An Intelligent Fuzzy-Neural Diagnostic System for Osteoporosis Risk Assessment

Chin-Ming Hong, Chin-Teng Lin, Chao-Yen Huang, and Yi-Ming Lin

Abstract—In this article, we propose an Intelligent Medical Diagnostic System (IMDS) accessible through common web-based interface, to on-line perform initial screening for osteoporosis. The fundamental approaches which construct the proposed system are mainly based on the fuzzy-neural theory, which can exhibit superiority over other conventional technologies in many fields. In diagnosis process, users simply answer a series of directed questions to the system, and then they will immediately receive a list of results which represents the risk degrees of osteoporosis. According to clinical testing results, it is shown that the proposed system can provide the general public or even health care providers with a convenient, reliable, inexpensive approach to osteoporosis risk assessment.

Keywords—BMD, osteoporosis, IMDS, fuzzy-neural theory, web interface.

I. INTRODUCTION

Osteoporosis is a disease characterized by low bone mass and microarchitectural deterioration of bone tissue, leading to enhanced bone fragility and a consequent increase in fracture risk [1]. It is the major cause of wrist, vertebral, and hip fractures, and the associated morbidity and mortality in fracture risk [1]. It is the major cause of wrist, vertebral, and hip fractures, and the associated morbidity and mortality of osteoporotic fractures are very significant [2]. Nowadays, there are a variety of methods such as bone mineral density (BMD) tests, blood tests, urine tests, and other biomarkers analyses, available for diagnosing osteoporosis. Generally speaking, BMD tests are the most popular and important methods in osteoporosis screening.

BMD tests can determine the severity of bone loss which highlights the risk of developing osteoporosis, predict the risk of future fractures and can also be used to monitor changes in BMD [3]. Currently, there are several common technologies such as dual-energy x-ray absorptiometry (DEXA), quantitative computerized technology (QCT), quantitative ultrasound technology (QUS), radiographic absorptiometry (RA), and single x-ray absorptiometry (SXA), available for measuring BMD [4].

Applying such technologies, however, often encounters several problems such as long testing time, high examination fees, expensive equipment expense, and probable radiation exposure. Besides, all of these are carried out either in hospitals or in laboratory environments. These negative factors not only decrease their accessibility, but also increase the complexity of BMD examinations. As a result, it is difficult to extensively perform BMD screening for specific groups with high osteoporosis risk, for example, postmenopausal women over 50 years old or the elderly. According to the latest population statistics, the number of women over 50 years old is estimated to be around two million in Taiwan [5]. This situation means that by means of conventional BMD tests, mass screening for osteoporosis to the groups in Taiwan (or most of the developed or developing countries) is practically infeasible and not cost-effective.

The system was automatically constructed by learning available numerical training data comprised of several biomarkers including patient’s age, serum calcium, serum phosphates, estradiol, progestosterone, and other laboratory measurements. Utilizing hematic, chemical, and mineralometric parameters, and risk factors, Binaghi et al. [16] built a fuzzy medical expert system for the representation and manipulation of medical knowledge and strategies so as to detect postmenopausal osteoporosis. Furthermore, there was recently a web-based medical expert system constructed mainly based on current medical knowledge, statistical information and physicians’ opinions, capable of estimating the probabilities of three BMD statuses (normal, osteopenia, and osteoporosis) by means of various risk factors concerning osteoporosis [17].

In view of the present literature, however, it can evidently be seen that the realization of most of the computer-based diagnostic systems for osteoporosis has two limitations. First, some of the patient’s clinical signs (such as results of BMD tests [13], [16], or other invasive and laboratory measurements [12]–[16]) on which the systems rely for deducing diagnostic conclusions are difficult to be acquired outside hospitals or laboratory environments. As mentioned above, this not only decreases the accessibility of the systems to the general public, but also increases the cost of diagnosing osteoporosis such that mass screening using such systems is practically infeasible. Second, during construction process, some of the systems need a great amount of consultation with medical experts [16], [17]. This involves the acquisition of diagnostic knowledge from experts. In many medical domains, however, this knowledge may be incomplete, since the relationship between clinical signs and medical meanings are not always transparent — osteoporosis diagnosis is exactly the case.

To overcome the drawbacks mentioned above, in this paper, we propose a novel learning-driven system — Intelligent Medical Diagnostic System (IMDS) accessible through web interface — for performing on-line initial screening for osteoporosis. The fundamentals of the IMDS are mainly based on the well-known theory of fuzzy-neural networks (FNNs). FNNs possess the ability to automatically learn useful knowledge from available examples without the active participation of domain experts during the process of constructing system, and can exhibit superiority over other conventional technologies (e.g. statistical techniques and regression analyses) due to its superior nonlinear modeling.
capability. Through the internet, users simply answer a series of directed questions to the IMDS, and then they will immediately receive the results of osteoporosis risk assessment. According to the results of clinical testing, it is shown that the proposed system does provide the general public or even health care providers with a convenient, reliable alternative to the assessment of BMD and to the risk assessment of osteoporosis.

The remainder of this paper is organized as follows. Section II introduces a new four-layered FNN and then explains its role in medical diagnosis. The architecture of the IMDS and the proposed learning algorithm are described in Section III. In Section IV, we use a BMD database to construct the IMDS, and then evaluate its performance in osteoporosis risk assessment. Finally, Section V concludes the paper.

II. FUZZY-NEURAL NETWORK

A diagnostic process can be formalized by evaluating available clinical signs (information drawn from the patient’s history and from clinical and laboratory findings) and then deducing appropriate conclusions according to a set of diagnostic rules which describe the relationship between signs and diseases [16]. Diagnosis work can essentially be regarded as a classification task; in other words, a diagnostic system is practically equivalent to a classifier.

It is well known that FNNs possess the ability to automatically acquire useful knowledge from a large number of examples, and can exhibit superior nonlinear modeling capability in many fields such as pattern classification, object recognition, model prediction, and system identification. A FNN is named by the fact that a fuzzy model described by if-then rules and inference mechanisms can completely be transformed into a multi-layered artificial neural network (MNN) composed of neurons and connecting links as long as each neuron in a MNN performs a corresponding function in a fuzzy model. This fact makes standard gradient learning methods in MNNs available to adjust the parameters of FNNs [18-19], and further facilitates the design of learning algorithms for FNNs. In this paper, therefore, we employ a FNN classifier as the kernel of the IMDS to perform the diagnosis works.

It is worth noting that differing from conventional complicated five-layered FNNs, a new four-layered one is proposed here so that we can construct a diagnostic system with the properties of lower complexity and easier implementation. As shown in Fig. 1, the proposed FNN consists of the following four layers in sequential order: an input layer, a fuzzification layer, a rule layer, and a class layer. This architecture can completely represent a Mamdani-like fuzzy model which is described as the following rules:

Rule j: If \( x_i \) is \( A_j^{(i)} \) and ... and \( x_k \) is \( A_j^{(k)} \)

then the input \( x \) belongs to class \( i \). \( 1 \)

where \( j \) is a rule index, \( N \) denotes the number of input features, \( A_j^{(i)} \) represents the fuzzy set (linguistic term) in \( i \)th subpremise of rule \( j \), \( x=[x_1, x_2, ..., x_N] \) indicates the set of all input features, and \( i \) is a class label which is a crisp index rather than a fuzzy set. Adopting such a model has two main merits. On one hand, we can interpret the fuzzy model by means of natural language. This not only provides us some insights into the working of the diagnostic system, but can also facilitate the adaptation of design parameters and operating conditions of the system.

On the other hand, since the consequent part of each if-then rule only consists of a crisp class label, we do not require a complicated defuzzification method (such as center of gravity, or mean of maximum commonly used in Mamdani fuzzy models) to derive a crisp output for the result of fuzzy inference. Instead, we use the simplest winner-takes-all method to get the crisp output of the model. This merit largely simplifies the derivation of parameter learning algorithm as well as reducing the complexity of the model.

Considering Fig. 1, we subsequently explain in more detail how the proposed FNN can implement the Mamdani-like fuzzy rules and inference.

A. Representation of Fuzzy Rule

The input layer consists of input neurons \( k, k=1,2,...,N \), each of which is respectively associated with a feature \( x_i \) representing a clinical sign or finding. The fuzzification layer consists of membership function neurons \( \mu_{k,N} \), \( k=1,2,...,F_k \), each of which corresponds to a linguistic term \( A_k \) such as small, middle, large, or other terms. The rule layer consists of rule neurons \( j, j=1,2,...,R \), each of which together with all the neurons connecting to it represent a fuzzy diagnostic rule.

The class layer consists of class neurons \( i, i=1,2,...,C \), each of which corresponds to a diagnostic class such as normal, osteopenia, or osteoporosis in this paper. The meanings of the remaining symbols used in Fig. 1 are clarified as follows: \( N \) is the number of input features, \( C \) is the number of classes, \( R \) is the number of fuzzy rules, and \( F_k \) denotes the total number of fuzzy sets defined in the domain of the \( k \)th input feature; moreover, \( a_r \) denotes the activation degree of rule neuron \( j \), which is equivalent to the fulfillment degree of the antecedent of rule \( x_j \), and \( ar \) is the activation degree of class neuron \( i \), which corresponds to the aggregated fulfillment degree of class \( i \).

Taking the formation of equation (1) into account, we have
the following insights into the connections between the neurons in any two adjacent layers. The neurons in each layer of the FNN only connect to the neurons in the succeeding layer. In view of the connections between the class layer and the rule layer, each rule neuron can be connected with at most $N$ fuzzy set neurons, each of which originates from an individual domain. Moreover, each rule neuron merely connects to a class neuron; however, since different rules could possess the same consequent class label, each class neuron could be simultaneously connected by several different rule neurons (for instance, the class neuron 1 in Fig. 1). With reference to the connections between the rule layer and the fuzzification layer, each fuzzy set neuron could connect to several different rule neurons; however, two or more fuzzy set neurons originating from a single input domain can not connect to the same rule.

**B. Realization of Inference Mechanism**

In this subsection, we explain how the FNN realizes fuzzy inference mechanism by forward propagating input signal through the network. In the beginning, the input layer receives input feature data from the external environment, and directly passes them to the fuzzification layer in which each fuzzy set is described by a triangular membership function

$$
\mu(x,a,b,c) = \max \left[ \min \left( \frac{x-a}{b-a}, \frac{c-x}{c-b} \right), 0 \right] \tag{2}
$$

where $a$, $b$, and $c$ represent the $x$-coordinates of the lower bound, vertex, and upper bound of a membership function $\mu$, respectively. The output signals of the fuzzy set neurons subsequently propagate into the rule layer in which each rule neuron uses a nonparametric $T$-norm operator to compute the fulfillment degree of a fuzzy rule. In general, algebraic operators result in better approximation results than intersection operators; however, it was concluded by experience that intersection operators perform better in better approximation results than intersection operators; therefore, we prefer algebraic ones. The aggregated fulfillment degree

$$
\mu_i = \prod_{i=1}^{N} \mu_i^{(j)} = \min \left\{ \mu_i^{(j)} \right\} \tag{3}
$$

where $\mu_i^{(j)}$ are membership functions associated with the fuzzy sets $A_i^{(j)}$ in the antecedent of rule $j$, and the fuzzy set neuron that wins the competition (i.e., possesses the highest fulfillment degree) is called a winning fuzzy set neuron. Note here that we assume the number of input features in the antecedent of rule $j$ is $N$; however, if we use some simplifying strategies (e.g., the feature and partial antecedent deletion presented in Section III) to improve the interpretability of the FNN, the number of input features is usually less than $N$.

In the class layer, each class neuron receives the signals originating from the rule layer, and then is responsible for aggregating the fulfillment degrees of the rules with the same consequent by means of some nonparametric $T$-conorm operator. Similarly, for interpretability, we prefer union operators to algebraic ones. The aggregated fulfillment degree of class $i$ can therefore be computed by

$$
ac_i = \prod_{j=1}^{N} ar_i^{(j)} = \max_{i=1,2,...,R} \left\{ ar_i^{(j)} \right\} \tag{4}
$$

where $ar_i^{(j)}$ stand for the outputs of all the rules with consequent class $i$, and $R$ indicates the number of such rules. Likewise, the rule neuron that wins the competition (i.e., possesses the highest fulfillment degree) is named a winning rule neuron. Finally, the output signals $ac_i$, $i=1,2,...,C$, represent the diagnostic responses (inference results) produced by the FNN with respect to some input feature vector $x$.

As for the complicated defuzzification methods commonly used in conventional FNNs, we ignore them in the proposed FNN; instead, we regard the class with the highest fulfillment degree as the most probable class to make a crisp interpretation for the diagnostic results.

**III. INTELLIGENT MEDICAL DIAGNOSTIC SYSTEM**

The IMDS system is mainly comprised of three parts: a learning module, a knowledge base, and a web-based user interface. The architecture of the system is shown in Fig. 2. The objective of the learning module is to construct a FNN classifier not only with high classification accuracy but also with good interpretability. This objective is accomplished by means of a series of learning algorithms including k-means clustering, gradient descent parameter tuning, and similarity-based model simplification. The knowledge base is practically a FNN classifier. It consists of a fuzzy inference engine and a fuzzy medical rule database. The former performs fuzzy inference with respect to the presented input feature vectors so as to provide reasonable diagnostic results; the latter contains a collection of diagnostic rules acquired from available medical databases which store a great amount of data relating to patient’s records or medical history. The web-based user interface provides users with a convenience path to access the service offered by the IMDS. Through this interface, users can submit their personal information and receive the related diagnoses in an easy manner. As the principles underlying the proposed FNN have been discussed in the previous section, we only introduce the other two parts in the following subsections.

**A. Initial Rule Learning Algorithm**

We apply the $K$-means clustering algorithm, which is one of the simplest algorithms in the field of pattern recognition, to acquire representative information from the given patient’s records or medical history. Each of the obtained clusters stands for a prototype for a particular behavior of the medical
knowledge under consideration; that is, each cluster can be used to define a diagnostic rule. Moreover, directly projecting each cluster onto individual inputs can obtain initial membership functions for each rule [19]

B. Parameter Learning Algorithm

The parameter learning algorithm is developed based on the well-known gradient descent method. According to our previous works [22], the algorithm is summarized as follows. Let $d_i$ be the desired class index, $e_i = d_i - a_i$, $c_i$ indicate the class error, and learning rate $\eta > 0$. The parameter vector $\theta_{\mu_i} = [a_{\mu_i}, b_{\mu_i}, c_{\mu_i}]^T$ of the membership function $\mu_{\mu_i}$ can be updated by the well-known delta learning rule:

$$\theta_{\mu_i} = \theta_{\mu_i} + \eta \Delta \theta_{\mu_i}$$  

(5)

where

$$\Delta \theta_{\mu_i} = \begin{cases} 
    e_i \frac{\partial \mu_{\mu_i}}{\partial \theta_{\mu_i}}, & \text{if } a_i = a^*_i = \mu_{\mu_i}, \\
    0, & \text{otherwise}
\end{cases}$$  

(6)

and

$$\frac{\partial \mu_{\mu_i}}{\partial \theta_{\mu_i}} = \begin{cases} 
    v_i, & \text{if } x_i < b_{\mu_i} \\
    \frac{v_i}{b_{\mu_i} - b_{\mu_i}}, & \text{if } x_i < b_{\mu_i}
\end{cases}$$  

(7)

and

$$v_i = \begin{bmatrix} (x_i - b_{\mu_i}), (x_i - a_{\mu_i}), 0 \end{bmatrix}^T$$  

(8)

Considering equation (3), we can find that in each time step, only the parameters of the membership function embedded in the winning fuzzy set neuron connected with the winning rule neuron should be adjusted; in other words, for a $C$-class classification problem, the number of parameter vectors that in each time step should be adjusted is less than or equal to $C$. This fact exhibits that the truncation operators employed in equation(3) and equation (4) do effectively reduce the computational burden those algorithms employing algebraic operators will encounter.

C. Model Simplification Algorithm

We have developed a strategy to simplify the fuzzy diagnostic rules of the IMDS in order to enhance the interpretability of the obtained rules [23]. Based on the set-theoretic similarity measurement derived by the well-known determinantal formula of area, the strategy consists of three steps: 1) fuzzy set mergerge, 2) antecedent feature pruning, and 3) rule number simplification. According to these steps, we simplify the fuzzy medical rule base step by step, and meanwhile enhance its interpretability and readability. It is worth noting that the set-theoretic similarity measurement proposed in the paper is different from the conventional ones, since it unifies the procedure for computing intersection area between two fuzzy sets. Fig. 3 shows the entire flowchart of the model simplification algorithm.

D. Web-Based Interface Implementation

In order to enhance the accessibility and convenience of the IMDS, we design a web-based user interface to facilitate the osteoporosis risk assessment for the general public. Taking advantage of the internet, users simply answer a series of directed questions to the IMDS, and then they will immediately receive the assessment results.

The IMDS server contains a webpage database which stores all of the webpage design data, a fuzzy medical rule database which collects all of the acquired diagnostic rules and parameters, and a web-based application which integrates the FNN classifier and all of the algorithms mentioned above. The IMDS uses client-server architecture to carry out diagnosis service. Through common web browser software, client users can send data requests to the IMDS servers, and then the servers can accept these requests, process them, and return the requested information to the clients.
gent Medical Diagnostic System is freely available to access via the IMDS website http://140.122.79.60/IMDS at National Taiwan Normal University (NTNU), Taipei, Taiwan, R. O. C.

**IV. EXPERIMENTS OF OSTEOPOROSIS RISK ASSESSMENT**

In the following, we will use a bone mineral density (BMD) data set to evaluate the performance of the IMDS. The BMD database was provided by Prof. Shu-Fang Chang from National Taipei College of Nursing, Taipei. The data set contains 274 cases and 29 of these cases have missing values. Since the IMDS can not yet deal with missing values, we only used 245 cases from the complete data set for evaluating the IMDS. We use 200 cases for training and 45 cases for testing the system. The data is randomly split such that each set contains roughly the same number of patterns for corresponding classes.

The standard of osteoporosis is defined by the World Health Organization (WHO) in terms of BMD measured by DEXA scanning [1], [3]. The result of a BMD test is represented as a T-score derived by comparing one’s BMD to that of young normal mean value. A T-score greater than or equal to –1 is in the normal range. A score between –1 and –2.5 is considered osteopenia, and any score lower than –2.5 is regarded as osteoporosis. According to this standard, we divide the 245 cases into three classes. Each case belongs to one of the three classes; as a result, normal has 66 cases, osteopenia has 73 cases, and osteoporosis has 106 cases. Moreover, since the 19 features contain both numeric and nominal variables, we need to preprocess them so that the proposed algorithms can successfully perform. First, since the first three features respectively represent three individual numeric variables, we normalize their values into [0, 1] interval. Second, as the last 16 features are nominal variables, we define each positive answer as 1, and each negative answer as 0.

The proposed algorithms are applied to automatically construct a FNN classifier from the BMD dataset. We designate the number of clusters as three so as to acquire three initial diagnostic rules, each of which respectively describes a single BMD class. After 300 iterations for parameter learning, the resultant FNN classifier generates 52 misclassifications (74% correct) on 200 training cases, and 14 misclassifications (69% correct) on 45 testing cases. This leads to a recognition rate of 73% (66 misclassifications) on the total data set. The obtained membership functions are shown in Fig. 5. Note here that the blue graphs indicate the antecedent membership functions of the rule describing normal situation, green graphs denote the membership functions of the rule describing osteopenia, and red graphs indicate the membership functions of the rule describing osteoporosis.

Due to heavy overlap between fuzzy sets in Fig. 5, however, it is obviously difficult to assign an appropriate linguistic term to each fuzzy set. We further utilize the model simplification algorithm to enhance the interpretability of the fuzzy sets. After model simplification, the classifier generates 58 misclassifications (71% correct) on 200 training cases and 15 misclassifications (66.7% correct) on 45 testing cases. This leads to a recognition rate of 70.2% (73 misclassifications) on the total data set. Fig. 6 shows the simplified fuzzy sets. Although the classification accuracy decreases about 4%, it is evident that the latter has much better interpretability than the former. Considering both accuracy and interpretability, we employ the simplified FNN as the kernel of the IMDS.

Finally, we gather additional 10 cases for clinical testing, among which normal has two cases, osteopenia has five cases, and osteoporosis has three cases. After presenting all the cases into the IMDS, the system generates only two misclassifications (80% correct). Figs. 7 and 8 show one of the cases and its results of risk assessment, respectively. Given the T-score of this case is –2.9, as shown in Fig. 8, the IMDS evidently provides correct results of osteoporosis risk assessment.

![Fig. 5 FNN after parameter learning](image)

![Fig. 6 FNN after model simplification](image)
In summary, the advantages of the IMDS are four-fold: (1) automatically generating diagnostic rules without medical experts’ active participation can significantly reduce the construction cost of the system; (2) the learning-from-example ability of the IMDS can extract implicit, previously unknown, and potentially useful medical knowledge from considerable patient’s data or medical history; (3) using a questionnaire of osteoporosis risk factors rather than other invasive methods or laboratory measurements to assess osteoporosis risk not only can considerably reduce the cost of mass screening, but can also speed up screening process; and (4) the convenient, interactive, web-based user interface can effectively increase the accessibility of the system.

Moreover, it should be noted that the reliability of the system mainly depends on the information supplied by the users; in other words, entering incorrect or misleading information will cause erroneous or unreliable diagnosis results.

REFERENCES