is equal to ٠.٠٠٠٢٤ ,which is traditionally less than error value at a first experiment ,after ٧٦٨ iterations ,through ١٢٨ epochs .Note that the decreasing in iterations and epochs numbers than at a previous first .

Worst error value that a network reached on training on this category is ٠.٠٢٢١٧ ,and the average error value is equal to ٠.٠٠٣٦٩ ,these worst and average error values are also less than that values in a first experiment .

Leaning rate value is updated during training on this category until be ٠.٠٩٧٩٥ ,is larger than value of learning rate on the same category in a previous experiment.



Figure(5-23): Change error value over iterations number for the fifth category at the second experiment

Outputs values for a patterns in fifth category at this experiment are as ;

output[١] = ٠.٤٧٣٩٨
output[٢] = ٠.٤٧٧٠٣

output[٣] = ٠.٤٨٠٠١
output[٤] = ٠.٤٧٥٦٦
output[٥] = ٠.٤٨٩٥٣
output[٦] = ٠.٤٧٨٢٨

This outputs values are appear for patterns as it ordered in fifth category .

Standard output value is;

output[١] = ٠.٤٧٣٩٨

To determine a closed interval for patterns outputs values in this category by ordering outputs with increasing order also, as ;

output[١] = ٠.٤٧٣٩٨
output[٤] = ٠.٤٧٥٦٦
output[٢] = ٠.٤٧٧٠٣
output[٦] = ٠.٤٧٨٢٨
output[٣] = ٠.٤٨٠٠١
output[٥] = ٠.٤٨٩٥٣

Note that a standard value is a minimum value and then will be a minimum boundary in a closed interval and all patterns values is larger than it .

Differences values for patterns larger than the standard value through a previous order for patterns are;

٠.٠٠١٦٨
٠.٠٠٣٠٥
٠.٠٠٤٣٠
٠.٠٠٦٠٣
٠.٠١٥٥٥

Differences values are also only for output values that larger than a standard one

A real closed interval for a fifth category ,that determined by patterns outputs values in this category as ;

[٠.٤٧٣٩٨ , ٠.٤٨٩٥٣]

,which is a decision region for take a decision for a pattern(s) recognition determine if a pattern(s) in this category or not .

Then the network is able to recognize any pattern(s) its output(s) value(s) within this interval to be in its patterns .

The range of a closed interval for a fifth category at a second experiment is ٠.٠١٥٥٥ ,which is larger than a range of this category interval at a first experiment .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٢ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Sixth Category :

Once again ,the network trained on this category with a new proposed values , and a modification in number of hidden neurons as in table(٥-٣) .

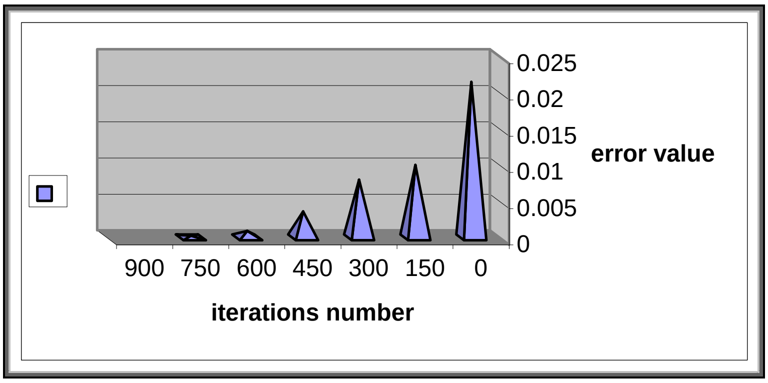
The Results :

The results of a network training on a sixth category depending on a proposed parameters values at a second experiment are changed than a first experiment .

Mean squared error value is equal to ٠.٠٠٠٢٣ ,which is traditionally less than error value at a first experiment ,this error value is a minimum error value that got on training on all categories ,after ٧٦٢ iterations , through ١٢٧ epochs , note that an iterations and epochs numbers are less than that at a previous experiment .

Worst error value at training on this category is equal to ٠.٠٢١٥٠ ,it is a least worst error value for training on categories of a training set . Average error value is equal to ٠.٠٠٣٥٨.

Leaning rate value is updated during training on this category until be ٠.٠٩٧٩٨ ,which is also larger than value for training on the same category in a first experiment.



Figure(5-24): Change error value over iterations number for the sixth category at the second experiment

Outputs values for a patterns in a sixth category at a second experiment are ;

output[١] = ٠.٦٣٠٥٨

output[٢] = ٠.٦٣٣١٠
output[٣] = ٠.٦٣٤٥١
output[٤] = ٠.٦٢٧٨١
output[٥] = ٠.٦٢٩٦٤
output[٦] = ٠.٦٢٨٩٩

this order for outputs values is appears as patterns ordered in a sixth category .

Standard output value is;

output[١] = ٠.٦٣٠٥٨

Also, to determine a boundaries of a closed interval for output values in this category ,it were ordered increasingly as;

output[٤] = ٠.٦٢٧٨١
output[٦] = ٠.٦٢٨٩٩
output[٥] = ٠.٦٢٩٦٤
output[١] = ٠.٦٣٠٥٨
output[٢] = ٠.٦٣٣١٠
output[٣] = ٠.٦٣٤٥١

Differences values for patterns around a standard value are:

٠.٠٠٢٧٧
٠.٠٠١٥٩
٠.٠٠٠٩٤
٠.٠٠٢٥٢
٠.٠٠٣٩٣

The real closed interval for a sixth category patterns at this experiment as ;

[٠.٦٢٧٨١, ٠.٦٣٤٥١]

that is a decision region to take a decision for a pattern(s) recognition determine if a pattern(s) in this category or not . Then the network is able to recognize any pattern(s) its output(s) value(s) within this interval to be in its patterns category .

Range of a closed interval for a sixth category at a second experiment is ٠.٠٠٦٧٠ ,this range less than range of this category interval at a first experiment ,which refer to that capacity of this category interval to a pattern(s) converge to it is decreased in this experiment in spite it increased for other categories intervals .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category,

the network was able to recognize ٥ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Seventh Category :

Once again ,the network trained on this category with a new proposed values ,and a modification in number of hidden neurons as in table(٥-٣) .

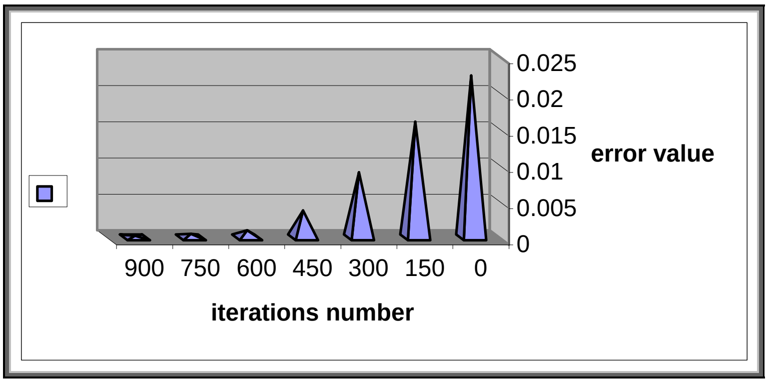
The Results :

The results of a network training on a seventh category depending on a proposed parameters values at a second experiment are changed than a previous experiment .

Mean squared error value is equal to ٠.٠٠٠٠٢٥ ,note that this value is also less than that one in a first experiment on the same category ,after ٨٣٢ iterations , through ١٣٩ epochs ,these numbers are less than that numbers at first experiment .

Worst error value at training on this category is equal to ٠.٠٢٢٣٧, while average error value is equal to ٠.٠٠٣٧٢ ,these values are also less than that values in a first experiment ,which means minimize error value at this experiment .

Leaning rate value is updated during training on this category until be ٠.٠٩٧٥٩ ,which is also larger than a value for training on the same category in a first experiment .



Figure(5-25): Change error value over iterations number for the seventh category at the second experiment

Outputs results values for a patterns in a seventh category at a second experiment are ;

output[١] = ٠.٧١٩٩٥
output[٢] = ٠.٧٢٩٦٤
output[٣] = ٠.٧٢٩١٨
output[٤] = ٠.٧٢٩٨٥
output[٥] = ٠.٧٢٨٤٣
output[٦] = ٠.٧٢٨٠٩

Also, this order for outputs values is appears as patterns ordered in a seventh category .

Standard output value is;
output[١] = ٠.٧١٩٩٥

To determine a closed interval for output values in this category , it were ordered increasingly as;

output[١] = ٠.٧١٩٩٥
output[٦] = ٠.٧٢٨٠٩
output[٥] = ٠.٧٢٨٤٣
output[٣] = ٠.٧٢٩١٨
output[٢] = ٠.٧٢٩٦٤
output[٤] = ٠.٧٢٩٨٥

Note that a standard value is a minimum value .

The differences values for a patterns that larger than standard value through a previous order as;

٠.٠٠٨١٤
٠.٠٠٨٤٨
٠.٠٠٩٢٣
٠.٠٠٩٦٩
٠.٠٠٩٩٠

Differences values are only for outputs values larger than a standard output value the pattern(s) output(s) value(s) must not exceed on this range for differences values, other wise it be out of the determined closed interval for this category .

The real closed interval for a seventh category patterns as ;

[٠.٧١٩٩٥, ٠.٧٢٩٨٥]

,which is a decision region for take a decision for a pattern(s) recognition, that determine if a pattern(s) in this category or not .

Then the network is able to recognize any pattern(s) its output(s) value(s) within this interval to be in its patterns category .

Range of this closed interval for a seventh category at a second experiment is ٠.٠٠٩٩٠ . Which is larger than range of this category interval at a first experiment , that refer to capacity of this category interval for a pattern(s) converge to it is increased in this experiment .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٤ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Eighth Category :

Once again ,the network trained on this category with a new proposed parameters values , for training in the second experiment as occurred in a table(٥-٣) .

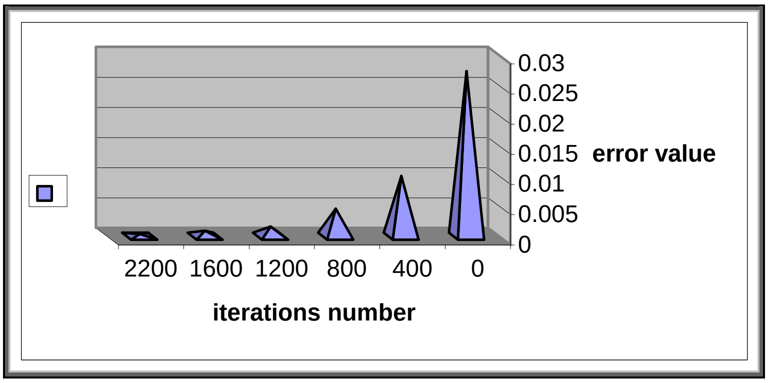
The Results :

Results of a network training on the eighth category are changed than a previous experiment .

Mean squared error value is equal to ٠.٠٠٠٠٣٧ ,which is less than a value in a first experiment ,after ٢١٤٨ iterations ,through ٣٥٨ epochs ,also note that the increasing in numbers of iterations and epochs than that in a previous one .

Worst error value at training on this category is equal to ٠.٠٢٧٤١ ,while average error value is equal to ٠.٠٠٤٥٦ , these values are less than that in a first experiment ,and average error value is a maximum one for training a network on a categories in a training set which means minimize error value at this experiment .

Leaning rate value is updated through training on this category until be ٠.٠٨٥١٥ ,which is less than a value for training on same category in a first experiment .



Figure(5-26): Change error value over iterations number for the eighth category at the second experiment

Outputs results values for a patterns in a seventh category at a second experiment are ;

output[١] = ٠.٩٧٥٢٦
output[٢] = ٠.٩٧٦٩٧
output[٣] = ٠.٩٧٥٩٦
output[٤] = ٠.٩٧٥٥٦
output[٥] = ٠.٩٧١٦٧
output[٦] = ٠.٩٧٢٦٦

This order for patterns outputs values is appears as patterns ordered in an eighth category .

Standard output value is;

output[١] = ٠.٩٧٥٢٦

To determine a closed interval for outputs values in this category through increasing order as ;

output[٥] = ٠.٩٧١٦٧
output[٦] = ٠.٩٧٢٦٦
output[١] = ٠.٩٧٥٢٦
output[٤] = ٠.٩٧٥٥٦
output[٣] = ٠.٩٧٥٩٦
output[٢] = ٠.٩٧٦٩٧

Differences values for a patterns outputs values around a standard output value with previous order as;

٠.٠٠٣٥٩
٠.٠٠٢٦٠
٠.٠٠٠٣٠
٠.٠٠٠٧٠
٠.٠٠١٧١

The real closed interval for an eighth category, that it determined by a patterns outputs values in this category is;

[٠.٩٧١٦٧, ٠.٩٧٦٩٧]

,which is a decision region for take a decision on pattern(s) recognition, that determine if a pattern(s) in this category or not .

Then the network is able to recognize any pattern(s) its output(s) value(s) within this interval to be in its patterns category .

Range of a closed interval of the eighth category at a second experiment is ٠.٠٠٥٣٠ ,which less than range of this category interval at a first experiment .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٣ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Discussion of the Second Experiment :

The categories intervals of a training set at the second experiment are;

First category	[٠.٠٠٤٤٤ ,٠.٠٣٦٠١]
Second category	[٠.١٧٥٦٣ ,٠.٢٠٢٠٨]
Third category	[٠.٢٧٣٧٢ ,٠.٢٧٦٤٥]
Fourth category	[٠.٣٢٢٨٧ ,٠.٣٢٧٤٩]
Fifth category	[٠.٤٧٣٩٨ ,٠.٤٨٩٥٣]
Sixth category	[٠.٦٢٧٨١ ,٠.٦٣٤٥١]
Seventh category	[٠.٧١٩٩٥ ,٠.٧٢٩٨٥]
Eighth category	[٠.٩٧١٦٧ ,٠.٩٧٦٩٧]

The ranges of these intervals at this order of categories intervals are;

٠.٠٣١٥٧
٠.٠٢٦٤٥
٠.٠٠٢٧٣
٠.٠٠٤٦٢
٠.٠١٥٥٥
٠.٠٠٦٧٠
٠.٠٠٩٩٠
٠.٠٠٥٣٠

A least range is for third category interval equal to ٠.٠٠٢٧٣ ,while a largest one is for a first category interval and equal to ٠.٠٣١٥٧ .

Differences of the regions between these intervals -error regions- for a second experiment are as;

٠.١٣٩٦٢
٠.٠٧١٦٤
٠.٠٤٦٤٢
٠.١٤٦٤٩
٠.١٣٨٢٨
٠.٠٨٥٤٤
٠.٢٤١٨٢

A least difference occurred between a third category interval and the fourth category interval equal to ٠.٠٤٦٤٢ .This may reflect that these patterns categories are rather way alike or may be similar in some features values but the network was capable on recognize it at a second experiment, and error in recognition is least ,the same view for other small differences values .

Largest difference(range) is occurred between a seventh and eighth category interval and it is equal to ٠.٢٤١٨٢ ,it means that the error is larger between these two categories .

Note :

١. The output value for any pattern(actual output) do not belongs to two(or more) categories interval in the same time at a second experiment .

٢. The outputs values at categories intervals were ordered increasingly ,this may because the choice of desired outputs values with increasing order ,and then it appear as sequential intervals in numbers line(if it is ordered) ,see figure(٥-٢٧) for a decision line at a second experiment ,this sequence began with a first category interval and then the second category interval and the third and the fourth ,and so on until ending of all the eights categories intervals.

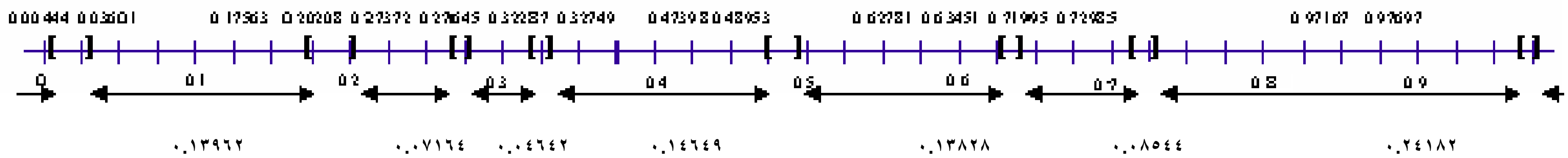
٣. In addition to regions between these closed intervals of categories ,there are regions out over the maximum value for all intervals(a maximum boundary of the eighth category interval in this experiment) ,and a less than a minimum value for all intervals categories(minimum boundary of a first category interval in this experiment).

It are also an error regions but it be larger ,since it expended *to infinity with the two directions ,a minus and plus infinity ,to $-\infty$ and ∞ ,respectively .*

The resultant of testing through this experiment, was at recognize ٦٧ *pattern* for all categories from the number of patterns ١٦٠ *patterns* that were used .

Categories	Epoch	It .No.	MSR	Average err.	Worst err.	Eta. updated	Min. boundary	Max. boundary	Range
١st category	١١٩٥	٧١٧٠	٠.٠٠٠٣٩	٠.٠٠٤٧٠	٠.٠٢٨٢٢	٠.٠١٦٧٢	٠.٠٠٤٤٤	٠.٠٣٦٠١	٠.٠٣١٥٧
٢n category	١٣٢	٧٩٢	٠.٠٠٠٣٣	٠.٠٠٤٣٠	٠.٠٢٥٨٣	٠.٠٩٧٨٢	٠.١٧٥٦٣	٠.٢٠٢٠٨	٠.٠٢٦٤٥
٣rd category	٢٨	١٦٨	٠.٠٠٠٢٩	٠.٠٠٤٠١	٠.٠٢٤٠٨	٠.٠٩٩٨٩	٠.٢٧٣٧٢	٠.٢٧٦٤٥	٠.٠٠٢٧٣
٤th category	٦٣	٣٧٨	٠.٠٠٠٢٧	٠.٠٠٣٩٤	٠.٠٢٣٦٤	٠.٠٩٩٤٩	٠.٣٢٢٨٧	٠.٣٢٧٤٩	٠.٠٠٤٦٢
٥th category	١٢٨	٧٦٨	٠.٠٠٠٢٤	٠.٠٠٣٦٩	٠.٠٢٢١٧	٠.٠٩٧٩٥	٠.٤٧٣٩٨	٠.٤٨٩٥٣	٠.٠١٥٥٥
٦th category	١٢٧	٧٦٢	٠.٠٠٠٢٣	٠.٠٠٣٥٨	٠.٠٢١٥٠	٠.٠٩٧٩٨	٠.٦٢٧٨١	٠.٦٣٤٥١	٠.٠٠٦٧٠
٧th category	١٣٩	٨٣٢	٠.٠٠٠٢٥	٠.٠٠٣٧٢	٠.٠٢٢٣٧	٠.٠٩٧٥٩	٠.٧١٩٩٥	٠.٧٢٩٨٥	٠.٠٠٩٩٠
٨th category	٣٥٨	٢١٤٨	٠.٠٠٠٣٧	٠.٠٠٤٥٦	٠.٠٢٧٤١	٠.٠٨٥١٥	٠.٩٧١٦٧	٠.٩٧٦٩٧	٠.٠٠٥٣٠

Table(5-4): Conclusions values for the eight categories from the second experiment



Figure(5-27): Decision Line for the second experiment

٥.٥ Third Experiment :

The third experiment idea is proposed to show effect of increasing number of patterns in each category than previous number which proposed in a first and second experiments .This effect will shown clearly through implementing on a categories of a training set with a same proposed training parameters values , and other essential parameters in first experiment ,as in table (٥-١) .The number of patterns were increased to one pattern into each category .Then number of patterns will be ٥٦ *patterns* , with ٧ *patterns* in each category .These patterns were extracted in same way and features also extracted in a same file for a category .

Inputs values were normalized to be inputs to a network ,later results will illustrate for each category .

Same steps would accomplish on outputs values such in a first experiment .Outputs values would be ordered increasingly to determine a boundaries of a closed interval of outputs values for each category ,and then compute a range(which represent a scope capacity) for each category interval to know how much values can this interval converge on outputs values to a standard output value for that category .

The ranges of error regions were also computed ,and compared with the others in a previous experiments .

Note :The seventh pattern that added to each category through this experiment is also a version from a standard figure for that category ,that is versions number is increased at each category .

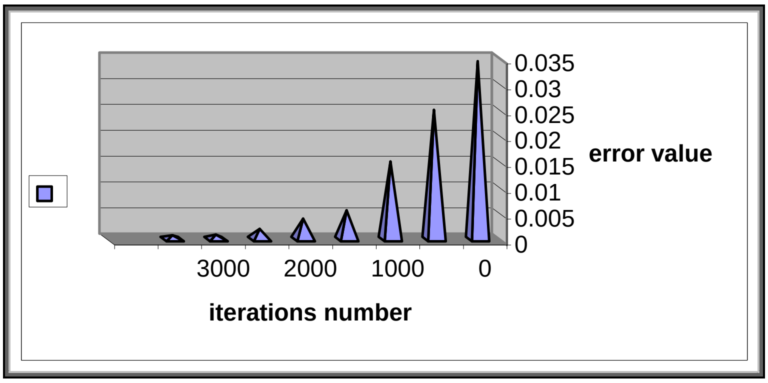
Note :Steps of a suggested method is also implemented , it is extracted from the resource and then extract features values to get a numerical file of a features values for seven patterns at each category .

First Category :

The results of training on this category were illustrated in figure(٥-٢٨) for changing error value over an iterations number .

Mean square error value is equal to ٠.٠٠٠٧٧ after ٣٧٠٣ iterations, through ٥٢٩ epochs at training .The worst error value over a full epochs number equal to ٠.٠٣٩٤٣ ,and the average error value over a full epochs number is ٠.٠٠٥٦٣ .

Learning rate is updated through training ,its value after end training is ٠.٠٧٠٤٢ .The error value here is a maximum error value for training over all categories at this experiment .



Figure(5-28): Change error value over iterations number for the first category at the third experiment

Outputs values for each pattern in this category under work in third experiment are;

output[٧] = ٠.٠٤٠٥٩
output[١] = ٠.٠٢٦٤٨
output[٢] = ٠.٠٢٠٦١
output[٣] = ٠.٠٤٥٣١
output[٤] = ٠.٠٢٦٥٦
output[٥] = ٠.٠٤١٨٩
output[٦] = ٠.٠٣٩٣٩

This order for patterns outputs values is appear here as patterns order in the first category after adding the seventh pattern(copy) .

Note the outputs values changed for patterns than a first experiment ,also changing a standard output value at this experiment .

Standard output value is;

output[١] = ٠.٠٢٦٤٨

After reordering outputs values ,as increasing order will be ;

output[٢] = ٠.٠٢٠٦١
output[١] = ٠.٠٢٦٤٨
output[٤] = ٠.٠٢٦٥٦
output[٦] = ٠.٠٣٩٣٩
output[٧] = ٠.٠٤٠٥٩
output[٥] = ٠.٠٤١٨٩
output[٣] = ٠.٠٤٥٣١

Differences values for patterns outputs values around a standard output value through the order be ;

٠.٠٠٥٨٧
٠.٠٠٠٠٨
٠.٠١٢٩١
٠.٠١٤١١
٠.٠١٥٤١
٠.٠١٨٨٣

Then the real closed interval for a first category ,at this experiment ,that determined by a patterns outputs values after ordering be ;

[٠.٠٢٠٦١ , ٠.٠٤٥٣١]

which is a decision region for pattern(s) recognition at this category through it a network be able to recognize a pattern(s) if it in a patterns of this category .

The range of a closed interval for this category at third experiment is ٠.٠٢٤٧٠ .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٩ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

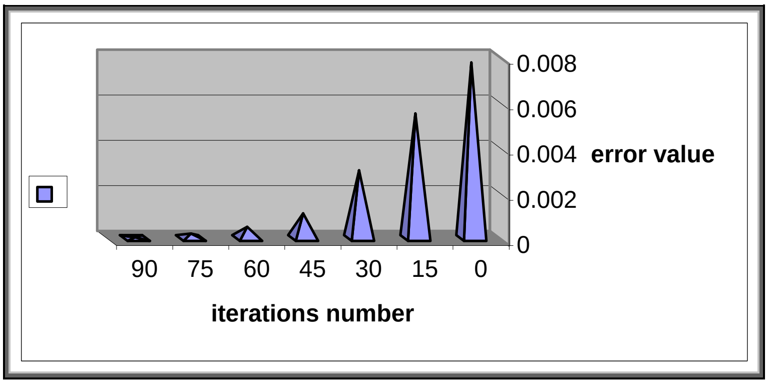
Second Category :

The results of training on this category were illustrated in figure(٥-٢٩) for changing error value over an iterations number .

Mean square error value for training is ٠.٠٠٠٠٣ ,which is a minimum error value for training on all categories in a training set at this experiment ,after ٩١ iterations through ١٣ epochs. Note that this number of iterations and epochs is a minimum number at work of a network at training .Worst error value over a full epochs is ٠.٠٠٧٧٦ ,and average error value over a full epochs is ٠.٠٠١١٠ .

These error values for average and worst error are a minimal values for training on all other categories .

Learning rate at this category is updated through training ,to be at value ٠.٠٩٩٩٧,it is also a maximum learning rate value for training on all categories ,that means the choice of this value is appropriate for this category .



Figure(5-29): Change error value over iterations number for second category at the third experiment

Outputs values for all patterns in this category under work in third experiment are;

output[٧] = ٠.١٤٨١٤
output[١] = ٠.١٤٧٠٠
output[٢] = ٠.١٤٨٦٥
output[٣] = ٠.١٤٤٨٠
output[٤] = ٠.١٦٠٩٩
output[٥] = ٠.١٥٣٨٠
output[٦] = ٠.١٤٢٢٨

This order for patterns outputs values is appear here as patterns ordered in second category after adding the seventh pattern to a category .Note outputs values changed for patterns than in a first experiment ,also changing a standard output value at this experiment .

Standard output value is;

output[١] = ٠.١٤٧٠٠

Now by reordering outputs values to determine a closed interval for this category ,with increasing order as ;

output[٦] = ٠.١٤٢٢٨
output[٣] = ٠.١٤٤٨٠
output[١] = ٠.١٤٧٠٠
output[٧] = ٠.١٤٨١٤

output[٢] = ٠.١٤٨٦٥
output[٥] = ٠.١٥٣٨٠
output[٤] = ٠.١٦٠٩٩

The differences values for patterns outputs values around a standard output value after ordering are ;

٠.٠٠٤٧٢
٠.٠٠٢٢٠
٠.٠٠١١٤
٠.٠٠١٦٥
٠.٠٠٦٨٠
٠.٠١٣٩٩

A real closed interval for a second category at third experiment , which determined by ordering a patterns outputs values in this category is ;

[٠.١٤٢٢٨, ٠.١٦٠٩٩]

which is a decision region for pattern(s) recognition at this category ,through it a network be able to recognize a pattern(s) if it in a patterns of this category .

The range of a closed interval for a second category at this experiment is ٠.٠١٨٧١ , is maximum than a range for this category in a previous experiments, also it is a maximum value for range of a closed interval category than all other intervals categories at this experiment ,that is wide capacity for pattern(s) output(s) value(s) .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٣ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

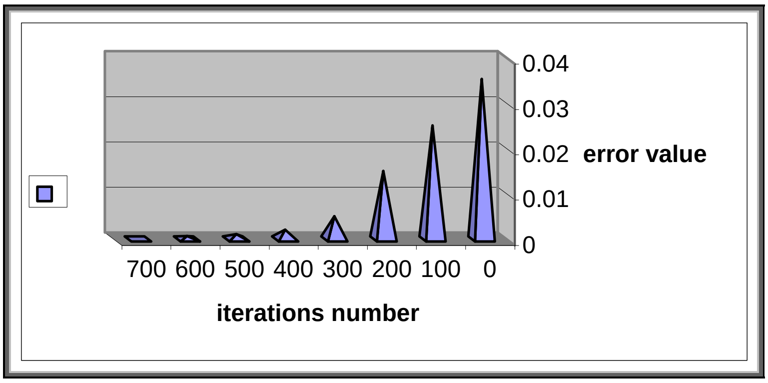
Third Category :

The results of training on this category were illustrated in figure(٥-٣٠) for changing error value over an iterations number .

Mean square error value for training on this category is equal to ٠.٠٠٠٦٢ ,after ٦٥١ iterations through ٩٣ epochs .

Worst error value over a full epochs is ٠.٠٣٥٢٦ , and the average error value over a full epochs is ٠.٠٠٥٠٣ .This average error values is a maximum value through training on all other categories .

A learning rate is updated through training ,and it is value after end the training is ٠.٠٩٨٩١ .



Figure(5-30): Change error value over iterations number for the third category at the third experiment

Outputs values for all patterns in this category and under work in a third experiment are;

output[٧] = ٠.٢٠٩٧١
output[١] = ٠.٢١٥٠٥
output[٢] = ٠.٢١٦٦٩
output[٣] = ٠.٢١٦١٣
output[٤] = ٠.٢١٧٠١
output[٥] = ٠.٢١٥٣٤
output[٦] = ٠.٢١٥٠٧

This order for patterns outputs values is appear as a patterns ordered in third category after adding the seventh pattern to a category ,also note that outputs values changed for patterns than that in a first experiment ,and changing a standard output value at this experiment .

Standard output value is;
output[١] = ٠.٢١٥٠٥

Outputs values were ordered to determine a closed interval for this category ,increasingly as ;

output[٧] = ٠.٢٠٩٧١
output[١] = ٠.٢١٥٠٥
output[٦] = ٠.٢١٥٠٧
output[٥] = ٠.٢١٥٣٤
output[٣] = ٠.٢١٦١٣
output[٢] = ٠.٢١٦٦٩
output[٤] = ٠.٢١٧٠١

Differences values for patterns outputs values around a standard output value after ordering be ;

٠.٠٠٥٣٤

٠.٠٠٠٠٢
٠.٠٠٠٢٩
٠.٠٠١٠٨
٠.٠٠١٦٤
٠.٠٠١٩٦

The real closed interval for a third category at this experiment , by ordering patterns outputs values increasingly in this category is;

[٠.٢٠٩٧١, ٠.٢١٧٠١]

which is a decision region for pattern(s) recognition at this category ,that is a network is able to recognize a pattern(s) its output(s) value(s) in this interval ,in its patterns category .

The range of a closed interval for a third category at this experiment is ٠.٠٠٧٣٠ .

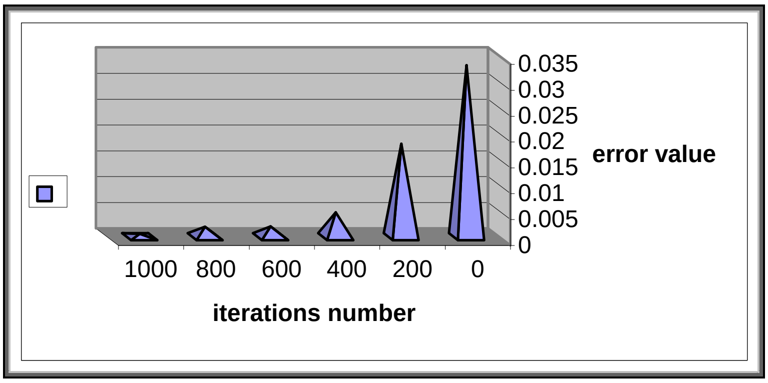
When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٣ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Fourth Category :

The results of training on this category were illustrated in figure(٥-٣١) for changing error value over an iterations number .

Mean square error value is ٠.٠٠٠٥٥ ,after ١٠٠٨ iterations through ١٤٤ epochs ,the worst error value over a full epochs is ٠.٠٣٣٢٢ , and average error is ٠.٠٠٤٧٤ .

A learning rate is updated through training ,and its value after end training is ٠.٠٩٨٩١ .



Figure(5-31):Change error value over iterations number for the fourth category at the third experiment

Outputs values for all patterns in this category at the third experiment are;

output[٧] = ٠.٣٠٧٢٢
output[١] = ٠.٣١٤٣٢
output[٢] = ٠.٣١٥٤٩
output[٣] = ٠.٣١٦٩٥
output[٤] = ٠.٣١٧٨٧
output[٥] = ٠.٣١٣٩٦
output[٦] = ٠.٣١٧٢٤

Also, this order for patterns outputs values is appear here after adding the seventh pattern to a category ,the outputs values changed for patterns that in a first experiment ,and changing a standard output value at this experiment .

Standard output value is;

output[١] = ٠.٣١٤٣٢

The outputs values were ordered to determine a boundaries of a closed interval for this category , as ;

output[٧] = ٠.٣٠٧٢٢
output[٥] = ٠.٣١٣٩٦
output[١] = ٠.٣١٤٣٢
output[٢] = ٠.٣١٥٤٩
output[٣] = ٠.٣١٦٩٥
output[٦] = ٠.٣١٧٢٤
output[٤] = ٠.٣١٧٨٧

The differences values for patterns outputs values around a standard output value after previous order be ;

٠.٠٠٧١٠
٠.٠٠٠٣٦

٠.٠٠١١٧
٠.٠٠٢٦٣

٠.٠٠٢٩٢
٠.٠٠٣٥٥

The real closed interval for a fourth category, at this experiment, which is determined by ordering increasingly a patterns outputs values in this category is ;

[٠.٣٠٧٢٢, ٠.٣١٧٨٧]

which is a decision region for pattern(s) recognition at this category ,and a network be able to recognize a pattern(s) its output(s) value(s) in this interval ,in its patterns category .

The range(scope) of this closed interval for the fourth category at this experiment is ٠.٠١٠٦٥ .

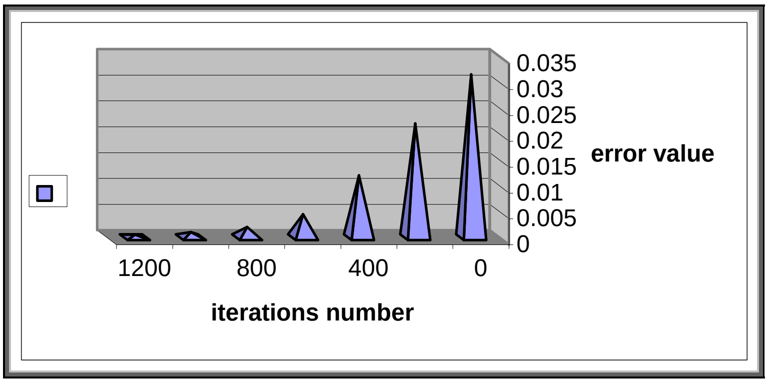
When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٤ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Fifth Category :

The results of training on this category were illustrated in figure(٥-٣٢) for changing error value over an iterations number .

Mean square error value for training in this category is ٠.٠٠٤٥٠ , after ١١٤٨ iterations through ١٦٤ epochs ,a worst error value over a full epochs is ٠.٠٣١٥١ , and the average error value over a full epochs is ٠.٠٠٠٤٩ .

Learning rate at this category is updated through training until be ٠.٠٩٦٦٧ .



Figure(5-32): Change error value over iterations number for the fifth category at the third experiment

The outputs values for all patterns in this category at the third experiment are;

output[٧] = ٠.٤٦٠٣٦
output[١] = ٠.٤٦٥٩١
output[٢] = ٠.٤٦٨١٠
output[٣] = ٠.٤٧٠٤٧
output[٤] = ٠.٤٦٦٩٧
output[٥] = ٠.٤٧٨٤٤
output[٦] = ٠.٤٦٩٠٦

Also, this order for patterns outputs values is appear as a patterns ordered in the fifth category after adding the seventh pattern to a category .Outputs values were changed for patterns than in a first experiment ,and changing a standard output value at this experiment .

Standard output value is;
output[١] = ٠.٤٦٥٩١

After ordering the outputs values to determine a closed interval for this category , with increasing order as ;

output[٧] = ٠.٤٦٠٣٦
output[١] = ٠.٤٦٥٩١
output[٤] = ٠.٤٦٦٩٧
output[٢] = ٠.٤٦٨١٠
output[٦] = ٠.٤٦٩٠٦
output[٣] = ٠.٤٧٠٤٧
output[٥] = ٠.٤٧٨٤٤

Differences values for patterns outputs values around a standard output value after previous order are;

٠.٠٠٥٥٥
٠.٠٠١٠٦
٠.٠٠٢١٩
٠.٠٠٣١٥
٠.٠٠٤٥٦
٠.٠١٢٥٣

The real closed interval for the fifth category at this experiment , which determined by ordering a patterns outputs values in this category is ;

[٠.٤٦٠٣٦, ٠.٤٧٨٤٤]

,which is a decision region for pattern(s) recognition at this category ,through it a network be able to recognize a pattern(s) it in a patterns of this category .

The range of a closed interval for the fifth category and at this experiment is ٠.٠١٨٠٨ ,which is larger than a range for this category in a previous experiments .

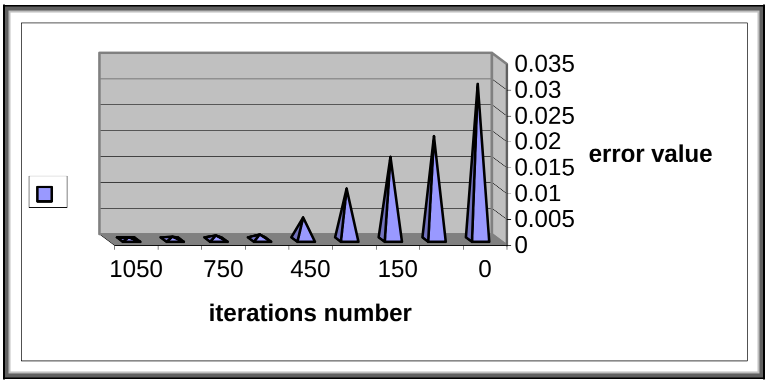
When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٩ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Sixth Category :

The results of training on this category were illustrated in figure(٥-٣٣) for changing error value over an iterations number .

Mean square error value for training is equal to ٠.٠٠٠٤٥ , after ١٠٥٧ iterations through ١٥١ epochs at end training .Worst error value over a full epochs number is ٠.٠٣٠٠٧ , and the average error value over a full epochs number is ٠.٠٠٤٢٩ .

Learning rate is updated through training ,and its value after training is ٠.٠٩٧١٧ .



Figure(5-33): Change error value over iterations number for the sixth category at the third experiment

The outputs values for patterns in this category at the third experiment are;

output[٧] = ٠.٦٠٨٦٩
output[١] = ٠.٦٢١٨٠
output[٢] = ٠.٦٢٤٦٧

output[٣] = ٠.٦٢٦٧٠
output[٤] = ٠.٦١٩٣١
output[٥] = ٠.٦٢١٠٣
output[٦] = ٠.٦٢٠٥٥

Also, this order for patterns outputs values is appear as a patterns ordered in the first experiment for this category ,but after adding the seventh pattern. The outputs values were changed for patterns than in a first experiment ,and changing a standard output value at this experiment .

Standard output value is;
output[١] = ٠.٦٢١٨٠

After reordering the outputs values with increasing order as ;

output[٧] = ٠.٦٠٨٦٩
output[٤] = ٠.٦١٩٣١
output[٦] = ٠.٦٢٠٥٥
output[٥] = ٠.٦٢١٠٣
output[١] = ٠.٦٢١٨٠
output[٢] = ٠.٦٢٤٦٧
output[٣] = ٠.٦٢٦٧٠

Differences values for a patterns outputs values around standard output value under previous order be ;

٠.٠١٣١١
٠.٠٠٢٤٩
٠.٠٠١٢٥
٠.٠٠٠٧٧
٠.٠٠٢٨٧
٠.٠٠٤٩٠

Then the real closed interval for sixth category at this experiment after the ordering is ;

[٠.٦٠٨٦٩, ٠.٦٢٦٧٠]

,which is a decision region for pattern(s) recognition at this category ,through it a network be able to recognize a pattern(s) if it in a network outputs values for this category to be in its patterns category .

The range of a closed interval for the sixth category and at this experiment is ٠.٠١٨٠١ ,is refer to that a capacity of this category interval is increased in this experiment than a range previous experiments .

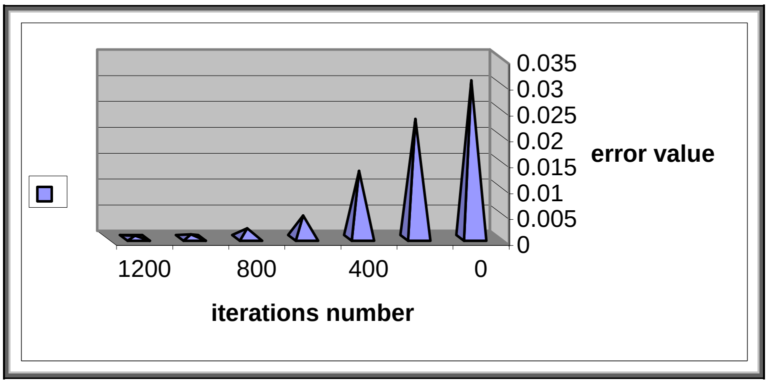
When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ١٣ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Seventh Category :

The results of training on this category were illustrated in figure(٥-٣٤) for changing error value over an iterations number .

Mean square error value is equal to ٠.٠٠٠٤٦ , after ١١٠٦ iterations through ١٥٨ epochs at training .The worst error value over a full epochs number is equal to ٠.٠٣٠٤٨, and average error value over a full epochs is ٠.٠٠٤٣٥ .

Learning rate is updated through training ,and its value after end training is ٠.٠٩٦٩٠ .



Figure(5-34): Change error value over iterations number for the seventh category at the third experiment

Outputs values for patterns in this category at the third experiment are;

output[٧] = ٠.٧٠٥٦٨

output[١] = ٠.٧١٢٢٨

output[٢] = ٠.٧٢١٥٦

output[٣] = ٠.٧٢١٠٤
output[٤] = ٠.٧٢١٧٢
output[٥] = ٠.٧٢٠٣٢
output[٦] = ٠.٧٢٠١٠

This order for patterns outputs values is appear as a patterns ordered in the first experiment for this category after adding the seventh pattern to a category .Outputs values were changed for patterns than in a first experiment ,and changing a standard output value at this experiment .

Standard output value is;
output[١] = ٠.٧١٢٢٨

After reordering the outputs values to determine a closed interval for this category ,with increasing order as ;

output[٧] = ٠.٧٠٥٦٨
output[١] = ٠.٧١٢٢٨
output[٦] = ٠.٧٢٠١٠
output[٥] = ٠.٧٢٠٣٢
output[٣] = ٠.٧٢١٠٤
output[٢] = ٠.٧٢١٥٦
output[٤] = ٠.٧٢١٧٢

The differences values for patterns outputs values around a standard output value under previous order be ;

٠.٠٠٦٦٠
٠.٠٠٧٨٢
٠.٠٠٨٠٤
٠.٠٠٨٧٦
٠.٠٠٩٢٨
٠.٠٠٩٤٤

Then the real closed interval for the seventh category at this experiment ,which is determined by a patterns outputs values in this category ,through previous order is ;

[٠.٧٠٥٦٨, ٠.٧٢١٧٢]

,which is a decision region for pattern(s) recognition at this category ,through it a network be able to recognize a pattern(s) if it output(s) value(s) be in this category interval .

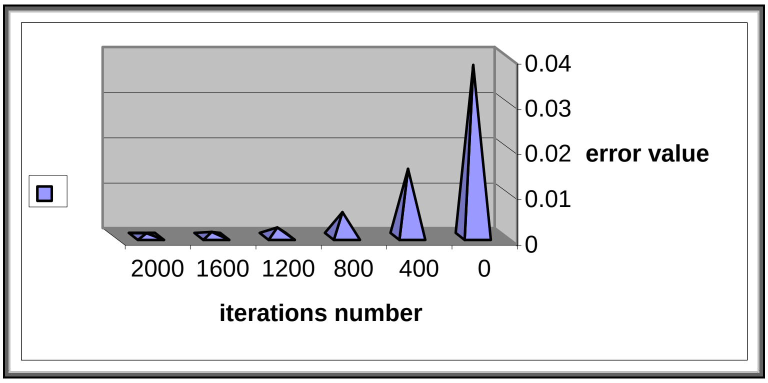
The range of a closed interval for a seventh category and at this experiment is ٠.٠١٦٠٤ ,is refer to that a capacity of this category interval is increased in this experiment than previous experiments .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٦ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Eighth Category :

The results of training on this category were illustrated in figure(٥-٣٥) for changing error value over an iterations number .

Mean square error value for training is equal to ٠.٠٠٠٧٢ , after ٢١٤٢ iterations through ٣٠٦ epochs. The worst error value that a network reached over a full epochs is equal to ٠.٠٣٧٩٦ ,which is a maximum value for worst error for training a network on all categories at training set ,average error value over a full epochs is ٠.٠٠٥٤٢ .The learning rate is updated through training ,and its value after end training is ٠.٠٨٨٩١ .



Figure(5-35): Change error value over iterations number for the eighth category at the third experiment

Outputs values for patterns in this category at the third experiment are;

output[٧] = ٠.٩٤٢٣٤
output[١] = ٠.٩٦٤٤٦
output[٢] = ٠.٩٦٦٠٨
output[٣] = ٠.٩٦٥١٤
output[٤] = ٠.٩٦٤٦٨
output[٥] = ٠.٩٦١٣٥
output[٦] = ٠.٩٦٢١٧

This order for patterns outputs values is appear as a patterns ordered in the eighth category after a after adding the seventh pattern to a category .Outputs values were changed for patterns than in a first experiment ,also changing a standard output value at this experiment .

Standard output value is;
output[١] = ٠.٩٦٤٤٦

Then reordering the outputs values to determine a closed interval for this category ,with increasing order as ;
output[٧] = ٠.٩٤٢٣٤
output[٥] = ٠.٩٦١٣٥
output[٦] = ٠.٩٦٢١٧
output[١] = ٠.٩٦٤٤٦
output[٤] = ٠.٩٦٤٦٨
output[٣] = ٠.٩٦٥١٤
output[٢] = ٠.٩٦٦٠٨

The differences values for patterns outputs values around a standard output value through order be ;

٠.٠٢٢١٢
٠.٠٠٣١١
٠.٠٠٢٢٩

٠.٠٠٠٢٢
٠.٠٠٠٦٨
٠.٠٠١٦٢

Then the real closed interval for the eighth category at this experiment ,that it determined by a patterns outputs values in this category ,through last order is ;

[٠.٩٤٢٣٤, ٠.٩٦٦٠٨]

,which is a decision region for pattern(s) recognition at this category ,that is a network be able to recognize a pattern(s) if it output(s) value(s) be in this category interval to be in a patterns of this category .

The range of a closed interval for the eighth category and at this experiment is ٠.٠٢٣٧٤ ,is refer to a capacity of this category interval at this experiment .

When a testing step performed through a testing set that also contained a fuzzy versions for a standard figure for this category, the network was able to recognize ٥ *patterns* from a testing set with ٢٠ *patterns* , for this category within a determined training interval .

Discussion of the Third Experiment :

The intervals of a training set categories at third experiment are;

First category	[٠.٠٢٠٦١ ,٠.٠٤٥٣١]
Second category	[٠.١٤٢٢٨ ,٠.١٦٠٩٩]
Third category	[٠.٢٠٩٧١ ,٠.٢١٧٠١]
Fourth category	[٠.٣٠٧٢٢ ,٠.٣١٧٨٧]
Fifth category	[٠.٤٦٠٣٦ ,٠.٤٧٨٤٤]
Sixth category	[٠.٦٠٨٦٩ ,٠.٦٢٦٧٠]
Seventh category	[٠.٧٠٥٦٨, ٠.٧٢١٧٢]
Eighth category	[٠.٩٤٢٣٤, ٠.٩٦٦٠٨]

The ranges of these intervals through this order are;

٠.٠٢٤٧٠
٠.٠١٨٧١
٠.٠٠٧٣٠
٠.٠١٠٦٥
٠.٠١٨٠٨
٠.٠١٨٠١
٠.٠١٦٠٤
٠.٠٢٣٧٤

Note that a least range(scope) is for third category-once again-interval and it is equal to ٠.٠٠٧٣٠ ,while largest range is for first category interval and it is equal to ٠.٠٢٤٧٠ .

The differences(ranges) of regions between these intervals(error regions) are for two sequential intervals at the third experiment as;

٠.٠٩٦٩٧
٠.٠٤٨٧٢
٠.٠٩٠٢١
٠.١٤٢٤٩
٠.١٣٠٢٥
٠.٠٧٨٩٨
٠.٢٢٠٦٢

The least range occurred between second and third category interval that equal to ٠.٠٤٨٧٢ .This may reflect that error region between these two categories intervals is small ,while a largest range is between a seventh and eighth category interval as ٠.٢٢٠٦٢ .It means that an error region is larger here than other regions .

Note :

١. *The output value for any pattern(actual output)do not belongs to two(or more)categories intervals in the same time at the third experiment .*

٢. *The outputs values at categories intervals were ordered increasingly ,this may because the choice of desired outputs values with increasing order (as in previous experiments),and then it also appear as sequential intervals in numbers line (if it is ordered),see figure(٥-٣٦) for a decision line at the third experiment .Also this sequence began with first category interval and then second category interval and the third and fourth ,and so on until ending with the eight category interval .*

٣. *Also, in addition to regions between these closed intervals of categories ,there are regions out over the maximum value for all intervals categories(maximum boundary of the eighth category interval in this experiment) ,and a less than minimum value for all categories intervals(minimum boundary of a first category interval in this experiment) .*

It are also an error regions but it be larger ,since it expended to infinity with the two directions ,a minus and plus infinity ,to

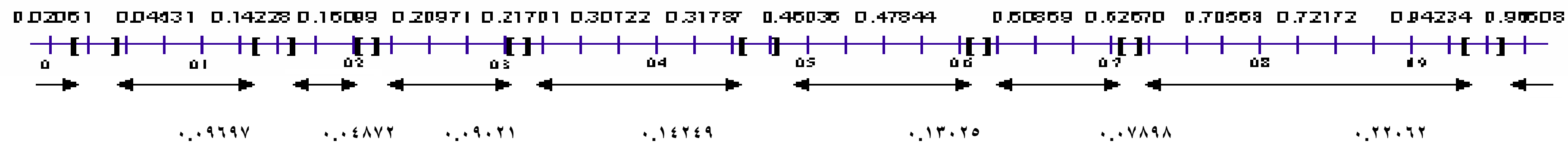
$-\infty$ *and* ∞ *,respectively .*

The resultant of testing through this experiment, was at recognize ١٠٢ *patterns* for all categories from the number of patterns ١٦٠ *patterns* that were used.

Not that the increasing in recognition ability in third experiment for most categories than a previous experiment ,so this mean improve the performance of network under conditions at this experiment .

Categories	Epoch	It .No.	MSR	Average err.	Worst err.	Eta. updated	Min. boundary	Max. boundary	Range
١st category	٥٢٩	٣٧٠٣	٠٫٠٠٠٠٧٧	٠٫٠٠٠٥٦٣	٠٫٠٣٤٤٣	٠٫٠٧٠٤٢	٠٫٠٢٠٦١	٠٫٠٤٥٣١	٠٫٠٢٤٧٠
٢n category	١٣	٩١	٠٫٠٠٠٠٠٣	٠٫٠٠٠١١٠	٠٫٠٠٠٧٧٦	٠٫٠٩٩٩٧	٠٫١٤٢٢٨	٠٫١٦٠٩٩	٠٫٠١٨٧١
٣rd category	٩٣	٦٥١	٠٫٠٠٠٠٦٢	٠٫٠٠٠٥٠٣	٠٫٠٣٥٢٦	٠٫٠٩٨٩١	٠٫٢٠٩٧١	٠٫٢١٧٠١	٠٫٠٠٧٣٠
٤th category	١٤٤	١٠٠٨	٠٫٠٠٠٠٥٥	٠٫٠٠٠٤٧٤	٠٫٠٣٣٢٢	٠٫٠٩٧٤٢	٠٫٣٠٧٢٢	٠٫٣١٧٨٧	٠٫٠١٠٦٥
٥th category	١٦٤	١١٤٨	٠٫٠٠٠٠٤٩	٠٫٠٠٠٤٥٠	٠٫٠٣١٥١	٠٫٠٩٦٦٧	٠٫٤٦٠٣٦	٠٫٤٧٨٤٤	٠٫٠١٨٠٨
٦th category	١٥١	١٠٥٧	٠٫٠٠٠٠٤٥	٠٫٠٠٠٤٢٩	٠٫٠٣٠٠٧	٠٫٠٩٧١٧	٠٫٦٠٨٦٩	٠٫٦٢٦٧٠	٠٫٠١٨٠١
٧th category	١٥٨	١١٠٦	٠٫٠٠٠٠٤٦	٠٫٠٠٠٤٣٥	٠٫٠٣٠٤٨	٠٫٠٩٦٩٠	٠٫٧٠٥٦٨	٠٫٧٢١٧٢	٠٫٠١٦٠٤
٨th category	٣٠٦	٢١٤٢	٠٫٠٠٠٠٧٢	٠٫٠٠٠٥٤٢	٠٫٠٣٧٩٦	٠٫٠٨٨٩١	٠٫٩٤٢٣٤	٠٫٩٦٦٠٨	٠٫٠٢٣٧٤

Table(5-5): Conclusions values for the eight categories from the third experiment



Figure(5-36): Decision Line for the third experiment

٥.٩ *Future Works*

I wish to this work to be from a literature reference for more and more incoming works ,as it was builds on a previous works ,so here would available several ideas and suggested schemes to any one want to work on the same subject .These ideas extracted during the time of research and the hard problems which confront perform this work, to be as a researches subjects in future .

- Perform the suggested method with functions depends on time as depended variable ,which also exploit continuous wavelets, on movable images such video images .
- Use a summation of more than two functions in the activation function .
- Use more than one wavelet function with one or more from membership functions
- Use an other continuous types of wavelets or membership functions .
- Use non linear functions as a relation between a wavelet(s) function(s) and membership function(s) .
- Sure ,try to implement the suggested method on a more complex images (not geometric curves) ,and/or with different colors .
- Perform the suggested method on colored images ,that is colored backgrounds with different colors and different coloring ways at coloring the closed area that determined by the curves in images at training set .

References

[١] Babuška R. ,"***Fuzzy Systems ,Modeling and Identification*** "
citeseer.ist.psu.edu/٣٥٥١٧٨.html

[٢] Jain L.C.,Jain R.K.,"***Hybrid Intelligent Engineering Systems***" ,world scientific publishing Co. pte .Ltd. ,in Singapore by Uto.print , (١٩٩٧)

[٣] العبيدي ,صلاح مهدي ,"نظام هجين باستعمال تحويل المويجة في الشبكات العصبية لتصنيف الصور" ,أطروحة ماجستير ,جامعة بابل ,٢٠٠٣

[٤] Newsletter of Coil-The Computational Intelligence and Learning Cluster " ***synergy*** " ,Issue ٣ ,Autumn (٢٠٠٠) , http:// www.dcs.napier.ac.uk/coil/ .

[٥] Klerfors D. ,"***Artificial Neural Networks*** ", http://hem.hj.sc/~de٩٦klda/NeuralNetworks.htm

[٦] Zurada J. M. ," ***Introduction to Artificial Neural Systems*** ", JAICO publishing House ,Bombay ,١٩٩٧, pgs.(٣١-٥٩) (١٧٥-١٩٠)(١٩٥-٢٢١).

[٧] OSSAR Y.,DREYFUS G. ,"***Initialization by Selection for Wavelet Network Training*** " , Neurocomputing ,vol. ٣٤ , pp.١٣١-١٤٣ ,(٢٠٠٠) .
www.neurones.espci.fr/Articles_PS/initialisation.**pdf**

[٨] Thuillard M. ,"***A REVIEW OF WAVELET NETWORKS, WAVENETS, FUZZY WAVENETS AND THEIR APPLICATIONS***",ESIT ٢٠٠٠, ١٤-١٥ September ٢٠٠٠, Aachen, Germany
citeseer.ist.psu.edu/thuillard٠٠review.html

[٩] Thuillard M. , " ***FUZZY LOGIC IN THE WAVELET FRAMEWORK*** ",Proc.Toolmet'٢٠٠٠ __Tool Environments and Development Methods for Intelligent System ,(٢٠٠٠)
www.hallinto.oulu.fi/viestin/t-pisto/arkisto٢٠٠٠/٨-٠٠/toolmet.htm

[١٠] Knapp R. B.," ***Fuzzy Sets and Pattern Recognition*** ,(١٩٩٨)
hci.sapp.org/lectures/knapp/**fuzzy/fuzzy.pdf**

[١١] Leondes C. T. ,"***Fuzzy Logic and Expert Systems Applications*** ",ACADEMIC PRESS,١٩٩٨,pgs.(١-٥)(٤٠-٥١)(٦٠-٧٧)

[١٢] Kasabov N. ,Kozma R. , Duch W.,
"***Rule Extraction from Linguistic Rule Networks and from Fuzzy Neural Networks: Propositional versus Fuzzy Rules*** "
http://www.phys.uni.torun.pl/~duch

[١٣] Mao Z. H. , Li Y.D. ,Sun Z.R. ," ***Comments on 'Fuzzy systems approximation by frames-SISO case'*** ",Elsevier Science B.V., Fuzzy Sets and Systems ١٠٠. ٣٧٣-٣٧٥ ,(١٩٩٨)
citeseer.ist.psu.edu/٣٥٥١٧٨.html

[١٤] Reyneri L.M. ," ***Unification of Neural and Wavelet Networks and Fuzzy Systems*** " .
http://polimage.polito.it/ l˜mr

[١٥] LI X., YAO Y. ,YUAN Z. ," ***On-line tool condition monitoring system with wavelet fuzzy neural network***",
Journal of Intelligent Manufacturing ٨, ٢٧١-٢٧٦ , (١٩٩٧)

[١٦] Yao Y., Li X., Yuan Z. ," ***Tool wear detection with fuzzy classification and wavelet fuzzy neural network***",
International Journal of Machine Tools & Manufacture ٣٩, ١٥٢٥–١٥٣٨, (١٩٩٩)

[١٧] Woodford B. J.,Kasabov N. K.,"***A Wavelet-Based Neural Network Classi£er for Temporal Data*** ",Australasia-Japan Joint Workshop ,University of Otago Dunedin New Zealand ,(٢٠٠١)

[١٨] Reznik L. , "***Fuzzy Controllers*** " ,Great Britain by Biddles Ltd,Guildford and King's Lynn ,١٩٩٧,pgs.(١-٥٧)(٦٣-٧٢)(١٢٤-١٥٢)

[١٩] Burges C. J.C. ,Simard P. Y. ,Malvar H. S.
,"***Improving Wavelet Images Compression with Neural Networks***"
,(٢٠٠١). http:// www.research.microsoft .com

[٢٠] Duch W. ,"***Uncertainty of data, fuzzy membership functions, and multi-layer perceptrons*** ",IEEE TRANSACTION

ON NEURAL NETWORKS, VOL. XX, NO. YY, (٢٠٠٤)

[٢١] Soto R. ,“ ***Uncertainty Systems Ia-٤٠٠٢ Fuzzifiers and Defuzzifiers***”,Centro de Sistemas Inteligentes ITESM, Campus Monterrey , (٢٠٠٣)
http://www-csi.mty.itesm.mx/~rsoto

[٢٢]Embrechts M. J.,”***DSES-٤٨١٠-٠١Intro to COMPUTATIONAL INTELLIGENCE & SOFT COMPUTING*** “,COURSE ANNOUNCEMENT --Spring ٢٠٠٣ .

[٢٣] Rao V. B. ,Rao H. V. ,” ***C^{++} Neural Networks and Fuzzy Logic***“ , Management Information Source,Inc. A subsidiary of Henry Holt and Company ,Inc,١٩٩٣, pgs. (٢٥-٣١)(١٠٣-١٠٦).

[٢٤] S. A.Pandya ,B.R.Macy,” ***Pattern Recognition with Neural Networks in C^{++}*** “,CRC Press,USA,١٩٩٦ ,pgs.(٨٧-١٠٧)

[٢٥] Pati Y. C., Krishnaparasad P.S., “***Analysis and Synthesis of Feedforward Neural Networks Using Discrete Affine Wavelet Transformations*** “,IEEE Trans. on Neural Networks ٤ (١) (١٩٩٣) ٧٣-٨٥., TR ٩٠-٤٤r١,systems research center,NSFD CD ٨٨٠٣٠١٢ .

[٢٦] KOBAYASHI K. ,TORIOKA T.,” ***A WAVELET NEURAL NETWORK FOR FUNCTION APPROXIMATION AND NETWORK OPTIMIZATION***“ ,Proceeding of ANNIE’٩٤ ,vol. ٤ ,
neuron-ai.tuke.sk/~hudecm/d/ANNIE٩٤.**pdf**

[٢٧] LeCunY. ,D. Jackel L. ,Boser B. ,S. Denker J. ,P. Graf H. , Guyon I. ,Henderson D. ,E. Howerd R. ,Hubbard W. , ” ***Handwritten Digits Recognition: Applications of Neural Network Chips and Automatic Learning*** “,IEEE Communications Magazine (١٩٨٩)

[٢٨] Srinivasa N. ,Ahuja N. , “***A Topological and Temporal Correlator Network for Spatiotemporal Pattern Learning, Recognition, and Recall*** ”,IEEE TRANSACTIONS ON NEURAL NETWORKS . vol. ١٠. No. ٢. March ,(١٩٩٩)

[٢٩] Echauz J. ,Vachtsevanos G., “***Radial Wavelet Neural Networks for Electroencephalographic Drug Detection***”, in Proc.٥th International Conference on Advances in Communication

& Control.Chania,Greece,(١٩٩٥).
[٣٠] Lekuati G. ," ***Adaptive Self –Tuning Neuro wavelet Network Controllers*** " ,Virginia Polytechnic Institute and State University, (١٩٩٧) ,
http://scholar .lib.vt.edu/theses/available/ etd-٥٤٥٠٢٤٣٩٧٤١١٣١ /unrestricted/ETD.pdf

[٣١] Wang P. ,Vachtsevanos G." ***Fault Prognosis Using Dynamic Wavelet Neural Networks*** ",Georgia Tech , AAAI Spring Symposium,Technical Report SS-٩٩-٠٤, ISBN ١-٥٧٧٣٥-٠٨١-٢, (١٩٩٩).

[٣٢] STARCK J.-L.,MURTAGH F.,BIJAOUI A.," ***Image Processing and Data Analysis***", Cambridge University Press,١٩٩٨,pgs.(٣-١٧)

[٣٣] Tun Li .Sheng,Ching C. Shu ,"***Function Approximation using Robust Wavelet Neural Networks*** ",NSC٨٩-٢٢١٣-e-٣٢٧-٠٠٧
www.cs.fiu.edu/~chens/PDF/ICTAI٠٢_WNN.pdf - ١٠٤k -

[٣٤] Oussar Y., Rivals I. , Personnaz L. , Dreyfus G.,
" ***Training Wavelet Networks for Nonlinear Dynamic Input-Output Modeling*****",** Neurocomputing, in press.
neurones.espci.fr/Francais.Docs/.../Articles_PS/Ondelettes.pdf

[٣٥] LUO FA-LONG,UNBEHAUEN R., " ***Applied Neural Networks for Singal Processing*** ",CAMBRIDGE UNIVERSITY Press ,١٩٩٧,pgs.(٣٥٥-٣٦٣).

[٣٦] Misiti M. , Misiti Y. , Oppenheim. Georges, Poggi. Jean-Michel ," ***Wavelet Toolbox***" ,User's Guide .Version ٢, For Use with MATLAB, COPYRIGHT ١٩٩٧ - ٢٠٠٢ by The MathWorks, Inc,
http://www.mathworks.com

[٣٧]For Use with MATLAB,"***Fuzzy Logic Toolbox***", User's Guide Version ٢, ١٩٩٥ - ٢٠٠٢ by The MathWorks, Inc.
http://www.mathworks.com

[٣٨] Noel Steven E. ,Szu Harold H. ,"***Proximity sensing with wavelet generated video***", Journal of Electronic Imaging,٧(٤),October (١٩٩٨)
http://www.isse.gmu.edu/~snoel/publications.htm

[٣٩] Noel Steven E. ,Szu. Harold H. ,Gohel Yogesh J., "***Doppler Frequency Estimation with Wavelets and Neural Networks***", in Proceedings of the ١٢th Annual International Symposium on Aerospace/Defense Sensing, Simulation, and Controls, Orlando, Florida, April (١٩٩٨) http://www.isse.gmu.edu/~snoel/publications.htm

[٤٠] القرغولي," سهاد احمد ,"**ضغط الصور الملونة باستخدام التحويل المويجي**" , أطروحة ماجستير, جامعة بابل, ٢٠٠٢

[٤١] Kaarna A. ," ***APPLICATION OF NEURAL NETWORKS TO WAVELET FILTERS SELECTION IN MULTISPECTRAL IMAGE COMPRESSION*** ",Proceedings of the Nordic Signal Processing Symposium, NORSIG ٢٠٠٠, Kolmården, Sweden, June ١٣ - June ١٥, ٢٠٠٠, pp. ٢٥-٢٨.

[٤٢] Thuillard M. ," ***New Perspectives for the Integration of Wavelet Theory into Soft Computing***", www.erudit.de/erudit/committees/Publications.htm

[٤٣] Bernard C. ,Mallat S. ,Slotine Jean-J. , "***Wavelet Interpolation Networks***", www.cmap.polytechnique.fr/~bernard/Publications/esann٩٨.**pdf**

[٤٤] Lewis James M. ,Burrus Sidney, "***APPROXIMATE CONTINUOUS WAVELET TRANSFORM WITH AN APPLICATION TO NOISE REDUCTION*** ", www-dsp.rice.edu/publicationsold/pub/lewi-icassp٩٨.pdf

[٤٥] Hellmann M. ," ***Fuzzy Logic Introduction*** " , www.epsilon.nought.de/tutorials/**fuzzy**/index.php

[٤٦] Robert F., "***Neural Fuzzy Systems*** ",Åbo Akademis tryckeri, Åbo,(١٩٩٥), ٢٤٩ pages.ISSN ٠٣٥٨-٥٦٥٤, ISBN ٩٥١-٦٥٠-٦٢٤-٠ http://www.abo.fi/~rfuller/nfs١.pdf

[٤٧] سد خان. ستار بدر , صالح .منى هادي , " **تقويم قابلية تطبيق المنطق المضبب**" ,مجلة جامعة بابل والعلوم الصرفة و التطبيقية, مجلد ٧, العدد ٣, ٢٠٠٢

[٤٨] Bauer P. , Nouak S., Winkler R. ," ***A brief course in Fuzzy Logic and Fuzzy Control*** ",version ١.٢ ,(١٩٩٦),the FTP server , ftp://ftp.flll.uni-linz.ac.at/pub/info/.

[٤٩] Bilgiç T. ,Türksen I. Burnhan,"***Measurements-Theoretic Frameworks for Fuzzy Set Theory*** ",M٥S ٢G٩ Canada ,Appeardin Trevor P. Martin and A.R,(١٩٩٥)
http://citeseer.ist.psu.edu/٣٠٣٢٤٣.html

[٥٠] Chen.C.H. ***"FUZZY LOGIC AND NEURAL NETWORK HANDBOOK "*** , McGraw-Hill ١٩٩٦ ,pgs.(٢.١-٢.٢٧) .

[٥١] Garibaldi Jonathan M., John Robert I.," ***Choosing Membership Functions of Linguistic Terms*** ",
www.cs.nott.ac.uk/~jmg/papers/fuzzieee-٠٣.pdf

[٥٢] جبار, نهلة إبراهيم," **معالجة الصور الطبية لتشخيص بعض الأمراض السرطانية**" ,
أطروحة ماجستير, جامعة بابل, ٢٠٠١.

[٥٣] For Use with MATLAB®, "***MATHLAB version 6.0***"
COPYRIGHT
١٩٩٢ - ٢٠٠٢ by The MathWorks, Inc.
http://www.mathworks.com

[٥٤] Odegard Jan E. , Burrus C. Sidney , "***SMOOTH BIORTHOGONAL WAVELETS FOR APPLICATIONS IN IMAGE COMPRESSION*** ",Proc. DSP Workshop, Loen, Norway,(١٩٩٦)
http://www-dsp.rice.edu

[٥٥] Wanlu J. , Shuqing Z. , Yiqun W. ,
"APPLYING MULTIRESOLUTION ANALYSIS FOR PROCESSING OF HYDRAULIC PUMP FAULT SIGNAL"
fluid.power.net/techbriefs/hanghzau/٥_٢_١.pdf