

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



The Artificial Intelligence Approach for Diagnosis, Treatment and Modelling in Orthodontic

Kazem Bahaa¹, Garma Noor² and Yousif Yousif³

¹*Mechatronics Eng. Dept., Baghdad University,*

²*Orthodontic Dept., Dentistry College, Baghdad University,*

³*Ministry of Higher Education and Scientific Research/ R&D Department,
Baghdad
Iraq*

1. Introduction

The discipline, science, and art of orthodontics are concerned with the face and ability to modify its growth. Orthodontists achieve their goals by manipulating the craniofacial skeleton, with particular emphasis on modifying the dentoalveolar region, external orthopedic forces are applied that mirror some techniques used in medical orthopedics. Most treatments, however, focus on modifying the occlusion and controlling dentoalveolar development and abnormal facial growth, thus, enormous amounts of designs and techniques invented in the diagnostic and treatment domains aiming at boolean etiological identification and optimized strategies of solution delivered. A valid problem assessment enables health providers to determine treatment need and priority, and as health care moves toward more stringent financial accountability. the inventory of the computer and its implementation in different medical field was of great interest, this interest are even greater with the artificial intelligence introduction (AI).

The best definition for the phrase “AI” calls for formalization of the term “intelligence”. Psychologist and cognitive theorists are of the opinion that intelligence helps in identifying the right piece of knowledge at the appropriate instances of decision making [1,2]. The phrase “AI” thus can be defined as the simulation of human intelligence on a machine. Thus, AI alternatively may be stated as a subject dealing with computational models that can think and act rationally [3-7].

The subject of AI spans a wide horizon. It deals with the various kinds of knowledge representation schemes, different techniques of intelligent search, various methods for resolving uncertainty of data and knowledge, different schemes for automated machine learning and many others. Among the application areas of AI, we have Expert systems, Game-playing, and Theorem-proving, Natural language processing, Image recognition, Robotics and many others. This chapter aims at bringing the insight of interest to the conjugation relatively recently happened between orthodontics discipline and AI subject.

2. Introduction to AI

The subject of AI was originated with game-playing and theorem-proving programs and was gradually progressed with theories from a number of parent disciplines. As a young discipline of science, the significance of the topics covered under the subject changes considerably with time. The subject of AI has been enriched with a wide discipline of knowledge from Philosophy, Psychology, Cognitive Science, Computer Science, Mathematics and Engineering. Thus in fig.1, they have been referred to as the parent disciplines of AI.

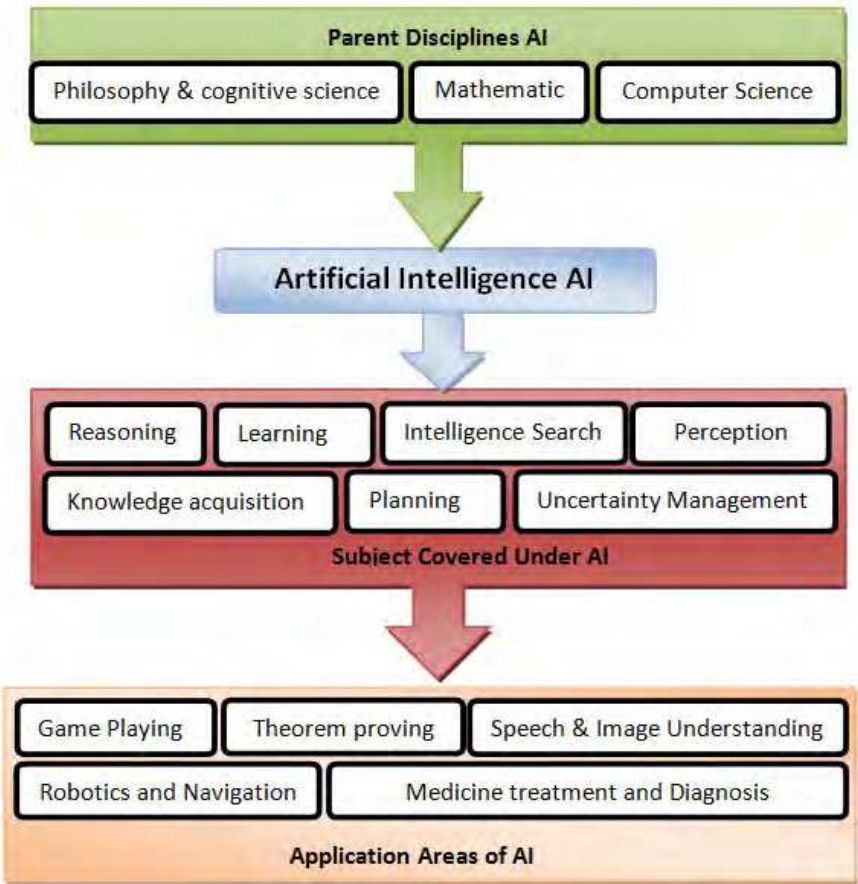


Fig. 1. Parent disciplines of AI

2.1 Artificial neural nets

Neural networks are composed of simple elements operating in parallel. These elements are inspired by biological nervous systems. As in nature, the network function is determined largely by the connections between elements. You can train a neural network to perform a particular function by adjusting the values of the connections (weights) between elements , fig. 2. Commonly neural networks are adjusted, or trained, so that a particular input leads to a specific target output. Such a situation can be shown as follows: there, the network is adjusted, based on a comparison of the output and the target, until the network output matches the target. Typically many such input/target pairs are needed to train a network [7,8]. One type of network sees the nodes as ‘artificial neurons’. These are called Artificial Neural Networks (ANNs). Natural neurons receive signals through synapses located on the dendrites or membrane of the neuron. When the signals received are strong enough (surpass

a certain threshold), the neuron is activated and emits a signal through the axon. This signal might be sent to another synapse, and might activate other neurons.

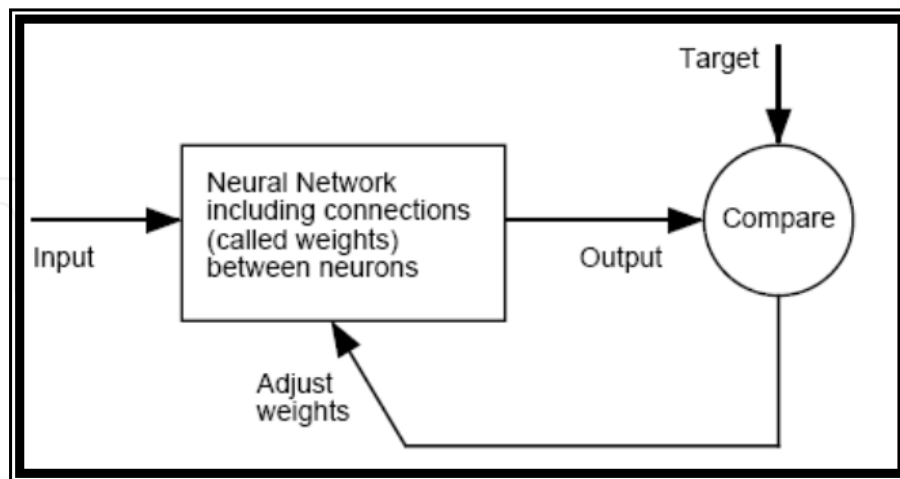


Fig. 2. Basic operation of ANN [9]

An interconnected assembly of simple processing elements, units or nodes, whose functionality is loosely based on the animal brain. The processing ability of the network is stored in the inter unit connection strengths, weights, obtained by a process of adaptation to, or learning from, a set of training patterns.

The benefits of using the neural network can be summarized as follows [10]:

1. **Nonlinearity:** an artificial neuron can be linear or non-linear a neural net made up of interconnection of non-linear neurons, is itself non-linear, note that even linear function could be modeled by non-linear neurons, while the inverse can't be done.
2. **Input and output mapping:** usual learning process of neural network carried out in a popular paradigm of learning called learning with teacher "supervised learning" here modification of synaptic weights of a neural network done by applying set of labeled training samples, each sample consist of a unique input signal and a corresponding desired response. The previous samples could be arranged in different manners so the network constructing an input output mapping for the problem.
3. **Adaptively:** neural networks have a built in capability to adapt their synaptic to change in the surrounding environment. This could be done by retraining of the model or make the network changes its synaptic weights in real time and this will be useful for pattern classification, signal processing, and control application.
4. **Fault tolerance:** a neural network, implemented in hard ware form, has the potential to be inherently fault tolerance, or capable of robust control. For example if a neuron or its connecting links are damaged and due to the nature of distributed information in neural network, this damage little effect on network response.

Neural Network Architecture: A neuron is an information-processing unit that is fundamental to the operation of a neural network. The block diagram of fig. 3 shows the model of a neuron, which forms the basis for designing (artificial) neural network.

The neuronal model of fig. (3) also includes an externally applied bias, denoted by (bk). The bias (bk) has the effect of increasing or lowering the net put of the activation function, depending on whether it is positive or negative, respectively [11], fig 4 shows common types of activation functions.

In mathematical terms:

$$f(x)=\left(\sum_{j=1}^nw_{jk}x_j+b_k\right)$$

(1)

Where $f(w,x,b)$ is the activation function.

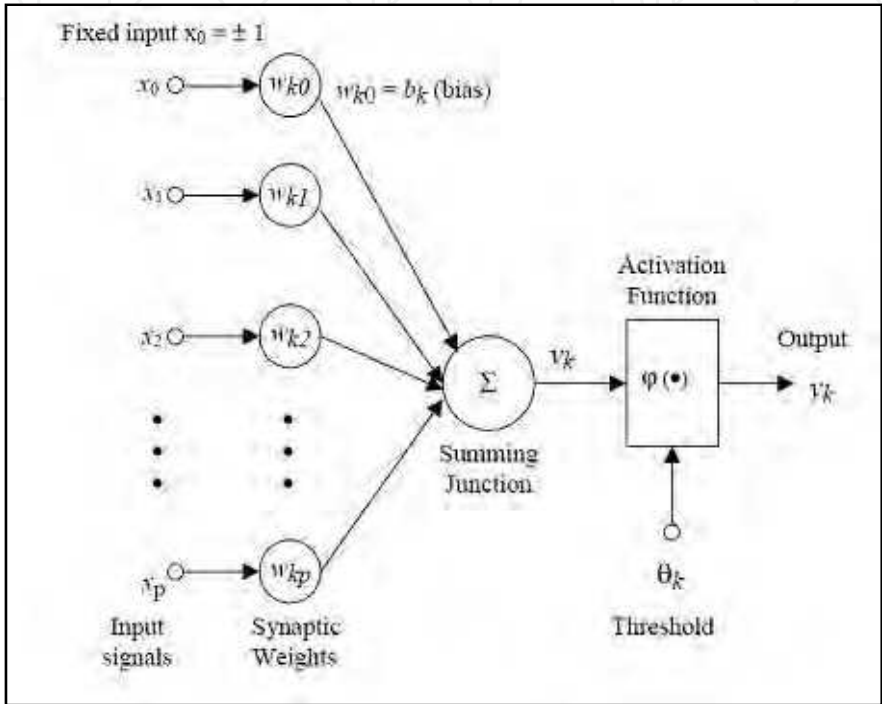


Fig. 3. Nonlinear model of a neuron

In general , there are four basic types of activation functions:

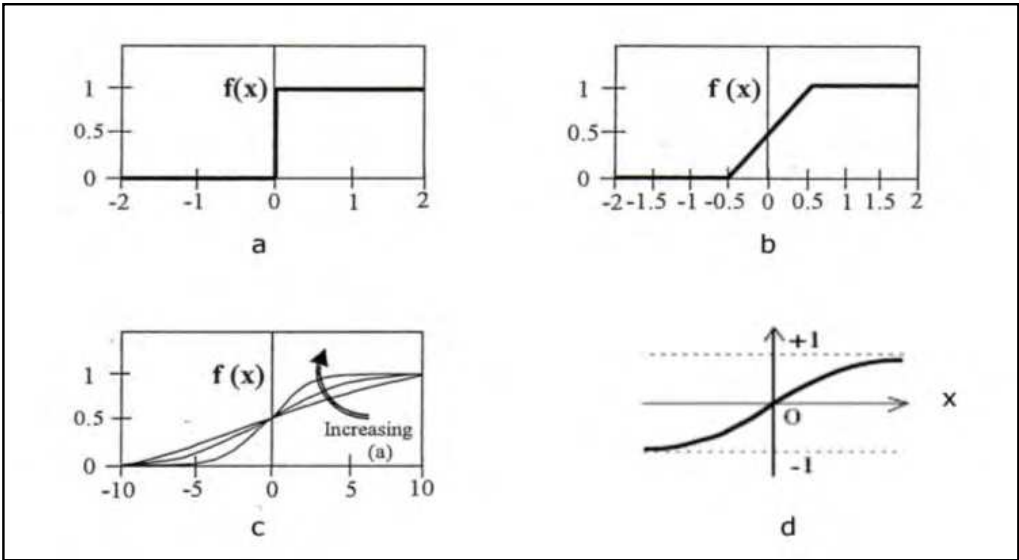


Fig. 4. Activation Functions

The management of neurons into layers and the connection patterns within and between layers is called the net architecture. The manner in which the neurons of a neural network are structured is intimately linked with the learning algorithms used to train the network [11]. Fig (5) shows different closes of network architecture

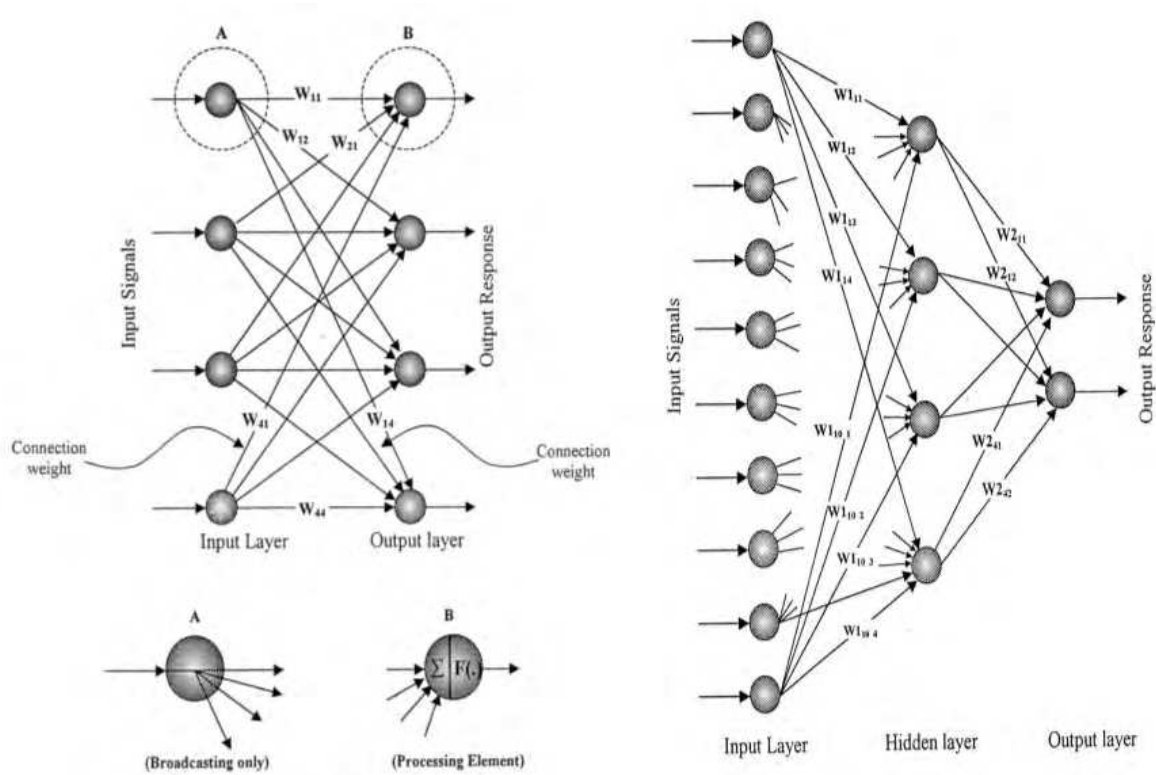


Fig. 5.a. Single-Layer Feedforward Networks b. Multilayer Feedforward Networks

2.2 Genetic algorithms

A genetic algorithm (GA) is a search heuristic that mimics the process of natural evolution. This heuristic is routinely used to generate useful solutions to optimization and search problems. Genetic algorithms belong to the larger class of evolutionary algorithms (EA), which generate solutions to optimization problems using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In a genetic algorithm, a population of strings (called chromosomes or the genotype of the genome), which encode candidate solutions (called individuals, creatures, or phenotypes) to an optimization problem, evolves toward better solutions. Traditionally, solutions are represented in binary as strings of 0s and 1s, but other encodings are also possible. The evolution usually starts from a population of randomly generated individuals and happens in generations. In each generation, the fitness of every individual in the population is evaluated, multiple individuals are stochastically selected from the current population (based on their fitness), and modified (recombined and possibly randomly mutated) to form a new population. The new population is then used in the next iteration of the algorithm. Commonly, the algorithm terminates when either a maximum number of generations has been produced, or a satisfactory fitness level has been reached for the population. If the algorithm has terminated due to a maximum number of generations, a satisfactory solution may or may not have been reached.

Genetic algorithms find application in bioinformatics, phylogenetics, computational science, engineering, economics, chemistry, manufacturing, mathematics, physics and other fields.

A typical genetic algorithm requires:

- a genetic representation of the solution domain,
- a fitness function to evaluate the solution domain.

A standard representation of the solution is as an array of bits. Arrays of other types and structures can be used in essentially the same way. The main property that makes these genetic representations convenient is that their parts are easily aligned due to their fixed size, which facilitates simple crossover operations. Variable length representations may also be used, but crossover implementation is more complex in this case. Tree-like representations are explored in genetic programming and graph-form representations are explored in evolutionary programming.[12]

The fitness function is defined over the genetic representation and measures the quality of the represented solution. The fitness function is always problem dependent. For instance, in the knapsack problem one wants to maximize the total value of objects that can be put in a knapsack of some fixed capacity. A representation of a solution might be an array of bits, where each bit represents a different object, and the value of the bit (0 or 1) represents whether or not the object is in the knapsack. Not every such representation is valid, as the size of objects may exceed the capacity of the knapsack. The fitness of the solution is the sum of values of all objects in the knapsack if the representation is valid, or 0 otherwise. In some problems, it is hard or even impossible to define the fitness expression; in these cases, interactive genetic algorithms are used.

Once we have the genetic representation and the fitness function defined, GA proceeds to initialize a population of solutions randomly, then improve it through repetitive application of mutation, crossover, inversion and selection operators as shown in fig 6.

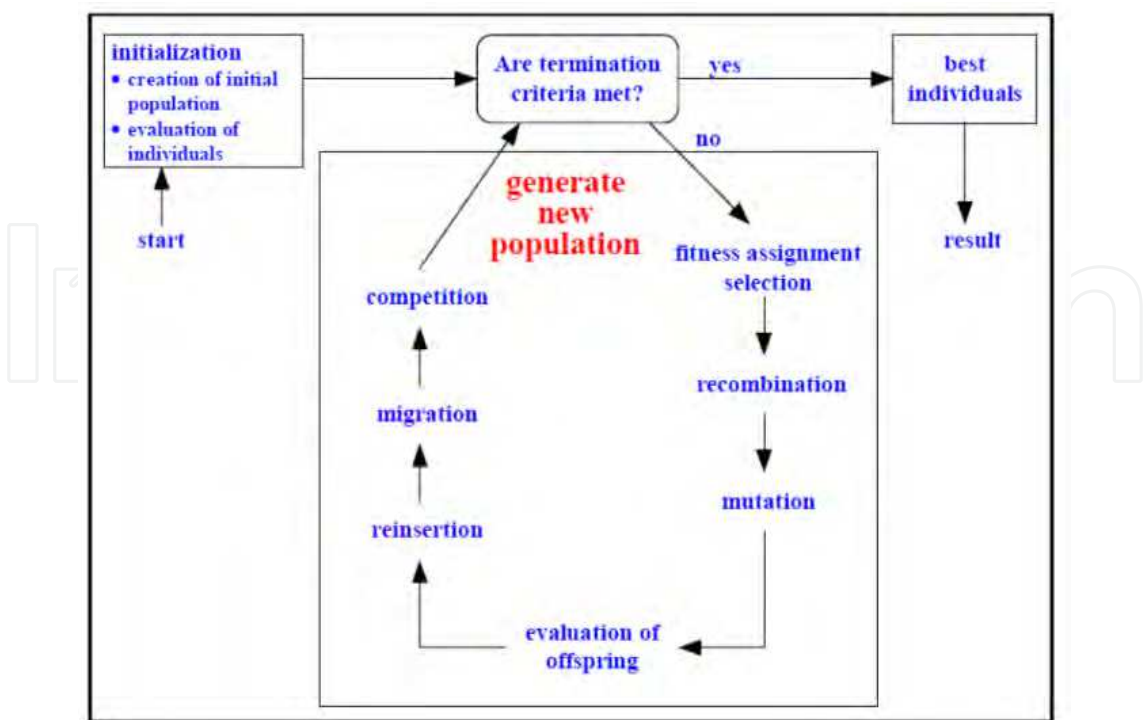


Fig. 6. Structure of an extended multi-population evolutionary algorithm

2.3 Fuzzy logic

Fuzzy logic [13] deals with fuzzy sets and logical connectives for modeling the human-like reasoning problems of the real world. A fuzzy set, unlike conventional sets, includes all elements of the universal set of the domain but with varying membership values in the interval $[0,1]$. It may be noted that a conventional set contains its members with a value of membership equal to one and disregards other elements of the universal set, for they have zero membership.

Fuzzy Sets and Crisp sets: The very basic notion of fuzzy systems is a fuzzy (sub)set. In classical mathematics we are familiar with what we call crisp sets. For example, the possible interferometric coherence g values are the set X of all real numbers between 0 and 1. From this set X a subset A can be defined, (e.g. all values $0 \leq g \leq 0.2$). The characteristic function of A , (i.e. this function assigns a number 1 or 0 to each element in X , depending on whether the element is in the subset A or not) is shown in fig7.[14]

The elements which have been assigned the number 1 can be interpreted as the elements that are in the set A and the elements which have assigned the number 0 as the elements that are not in the set A .

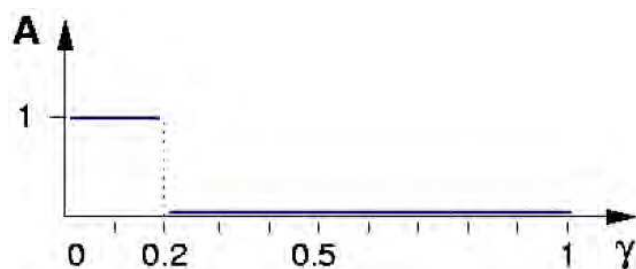


Fig. 7. Characteristic Function of a Crisp Set

This concept is sufficient for many areas of applications, but it can easily be seen, that it lacks in flexibility for some applications like classification of remotely sensed data analysis. For example it is well known that water shows low interferometric coherence g in SAR images. Since g starts at 0, the lower range of this set ought to be clear. The upper range, on the other hand, is rather hard to define. As a first attempt, we set the upper range to 0.2. Therefore we get B as a crisp interval $B=[0,0.2]$. But this means that a g value of 0.20 is low but a g value of 0.21 not. Obviously, this is a structural problem, for if we moved the upper boundary of the range from $g=0.20$ to an arbitrary point we can pose the same question. A more natural way to construct the set B would be to relax the strict separation between low and not low. This can be done by allowing not only the crisp decision Yes/No, but more flexible rules like "fairly low". A fuzzy set allows us to define such a notion. The aim is to use fuzzy sets in order to make computers more 'intelligent', therefore, the idea above has to be coded more formally. In the example, all the elements were coded with 0 or 1. A straight way to generalize this concept, is to allow more values between 0 and 1. In fact, infinitely many alternatives can be allowed between the boundaries 0 and 1, namely the unit interval $I = [0, 1]$.

The interpretation of the numbers, now assigned to all elements is much more difficult. Of course, again the number 1 assigned to an element means, that the element is in the set B and 0 means that the element is definitely not in the set B . All other values mean a gradual membership to the set B .

This is shown in Fig. 8. The membership function is a graphical representation of the magnitude of participation of each input. It associates a weighting with each of the inputs that are processed, define functional overlap between inputs, and ultimately determines an output response. The rules use the input membership values as weighting factors to determine their influence on the fuzzy output sets of the final output conclusion.

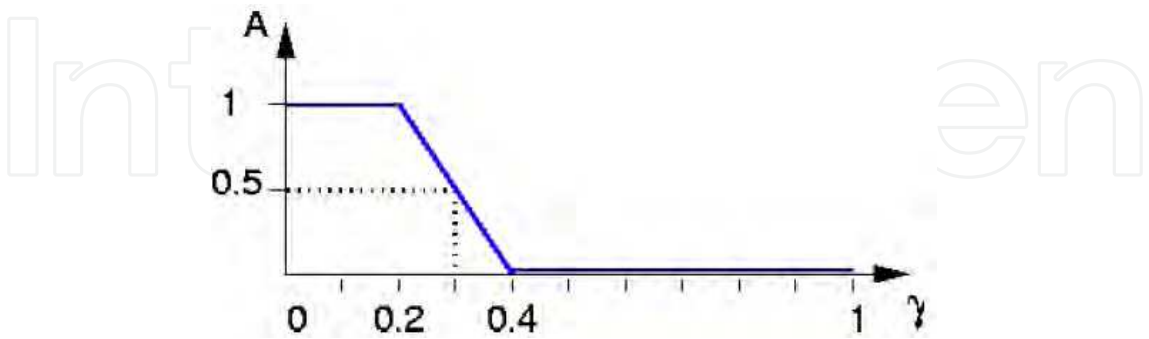


Fig. 8. Characteristic Function of a Fuzzy Set

Operations on fuzzy sets: We can introduce basic operations on fuzzy sets. Similar to the operations on crisp sets we also want to intersect, unify and negate fuzzy sets. In his very first paper about fuzzy sets [14], L. A. Zadeh suggested the minimum operator for the intersection and the maximum operator for the union of two fuzzy sets. It can be shown that these operators coincide with the crisp unification, and intersection if we only consider the membership degrees 0 and 1. For example, if A is a fuzzy interval between 5 and 8 and B be a fuzzy number about 4 as shown in the fig. 9.

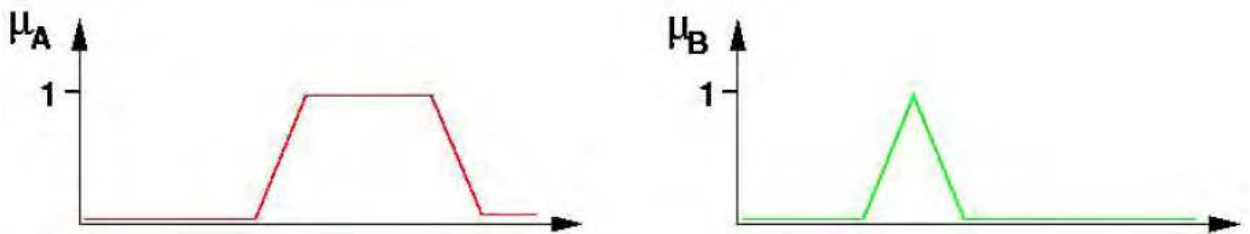


Fig. 9. Example fuzzy sets

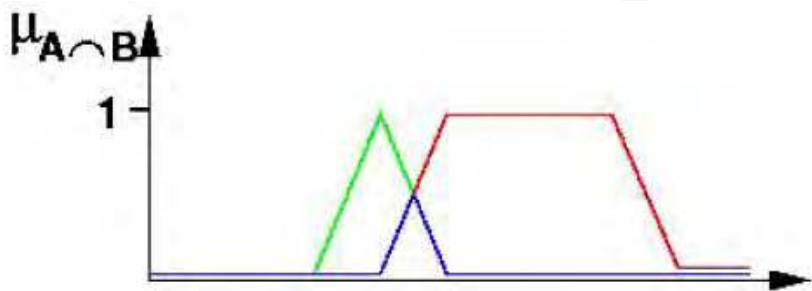


Fig. 10. Example: Fuzzy AND

In this case, the fuzzy set between 5 and 8 AND about 4 is set between 5 and 8 OR about 4 is shown in fig 11, and the the NEGATION of the fuzzy set A is shown in fig 12.

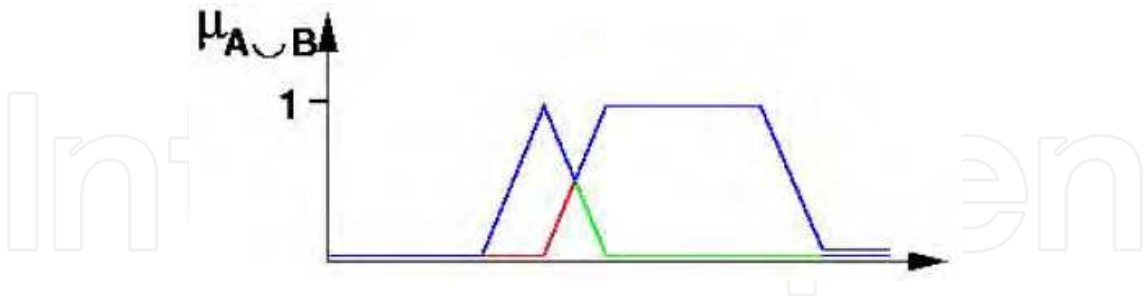


Fig. 11. Example: Fuzzy OR

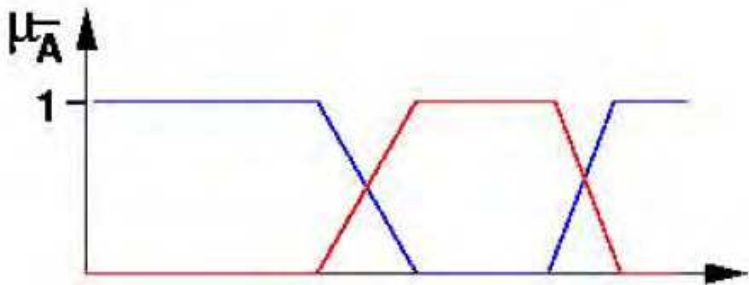


Fig. 12. Example: Fuzzy NEGATION

Fuzzy Classification: Fuzzy classifiers are one application of fuzzy theory. Expert knowledge is used and can be expressed in a very natural way using linguistic variables , which are described by fuzzy sets Now the expert knowledge for this variables can be formulated as a rules like

IF feature A low AND feature B medium AND feature C medium AND feature D medium THEN Class = class4

The rules can be combined in a table calls rule base [14]

R#	Feature A	Feature B	Feature C	Feature D	Class
1:	Low	Medium	Medium	Medium	Class 1
2:	Medium	high	Medium	Medium	Class 2
3:	Low	Medium	Medium	Medium	Class 3
4:	low	Medium	Medium	Medium	Class 1
5:	Medium	Medium	Medium	Medium	Class 4
...:
N:	Low	High	Medium	Low	Unknown

Linguistic rules describing the control system consist of two parts; an antecedent block (between the IF and THEN) and a consequent block (following THEN). Depending on the system, it may not be necessary to evaluate every possible input combination, since some may rarely or never occur.[14].

By making this type of evaluation, usually done by an experienced operator, fewer rules can be evaluated, thus simplifying the processing logic and perhaps even improving the fuzzy logic system performance. The inputs are combined logically using the AND operator to produce output response values for all expected inputs. The active conclusions are then combined into a logical sum for each membership function. A firing strength for each output membership function is computed. All that remains is to combine these logical sums in a defuzzification process to produce the crisp output. e.g for a for the rule consequents for each class a so-called singleton or a min-max interference can be derived which is the characteristic function of the respective set . e.g. For the input pair of $H = 0.35$ and $L = 30$ the scheme below (see Fig 14.) would apply.

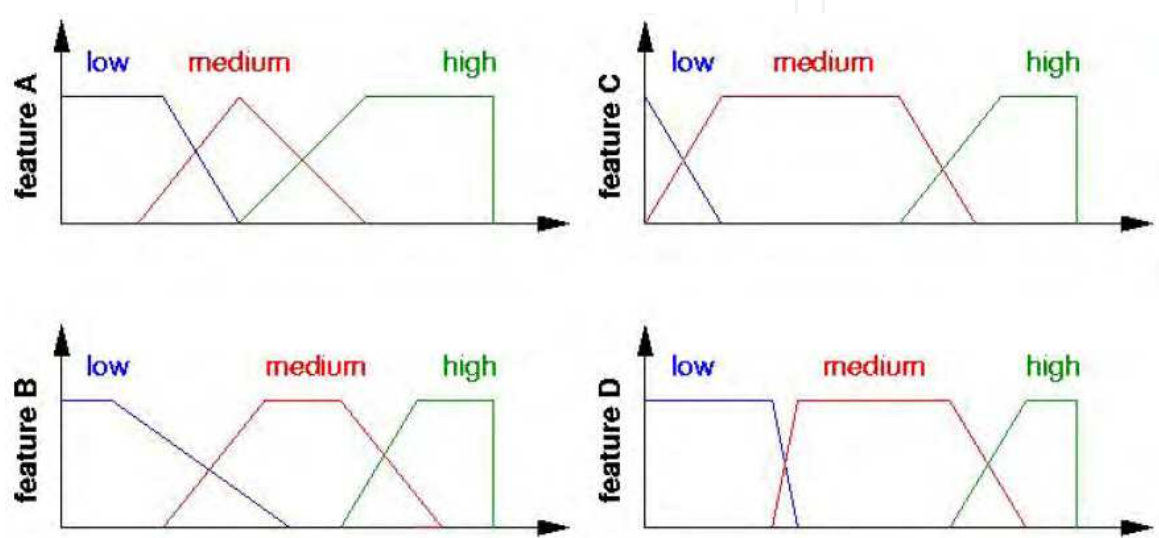


Fig. 13. Example: Linguistic Variables

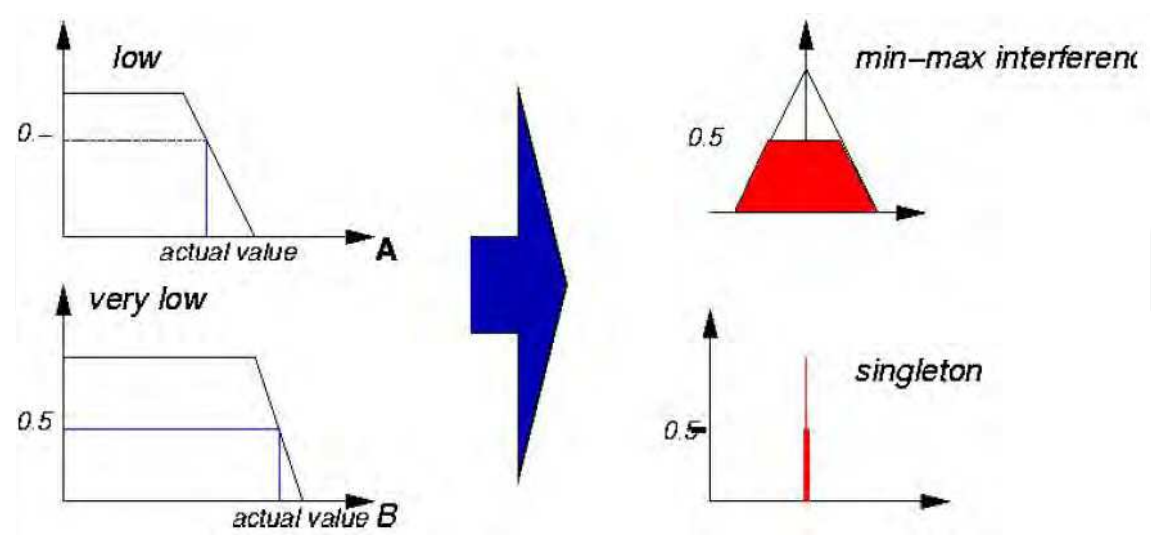


Fig. 14. Interference for rule IF H very low AND a low THEN Class = class 1

The fuzzy outputs for all rules are finally aggregated to one fuzzy set. To obtain a crisp decision from this fuzzy output, we have to defuzzify the fuzzy set, or the set of singletons. Therefore, we have to choose one representative value as the final output. There are several

heuristic methods (defuzzification methods), one of them is e.g. to take the center of gravity of the fuzzy set as shown in fig 15., which is widely used for fuzzy sets. For the discrete case with singletons usually the maximum-method is used where the point with the maximum singleton is chosen.

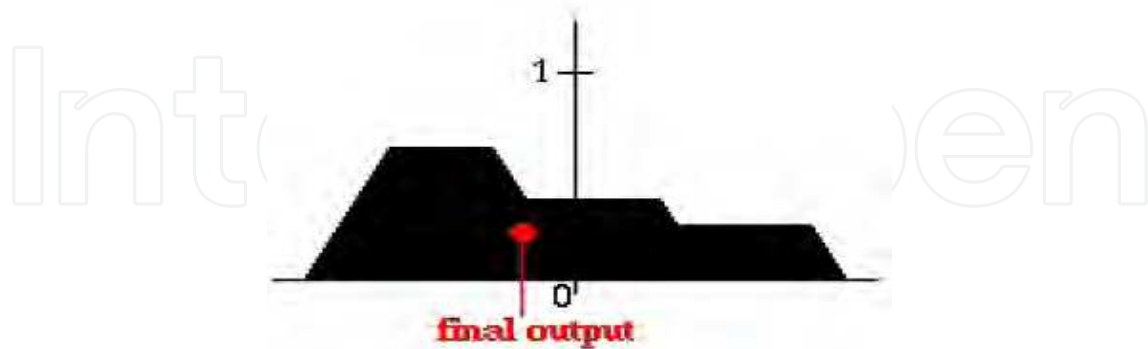


Fig. 15. Defuzzification using the center of gravity approach

3. Applications of AI techniques

Almost every branch of science and engineering currently shares the tools and techniques available in the domain of AI. However, we mention here a few typical applications, [15-20].

Expert Systems: An expert system consists of a knowledge base, database and an inference engine for interpreting the database using the knowledge supplied in the knowledge base.

Image Understanding and Computer Vision: A digital image can be regarded as a two-dimensional array of pixels containing gray levels corresponding to the intensity of the reflected illumination received by a video camera .

Navigational Planning for Mobile Robots: Mobile robots, sometimes called automated guided vehicles (AGV), are a challenging area of research where AI finds extensive applications. The navigational planning problem persists in both static and dynamic environments.

Speech Understanding: the main problem is to separate the syllables of a spoken word and determine features like amplitude, and fundamental and harmonic frequencies of each syllable. The words then could be identified from the Extracted features by pattern classification techniques.

Scheduling: In a scheduling problem, one has to plan the time schedule of a set of events to improve the time efficiency of the solution.

Intelligent Control: In process control, the controller is designed from the known models of the process and the required control objective. When the dynamics of the plant is not completely known, the existing techniques for controller design no longer remain valid.

Rule-based control is appropriate in such situations.

System Modeling and Optimization:

Optimization methods have been applied over years to generate solutions that solely maximize performance. In order to assess the performance variance of a solution, a few near optimal solutions are selected and studied under assumed stochastic parametric variations via simulation. There are reports on the use AI guide or bias search strategies. A new evolutionary algorithm that is capable of generating robust optimal solutions for constrained robust design problems.

3.1 Using AI for medical applications

The implementation of human intelligence in scientific equipment has been the subject of scientific research for a long time and of the medical research in the last decades. In the 1950's computer simulation of biological neural network was first introduced. In 1943 McCulloch and Pitts stated the definition of the first artificial neuron. In parallel with the evolution of computer technology, modeling of increasingly complicated neural functions and activity of simple neural clusters was defined. Mathematical models that could be applied for practical applications were developed between 1982 and 1987 based on the works of Hopfield [21], Kohonen [22] and Rummelhart and McLlland [23]. The advantage of neural networks over conventional programming lies in their ability to solve problems that do not have an algorithmic solution or the available solution is too complex to be found. Neural networks are well suited to tackle problems that people are good at solving, such as prediction and pattern recognition. Neural networks have been applied within the medical domain for clinical diagnosis [24-26], image analysis and interpretation [27,28] signal analysis and interpretation [29] and drug development [30].

Functional division of neural network applications in medicine; Papik et al, 1998) [31]:

1. Modelling: Simulating and modelling the functions of the brain and neurosensory organs.
2. Signal processing: Bioelectric signal filtering and evaluation.
3. System control and checking: Intelligent artificial machine control and checking based on responses of biological or technical systems given to any signals
4. Classification: Interpretation of physical and instrumental findings to achieve more accurate diagnosis.
5. Prediction: Neural network provide prognostic information based on retrospective parameter analysis.

Fuzzy logic [11] has been applied to dental and medical sciences (Sims-Williams et al, 1987) in order to construct systems that can infer precise recommendations for solving problems that have uncertain properties [32-36]. Brown et al. (1991) [37] applied fuzzy logic to solve orthodontic problems in an expert system, designed to provide advice for treatment planning of Class II division 1 malocclusions. They reported that their system produced more acceptable treatment plans than those used by general dental practitioners. Similarly, Tanaka et al. (1997) [36] applied fuzzy reasoning to their computer-assisted diagnostic system for ultrasonography for the purpose of providing a diagnostic aid for unskilled clinicians. (head gear)

4. Using AI for orthodontics

Many researchers intended to capture the outlines matching between pleasing smile and harmonically face. Usually the challenging face the orthodontist to figure out the orthodontic problems, there diagnosis and environment of origin keeping away distracting factors. The traditional regime for diagnosis include multiple steps for orthodontic problem identification, these steps generally categorized according to three sources (1) multiple questioning records including chief complaint, patient's dental and medical history; (2) clinical examination of the patient; and (3) assessment of diagnostic records, including dental casts, radiographs, and facial and intraoral images [38].

It is mandatory to contextualize data have been driven. All the data base collected gathered in an elaborate process to achieve the most appropriate treatment planning, treatment plan

is the second challenge facing the orthodontist, the enormous variation in dental malocclusion gathered with different facial pattern and the presence of large number of available treatment modalities, all these leading the decision process in orthodontics to challenging area even to the experienced orthodontist. Often more than one treatment plan can successfully resolve an orthodontic problem, and as a consequences form these two interrelated orthodontic processes (diagnosis and treatment plan) the treatment can be customized, orthodontic books and articles are profound with researches discussing protocols for decision making process regarding orthodontic problem definition and treatment option. Thus our chapter attempt to bring the insight to the invention of the artificial intelligence as a system aid in the essential orthodontic steps namely; diagnosis, treatment plan and treatment optimization.

4.1 Diagnosis using expert system

Expert system (ES) is an important branch of the field of artificial intelligence (AI).. ES is a computer program system that processes knowledge and information, which is composed primarily of a knowledge base and an inference machine. ES simulates the decision making and working processes of experts and solves actual problems in the field of a single specialty[39]. Generally, in a medical or dental expert system, a set of knowledge base is derived from experienced clinicians and represents their knowledge, which can be used for clinical consultations [40,41].With this type of system, uncertainty is a major problem in decision making because non-evidence-based knowledge has to be represented mathematically. Poon et al [42] were the first to use a new approach to knowledge acquisition known as Ripple-Down Rules in Dentistry to develop an ES in clinical orthodontics. This system comprises a knowledge base of 680 rules. Investigators found that such an ES has potential as an interactive advisory tool and is applicable in clinical orthodontic situations.

Hammond et al [43] pointed out in a review that traditional rule-based expert systems had some limitations when applied to orthodontic diagnosis and treatment planning. These limitations may be avoided by using a case-based system, which is a particular type of ES that uses a stored data bank of previously treated cases to provide knowledge for use in solving new treatment problems. Hammond et al [44] also investigated the application of this method in the field of orthodontic diagnosis and treatment planning. A case base of 300 cases was entered into a case based ES shell. A test set of 30 consecutive cases then was used to test the diagnostic capacity of the system. The computer-generated treatment plan matched the actual treatment plan in 24 of the 30 cases. In another work by Lux et al [45] the growth of 43 orthodontically untreated children was analyzed by lateral cephalograms taken at the ages of 7 and 15. For the description of craniofacial skeletal changes, the concept of tensor analysis and related methods were applied. Through the use of an ANN, namely, self-organizing neural maps (SOM), resultant growth data were classified, and relationships of the various growth patterns were monitored. This type of network provided a frame of reference for classifying and analyzing previously unknown cases with respect to their growth pattern, Brickley et al [46] concluded that ANN expert systems may be trained with clinical data only and therefore can be used in cases where "rule-based" decision making is not possible. This is the case in many clinical situations. ANN therefore may become important decision-making tools within dentistry, we'll discuss one of these expert systems;

An expert orthodontic index:

A valid initial assessment enables health providers to determine treatment need and priority, an accurate final diagnostic assessment assists patients and orthodontists to conclude if a worthwhile improvement is achieved. Orthodontists have developed several occlusal indices during the past few decades to evaluate treatment need, complexity, and success. Among the developed indices, Peer Assessment Rating (PAR) is one of the most common ones that is used to evaluate the quality of treatment. Richmond et al [47] developed PAR in 1992 to create consistency and standardization in assessing orthodontic treatment outcome. It is a weighted summation of health traits that influence the malocclusion. It summarizes data about the misalignment in a single score that reflects deviation from the ideal occlusion. The treatment success can then be evaluated by comparison of the pre- and post-treatment PAR scores. In spite of its extensive use, PAR suffers from several limitations. In summary, PAR is constrained by its strict linear mathematical expression with fixed coefficients, while a non-linearity may enhance the subjective opinions of orthodontists more accurately. Different versions of PAR index have been developed in order to improve the weighting system by using traditional regression techniques, but they are still restricted by the non-adjustable linear coefficient. A fuzzy index was developed by Zarei A. et al in 2007[48] using neural network with fuzzy approach. Zarei et al in 2009 [49] improved the quality of fuzzy index using union rule configuration. Further, an intelligent system that represent orthodontists' visual perception in assessing patients was developed.

Panel of orthodontists with randomized patients' files, each of which contained their data prior to treatment, during treatment progress, and at the end of treatment. profiles of 560 cases of malocclusion was used by the panel of orthodontist. each profile includes clinician's assessment that based on cephalometric tracing interpretations, visual perception, and clinical appearance. The panel assessed the cases using a visual analog scale. A visual analog scale is one of the most common measurement scales used in health care research and has also been used in dental studies [50].

Modeling was used for the identification of the expert system using the input-output data. Sugeno models [51] are good candidates for situations when a desired action can not necessarily be described verbally by experts. Therefore, Sugeno models provide a good way to model clinicians' assessment when using numerical data. a set of input-output data to first identify the fuzzy system for this collection of data and then optimize this model by adjusting the parameters. Input to this fuzzy inference model includes five variables that orthodontists associate with assessment of treatment outcome, The input variables for this model include the following linear and angular cephalometric measurements; overjet, ANB angle, Lower Incisor to Mandibular Plane angle (LI-MnPI), SNB angle and Upper Incisor to Sella Nasion angle (UI-SN), while the output parameter is the arithmetic mean of the panel's assessments. Subtractive clustering to identify the rule base was performed, clustering of data forms the basis of many system modeling algorithms.

Neural networks were utilized to learn the characteristics of the data and selected the parameters of input and output membership functions to best reflect those characteristics. The parameters of membership functions are modified during the learning process to minimize the fitness function. The adjustments of these parameters are facilitated by a gradient vector, which provides a measure of how well the fuzzy inference system is

modeling the data. Optimization of the parameters of the initial model with respect to training data by minimizing the sum of the squared difference between actual and desired outputs. Fig 16 depicts the performance of the model in predicting the assessment for training and testing patterns. It is evident from this figure that the model assessment is very close to the panel assessment for most of the patterns.

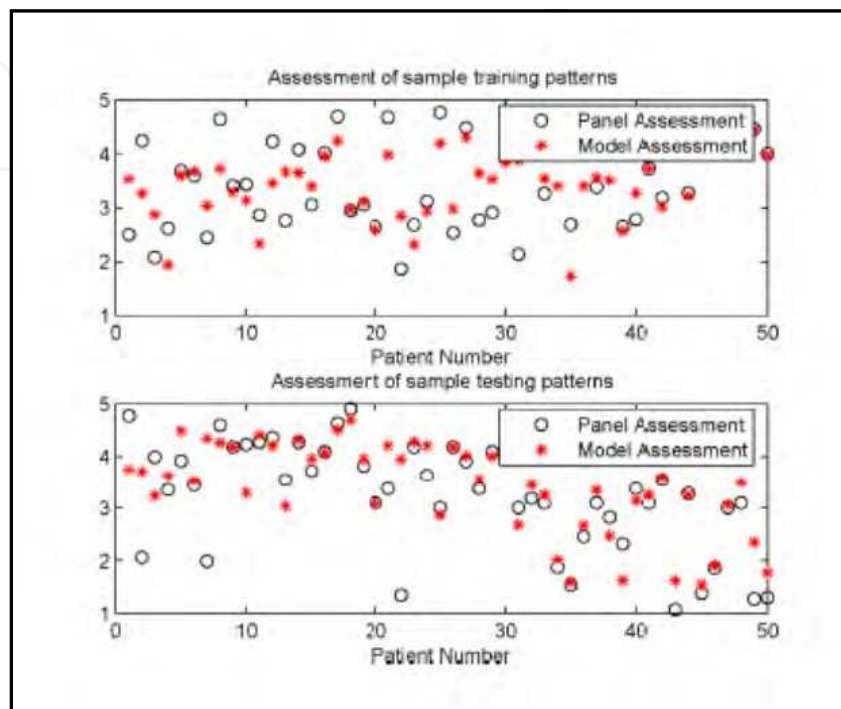


Fig. 16. Comparison of the panel assessment and the model assessment [49]

a Neuro-Fuzzy Assessment Index that is highly correlated with clinicians' opinion were successfully produced. neural network and fuzzy logic had been used for assessing orthodontic treatment outcome and developed a robust and realistic model that has a flexible interpretation of data. Applying Subtractive Clustering technique avoided the combinatorial explosion of rules in our model. hybridization of neural network and fuzzy logic improved the quality of the orthodontic index.[52].

Cephalometrics Analysis

Cephalometric analysis is a useful diagnostic tool to determine facial type and prediction of growth pattern, enabling clinician to determine facial disharmonies in order to centralize therapeutic measures during treatment and modify facial growth.

According to Graber and Vanarsdall [38], the commonly used radiographic views are:

Lateral or profile cephalograms: used to study anteroposterior and vertical relationships.

Frontal or postero-anterior cephalograms: used to evaluate the transversal and vertical relationships in the frontal plane. Submentovertebral or basal cephalograms: used to the balance in transversal plane.

Two approaches may be used to perform a cephalometric analysis: a manual approach, and a computer-aided approach. The manual approach is the oldest and most widely used. It consists of placing a sheet of acetate over the cephalometric radiograph, tracing salient features, identifying landmarks, and measuring distances and angles between landmark

locations. The other approach is computer-aided. Computerized cephalometric analysis uses manual identification of landmarks, based either on an overlay tracing of the radiograph to identify anatomical or constructed points followed by the transfer of the tracing to a digitizer linked to a computer, or a direct digitization of the lateral skull radiograph using a digitizer linked to a computer, and then locating landmarks on the monitor.[42-44].

Afterwards, the computer software completes the cephalometric analysis by automatically measuring distances and angles. The evolution from full manual cephalometrics to computer assisted-cephalometric analysis is aimed at improving the diagnostic ability of cephalometric analysis through errors reduction and time saving. Computerized or computer-aided, cephalometric analysis eliminates the mechanical errors when drawing lines between landmarks as well as those made when measuring with a protractor. However, the inconsistency in landmark identification is still an important source of random errors both in computer-aided digital cephalometry and in manual cephalometric analysis.[45-47] taking into account the imprecise, inconsistent, and paracomplete data inherent to the analytical process. There have been efforts to automate cephalometric analysis with the aim of reducing the time required to obtain an analysis, improving the accuracy of landmark identification and reducing the errors due to clinicians' subjectivity.

In an automated cephalometric analysis a scanned or digital cephalometric radiograph is stored in the computer and loaded by the software. The software then automatically locates the landmarks and performs the measurements for cephalometric analysis. The challenging problem in an automated cephalometric analysis is landmark detection, given that the calculations have already been automated with success. The first attempt at automated landmarking of cephalograms was made by Cohen in 1984,[53] followed by more studies on this topic. Automatic identification of landmarks has been undertaken in different ways that involve computer vision and artificial intelligence techniques.

The automated approaches can be classified into four broad categories, based on the techniques, Leonardia R et al[54] mentioned these categories with techniques examples for each approach recorded by different authors:

1. Image filtering plus knowledge-based landmark search; [55-58]
2. model- based approaches [59-64]
3. soft-computing approaches [65-68]
4. hybrid approaches. [69-73]

the relative advantages and disadvantages of these technical approaches used in the automated identification of cephalometric landmarks; Image filtering plus knowledge-based landmark search are list in table 1.

The informational importance of the cephalometric analysis was accompanied by many unnegligible sites of imprecision, this significant degrees of vagueness, and even inconsistency, was also making clinical application of the cephalometric data interpretation and driven information of less effective values than expected by the clinicians. to interpret how cephalometric variables behave in a complete contextualized scenario. Many trials was made to extract the ability of artificial intelligent techniques as favorable interpretational tool for the usual inconsistency of biological information. As a matter of fact artificial intelligence (AI) theories or techniques have few and recent applications in craniofacial biology, specifically in clinical application of cephalometrics, the multiple discussed studies was successfully produced at the level of researches taking in to account that the systems described in the literature are not accurate enough to allow their use for clinical purposes as errors in landmark detection were greater than those expected with manual tracing,

therefore; most of these methods have not been adopted in clinical practice [53]. The inconsistency of the informational driven cephalometric data gave the authors an additional challenging interface to overcome both cephalometric and modeling techniques inconsistencies, and yields sequential attempts of automated cephalometric analysis; we'll discuss some of these attempts in details hoping the enrichment of reader information of artificial intelligent approach for this diagnostic tool;

Techniques	Advantages	Disadvantages
Image filtering plus knowledge-based landmark search	Easy to implement Image filtering techniques are well studied and a large number are available By encoding proper anatomical knowledge better accuracy	Can fail to capture morphological variability in the radiographs Filtering results are highly dependent on image quality and intensity level Sensitive to noise in the image. Not all landmarks lie on edge and, moreover, the edges or curve are often unclear.
Model-based approach	Is invariant to scale, rotation, and translation (the structure can be located even if it is smaller or bigger than given model). Accommodates shape variability	Needs models that must be created by averaging the variations in shape of each anatomical structure on given set of radiographs. Model deformation must be constrained and is not always precise Cannot be applied to partially hidden regions Sensitive to noise in image.
Soft-computing or learning approach	Accommodates shape variability. Tolerant to noise. Techniques are well studies. Large selection of software tools available.	Results depend on the training set. Difficult to interpret some results. A number of network parameters, such as topology and number of neuron must be determined empirically.

Table 1. Technical approaches used to automatically identify cephalometric landmarks and their advantages and disadvantages [54].

Abe [74] and Mario et al [75]mentioned important limitations that conventional cephalometric holds, mostly due to the fact that the cephalometric variables are not assessed under a contextualized scope and carry on important variation when compared to samples norms. Because of that, its clinical application is subjective. Also discordance between orthodontists about diagnosis and treatments it is not uncommon, due to the inevitable uncertainties involved in the cephalometrics variables, and both suggest that this is a perfect scenario to evaluate the paraconsistent neural network capacity to perform with uncertainties, and inconsistencies in a practical problem.

Abe [74] develops an expert system in his work to support orthodontic diagnosis, the system based on the paraconsistent approach. Paraconsistent artificial neural network (PANN) was introduced in the Bulletin of Symbolic Logic [74], In the structure proposed the inferences that were based upon the degrees of evidence (favorable and unfavorable) of

abnormality for cephalometrics variables, which may have infinite values between “0” and “1”, the suggested PANN refined in Abe [74] work to produce an expert system to support orthodontic diagnosis, which may have infinite values between “0” and “1”. Therefore, the system may be refined with more or less outputs, depending upon the need. Such flexibility allows that the system can be modeled in different ways, allowing a finer adjusting. The system requires measurements taken from the head lateral radiography of the patient that will be assessed. The precision of the system increase as much as data is added.

Another work was made by Mario et al [75] to overcome these interpretational shortcomings, once again suggesting the contribution of Mathematics to Biology, better translating natural phenomena. Moreover, single correlations are insufficient for the assessment of facial patterns as many variables must be simultaneously considered in order to establish patterns. And once again the paraconsistent logic suggested as a model for detection and treatment of contradictions, enriching the use of soft mathematics tools in biology. research intends to test such model, it is reasonable to expect that the proposed model can well detect inconsistencies and better interpret craniofacial morphology [75], the cephalometric diagnostic model used logical states which represented in figure 7.

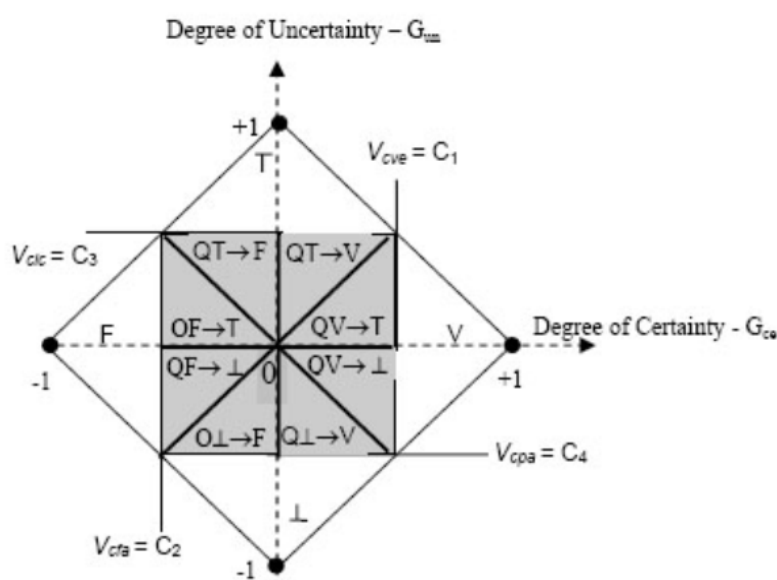


Fig. 17. Logical states: extreme and nonextreme states [75]

PANN

PANN was introduced in the *Bulletin of Symbolic Logic* (10). Its basis leans on paraconsistent annotated evidential logic $Et(10)$. Let us present it briefly. The atomic formulas of the logic Et are of the type

$p(\mu, \lambda)$, where $p(\mu, \lambda) \in [0, 1]^2$ and $[0, 1]$ is the real unitary interval (p denotes a propositional variable). The $p(\mu, \lambda)$ can be intuitively read: “It is assumed that p ’s favorable i.evidence is μ and contrary evidence is λ . Thus,

- $p(1.0, 0.0)$ can be read as a true proposition;
- $p(0.0, 1.0)$ can be read as a false proposition;
- $p(1.0, 1.0)$ can be read as an inconsistent proposition;
- $p(0.0, 0.0)$ can be read as a paracomplete (unknown) proposition;
- $p(0.5, 0.5)$ can be read as an indefinite proposition.

In the PANN, the main aim is to know how to determine the certainty degree concerning a proposition, if it is False or True. Therefore, the model took the certainty degree G_{ce} into account. The uncertainty degree G_{un} indicates the “measure” of the inconsistency or paracompleteness [78]. If the certainty degree is low or the uncertainty degree is high, it generates an indefinición; the basic scheme is shown in Fig. 18.

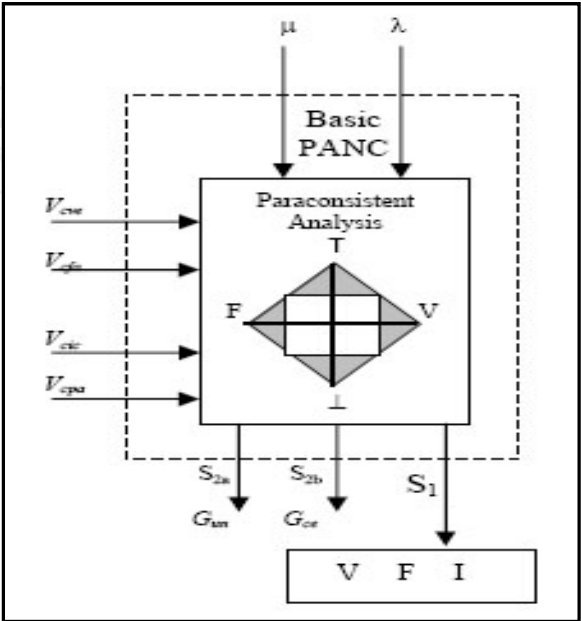


Fig. 18. The basic steps of a paraconsistent artificial neural cell

The model suggested by Mario et al [75] utilize selected set of cephalometric variables based on expertise (Figs. 9 a and b). These cephalometric variables are usually collected by experts [79] through characteristic points in a cephalometric X-ray. The selected cephalometric variables feed the PANN in the following three units: Unit I, considering the anteroposterior discrepancy; Unit II, considering vertical discrepancy; and Unit III, taking into account dental discrepancy (see Fig. 20).

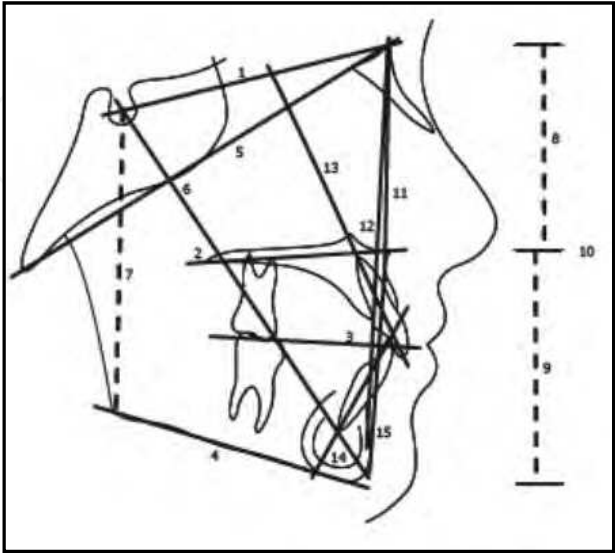


Fig. 19. a Cephalometric Variables [73]

(1. Basion 2. Sella 3. Nasion 4. Posterior Nasal Spine 5. Anterior Nasal Spine 6. Inter-Molars 7. Inter-Incisors 8. Gonion 9. Menton 10. Gnathion 11. A Point 12. B Point 13. Pogonion 14. Incisal Edge - Upper Incisor 15. Apex - Upper Incisor 16. Incisal Edge - Lower Incisor 17. Apex - Lower Incisor)

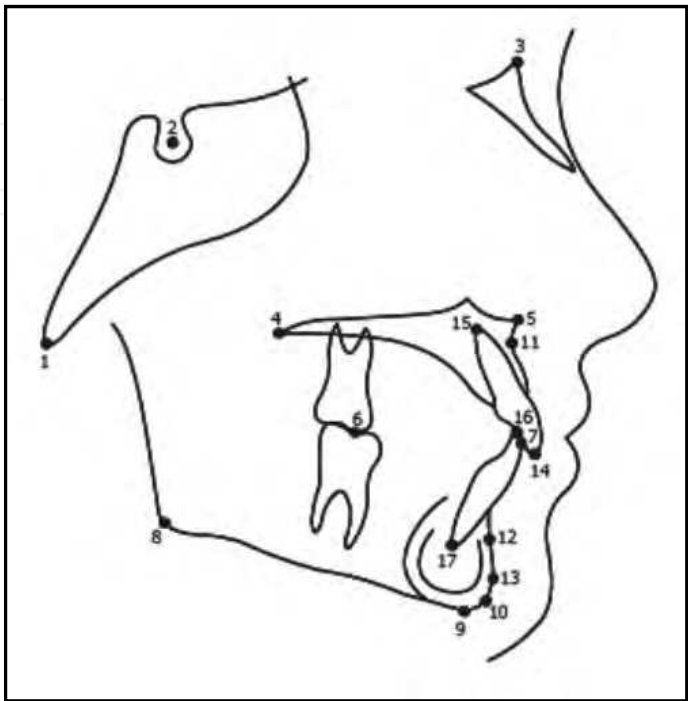


Fig. 19. b Proposed Cephalometric Analysis [73]

1. Anterior Cranial Base 2. Palatal Plane (PP) 3. Oclusal Plane (OP) 4. Mandibular Plane (MP) 5. Cranial Base 6. Y Axis 7. Posterior Facial Height 8. Anterior Facial Height - Median Third 9. Anterior Facial Height - Lower Third 10. Anterior Facial Height 11. SNA 12. SNB 13. Long Axis - Upper Incisor 14. Long Axis - Lower Incisor 15. A Point - Pogonion Line Wits: distance between the projections of the A and B Points on the occlusal plane.

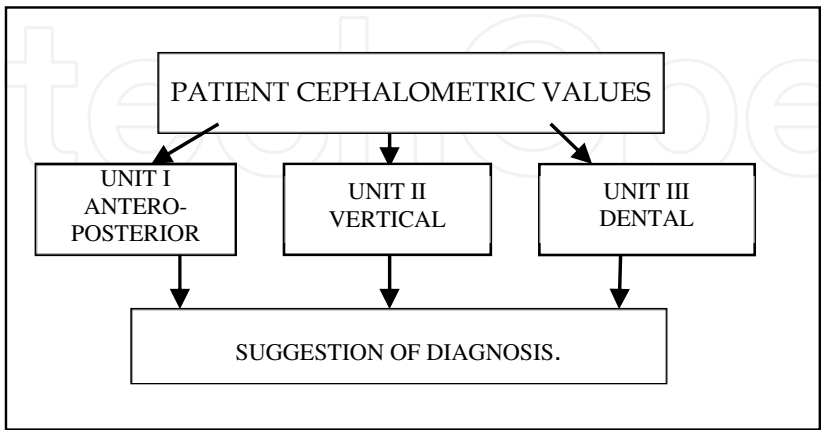


Fig. 20. Functional macroview of the paraconsistent artificial neural network architecture used by Mario et al [75]

Each unit has the specific following components, as shown in Fig 21:

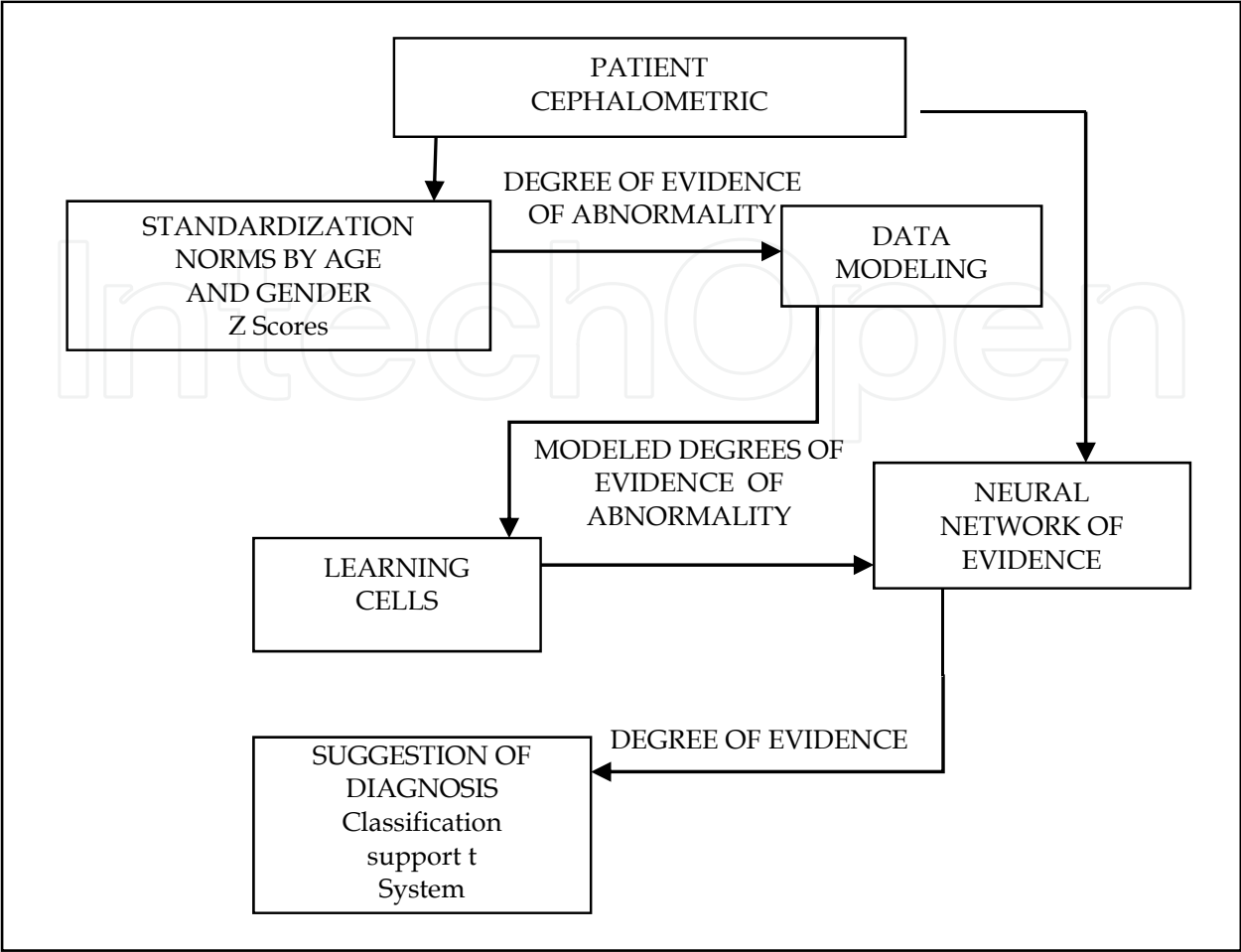


Fig. 21. Functional microview of the structure of each unit represented in Fig. 20 [75]

Tanikawa C et al [80] studied the reliability of a system that performs automatic recognition of anatomic landmarks and their surrounding anatomic structures in which the landmarks are located on lateral cephalograms using landmark-dependent criteria unique to each respective landmark. Recently, a system that recognizes general grayscale images using an automated psychologic brain model [81] has been developed, i.e., a hardware-friendly algorithm to accomplish real-time recognition by recalling a set of modeled data that is mathematically described using a finite number of traits and previously stored in the system. This system employs a new technique called the projected principal edge distribution (PPED) as a means for extracting features from an image, and it has been confirmed that the system demonstrates robust performance in recognizing images, including cephalograms [82-84]. Although experiments have suggested the efficacy of the system in recognizing images, it remains uncertain whether such a system will detect conventionally used landmarks with high precision. On the other hand, as mentioned before, that topographic variations exist in humans' subjective judgments of cephalometric landmarks, and the shapes and size of the variances are unique to each landmark [85]. Mathematical formulation of these landmark-dependent variations in measurement would help researchers to evaluate objectively the reliability of the automatic cephalogram recognition system.

Tanikawa C et al [80] system incorporates two major tasks: the "knowledge-generation" (system learning) phase and the "recognition" phase. In the knowledge generation phase,

image data extracted from learning samples are converted into PPED vectors consisting of 64 variables that feature contours of the anatomic structures [81,82,83]. From these vectors, template vectors, i.e., the principal information for identifying the landmarks, are generated using a generalized Lloyd algorithm [86] for each landmark, which are stored in the system as the system's knowledge. During the recognition phase, the system is designed to perform pixel-by-pixel film scanning with template-matching operations between PPED vectors that are generated from an input film and template vectors stored on the system. The system recognizes the most matched position as a landmark position. schematic representation can be seen in fig 22.

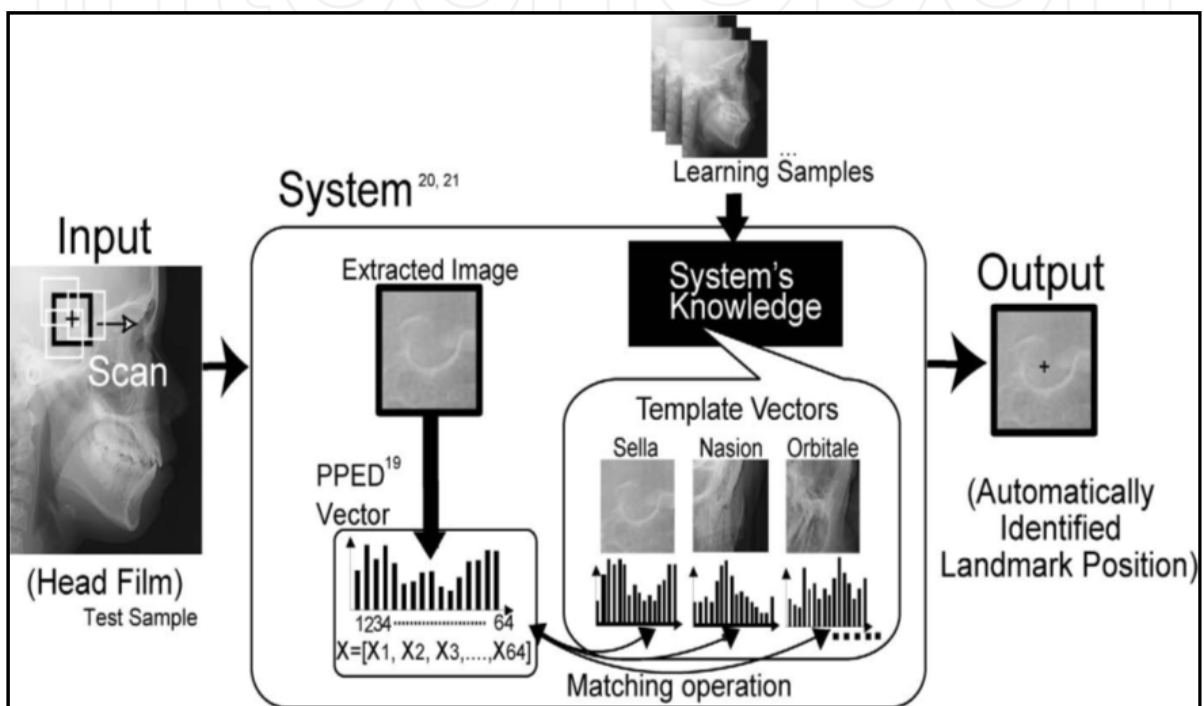


Fig. 22. Schematic representation for automatic recognition of anatomic landmarks [80]

To evaluate the system's performance reliability, scattergrams that designated errors for manual landmark identification when 10 orthodontists identified a landmark on 10 cephalograms were obtained according to the method reported by Baumrind and Frantz [85]. Confidence ellipses with a confidence limit of α were developed for each landmark from the scattergram, the system was evaluated using confidence ellipses with $\alpha = .01$. In short, when a system-identified point was located within a confidence limit of $\alpha = .01$, the landmark identification was judged to be successful.

To evaluate the accuracy of the landmark identification provided by the systems and if system's definition of a landmark position is clinically acceptable, it has been a critical issue in testing the performance reliability of such systems. Three major methods for such an evaluation have been employed so far. In the first method, an individual orthodontist makes a visual judgment as to whether or not the system's recommendation is acceptable [87]. The second approach involves describing mean recognition errors, i.e., the mean distance between the point provided by an orthodontist(s) and the point determined by the system [88,89]. The third method is to examine whether the system-identified landmark is located in a circle with a 2-mm radius [88-93], see fig13.

The fiducial zones established by the panel of experienced orthodontists are considered valid for evaluation of the ability of the automatic recognition system to recognize anatomic features. With the incorporation of the rational assessment criteria provided by confidence ellipses, the proposed system was confirmed to be reliable. The system successfully recognized anatomic features surrounding all the landmarks. The mean success rate for identifying the landmark positions was 88% with a range of 77% to 100%.

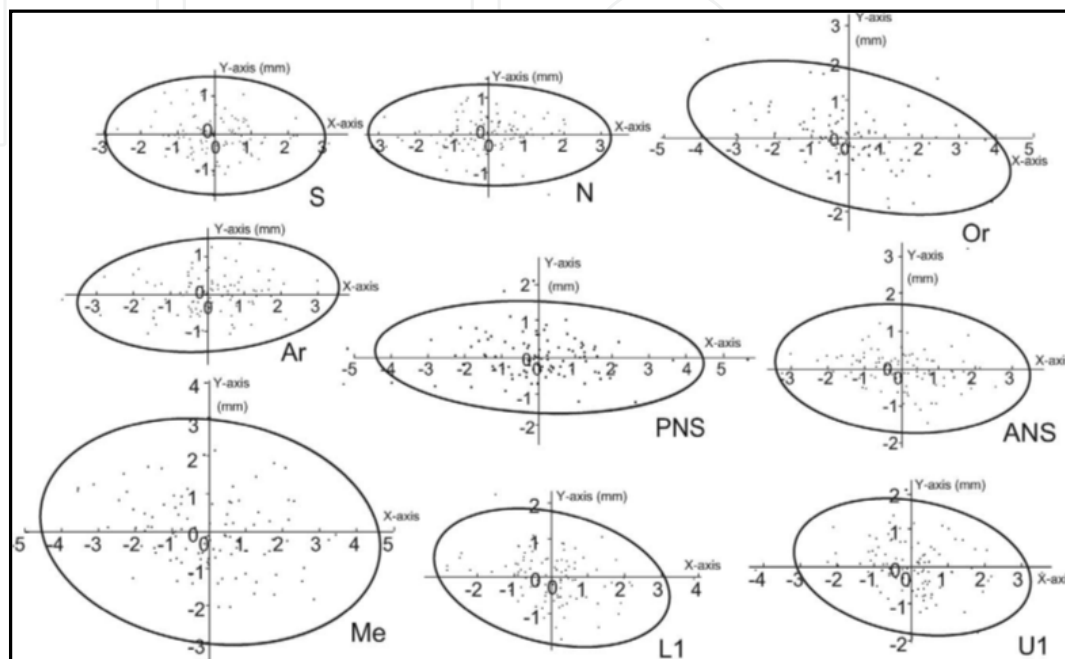


Fig. 23. Confidence ellipses obtained for cephalometric landmarks. Black points indicate coordinate values of landmarks identified by 10 orthodontists on 10 cephalograms. The black lines designate confidence ellipses with $\alpha=0.01$. Origin indicates the best estimate; x-axis, the line that passes through the origin and is parallel to the line S-N; and y-axis, the line that is perpendicular to the x-axis through the origin. [80].

In 2011 Banumathi A et al suggested [94] Another diagnostic model, Artificial intelligence role in dentofacial deformities diagnosis was discussed. The dentist must be familiar with morphological and functional maturity also oral surgeon and the orthodontist should be able to relate this knowledge to specific clinical problems such as skeletal mealocclusion and craniofacial anomalies. This understanding should influence the selection, planning and timing of treatment for patients who require orthognathic surgery. And the decision forming through the available choices, whether accepting the underlying deformity and taking the camouflage treatment as a choice or maybe surgical correction is the sole solution could be offered, of course full awareness of patient psychological aspects, underlying skeletal and/or dental malrelations and specific age, all should be taken into account. Cephalometric analysis was of important priority in deciding the acceptance and selection of appropriate orhognathic surgery for the underlying case. the diagnostic model was proposed by Edge sharpening of various bones in the lateral view of the face in cephalometric image referred to preprocessing, this is achieved through a histogram equalization process. From the literature, histogram equalization is enough to improve the contrast of the cephalometric image [95]. The edge features are then extracted from the enhanced cephalometric image and they are classified as landmark and non landmark

points using Support vector machine technique. Finally, angles between various landmark points are calculated to find out the deformities in the dento-facial growth. Banumathi A. et al [94] used in this study the Projected Principal Edge Distribution (PPED) vectors as a system for medical image recognition, and was used also in image recognition system discussed above described by Tanikawa C et al [80], as this techniques proved to provide better results.

4.2 Planning of treatment using AI

Enormous amount of variant subjects lies in the etiological list of orthodontic problems, we'll try at this section to show the artificial intelligent task in the planning of appropriate therapeutic goals can be achieved within certain boundaries for each problem, these boundaries considers the available problematic outcomes and its related factor as the backbone of system modeling and comparing these treatment plan with authors subjective opinions to simulate the treatment plan created by human brain.

Cranifacial Growth modification

Planning of treatment in the field of orthodontics and maxillo-facial surgery is largely dependent on the classification of individual growth of a patient. Work by Lux CJ et al[45] suggested the use of an artificial neural network, namely self-organizing neural maps, the growth of 43 orthodontically untreated children was analyzed by means of lateral cephalograms taken at the ages of 7 and 15. For the description of craniofacial skeletal changes, the concept of tensor analysis and related methods have been applied. Thus the geometric and analytical limitations of conventional cephalometric methods have been avoided, the resultant growth data were classified and the relationships of the various growth patterns were monitored by using an artificial neural network. As a result of self-organization, the 43 children were topologically ordered on the emerging map according to their craniofacial size and shape changes during growth. As a new patient can be allocated on the map, this type of network provides a frame of reference for classifying and analysing previously unknown cases according to their growth pattern. The morphometric methods applied as well as the subsequent visualization of the growth data by means of neural networks can be employed for the analysis and classification of growth-related skeletal changes in general.

Impacted canine

An impacted canine requires a complex therapeutic management, The therapeutic approach to impacted canines is interdisciplinary, with many factors accounting for the final orthodontic and periodontal outcomes. Pretreatment radiographic features of impacted canines— α -angle, d-distance, and sector of impaction according to Ericson and Kurol[96,97] have been shown to be predictive factors for the durations of orthodontic traction and comprehensive orthodontic treatment to reposition the impacted tooth. The more severely displaced the canine with regard to the adjacent maxillary incisors, the longer the orthodontic treatment[98], most investigations evaluated the relationships between factors accounting for treatment outcomes of impacted canines with descriptive statistics or linear regression on a priori identified variables; more recent studies used multilevel statistics to study associations among factors without determining causal relationships [99,100]. The multiple factors affecting the ultimate treatment approaches and duration should be included in the overall AI model. In 2010 Nieri M et al [101] used Bayesian networks (BN) to comprehensive surgical-orthodontic treatment of maxillary impacted canines to evaluate the relative role and the possible causal relationships among various factors affecting the clinical approach to this condition. BN adopt an intermediate approach

between statistics and artificial intelligence. An automatic structural learning algorithm of the BN was used as an explorative statistical technique for detecting possible causal relationships among these variables:

- demographic variables (sex and age);
- topographic variables (clinical and radiographic): site (buccal or palatal), side (left or right), unilateral or bilateral (patient), a-angle, d-distance,
- s-sector; treatment technique (tunnel); duration of traction, duration of treatment;
- periodontal variables ; Width of keratinized tissue (KT), from the gingival margin to the mucogingival junction; and Probing depth (PD) measurements. These were evaluated for the treated teeth through the therapeutic course.

In the BN analysis. the metric variables were transformed into binary variables by using the median values as a threshold as shown in fig 24.

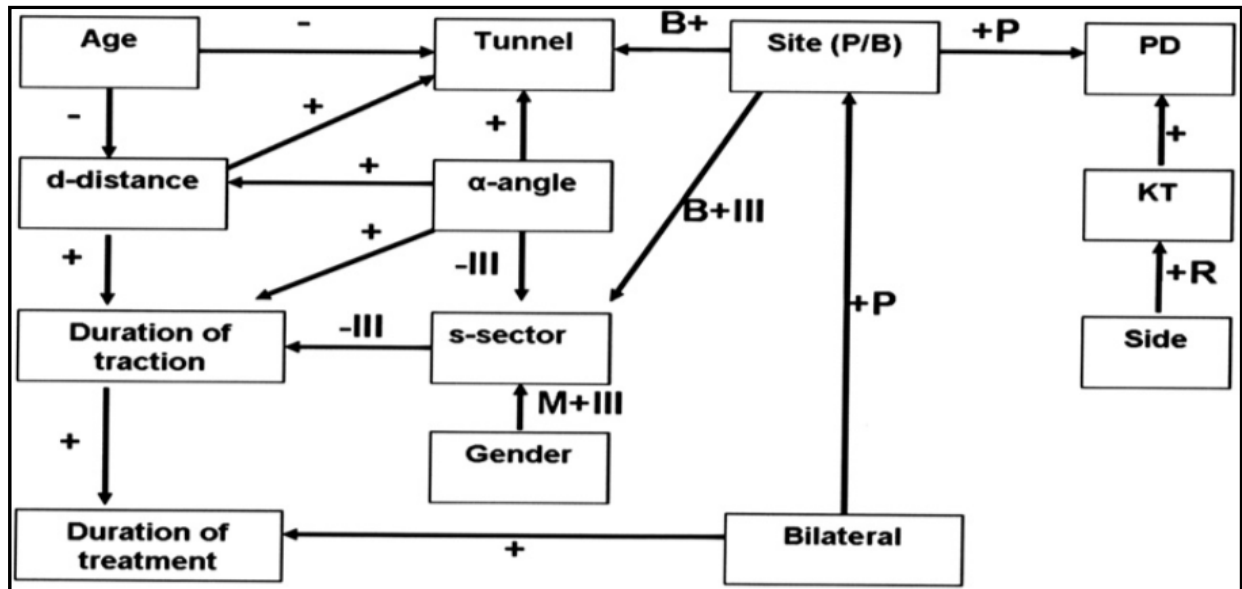


Fig. 24. The graph generated by the structural learning algorithm .P, Palatal; B, buccal; PD, probing depth; KT, keratinized tissue; R, right side; III, sector 3;M, male;1, the variable at the base of the arrow positively influences the variable at the arrowhead,the variable at the base of the arrow negatively influences the variable at the arrowhead, from Nieri M et al [101]

The BN approach confirmed the results of previous investigations on the same population in which the final periodontal outcomes after the surgical-orthodontic repositioning of maxillary impacted canines were unrelated to pretreatment diagnostic variables on the panoramic radiographs [100,102].The application of BN to diagnostic and therapeutic aspects of comprehensive surgical-orthodontic treatment of maxillary impacted canines identified several possible causal relationships among factors affecting the final outcomes of therapy.

Extraction demands in orthodontics

Early in the 20th century the maintaining of intact dentition became an important goal of orthodontic treatment. Angle and his followers strongly opposed extraction for orthodontic purposes. With the emphasis on dental occlusion that followed, however, less attention came to be paid to facial proportions and esthetics at that time. Small jaw size relative to the size of the teeth is an important factor in planning orthodontic therapy, as it implies that a significant percentage of patients will continue to require extractions to provide space for

aligning the remaining teeth. for over 100 years it has been a key question in planning orthodontic treatment. In orthodontics, there are two major reasons to extract teeth [103]:

1. to provide space to align the remaining teeth in the presence of severe crowding, and
2. to allow teeth to be moved (usually, incisors to be retracted) so protrusion can be reduced or so skeletal Class II or Class III problems can be camouflaged.

The alternative to extraction in treating dental crowding is to expand the arches; the alternative for skeletal problems is to correct the jaw relationship, by modifying growth or surgery. the majority of patients were treated with extractions to provide enough space for the other teeth. At present there again is great enthusiasm for expanding dental arches, on the theory that soft tissue adaptation will allow the expansion to be maintained, therefore; orthodontic treatments for malocclusion can be classified as extraction treatments and nonextraction treatments. The decision of extraction or not might be challenging and aimed to correct the malocclusion and enhancement of dental and facial appearance.

The decision to extract requires a multiple-factor analysis, which often includes the clinical experiences of the orthodontist. Currently, many multiple-factor analysis methods are available for use. Among these, the most frequently used is the statistical process known as fuzzy grouping analysis. Fuzzy grouping analysis regroups multiple factors based on their closeness in affecting the extraction decision. Classification by this algorithm is applicable to many patients. Xie X et al [39] study construct a decision-making expert system (ES) for orthodontic treatment by using a new approach. The ANN model was constructed to predict whether malocclusion patients between 11 and 15 years old required orthodontic extraction treatment. ANN model had 23 neurons in the input layer and 1 neuron in the output layer; this corresponded to the use of extraction or nonextraction treatments.

The model was implemented using the FORTRAN programming language, which is based on the principle of artificial neural networks. This Back Propagation (BP) ANN employs the error backward propagation learning algorithm. The basic principle of the BP algorithm is the propagation of errors from the output layer backward to the input layer by each layer that "shares" the error with neurons of each layer. Thus the reference errors of each layer of neuron are obtained for use in adjusting the corresponding connection weights, to make the error function diminish as far as possible. To enhance the performance of BP networks, a suitable learning parameter η and momentum parameter ε should be chosen properly. 25 indices were selected for screening of subjects. Two of these were nonquantification indices, which included the situation of heredity and protruded anterior teeth uncovered by incompetent lips. Among the quantifiable indices, 5 were derived from cast measurement, 13 from hard tissue cephalometrics, and 5 from soft tissue cephalometrics.

Contributions of the 23 input layer indices to the output layer index were analyzed through the method of neural network data processing. The connection strengths of each neuron in the input layer with each neuron in the hidden layer were used to represent the values of contribution from every input index. The values of a new index F (i) were calculated respectively to represent the contributions. These new indices were ordered by their magnitude, with the largest on top. The new index described the contributions from every input index to the result, as is shown in Table 2. After the data were preprocessed, all input indices were valued at between 0 and 1. the output index was extraction or nonextraction, quantification was processed as 0.99 for "yes" and 0.01 for "no."

Data from the 180 patients-in-training set were used to train the ANN model described above. Data from 20 patients were used to test the accuracy of the ANN model. When η was chosen as 0.9 and ε as 0.7, and the number of neurons in the hidden layer was 13, the model

had a nice learning effect. The 20 test samples proved successful in evaluating factors that affect the decision-making process. The rate of accuracy was 100%, which demonstrated that the constructed ANN could make correct decisions regarding the data of the trained 180 samples. Then, the data of 20 testing set samples that had not been trained were tested, and it demonstrated that the rate of accuracy was 80%. As for the marginal cases.

Input index	F(i)	order
Anterior teeth uncovered by incompetent lips	14042.44	1
IMPA (LI-MP)	11833.00	2
Overjet	6693.52	3
Crowding in the upper dental arch, mm	6135.12	4
Space for correction Sppe’s curve, mm	5948.78	5
ANB	5891.38	6
Overbite	5689.83	7
LI-NB	5640.86	8
LL-E plane, mm	4697.43	9
Soft tissue convexity (Ns-Sn-Pos)	4033.89	10
Interincisal angle (UI-LI)	4000.84	11
UI-NA, mm	3967.02	12
Z angle	2872.43	13
Wits, mm	2664.37	14
NLA (Cm-Sn-UL)	2620.67	15
UI-SN	2447.43	16
UL-E plane, mm	2250.07	17
UI-NA	2199.48	18
Heredity	2190.95	19
Crowding in the lower dental arch, mm	2173.83	20
FMIA (L1-FH)	1966.10	21
L1-NB, mm	1520.10	22
FMA (FH-MP)	410.39	23

Table 2. Analysis of Contributions of Every Input Index used inXie X et al [39] expert system

Using AI in selecting the appropriate treatment modalities

a computer-assisted inference model for selecting appropriate types of headgear appliance for orthodontic patients and act as a decision-making aid for inexperienced clinicians was developed by Akgam M.O and Takada K [104] Headgear is mainly used in orthodontic practice to deliver extra-oral forces to the upper dental arch for anchorage purposes, distalizing teeth and/or inhibiting forward maxillary growth. It has three main types, i.e. low, medium, and high-pull describing the direction of force applied to the upper molar teeth in the sagittal plane[105]. The choice of the precise type of headgear may not be difficult when considering its application in 'typical' cases, such as those exhibiting a Class II malocclusion with a deep overbite, large over jet, and a low mandibular plane angle. A problem may arise, however, particularly for orthodontists who have less clinical experience, with 'borderline' or 'marginal' subjects, such as those having a deep overbite, a moderate to severe over jet, and a high mandibular plane angle. This is because decision making in choosing an appropriate headgear type cannot be dealt with in a discrete, but

rather a continuous manner, i.e. *fuzzy logic*. the study incorporates three variables, namely, overjet, overbite, and mandibular plane angle, were used as input variables to the system. The mandibular plane angle was defined as the angle formed by the SN and mandibular planes. These variables were obtained from the lateral cephalograms.

For each input variable, three fuzzy sets for the low, medium, and high-pull types of headgear were defined on the basis of the authors' subjective judgment, which included their clinical experience and knowledge of the normative means and standard deviations for each variable. For each fuzzy set, the *fuzzy trapezoid function* was employed to construct membership functions. The fuzzy sets for each variable were determined with an assumption that the remaining two variables took normative values. For ease of understanding and simplicity, a graphic interpretation of the element and membership grade pairs which were created for the low, medium and high-pull types using each input variable is provided in Figures 25,26,27, Geometric mean aggregation was used for the inference, Geometric mean aggregation

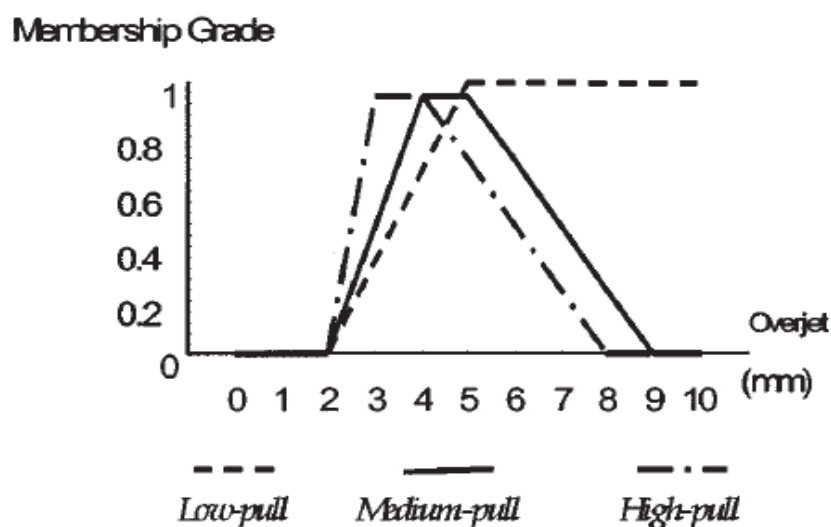


Fig. 25. Plot of membership functions for the input of overjet for each of three sets, i.e. the low-, medium-, and high-pull types of headgear.

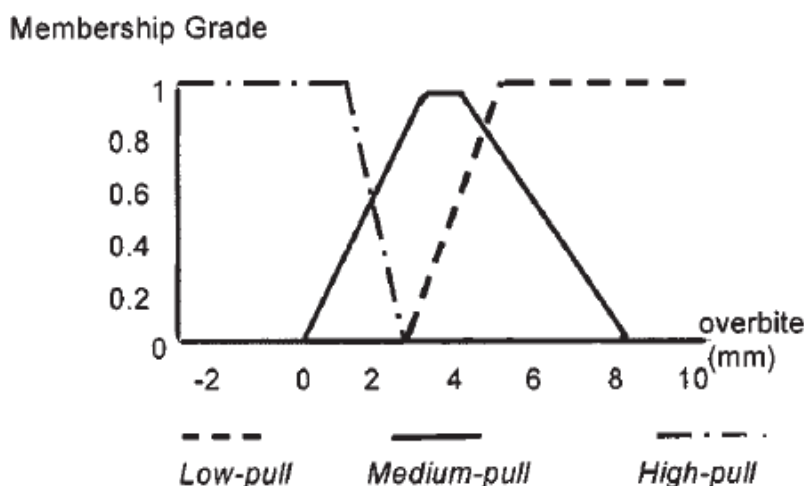


Fig. 26. Plot of membership functions for the input of overbite for each of three sets, . the low-, medium-, and high-pull types of headgear.

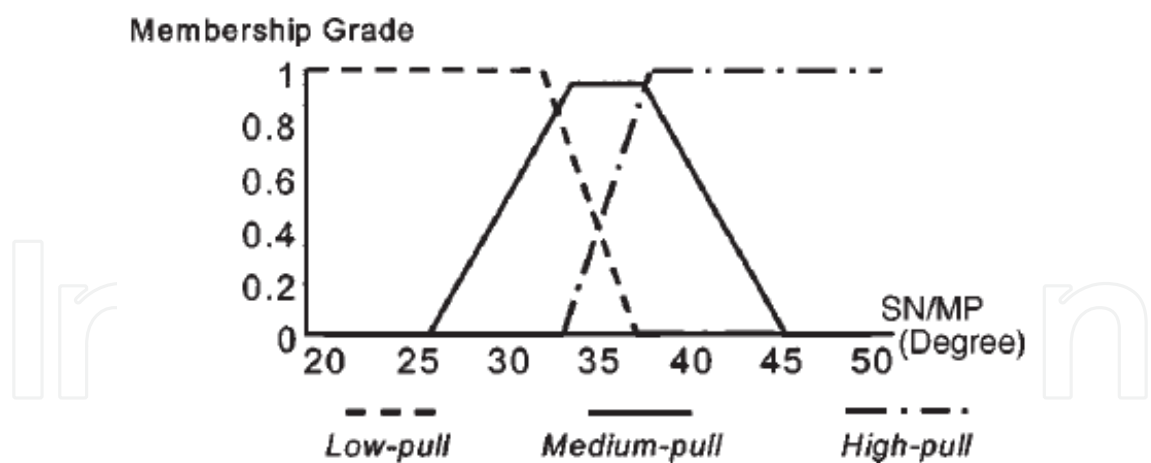


Fig. 27. Plot of membership functions for the input of mandibular for each of three sets, the low-, medium-, and high-pull types of headgear.

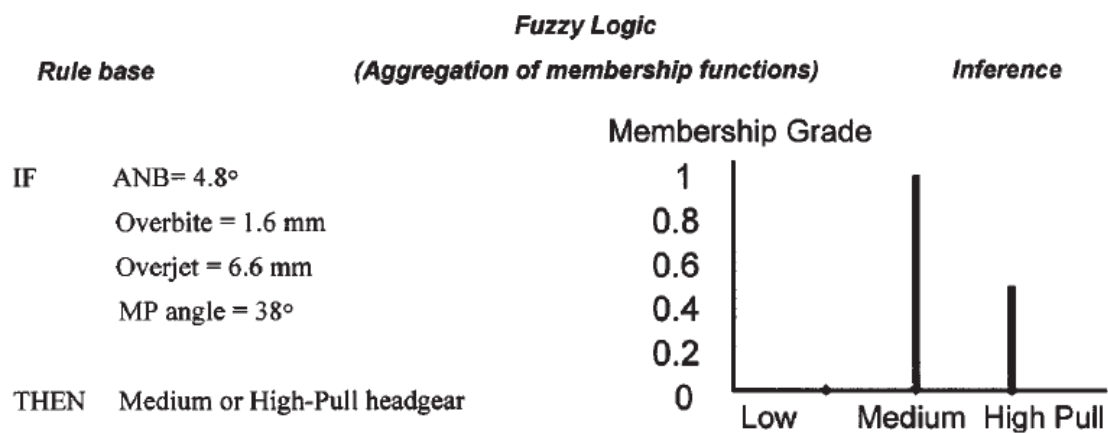


Fig. 28. The computer provides a selection of headgear types in which each choice is accompanied by a membership grades.

was the operation by multiple fuzzy sets were combined to produce single fuzzy set, the inference system was designed to calculate degrees of certainty for the use of each headgear type by means of membership grades fig 28.

The model was designed to calculate the degree of certainty for choosing low-, medium- or high-pull types of headgear. Eight orthodontic experts evaluated the decisions inferred by the system for 85 orthodontic cases. This group of clinicians was satisfied with the system's recommendations in 95.6 percent of the cases. In addition, the majority of the examiners (i.e. equal to or more than six out of eight) were satisfied with the system's recommendations in 97.6 per cent of the cases examined. Thus, the inference system developed was confirmed as being reliable and effective for clinical use in orthodontics.

4.3 The force system design for orthodontic treatments using AI

The most common aspects in the orthodontic treatment of extraction cases are the canine, the incisor and the mass retraction ,tooth retraction during space closure is achieved through two types of mechanics a. Sliding mechanics (friction mechanics) and b. Segmental or sectional mechanics (friction free technique). In the segmented arch technique, frictionless

springs are used to attract the segments of teeth on either side of an attraction site, there are different retraction springs that can be used. Many variables affect the force system they could produce; geometry, material, cross section, position, activation distance, etc. Tooth movement and orthopedic changes are the result of an applied force system and the tissue response to it. The force system is currently the major factor that the orthodontist can control to achieve desirable orthodontic tooth movement. Force system generated from complex geometric appliances produce forces and moments, it's important to control not only the magnitude of the force but also the moment to force ratio to produce the desired tooth movement. Force systems originated from orthodontic appliances have been studied by means of static systems for simple springs [106,107] or by experimental method [108-112], and numerical approaches [113] or by dynamic systems (typodont systems) [114]. The numerical methods are the most recent approaches having been merged with the medical area due to computer science, while in the experimental methods, the body of evidence is submitted to mechanical tests, which might determine the force system more accurately [115].

During an orthodontic space closure, the optimum response, both clinically and histologically, depends on the precision and calibration of the force systems to be used, therefore; a variety of prefabricated and precalibrated orthodontic loops are able to deliver precise and carefully controlled forces was utilized. Attempts to improve the force systems produced by this appliance have resulted in a number of different loop designs. Control of the force systems applied to the teeth is one of the main challenges in orthodontic biomechanics. Thus, the theoretical prediction of the forces and moments produced by the orthodontic appliance is important to control treatment.

It can be seen that if a reliable analytical or numerical method of the closing loop analysis is available, then any orthodontist can use this tool to calculate the characteristics of the closing loops theoretically without resorting to costly and time-consuming experiments. In many previous studies many researchers were developed a mathematical models for simulating the force system produced by orthodontic appliances based on small-deflection linear theory, large-deflection nonlinear theory and finite element methods, the last decade witness the innovative AI modeling using soft computing approaches, The advantage for using a black box of AI elements (like Neural networks, Genetic algorithms) in simulating the force system produced by the orthodontic appliances its ability to capture the real behavior of the orthodontic appliances (spring system).

In this section we will discuss some of the available AI models that can be used in modeling of the effective design for appliances for orthodontic treatment.

5. Force system prediction using artificial neural network

As we mentioned above the retraction loops force system namely, force, moment and moment to force ratio is affected by various parameter, Kazem et al [116] produced an artificial neural network based on an experimental force system evaluation of T-retraction springs, the experimental procedure includes studying the effect of cross section and activation distance on the force system produced by T-retraction springs, Forty T-looped stainless steel arch wires of three different cross sections were used in the testing procedure, their sizes were (0.018*0.025 in., 0.017*0.025 in., and 0.016*0.022 in.), A new test apparatus specially designed and operated by the researcher [117] (Garma NMH) for the measurement of the horizontal forces and the moments of sectional springs is used for teeth

retraction, Each one of the already prepared twenty sectional stainless steel T-loop of each group was activated by 1 mm, 2 mm, and 3 mm, respectively, and the readings from load cell outputs were recorded, .

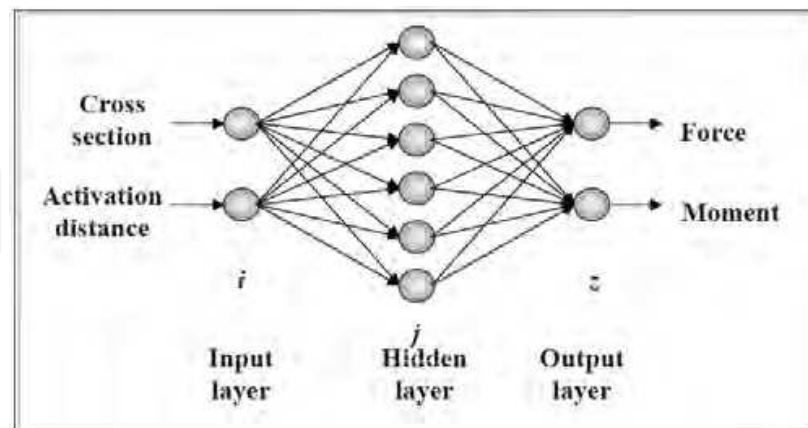


Fig. 29. NN architecture used for the force system modeling

After the results were obtained, they were used in the ANN modeling to evaluate its ability in the prediction process of the T-spring force system. Neural network training can be made more efficient if certain preprocessing steps are performed on the network inputs and targets. Figure 5 illustrates the preprocessing and postprocessing stage in prediction model. The neural network model is organized as a number of input neurons equal to the number of independent variables, which are spring properties (cross section and activation distance), and a number of output neurons equal to the number of dependent variable, which is the force system component (force and moment). Although the number of neurons in the input and output layers is specified depending on the problem, there is no hard and fast rules to specify the number of neurons in the hidden layer. Up to this day, the problem of specifying the optimal number of hidden neurons is an active area of research. Often a trial and error approach is starting with a modest number of hidden neurons and gradually increasing the number if the network fails to reduce its performance index (training error)[118,119]. The simulated annealing technique is used to capture the best weights and biases. The experimental results were used to train and test in neural network; seven measured results were used, from the total of nine, as data sets to train the network. Many neural networks architectures are used to train the data set to produce the least error; the neural network model was trained by using Levenberg - Marquardt Algorithm. many different trial numbers of the hidden layer neurons and types of the activation function were used at each time. Figure show optimized ANN training session. To evaluate the effect of increasing the hidden layer in to two layers on the ANN performance, many trial numbers of the two hidden layers training and accuracy prediction was done with different number of the two layer hidden neurons., the prediction accuracy for the testing patterns is based on the mean absolute percent error .

A network with one hidden layer include 6 neurons trained by Levenberg-Marquardt algorithm showed the best performance indication. Figure (20) shows the resulted network architecture, the network architecture consists of two input neurons (i), and one hidden layers contain six neuron (j) with nonlinear activation function (tangential sigmoid) and two output neurons (z) with linear activation function.

The prediction accuracy of the optimized ANN architecture are illustrated in (table), the mean error of data test set of the force prediction is (5.707%), while for the moment prediction is (4.048%),The multilayer feed forward neural network was successful in mapping the relationship among inputs parameters of the T-spring “cross section and activation distance” and output force system “horizontal force and moment.” The successful ANN mapping of the relationship between the spring properties and resultant force system can happen by utilizing other researche\ results, and it would be more beneficial as generalization increased with increasing input data set _spring properties_ in this situation.

Set No.	Actual force (gm)	ANN Prediction	Error %	Moment (gm mm)	ANN Prediction	Error %
1	291.3	299.3	2.7487	1311.11	1301.7	0.7183`
8	375.3	342.8	8.666	2177.14	2016.5	7.3785

Table 3. Test data sets and network prediction after optimization

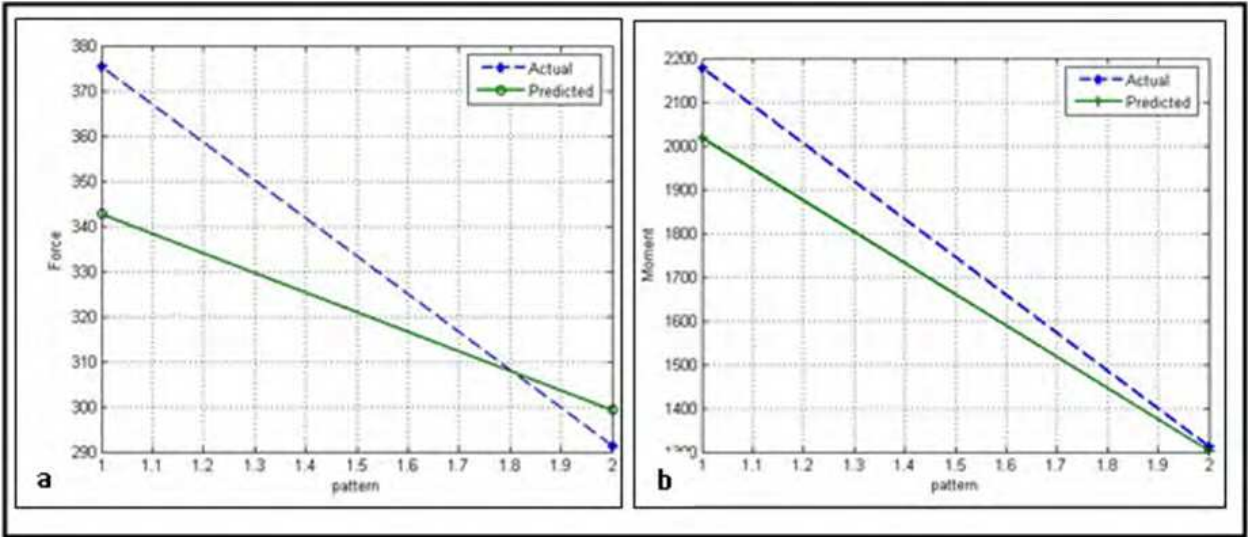


Fig. 30. Predicted and measured data in test set a-force, b- moment.

6. Multi-objective design optimization using GA

The multi-objective optimization is a vector of decision variables which satisfies constraints and optimizes vector functions whose elements represent the objective function. These functions form the mathematical description of performance criteria which are usually in conflict with each other. Hence, the term ‘optimizes’ means finding such a solution which would give the values of all objective functions acceptable to the designer, Osyczka [120]. The Genetic Algorithm is used in this work to optimize our engineering-orthodontic problem (select the best T-spring dimension and Material to get the required spring stiffness and moment to force ratio); i.e. to obtain an optimal force (spring stiffness) and a (M / H) ratio capable of pure translation together. The spring design parameters are encoded directly, using real codification, as strings (chromosomes) to be used for GA. For T-spring thirteen parameters should be optimized as shown in the following chromosome:

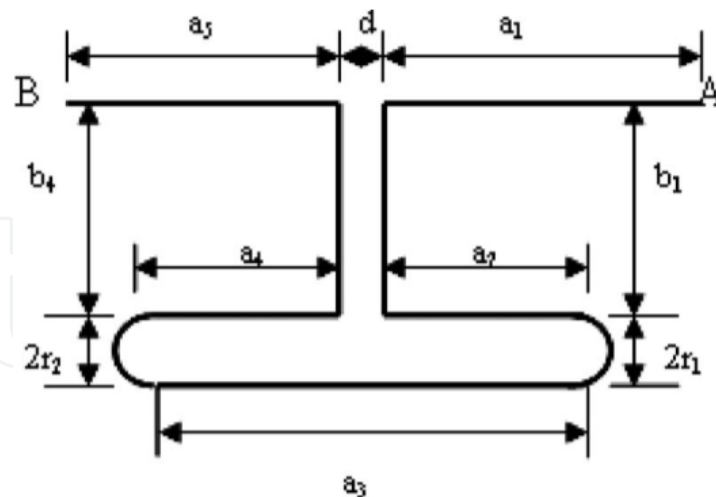


Fig. 31. T-spring Dimensions

$$x = [x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, x_{12}, x_{13}] \quad (2)$$

$$= [a_1, a_2, a_3, a_3, a_4, a_5, b_1, b_3, r_1, r_2, d, t, w, E] \quad (3)$$

Of these variables, $a_1, a_2, a_3, a_4, a_5, b_1, b_3, r_1, r_2$ are real-parameters with specific range (Lmin, Lmax), t and w are discrete variables having the standard cross section dimensions values of orthodontic arch wires, E is a discrete variable having the standard Young's modulus values of orthodontic arch wires.

The two objective optimization problems are as follows [121]:

$$\text{Minimize } f_1(x) = \text{springstiffness}(kx) \quad \text{Minimize } f_2(x) = |[(M/H) - (M/H)_D]| \quad (4)$$

Subjected to

$$\begin{aligned} g_1(x) &= (x_1 + x_2 + x_3 + x_4 + x_5 + x_6 + x_7 + \pi x_8 + \pi x_9 + x_{10}) \geq (L_T) \\ g_4(x) &= (x_1 + x_5 + x_{10}) \leq (L_T) \end{aligned} \quad (5)$$

$$g_4(x) = (x_1 + x_5 + x_{10}) \leq (L_T)) \quad (5)$$

$$g_2(x) = (x_6 + 2x_8) \leq (H_T) \quad (6)$$

$$g_3(x) = (x_7 + 2x_9) \leq (H_T) \quad (7)$$

$$g_5(x) = (x_2 + x_5 + x_{10}) \leq (L_T)) \quad (8)$$

$$g_6(x) = (x_2 + x_8 + x_9 + x_4) \leq (x_1 + x_5 + x_{10}) \quad (9)$$

$$g_7(x) = (x_3 + x_8 + x_9) \leq (x_1 + x_5 + x_{10}) \quad (10)$$

$$g_8(x) = x_2 + x_4 + x_{10} = x_3 \quad (11)$$

$$g_9(x) = (x_6 + 2x_8) - (x_7 + 2x_9) \leq \Delta \quad (12)$$

Where the kx and M/H are calculated using Castigliano's second theorem. $(M/H)D$ is the required moment to force ratio for orthodontic treatment. The nine constraints are chosen to make sure that the produced solutions are within the required total spring dimensions (total length (LT) and total height (HT)). The maximum allowable difference in total height between the left and right end is given by (Δ) . We can add any type of design constraint and to be sure that we will converge to some applicable design for the required application.

The presence or absence of a member in the spring structure is determined by comparing the length of the member with the designer defined small critical length, e . If a length is smaller than e , that member is assumed to be absent in the realized T-spring.

Operators in genetic algorithm

A new methodology for the optimization of the design parameters for T-spring arch wire had been developed by Kazem [121]. The proposed analytic model depending on the Castigliano's second theorem with an accurate boundary conditions and geometrical representation provides acceptable results for symmetric and asymmetric spring although it depends on the small deflection theory, in comparison with the results obtained using non-linear FEM. The multi-objective optimization for the spring design parameters is adopted successfully using GA method and the results show that depending on the above methodology, we can make good estimation of the required design parameters for the T-spring. Future work includes improving the analytical model for the spring system depending on the large deflection theorem and also, more inspection is needed by using other evolutionary algorithms like Strength Pareto Evolutionary Algorithm (SPEA) and SPEA 2 that update the ranking and selection criteria used in GA.

The initial population of strings (Real number coding) is generated at random and then the search is carried out among this population. The evolution of the population elements is non-generational, which means that the new replace the worst ones. The main different operators adopted in the GA are reproduction, crossover and mutation.

What concerns the reproduction operator is the successive generations of new strings which are generated based on their fitness values. In this case, a 5-tournament is used to select the strings for reproduction. For the tournament selection, only discrete values can be assigned and for higher range of selection intensity rather than ranking selection. About 50% of the population is lost at tournament size $Tour=5$. Tournament selection leads to high diversity for the same selection intensity compared to truncation selection [122]. At current search, tournament size less than 5 makes the solution progress slow toward the optimum solution and that which is more than 5 makes solution fall in the local optimum.

With a given probability P_c , the crossover operator adopts the single point technique and, therefore, the crossover point is only allowed between genes or, in other words, the crossover operator cannot disrupt genes. The mutation operator replaces one gene value x_t with another one generated randomly with a specified range by a given probability P_m .

According to our knowledge, such an approach has not been tested yet on orthodontic spring optimum design problem.

The size of the mutation step is usually difficult to choose. The optimal step-size depends on the problem considered and may even vary during the optimization process. It is known, that small steps (small mutation steps) are often successful, especially when the individual is already well adapted. However, larger changes (large mutation steps) can, when successful, produce good results much quicker. Thus, a good mutation operator should often produce small step-sizes with a high probability and large step-sizes with a low probability.

Two indices are used to qualify the evolving solution. All indices are translated into penalty functions to be minimized. Each index is computed individually and is integrated in the fitness function evaluation. The fitness function ff adopted for evaluating the candidate solutions is defined after Coello and Christiansen [123]:

$$f_f = \beta_1 f_1 + \beta_2 f_2$$

Where $\beta_1 + \beta_2 = 1$

The optimization goal consists of finding a set of design parameters that minimize ff according to the priorities given by the weighting factors β_i ($i = 1, 2$), where each different set of weighting factors must result in a different solution.

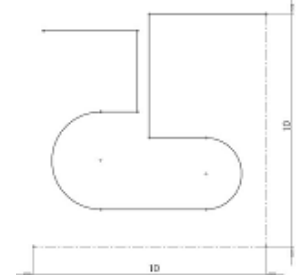
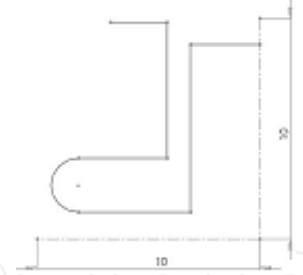
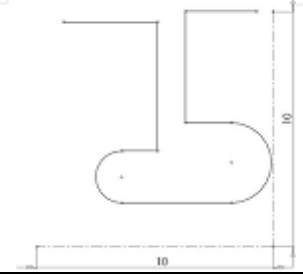
Optimizati on weight factor (β_1, β_2)	Spring shape and material for Optimized solution vector (x)	f1			f2		
		GA	From FEM	Error %	GA	From FEM	Erro r%
(0.75,0.25)		40.32	42.67	-5.8	1.31	1.11	15.2
	txw=0.40x0.55 E=172,000Mpa						
(0.5,0.5)		49.11	58.56	-19.2	1.66	1.4	15.6
	txw=0.40x0.55 E=172,000Mpa						
(0.25,0.75)		50.29	46.18	8.17	0.61	0.67	-9.8
	txw=0.482x0.635 E=192,000Mpa						

Table 4. Optimized Solutions for T-Spring design.

The optimized spring geometry and materials produced by GA were modeled by using FEM to calculate the spring stiffness ($=f1$) and the difference between resultant moment to force ratio and the user specified ratio ($=f2$) as given in Table 4.

7. Summary

In summary the subject of artificial Intelligence (AI) deals with symbolic processing than numeric computation. Knowledge representation, reasoning, planning, learning, intelligent search and uncertainty management of data and knowledge are the common areas covered under AI. Some of the applications areas of AI are speech and image understanding, expert systems, pattern classification, system optimization and navigational planning of mobile robots, the recent implementation of AI in the medical field and orthodontics was of special concern in this chapter, as the discipline of orthodontics task is to deal with the boolean etiological identification and optimized strategies of solution delivered to treat dentoalveolar and/or facial skeletal malrelation, the AI role used to achieve this task using variant techniques, AI incorporated many trial of changing the techniques used to simulate the clinical situations in the three essential sequences, diagnosis, treatment plan and treatment. The presence of differential problematic orthodontic problems, their origins and the consequence treatment makes the understanding of the AI aspect and its techniques essential to choose the techniques discriminating different problems and subsequent solution. Three aspects of AI;

- Artificial Neural Nets
- Genetic Algorithms
- Fuzzy logic

Artificial intelligence trials with these three techniques in different orthodontic steps yield numerous researches enrich the orthodontic domain with logical and economic tool substitutes the elongated sequential and sophisticated techniques, these trials yet are active area of research and need elaborated studies to reach their ultimate clinical assumption.

8. References

- [1] Konar A., Uncertainty Management in Expert Systems using Fuzzy Petri Nets, Ph.D. thesis, Jadavpur, 1999.
- [2] Newell, A. and Simon, H.A., Human Problem Solving, Prentice-Hall, Englewood Cliffs, NJ, 1972.
- [3] Luger, G. F. and Stubblefield, W. A., Artificial Intelligence: Structures and Strategies for Complex Problem Solving, Benjamin/Cummings, Menlo Park, CA, 1993.
- [4] Schalkoff, J., Culberson, J., Treloar, N. and Knight, B., "A world championship caliber checkers program," Artificial Intelligence, vol. 53, no. 2-3, pp. 273-289, 1992.
- [5] Winston, P. H., Artificial Intelligence, Addison-Wesley, 2nd ed., Reading, MA, 1994.
- [6] Newell, A. and Simon, H.A., Human Problem Solving, Prentice-Hall, Englewood Cliffs, NJ, 1972.
- [7] Charniak, E. and McDermott, D., Introduction to Artificial Intelligence, Addison-Wesley, Reading, MA, 1985.
- [8] E. A. Metzbowder, J. J. DeLoach, "Secondary effects in neural network analysis of the mechanical properties of welding alloys for HSLA shipbuilding steels" Mathematical

- Modeling of Weld Phenomena 6, eds H. Cerjak and H. K. D. H. Bhadeshia, published by Maney, London, 2002. pp. 231-242.
- [9] H Demuth, M Beale, M Hagan "Neural Network Toolbox User's Guide Ver5 " the mathworks ,Inc., Natick, MA, USA, 2006
- [10] D. T. Pham and L. Xing "Neural networks for identification, prediction and control" Springer publishing company (1997)
- [11] Haykan Simon, "Neural Network – a comprehensive foundation", prentice Hall, 1999.
- [12] Pohlheim, Hartmut, GEATbx: Introduction Evolutionary Algorithms: Overview, Methods and Operators GEATbx version 3.8 (2006)
- [13] Zadeh L A 1965 Fuzzy sets. Information and Control 8: 338-353
- [14] M. Hellmann, Fuzzy Logic Introduction Laboratoire Antennes Radar Telecom , F.R.E CNRS 2272, Equipe Radar Polarimetrie, Université de Rennes, France
- [15] Dean, T., Allen, J. and Aloimonis, Y., Artificial Intelligence: Theory and Practice, Addison-Wesley, Reading, MA, 1995.
- [16] Dickmans, E.D., Mysliwetz, B. and Christians, T., "An integrated spatio-temporal approach to vehicles," IEEE Trans. on Systems, Man and Cybernetics, vol. 20, no. 6, Dec. 1990.
- [17] Dougherty, E. R. and Giardina, C.R., Mathematical Methods for Artificial Intelligence and Autonomous Systems, Prentice-Hall, Englewood Cliffs, NJ, 1988.
- [18] Feigenbaum, E. A. and Feldman, J., Eds., Computers and Thought, McGraw-Hill, New York, 1963.
- [19] Haugeland, J., Ed., Artificial Intelligence: The Very Idea, MIT Press, Cambridge, 1985.
- [20] Jackson, P., Introduction to Expert Systems, Addison-Wesley, Reading, MA, 1986.
- [21] Hopfield J.J. Neural networks and physical systems with emergent collective computational abilities. Proc Nat Acad Sci. 1982; 2554- 2558.
- [22] Kohonen T. Self-organised formation of topologically correct feature maps. Biological Cybernetics, 1982; 43: 59-69.
- [23] Rummelhart DE, McClelland J.L. Parallel distributed processing: Explorations in the microstructure of Cognition. Vol. I. Foundation, MIT Press 1987.
- [24] Baxt, W. J., Application of artificial neural network to clinical medicine. Lancet. 346:1135-1138, 1995.
- [25] Kocyigit, Y., Alkan, A., and Erol, H., Classification of EEG recordings by using fast independent component analysis and artificial neural network. J. Med. Syst. 32 (1)17-20, 2008.
- [26] Silva Lopes H F & Abe J M. , Anghinah R. Application of Paraconsistent Artificial Neural Networks as a Method of Aid in the Diagnosis of Alzheimer Disease J Med Syst (2010) 34:1073-1081
- [27] Ventouras, E. M., Monoyou, E. A., Ktonas, P. Y., Paparrigopoulos, T., Dikeos, D. G., Uzunoglu, N. K., and Soldatos, C. R., Sleep spindle detection using artificial neural networks trained with filtered time-domain EEG: a feasibility study. Comput. Methods Programs Biomed. 78 (3)191-207, 2005.
- [28] Miller, A., The application of neural networks to imaging and signal processing in astronomy and medicine, Ph.D. thesis, Faculty of Science, Department of Physics, University of Southampton, 1993.
- [29] Prado, J. C. A. Constelação fônica e redes neurais artificiais: aplicabilidade na análise computacional da produção da fala (in Portuguese). Ph.D. Thesis. FFLCH—USP; São Paulo, SP, BR; 2007.

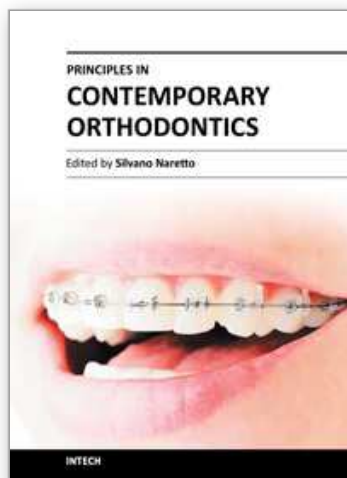
- [30] Weinstein, J., Kohn, K., and Grever, M., Neural computing in cancer drug development: predicting mechanism of action. *Science*. 258:447–451, 1992.
- [31] Papik K, Molnar B, Schaefer R, Dombovari Z, Tulassay Z, Feher J. Application of neural network in medicine. *Med Sci Monit*. 1998; 4(3): 538-46.
- [32] Sims-Williams J H, Brown I D, Matthewman A, Stephens C D 1987 A computer controlled expert system for orthodontic advice. *British Dental Journal* 163: 161-166
- [33] Mackin N, Sims-Williams J H, Stephens C D 1991 Artificial intelligence in dental surgery: an orthodontic expert system, a dental tool of tomorrow. *Dental Update* 18: 341-343
- [34] Hudson D L, Cohen M E 1994 Fuzzy logic in medical expert systems. *IEEE Engineering in Medicine and Biology* 3: 693-698
- [35] Stephens C D, Mackin N, Sims-Williams J H 1996 The development and validation of an orthodontic expert system. *British Journal of Orthodontics* 23: 1-9
- [36] Tanaka T, Miwa K, Kanda S 1997 Application of fuzzy reasoning in an expert system for ultrasonography. *Dentomaxillofacial Radiology* 26: 125-131
- [37] Brown I D, Adams S R, Stephens C D, Erritt S J, Sims-Williamis J H 1991 The initial use of a computer controlled expert system in the treatment planning of Class II division 1 malocclusion. *British Journal of Orthodontics* 18: 1-7
- [38] Graber TM, Vanarsdall RL. *Orthodontics: Current principles and techniques*. 2nd ed. St. Louis: Mosby, 1994.
- [39] Xie X; Wang L; Wang A. Artificial Neural Network Modeling for Deciding if Extractions Are Necessary Prior to Orthodontic Treatment *Angle Orthod*. 2010;80:262–266.
- [40] Stheeman S E, Van der Stelt P F, Mileman P A 1992 Expert systems in dentistry past performance – future prospects. *Journal of Dentistry* 20: 68-73
- [41] Takada K, Sorihashi Y, Akcam M O 1998 Orthodontic treatment planning: its rationale for inference. In: Carels C, Willems G (eds) *The future of orthodontics*. Leuven University Press, Belgium, pp. 203-211
- [42] Poon KC, Freer TJ. EICO-1: an orthodontist-maintained expert system in clinical orthodontics. *Aust Orthod J*. 1999; 15:219–228.
- [43] Hammond RM, Freer TJ. Application of a case-based expert system to orthodontic diagnosis and treatment planning: a review of the literature. *Aust Orthod J*. 1996;14:150–153.
- [44] Hammond RM, Freer TJ. Application of a case-based expert system to orthodontic diagnosis and treatment planning. *Aust Orthod J*. 1997;14:229–234.
- [45] Lux CJ, Stellzig A, Volz D, Jäger W, Richardson A, Komposch G. Department of Orthodontics, Dental School, University of Heidelberg, Germany. Growth dev.aging. A neural network approach to the analysis and classification of human craniofacial growth. 1998;(6) 295-106
- [46] Brickley MR, Shepherd JP, Armstrong RA. Neural networks: a new technique for development of decision support systems in dentistry. *J Dent*. 1998;26:305–309.
- [47] S. Richmond, WC. Shaw, KD. O'Brien, IB. Buchanan, R. Jones ,CD. Stephens, CT. Roberts, M. Andrews, "The development of the par index (peer assessment rating): reliability and validity," *European J of Orthodontics*, vol. 14, no. 2, pp. 125-129, 1992.
- [48] Zarei A, El-Sharkawi M, Hairfield M. An intelligent system for assessment of orthodontic treatment outcome using fuzzy union rule configuration. in *Proc. IASTED*, May 2007, 428-433.

- [49] Zarei A, Hairfield M, Mirsaeidghazi SS. A Novel Neuro-Fuzzy Assessment Index for Orthodontics. *IEEE Xplore*. 2009;3493-3497.
- [50] Vieira A, Lawrence H, Limeback H, Sampaio F, and Gryn timer M. A visual analog scale for measuring dental fluorosis severity. *American Dental Association*, vol. 136, pp. 895-901, 2005.
- [51] Takagi T and Sugeno M. Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. Systems, Man. And Cybernetics*, 1985;15: 116-132,.
- [52] Jyh-Shing Roger Jang and Chuen-Tsai Sun, Neuro-fuzzy and control. in *Proc. IEEE*.1995; 83(3): 378-406.
- [53] Cohen AM, Ip HH, Linney AD. A preliminary study of computer recognition and identification of skeletal landmarks as a new method of cephalometric analysis. *Br J Orthod*. 1984;
- [54] Leonardi R, Giordano D, Maiorana F, Spampinato C. Automatic cephalometric analysis. *Angle Orthod*. 2008;78:145-15.
- [55] Parthasarathy S, Nugent ST, Gregson PG, Fay DF. Automatic landmarking of cephalograms. *Comput Biomed Res*.1989;22:248-269.
- [56] Tong W, Nugent ST, Jensen GM, Fay DF. An algorithm for locating landmarks on dental X-rays. *Conference proceedings: 11th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE 1989(2): 552-554.
- [57] Forsyth DB, Davis DN. Assessment of an automated cephalometric analysis system. *Eur J Orthod*. 1996;18:471-478.
- [58] Ren J, Lid D, Feng D, Shao J. A knowledge-based automatic cephalometric analysis method. *Conference proceedings: 20th International Conference of the IEEE Engineering in Medicine and Biology Society*. IEEE 1998;20(2):723-727.
- [59] Cardillo J, Sid-Ahmed MA. An image processing system for locating craniofacial landmarks. *IEEE T Med Imaging*. 1994; 13:275-289.
- [60] Rudolph DJ, Sinclair PM, Coggins JM. Automatic computerized radiographic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop*. 1998;113:173- 179.
- [61] Hutton TJ, Cunningham S, Hammond P. An evaluation of active shape models for the automatic identification of cephalometric landmarks. *Eur J Orthod*. 2000;22:499-508.
- [62] Romaniuk B, Desvignes M, Revenu M, Deshayes MJ. Shape variability and spatial relationships modeling in statistical pattern recognition. *Pattern Recogn Lett*. 2004;25: 239-247.
- [63] Saad AA, El-Bialy A, Kandil AH, Sayed AA. Automatic cephalometric analysis using active appearance model and simulated annealing. *GVIP 05. Conference proceedings: InterAUTOMATIC*
- [64] CEPHALOMETRIC ANALYSIS 151 *Angle Orthodontist*, Vol 78, No 1, 2008 national Conference on Graphics, Vision and Image Processing, 2005 Dec 19-21, Cairo, Egypt (2005). ICGST; 2005.
- [65] Innes A, Ciesielski V, Mamutil J, Sabu J. Landmark detection for cephalometric radiology images using pulse coupled neural networks. In: Arabnia H, Mun Y, eds. *Conference proceedings: IC-AI '02, International Conference on Artificial Intelligence*, 2002 June 24-27, Las Vegas, Nevada. CSREA Press; 2002;2:511-517.
- [66] Chakrabartty S, Yagi M, Shibata T, Cauwenberghs G. Robust cephalometric identification using support vector machines. *Conference proceedings: ICME '03*,

- International Conference on Multimedia and Expo, 2003 Jul 6–9. IEEE 2003;III:429–432. 32.
- [67] Ciesielski V, Innes A, Sabu J, Mamutil J. Genetic programming for landmark detection in cephalometric radiology images. *Int J Knowl Based Intell Eng Syst.* 2003;7:164–171.
- [68] El-Feghi I, Sid-Ahmed MA, Ahmadi M. Automatic localization of craniofacial landmarks for assisted cephalometry. *Pattern Recognit.* 2004;37:609–621.
- [69] Liu J, Chen Y, Cheng K. Accuracy of computerized automatic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop.* 2000;118:535–540.
- [70] Grau V, Alcaniz M, Juan MC, Monserrat C, Knoll C. Automatic localization of cephalometric landmarks. *J Biomed Inform.* 2001;34:146–156. 34.
- [71] Yang J, Ling X, Lu Y, Wei M, Ding G. Cephalometric image analysis and measurement for orthognathic surgery. *Med Biol Eng Comput.* 2001;39:279–284.
- [72] Giordano D, Leonardi R, Maiorana F, Cristaldi G, Distefano M. Automatic landmarking of cephalograms by CNNs. *Lect Notes Artif Int.* 2005;3581:342–352.
- [73] Yue W, Yin D, Li C, Wang G, Xu T. Automated 2-D cephalometric analysis on X-ray images by a model-based approach. *IEEE Trans Biomed Eng.* 2006;53:1615–1623.
- [74] Abe JM. Paraconsistent Artificial Neural Networks: an Introduction, *Lecture Notes in Artificial Intelligence.* Carbonell & J. Siekmann, 2004.
- [75] Mario M C, Abe J M., Ortega N R S., Santo Jr M D. Paraconsistent Artificial Neural Network as Auxiliary in Cephalometric Diagnosis. *Artificial Organs* 2010 34(7):E215–E221, Wiley Periodicals, Inc..
- [76] Sylvan R, Abe JM. On general annotated logics, with an introduction to full accounting logics. *Bull Symb Log* 1996;2:118–9.
- [77] Mario MC. Modelo de análise de variáveis craniométricas através das Redes Neurais Artificiais Paraconsistentes. PhD thesis, University of São Paulo, 2006.
- [78] Da Silva Filho JI, Abe JM. Fundamentos das Redes Neurais Paraconsistentes – Destacando Aplicações em Neurocomputação, (in Portuguese) Editôra Arte & Ciência: ISBN 85-7473-045-9. 2001;247.
- [79] Ricketts RM. Cephalometric analysis and synthesis. *AngleOrthod* 1961;31:141–56.
- [80] Tanikawa C, Yagi M, Takada K. Automated Cephalometry: System Performance Reliability Using Landmark-Dependent Criteria. *Angle Orthod.* 2009;79:1037–1046.
- [81] Shibata T. Intelligent VLSI systems based on a psychological brain model. *Proc IEEE Int Sym on Intelligent Signal*
- [82] Yagi M, Shibata T. An image representation algorithm compatible to neural-associative-processor-based hardware recognition systems. *IEEE Trans Neural Netw.* 2003;14(5):1144–1161.
- [83] Levy-Mandel AD, Venetsanopoulos AN, Tsotsos JK. Knowledge- based landmarking of cephalograms. *Comput Biomed Res.* 1986;282–309.
- [84] Takada K, Yagi M, Tanikawa C. Dependable systems for decisions in clinical dentistry-how electronics implements what dentists know. *Proc Int Sym on Advanced Electronics for Future Generations, “Secure-Life Electronics” for Quality Life and Society.* 2005 Oct 11–12, Tokyo, Japan. 2005:15– 21.
- [85] Baumrind S, Frantz RC. The reliability of head film measurements. 1. Landmark identification. *Am J Orthod.* 1971; 60:111–127
- [86] Fujita K, Takada K, QianRong G, Shibata T. Patterning of human dental arch wire blanks using a vector quantization algorithm. *Angle Orthod.* 2002;72:285–294.

- [87] Levy-Mandel AD, Venetsanopoulos AN, Tsotsos JK. Knowledge- based landmarking of cephalograms. *Comput Biomed Res.* 1986;282-309.
- [88] Rudolph DJ, Sinclair PM, Coggins JM. Automatic computerized radiographic identification of cephalometric landmarks. *Am J Orthod Dentofacial Orthop.* 1998;113:173- 179.
- [89] Romaniuk B, Desvignes M, Revenu M, Deshayes MJ. Shape variability and spatial relationships modeling in statistical patter recognition. *Pattern Recogn Lett.* 2004;25: 239-247.
- [90] Grau V, Alcaniz M, Juan MC, Monserra C, Knoll C. Automatic localization of cephalometric landmarks. *J Biomed Inform.* 2001;34:146-156.
- [91] Ciesielski V, Innes A, Mamutil J. Genetic programming for landmark detection in cephalometric radiology images. *Int J Knowl Base Intell Eng Syst.* 2003;7:164-171.
- [92] Giordano D, Leonardi R, Maiorana F, Cristaldi G, Distefano ML. Automatic landmarking of cephalograms by cellular neural networks. *Lect Notes Artif Int.* 2005;3581:342-352.
- [93] Yue W, Yin D, Wang G, Xu T. Automated 2-D cephalometric analysis on X-ray images by a model-based approach. *IEEE Trans Biomedical Eng.* 2006;53:1615-1623.
- [94] Banumathi A, Raju S. , Abhaikumar V. Diagnosis of Dental Deformities in Cephalometry Images Using Support Vector Machine *J Med Syst* (2011) 35:113-119
- [95] Romaniuk, B., Desvignes, M., Robiaille, J., Revenu, M., Deshayes, M. J., Augmented reality and semi automated landmarking of cephalometry radiographs. *Comput. Anal. Images Patterns*, 9th International Conference CAIP 2001. 410-418, 2001. Sep 5-7
- [96] Ericson S, Kurol J. Radiographic assessment of maxillary canine eruption in children with clinical signs of eruption disturbance. *Eur J Orthod* 1986;8:133-40.
- [97] Ericson S, Kurol J. Early treatment of palatally erupting maxillary canines by extraction of the primary canines. *Eur J Orthod* 1988; 10:283-95.
- [98] Stewart JA, Heo G, Giover KE, Williamson PC, Lam EWN, Major PW. Factors that relate to treatment duration for patients with palatally impacted maxillary canines. *Am J Orthod Dentofacial Orthop* 2001;119:216-25.
- [99] Crescini A, Nieri M, Buti J, Baccetti T, Mauro S, Pini Prato GP. Short- and long-term periodontal evaluation of impacted canines treated with a closed surgical-orthodontic approach. *J Clin Periodontol* 2007;34:232-42.
- [100] Crescini A, Nieri M, Buti J, Baccetti T, Pini Prato GP. Pre-treatment radiographic features for the periodontal prognosis of treated impacted canines. *J Clin Periodontol* 2007;34:581-7.
- [101] Nieri M, Crescini A, Rotundo R, Baccetti T, C Pierpaolo, Pratoe GPP. Factors affecting the clinical approach to impacted maxillary canines: A Bayesian network analysis. *Am J Orthod Dentofacial Orthop* 2010;137:755-62.
- [102] Crescini A, Nieri M, Buti J, Baccetti T, Pini Prato GP. Orthodontic and periodontal outcomes of treated impacted maxillary canines. *Angle Orthod* 2007;77:571-7.
- [103] Proffit W R, Fields H W, Sarver DM, Contemporary orthodontics. Mosby Elsevier, 4th ed., 2007.
- [104] Akgam M. O, Takada K. Fuzzy modelling for selecting headgear types. *European Journal of Orthodontics.* 2002(24):99-106.
- [105] Williams J K, Cook P A, Isaacson K G, Thorn A R Fixed orthodontic appliances— principles and practice. 1996, Wright, Oxford

- [106] Braun, S., and Garcia, J. L., 2002, "The Gable Bend Revisited," *Am. J. Orthod. Dentofacial Orthop.*, 122, pp. 523-527.
- [107] Choy, K., Pae, E., Park, Y. C., and Burstone, C. J., 2002, "Controlled Space Closure With Statically Determinate Retraction System," *Angle Orthod.*, 72, pp. 191-198.
- [108] Solonche, D. J., Burstone, C. J., and Vanderby, R., 1977, "A Device for Determining the Mechanical Behavior of Orthodontic Appliances," *IEEE Trans. Biomed. Eng.*, BME-24, pp. 538-539.
- [109] Nagerl, H., Burstone, C. J., Beacher, B., and Messenburg, D. K., 1991, "Center of Rotation With Transverse Forces," *Am. J. Orthod. Dentofacial Orthop.*, 99, pp. 337-345.
- [110] Braun, S., Sjrursen, R. C., and Legan, H. L., 1997, "On the Management of Extraction Sites," *Am. J. Orthod. Dentofacial Orthop.*, 112, pp. 645-655.
- [111] Siatkowski, R. E., 1997, "Continuous Arch Wire Closing Loop Design, optimization, and Verification Part I," *Am. J. Orthod. Dentofacial Orthop.*, 112, pp. 393-402.
- [112] Ferreira, M. A., 1999, "The Wire Material and Cross Section Effect on Double Delta Closing Loops Regarding Load and Spring Rate Magnitude: An In Vitro Study," *Am. J. Orthod. Dentofacial Orthop.*, 115, pp. 275-282.
- [113] Raboud, D. W., Faulkner, M. G., Lipsett, A. W., and Habersstock, D. L., 1997, "Three Dimensional Effects in Retraction Appliance Design," *Am. J. Orthod. Dentofacial Orthop.*, 112, pp. 378-392.
- [114] Rhee, J., Chu, Y., and Row, J., 2001, "A Comparison Between Friction and Frictionless Mechanics With a New Type Simulation System," *Am. J. Orthod. Dentofacial Orthop.*, 119, pp. 292-299.
- [115] Ferreira, M. A., Oliveira, F. T., Ignacio, S. A., and Borges, P. C., 2005, "Experimental Force Definition System for a New Orthodontic Retraction Spring," *Angle Orthod.*, 75_3_, pp. 368-377.
- [116] Kazem B. I. ; Ghaib N. H.; Garma N. M. H., "Experimental Investigation and Neural Network Modeling for Force System of Retraction T-Spring for Orthodontic Treatment"; *Journal of Medical Devices ,ASME*, Vol 4, 2010, pp 021001-1-7.,
- [117] Garma N. M. H., "An experimental and artificial neural network prediction of cross section and activation distance effect on T-spring force system. Master's thesis. Department of orthodontics, Baghdad.
- [118] Sukthomya, W., and Tannock, J., 2005, "The Training of Neural Networks to Model Manufacturing Processes," *J. Intell. Manuf.*, 16_1_, pp. 39-51.
- [119] Yousif, K. Y., Daws, K. M., and Kazem, B. I., 2008, "Prediction of Friction stir welding Characteristics Using Neural Network," *Jordan Journal of Mechanical and Industrial Engineering _JJMIE_*, 2_3_, pp. 151-155.
- [120] Osyczka, A., 1985, "Multicriteria Optimization for Engineering Design," in *Design Optimization*, J. S. Gero, ed., Academic, New York, pp. 193-227.
- [121] Bahaa I Kazem "Multi-Objective Optimization for the Force System of Orthodontic Retraction Spring Using Genetic Algorithms ";*Journal of Medical Devices ,ASME*, Vol 3, 2009, pp 041006-1-7.,
- [122] Pohlheim, H., 2004, "GEATbx Introduction Evolutionary Algorithms: Overview, Methods and Operators," Version 3.5a, July 2004, [http:// www.geatbx.com](http://www.geatbx.com)
- [123] Coello, C. A., and Christiansen, A. D., 2000, "Multiobjective Optimization for Trusses Using Genetic Algorithms," *Comput. Struct.*, 75, pp. 647-660.



Principles in Contemporary Orthodontics

Edited by Dr. Silvano Naretto

ISBN 978-953-307-687-4

Hard cover, 584 pages

Publisher InTech

Published online 25, November, 2011

Published in print edition November, 2011

Orthodontics is a fast developing science as well as the field of medicine in general. The attempt of this book is to propose new possibilities and new ways of thinking about Orthodontics beside the ones presented in established and outstanding publications available elsewhere. Some of the presented chapters transmit basic information, other clinical experiences and further offer even a window to the future. In the hands of the reader this book could provide an useful tool for the exploration of the application of information, knowledge and belief to some orthodontic topics and questions.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Kazem Bahaa, Garma Noor and Yousif Yousif (2011). The Artificial Intelligence Approach for Diagnosis, Treatment and Modelling in Orthodontic, Principles in Contemporary Orthodontics, Dr. Silvano Naretto (Ed.), ISBN: 978-953-307-687-4, InTech, Available from: <http://www.intechopen.com/books/principles-in-contemporary-orthodontics/the-artificial-intelligence-approach-for-diagnosis-treatment-and-modelling-in-orthodontic>

INTECH
open science | open minds

InTech Europe

University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China

Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

© 2011 The Author(s). Licensee IntechOpen. This is an open access article distributed under the terms of the [Creative Commons Attribution 3.0 License](https://creativecommons.org/licenses/by/3.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

IntechOpen

IntechOpen