

DESIGN OF CONSISTENT SYSTEM FOR RADIOLOGISTS TO SUPPORT BREAST CANCER DIAGNOSIS

Boris Kovalerchuk,

Department of Computer Science, Central Washington University, Ellensburg, WA 98926-7520, USA
borisk@cwu.edu

Evgenii Vityaev,

Institute of Mathematics, Russian Academy of Science, Novosibirsk, 630090, Russia
vityaev@math.nsk.ru

James F. Ruiz

Department of Radiology, Woman's Hospital, Baton Rouge, LA 70895-9009, USA

ABSTRACT

The overall purpose of this study is to develop a prototype radiologic consultation system. The system should provide a second diagnostic opinion based on similar cases, incorporating the experience of many radiologists, their diagnostic rules and a data base of previous cases. The system allows a radiologist to enter the description of a particular case using the lexicon such as BI-RADS of American College of Radiology and retrieve the second diagnostic opinion (probable diagnosis) for a given case. The system also allows a radiologist to get more important information in comparison with known Computer-Aided Diagnostic (CAD) systems. These advances are based on a new computational intelligence technique.

We implemented a rule-based prototype diagnostic system. The diagnosis is based on the opinions of radiologists in combination with the statistically significant diagnostic rules extracted from the available data base.

1. INTRODUCTION

We develop a system which should meet following requirements. The radiologist can retrieve the second diagnostic opinion (probable diagnosis) for a given case with:

- (1) BI-RADS description of similar cases;
- (2) diagnostic opinions of other radiologists about these cases;
- (3) clinical and pathological data for these cases and
- (4) potential for viewing digital reproduction of the

A combination of these technological innovations can significantly improve the effectiveness

sample mammograms.

In an advanced mode the radiologist should be able to analyze:

(5) diagnostic rules from a Diagnostic Rule Base, applicable for a given case and the system should deliver a probable diagnosis. These rules should be understandable by any radiologist without sophisticated knowledge of the mathematical methods inside of the system;

Also a radiologist should have the option to enter his/her diagnosis and to obtain:

(6) a comparison of his/her diagnosis with simulated opinions of other radiologists.

(7) rationale of diagnostic rules by these experienced radiologists;

(8) significance of rules from statistical perspective.

(9) a comparison of radiologist's diagnostic opinion with data-based diagnosis. (This diagnosis is inferred from rules discovered by a "data mining" software from data base. This breast cancer data base is developed as a part of our system and is open to incorporate other data bases).

Next the radiologist should have the option for more sophisticated interaction with the Consultation System. The system should be able:

(10) to interview him/her to extract his/her diagnostic rules

(11) to compare the set of the radiologist's diagnostic rules with rules of other radiologists

(12) to compare the set of the radiologist's rules with rules extracted from a data base

(13) to show cases from the data base with pathologically confirmed diagnosis which may contradict his/her diagnostic rules.

of breast cancer diagnosis and open new perspectives in telemedicine.

2. BACKGROUND

In the U.S. breast cancer is the most common cancer in women, with an estimated 182,000 cases in 1995 [Wingo, et al., 1995]. The most effective tool in the battle against breast cancer is screening mammography. However, it has been found by many authors that intra- and inter- observer variability in mammographic interpretation is significant (about 25%). These data show that there is a problem with the reliability of mammographic interpretation and clearly demonstrate the need to improve the reliability of these interpretations.

There are Computer-Aided Diagnostic (CAD) systems based on neural networks, nearest neighbor methods, discriminant analysis, cluster analysis, linear programming based methods and genetic algorithm ([F. Shtern, 1996], [SCAR, 1996], [TIWDM, 1996] and [CAR, 1996].

3. COMPONENTS OF SYSTEM

Below we present a list of components of the consultation system and their major functions.

Component 1. Data/Image Base:

Supports extracted features of mammograms,
Supports patient clinical records,
Supports digital images of mammograms

Component 2. Diagnostic Rule Base:

Support rules extracted from Data Base
Support rules obtained by interviewing of radiologists.

Component 3. Diagnosis simulator:

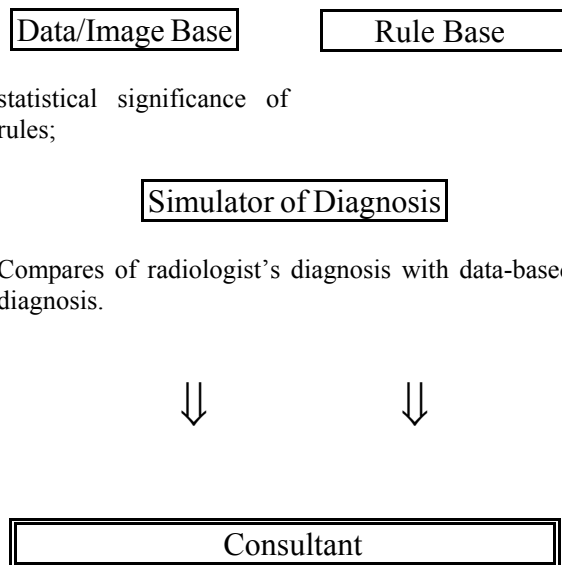
Generates diagnosis for a particular case based on the Diagnostic Rule Base.

Component 4. Consultant:

Retrieves simulated diagnosis for a given case;
Retrieves similar cases (clinical and pathological data for these cases);
Retrieves simulated diagnosis of other radiologists for similar cases;
Displays digital reproduction of the mammograms from the Data Base;
Retrieves diagnostic rules applicable for a given case from the Rule Base;
Compares user's (radiologist's) diagnosis with simulated opinions of other radiologists;
Retrieves rationale of applicable diagnostic rules by

Techniques used to find similar cases (property (1)) can be different from one CAD system to another. Each CAD system can be enhanced to have this property. We use a rule based approach in which a

experienced radiologists from the Rule Base; Retrieves



statistical significance of rules;

Compares of radiologist's diagnosis with data-based diagnosis.

Figure 1. Components of Consultation System.

4. METHOD

A radiologist enters into the CONSULTATION SYSTEM the description of a particular case. This can be done, for example in terms of the BI-RADS lexicon of American College of Radiology. Using a CAD system, the radiologist retrieves on the screen the second diagnostic opinion (probable diagnosis) for a given case. There are many promising CAD approaches [SCAR, 1996], [TIWDM, 1996] and [CAR, 1996].

Diagnostic opinions of other radiologists about similar cases are directly stored in the data base and will be available on the Internet.

A radiologist may have a local CD ROM collection of similar cases (e.g. 12 CDS are currently available from Livermore Lab. and University of California. It is also possible to obtain locally mammographic images from Internet Mammographic Data Base (University of South Florida) and other growing sources (University of Chicago).

diagnostic rule applicable to the study can be identified in the knowledge base. Then all cases in the data base are retrieved for which the premise of the rule is true. For example for a rule: "IF x and y then z", all cases

with x and y are displayed for the radiologist.

An advanced mode is designed to allow the radiologist to analyze diagnostic rules from a Computer-Aided Diagnostic system. These rules are **applicable** for a given case, and used by the system to deliver a probable diagnosis. These **rules are understandable** by any radiologist without sophisticated knowledge of the mathematical methods inside of the CAD system.

This analysis can become the most effective way for radiologists to share their experience to improve reliability of interpretation. These rules are stored in the rule base. The method to find these rules in rule base is based on comparison of premises of rules and a case entered by a radiologist into the system. Also a radiologist is able to enter his/her diagnosis for a studied case and to obtain a comparison of his/her diagnosis with computer simulated opinions of other radiologists for the same case. (Consultation system does not have actual opinions of other radiologists for this case, but using their rules, the system can simulate the opinion, i.e. the consultation system will apply diagnostic rules of other radiologists for this case description.

The user also has access to:

- the rationale of diagnostic rules by these experienced radiologists;
- the significance of rules from statistical perspective.
- a comparison of radiologist's diagnostic opinion with data-based diagnosis.

There are no serious obstacles to achieving these goals. The technique for obtaining the rationale behind the rules used by radiologists has been developed. Also we use the Consultation System to broaden the base of experience in this part of the data base.

Determining **statistical significance** is difficult for many CAD systems [Kovalerchuk *et al*, 1996]. The most popular methods are based on Neural Networks, but do not have a mechanism to evaluate statistical significance of diagnosis. Therefore the reliability of diagnosis is based only on the performance on training and testing data [Gurney, 1994]. These populations may or may not be sufficiently representative for the entire population [Kovalerchuk *et al*, 1996]. We use an original method [Vityaev, Moskvitin, 1993], which allows the derivation of diagnostic rules and evaluates statistical significance of

To restore all diagnostic rules thousands of questions might be needed if questions are not specially organized. For 11 diagnostic features of clustered calcifications there are ($2^{11}=2,048$) feature combinations, representing cases. The questioning

these rules. Examples of rules extracted from 156 cases (77 malignant and 79 benign) are given below. A radiologist may check his/her diagnostic opinion by comparing this opinion with the diagnosis made with rules derived from the data base of the Consultation System. The system diagnosis is inferred from rules discovered in the data base by a "data mining" software [Vityaev, Moskvitin, 1993]. The breast cancer data base is developed as a part of consultation system and is open to incorporate other data bases.

The interview of a radiologist to extract rules is realized using an original method [Kovalerchuk *et al*, 1996 a,b] and the comparison of rules is performed by translating the rules into a monotone Boolean functions and then comparing these functions [Kovalerchuk *et al*, 1996 a,b]. The demonstration of cases is most important and is realized by comparing the simulated diagnosis of a given radiologist and pathologically confirmed diagnosis in the data base.

5. RESULTS

Diagnostic Rules Acquisition. Examples of diagnostic rules extracted in a pilot study are presented below. Expert diagnostic rules were extracted from specially organized interviews of a radiologist (J. Ruiz, MD). For details of the method see [Kovalerchuk, *et al*, 1996 a,b]. This method is based on the theory of Monotone Boolean Functions and hierarchical approach. One of the extracted rules is presented below. RULE 1: IF NUMBER of calcifications per cm^2 (w_1) is large AND TOTAL number of calcifications (w_3) is large AND irregularity in SHAPE of individual calcifications is marked THEN highly suspicious for malignancy. The mathematical expression for this rule is $w_1 w_3 y_1 = \Rightarrow$ "highly suspicious for malignancy".

In this study of calcifications found on mammograms we used the following features: 1) the number of calcifications/ cm^2 ; 2) the volume (in cm^3); 3) total number of calcifications; 4) irregularity in shape of individual calcifications; 5) variation in shape of calcifications; 6) variation in size of calcifications; 7) variation in density of calcifications; 8) density of calcifications; 9) ductal orientation; 10) comparison with previous exam; 11) associated findings.

procedure required only about 40 questions. i.e. 50 times fewer questions than the full set of feature combinations [Kovalerchuk *et al*, 1996 a,b]. Note that practically all studies in CAD systems derive diagnostic rules using significantly less than 1,000 cases [Gurney,

1994]. This is the first attempt to work with such a large number of cases (2,000).

Diagnostic Rules extracted from Data Base. A study used 156 cases (77 malignant, 79 benign), described with 11 features of clustered calcification listed above and extended with two features: Le Gal type and density of parenchyma with the diagnostic classes: "malignant" and "benign".

With Logical Analysis of Data method [Vityaev, Moskvitin, 1993] 44 statistically significant diagnostic rules were extracted with the conditional probability greater than 0.7. There were 30 regularities with the conditional probability greater than 0.85, 18 rules with conditional probability more than 0.95. The total accuracy of diagnosis was 82%. The False/negative rate was 6.5% (9 malignant cases were diagnosed as benign) and false/positive rate was 11.9% (16 benign case were diagnosed as malignant). For the 30 more reliable rules we obtained 90% total accuracy, and for the 18 most reliable rules we obtained 96.6% accuracy with only 3 false positive cases (3.4%). Neural Network ("Brainmaker") software had given 100% accuracy on training data but for Round-Robin test the total accuracy fell to 66%. The main reason for this low accuracy is that NN *do not have a mechanism to evaluate statistical significance /reliability of the performance on training data*. Poor results (76% on training data test) were also obtained with Linear Discriminant Analysis ("SIGAMD" software). Decision Tree approach ("SIPINA" software) has performed with accuracy of 76%-82% on training data. This is worse than what we obtained for the LAD method with the much more difficult Round-Robin test. The extremely important false-negative rate was 3-8 cases (LAD), 8-9 cases (Decision Tree), 19 cases (Linear Discriminant Analysis) and 26 (NN).

Note also that only LAD and decision trees produce diagnostic rules. These rules make a CAD decision process visible to radiologists. With these methods radiologists can control the decision making process. Linear discriminant analysis gives an equation, which separates benign and malignant classes. For

example, $0.0670x_1 - 0.9653x_2 + \dots$ represents a case. How would one interpret the weighted number of calcifications/cm² (x1) plus weighted volume (cm³)(x2)? There is no direct medical sense in this formula.

6. CONCLUSION

Our study has shown that used Logical Data analysis approach is appropriate for designing a consultation diagnostic system under requirements presented in section 1. This approach can be used for development of a full-size consultation system.

REFERENCES

1. Kovalerchuk, B., Triantaphyllou, E., Despande, A., Vityaev, E. Interactive learning of Monotone Boolean Functions, *Information Systems*, Vol. 94, Issue 1-4, 1996, pp. 87-118.
2. Kovalerchuk B., Triantaphyllou, E, Ruiz, J. Monotonicity and logical analysis of data: a mechanism for evaluation of mammographic and clinical data. In: *Computer applications to assist radiology*, Symposia Foundation, Carlsbad, CA, 1996, 191-196.
3. *Third International Workshop on Digital Mammography*, 1996, University of Chicago.
4. Shtern, F. Novel digital technologies for improved control of breast cancer. In: *Computer Assisted Radiology*, Paris, June 1996, Elsevier, NY, 1996. pp.357-361.
5. Wingo, P.A., Tong, T, Bolden, S. Cancer Statistics, *Ca-A Cancer Journal for Clinicians*, v. 45, n. 1, pp. 8-30, 1995.
6. *CAR'96, Computer Assisted Radiology*, International Conference, Paris, 1996, Elsevier.
7. Vityaev, E., Moskvitin, A. Introduction to discovery theory. Discovery: software system. *Computational Systems*, Institute of Mathematics, Novosibirsk, 1993, n. 148, pp.117-163 (in Russian).