Design and Development of an Expert System in Differential Diagnosis of Maxillofacial Radio-lucent Lesions

Afshin Ameri D.D.S.
Malardalen University, Sweden
aai08001@student.mdh.se

Hessam Moshtaghi D.D.S. *
Shahid Beheshti University, Iran

ABSTRACT
In this paper we discuss different approaches in designing of Medical Diagnosis Expert Systems with focus on Maxillofacial region and present a new method for developing such systems. The influence of unreported symptoms of a lesion is an important factor in Medical Diagnosis, which is ignored in current designs; Also the low number of selected features and the low number of diagnosable lesions are two main disadvantages of current approaches in today’s systems. Our Default Sum criterion tries to overcome these problems while defining a new algorithm which allows for numerous amounts of lesions and their respective symptoms. Then we introduce CDSX, a differential diagnosis expert system which is based on our algorithm. The system is still under training, but is being used for educational purposes in Shahid Beheshti University of Medical Sciences.

1. INTRODUCTION

It is estimated that preventive medical errors are the main cause of 98000 deaths each year just in the US,[11] If this level of proficiency was applied in airline and banking industries it would be enough for two dangerous airplane landings at airports and 32000 wrong banking transactions per hour.[13] Some common examples of Medical Errors involve misdiagnosis, giving the wrong drug, giving two or more contradictory drugs, wrong site surgery and Gossypiboma (a surgical sponge left in body of patient after surgery). Autopsy examinations have revealed that 35%-40% of these errors are due to misdiagnosis. [13]

The whole process of treatment depends on correct diagnosis of a condition. Diagnosing a disease or lesion requires years of experience in medicine, especially when a patient is showing symptoms of rare conditions that physician may have never encountered before or has limited knowledge about.[8] Several factors may lead to misdiagnosis:

Fatigue related errors. Stress and high workload which are characteristics of a physician’s job may lead to fatigue which in turn may cause error in human cognitive process through distraction, confirmation bias, memory bias, oversimplification, premature closure and short term memory lapses.[13][15]

Inadequate Knowledge. Which may be a result of inadequate experience (i.e. interns) or rarity of the condition.[12]

Human performance. Which is affected by environmental conditions, psychological, perceptual and cognitive processes.[10]

With the processing and storage abilities of computers, it is very useful to develop an expert system, which can help physicians with their diagnosis. Expert systems can help physicians by informing them about unrecognized information needs of a diagnosis, standardizing diagnostic and treatment procedures and even as a training tool with detailed information about symptoms, conditions and diagnosis.[8]

Medical diagnosis expert systems have been an interesting topic for many researchers since introduction of MYCIN [17] in 1970 decade. They have been used in psychiatric treatment [7], prostate cancer diagnosis [3], lung disease diagnosis [1], oncology [2] and even selecting surgical candidates [16]. There are some expert system in Maxillofacial diagnosis too, most notably; ORAD which is a Bayesian belief network oral lesions diagnostic expert system. [19] Some other oral expert systems are COMRADD [9] and ICOHR [14].

Usual approaches in designing expert systems consist of rule-based expert systems, Bayesian belief networks, fuzzy logic expert systems and artificial neural networks. We will have a short overview of each approach and discuss their weak and strong points.

Rule-based Expert Systems are the first approaches taken in developing a medical expert system. However their mandatory questions which are crucial for diagnosis, put limitations on their usage.[8]
Artificial Neural Networks which resemble a raw representation of human brain have shown great potential in diagnosing medical conditions.\[4\] These kind of expert systems present good results for problems which cannot be formalized very well.\[6\] The major problem in developing a medical decision support neural network is its dependency on large number of training cases which are required to gain a good diagnostic ability.\[8\] These large number of training cases may not always be available.

2. METHODS

2.1 System Architecture

Most of current medical expert system implementations are limited to few number of lesions–sometimes just one special condition– and symptoms. This makes them impractical for use in a real medical or educational facility. In the other hand some expert system designs have an unrealistic approach to decision making by medical experts; Although it may seem that diagnosis of a lesion can be done by a set of rules and facts, it is not always applicable. In medical society it is always acceptable that by a given set of symptoms, two different experts may come to two different conclusions. Even some difference between highly accredited medical text books is normal in the field.

In order to make a more practical diagnosis system with respect to its medical and educational use we concluded that the system should have the following characteristics:

1. Be applicable to numerous lesions and symptoms.
2. Can use all the symptoms in its decision making regardless of its importance in diagnosing a specific lesion.
3. Can suggest its differential diagnosis based on unknown information of a patient (i.e. laboratory test results which are not ready at the time).

We selected radio-lucent lesions of the jaw bones as the system’s field of expertise, because radio-lucent lesions are more common than other types of radiographic lesions in the jaw bones and because some of them have similar symptoms which makes them difficult to diagnose; also there are some rare conditions in this set too. This set of lesions covers over 140 lesions in the jaws according to Wood & Goaz reference book in Dentistry[21].

One of the biggest obstacles that we had in our way, was little amount of real patient information available to us; Because of weak patient screening system, there was little patient and pathologic finding data available to be used in implementation of the system. Therefore we could not use any learning algorithm to develop the system. In the other hand, the number of lesions to be diagnosed by the system made it impossible to choose rule-based systems as the main approach in system development. Considering simplicity of implementation and inadequate training data we decided to use a simplified version of neural networks to develop the system.

2.2 Sum Criterion

Sum Criterion is the heart of our system. It helps to calculate a "Score" for each lesion in the system. Based on these scores, we can choose those with higher ranks as lesions which should be considered in differential diagnosis. In order to define this score, let’s start with some few basic definitions. First consider L as the set of all lesions and F as the set of all diagnostic features and $l \in L \land f \in F$. We define a value $DV(l, f)$ which shows diagnostic value of feature f in diagnosing lesion l. Now consider PF as a set of all the features which are present in a patient. It is obvious that:

$$PF \subseteq F$$

Now given a set of all present symptoms in a patient we can calculate "Score" as follows:

$$Score(l) = \sum_{i=1}^{n} DV(l, f_i) \text{ where } f_i \in PF$$

Please note that this definition can be easily considered as a simple neural network with no hidden layers and the diagnostic features as inputs and lesions as outputs. $DV(l, f_i)$ simply defines the weight of each edge in this simple neural network.

To build F set we decided to gather all the information from dentistry text books and several experts in the field. All the symptoms for selected lesions where gathered from well-known Maxillofacial diagnosis reference books[21][20][18]. In order to gather experts ideas, two Maxillofacial Diagnosis experts and two Maxillofacial Radiology experts were provided with a list of lesions and were asked to write the symptoms they value for diagnosing the lesions. The list features were then categorized in 28 different groups. Using Microsoft Access 2000 we built a database and placed each diagnosis group in a separate table.

The next step was to score each feature against each lesion, those features which were not considered in diagnosis of a lesion were scored zero. Other features were scored based on their importance as it was mentioned by the reference books. Score values were selected from a range of 0 to 100 with two exceptions: Pathognomonic features of a lesion...
were assigned a very high value (1000) and those features which have negative effect in diagnosis of a lesion were assigned negative value of -100. These crude weights were converted to charts and handed to the expert cooperating with the project. After various meetings and discussion with the experts, diagnostic values were refined.

2.3 Results with Sum Criterion

We created the main diagnostic application using Microsoft Visual Basic 6 and Access 2000. Because the number of real world cases to test the system was not applicable (12 cases), we decided to use fabricated patient data for testing. We asked three experts, which were not a part of data analyzing, to deliver us with some fabricated cases, we should mention that these cases were all from a subset of the lesions which contained 12 different conditions. From the 60 different cases that we collected the system was successful in 62% of the cases.

Although a 62% rate is not bad for an untrained system, it was strange that the wrong results happened for some lesion more. We decided to search for possible errors in our work and after various meetings with the experts we came to the conclusion that the main cause of this result is that the system does not consider unpresent features of a lesion in patient in its calculation. Clearly speaking there may be some features which have low importance in diagnosing a lesion, but if they are not present, they have a bigger negative effect on diagnosing that lesion.

2.4 Default Sum Criterion

In order to calculate the effect of unpresent features in our diagnosis. We refined the Sum Criterion. We mentioned $DV(l, f)$ and $PF$ before (2.2), now we define some new terms: first let $UF$ be the set of all features which are not present in the patient and

$$UF \subseteq F, \quad UF \cap PF = \emptyset$$

Also let $DU(l, f)$ be diagnostic value of $f$ in diagnosing $l$ if $f$ is not present in the patient. Now we can define a lesion’s “Score” as follows:

$$Score_{dsc}(l) = \sum_{i=1}^{n} DV(l, f_i \in PF) + \sum_{j=1}^{m} DU(l, f_j \in UF)$$

This new score function meant that another round of data collection was necessary. Therefore, we gathered new information for diagnostic value of unpresent features. But in order to reduce the development time, we reconfigured the last equation by assuming that $PF \cup UF = F$. Based on this assumption and the following definitions:

$$DS(l) = Score_{dsc}(l) \quad \text{where} \quad PF = F$$

$$\Delta(l, f) = DV(l, f) - DU(l, f)$$

It can be proved that

$$Score_{dsc}(l) = \sum_{i=1}^{n} \Delta(l, f_i) + DS(l)$$

This revised version of score calculation allowed us to develop the system with minimal changes in database design. However we accept that our assumption makes the system less practical, but in this step our main priority was to see the effect of unpresent symptoms that we had added to the system.

3. RESULTS

The final system which we call it CDXS (Comprehensive Maxillofacial differential Diagnosis eXpert System - Fig.1) contains 143 different lesions of the jaw-bones which have a radio-lucent radiographic image and 337 different diagnostic features. Diagnostic features are categorized in 28 groups to ease data entry for the user (Fig.2). The first test results which were conducted against 50 new fabricated cases for another subgroup of the lesions showed 82% accuracy with the new system and 64% accuracy with the previous system. CDXS is going under real case examinations right now.

![Figure 1: CDXS Start Page](image)

In order to show you an example of CDXS diagnosis, we will consider a female 19 years old patient with small swelling on her right canine tooth area, which is a little bit rubbery when touched by hand. The tooth seems a little bit out of its place but is vital. The radiographic appearance of the lesion shows a radiolucency which has an oval shape and sharp borders.

The above definitions should guide an expert to diagnosis of a Dentigerous Cyst, but it may be misdiagnosed by a young intern with Traumatic Bone Cyst or Pericoronal Abscess. 1 For a full review of these lesions and their differential diagnosis please refer to [21] chapter 18.
but we omitted the aspiration\textsuperscript{2} result in order to give less information to the system. As you may have noticed, all the symptoms which are selected, are those who do not need any special kind of tests or knowledge and can be achieved at the first visit of the patient. The system results are shown in figure 3.

Please note that, Paradental Cyst and Eruption Cyst are special kinds of Dentigerous Cyst which happen at non tooth bearing areas, and that’s why they are at lower ranks. Also, as it seems obvious, Circumferential/Lateral Dentigerous Cyst is another subtype of Dentigerous Cyst. Finally it is noticeable that the nearest candidate for misdiagnosis which is Traumatic Bone Cyst stands at the 5\textsuperscript{th} rank.

The final version of the system, does not report any diagnostic score values to the user, because we did not want to give any prejudice to the doctor about the lesions which should be considered for differential diagnosis.

As a result of using rough values for training the system, it needs to be trained to reach its peak diagnostic ability. Therefore it is now installed on Shahid Beheshti University computer network and currently performs as a source for training undergraduate students with their educations. Each new case also acts as a training example for the system too, and we hope that with these training, through time we can get better performances from the system.

\textsuperscript{2}Using a needle to extract the inner fluids of a lesion and examine them.

4. SUMMARY AND CONCLUSION

Because of the weak patient screening system, it is almost impossible to build, train or test these kind of systems by using real patient symptoms and diagnosis. This was the main reason that we decided to develop CDXS mainly as a educational training tool. But in places where large amount of patient data is available, using data mining techniques will surely improve data acquisition efficiency and development time of the system.

As we mentioned in the beginning of this paper, medical diagnosis is not as straightforward as it seems. Our results show that sometimes the importance of unpresent symptoms can have a great effect on outcome result of a diagnosis by an expert. In differential diagnosis there are some features of a lesion which if present do not have great diagnostic value to an expert, but if they are not present then the expert may out rule a disease. We believe that the feature medical expert systems should take this factor into consideration to improve their diagnostic capabilities.

5. ACKNOWLEDGEMENTS

We gratefully acknowledge Prof. Ahmad Ghaem Maghami, Dr. Amin Tavakoli and Dr. Aghdasi for their great cooperation with this study.

6. REFERENCES

Figure 3: Results for a 19 years old female with a rubbery swelling on right canine. Tooth displacement and well-defined oval radiolucency are also present.

results of a simulation test with actual clinical cases. 

Academic Radiology, 1, Jan 2004.


