EO-2

Computer-Aided Diagnosis in Medical Imaging
K. Doi, H. MacMahon, M.L. Giger and K.R. Hoffmann, editors

Computer-Aided Diagnosis and its Potential Impact on Diagnostic Radiology

Kunio Doi
Department of Radiology, The University of Chicago, Chicago, IL, USA

1. INTRODUCTION

Over the last ten years or so, many investigators have attempted seriously to develop computerized schemes for automated detection and quantitative analysis of abnormalities in radiologic images. These efforts have generally been referred to as computer-aided diagnosis (CAD). It is believed that there are several hundred institutions around the world which have been involved in research and development of some aspects of CAD. These include both academic institutions for basic studies and industrial organizations for commercialization efforts. Most of these studies included computer modeling, analysis of image features, development of algorithms and prototype systems, observer performance studies, and clinical evaluations, which have been carried out at many departments involving Diagnostic Radiology, Medical Physics, Computer Sciences, and Electrical Engineering. My purpose in this paper is to provide a brief overview of CAD which is related to chest radiography and to discuss its potential impact on diagnostic radiology and medical physics.

2. BASIC CONCEPT OF COMPUTER-AIDED DIAGNOSIS

CAD may be defined as a diagnosis made by a radiologist who takes into account the results of the computer output as a "second opinion."[1-6] The computer output is derived from quantitative analysis of radiologic diagnostic images. It is important to note that the computer is used only as a tool to provide additional information to clinicians, who will make the final decision as to the diagnosis of a patient. Therefore, the basic concept of CAD is clearly different from that of "automated diagnosis," which had been investigated in the 1960s and 1970s, as will be discussed later.

The purpose of CAD is to improve the diagnostic accuracy and also the consistency of radiologists' image interpretation by using the computer output as a guide. The computer output can be very helpful because a radiologist's diagnosis is made based on subjective judgment and because radiologists tend to miss lesions such as lung nodules in chest radiographs, and microcalcifications and masses in mammograms. In addition, variations in diagnosis, such as inter-observer and intra-observer variation, can be large.

Usually, two types of general approaches are employed in computerized schemes for CAD. One is to find the location of lesions such as lung nodules in chest images by searching isolated abnormal patterns with a computer. Another is to quantify the image features of normal and/or abnormal patterns

such as lung textures related to interstitial infiltrates in chest images and vessel sizes related to stenotic lesions in angiograms.

Computerized schemes for CAD generally include three basic components which are based on three different technologies. The first component is image processing for enhancement and extraction of lesions. It is important to note that the image processing involved in CAD schemes is aimed at facilitating that the computer, but not the human observer, picks up the initial candidates of lesions and suspicious patterns. Various image-processing techniques have been employed for different types of lesions. Some of the commonly used techniques include filtering based on Fourier analysis, wavelet transform, morphological filtering, the difference image technique, and artificial neural networks (ANN).

The second component is the quantitation of image features such as the size, contrast, and shape of the candidates selected in the first step. It is possible to define numerous features based on some mathematical formula that may not be easily understood by the human observer. However, it is generally useful to define, at least at the initial phase of CAD development, image features that have already been recognized and described subjectively by radiologists. This is because radiologists' knowledge is based on their observations on numerous cases over the years, and their diagnostic accuracy is generally very high and reliable. One of the most important factors in the development of CAD schemes is to find unique features that can distinguish reliably between a lesion and other, normal anatomic structures.

The third component is data processing for distinction between normal and abnormal patterns, based on the features obtained in the second step. A simple and common approach employed in this step is a rule-based method, which may be established based on the understanding of lesions and other normal patterns. Therefore, it is important to note that the rule-based method may provide useful information to improve the CAD schemes. Other techniques used include discriminant analysis, ANN, and the decision-tree method. It is our experience that the combination of the rule-based method with other methods such as ANN tends to produce the best results in terms of high performance of CAD schemes.

Because the basic concept of CAD is broad and general, CAD is applicable to all imaging modalities, including conventional projection radiography, computed tomography (CT), magnetic resonance imaging (MRI), ultrasound imaging, and nuclear medicine imaging. In addition, computerized schemes for CAD can be developed for all kinds of examinations on every part of the body, including the skull, chest, abdomen, bone, and vascular system. However, the research subjects investigated for CAD in the past have been relatively limited. The major focus on recent research subjects has been in the field of mammography for early detection of breast cancer. Many CAD schemes are related to detection and characterization of masses and clustered microcalcifications in mammograms.[7-14] In chest radiography [15-22], computerized schemes have been developed for detection of lung nodules, interstitial infiltrates, cardiomegaly, and pneumothoraces.

Computerized schemes in angiography [23-25] include the quantitative analysis of stenotic lesions and the determination of pulsatile blood flow rates.

3. COMPUTERIZED SCHEME FOR DETECTION OF LUNG NODULES IN CHEST RADIOGRAPHS

It is well known that radiologists may miss about 30% of lung nodules in chest images. The cause of this oversight is believed to be the overlap with normal anatomic structures, which tend to camouflage subtle lung nodules. However, it has been reported that many of the missed nodules may be detected retrospectively or by use of a double reading procedure. Therefore, it has been assumed that the prompt given by the computer output indicating the potential sites of nodules would improve the radiologists' performance in detecting lung nodules.

We have been developing a computerized scheme for automated detection of lung nodules in PA chest images. The basic scheme for the major steps is shown in Fig. 1. First, digital chest images or a digitized version of conventional chest radiographs are subjected to two image-filtering operations which are operated in parallel. One produces an enhanced image of nodules by using a filter such as a matched filter, and the other provides a suppressed image of nodules by using another filter such as a ring average filter. A difference image is then obtained by subtraction of the suppressed image from the enhanced image. The difference image contains strongly enhanced nodules and strongly suppressed background, which is due to the removal of the majority of the normal anatomic background structures, and therefore is useful for identifying candidates of nodules. This difference image technique is considered a generalization of edge enhancement techniques, and in fact the difference image looks very similar to edge-enhanced images. If the two filters are linear filters, then the two can be combined into one filter operation. However, if a non-linear filter is applied, two filtering operations need to be applied separately in parallel. The difference image technique is applicable to computerized detection of isolated abnormal patterns such as microcalcifications and masses in mammograms.

The initial candidates of lung nodules are identified from relatively round patterns with large pixel values in the difference image, which may include nodules as well as normal structures such as ribs and pulmonary vessels. Therefore, image features of these candidates are quantified in terms of the size, contrast, and other parameters related to the shape of the candidates. A rule-based method is then applied which removes some candidates as false positives, when their features are matched to those of normal anatomic structures such as end-on vessels, rib-rib crossings, rib-vessel crossings, aggregates of vessels, and rib-clavicle crossings. Finally, the ANN is applied for further removal of some false positives by training the ANN by use of a number of image features on the remaining candidates.

Figure 2 shows a chest image with a lung nodule and the computer output (arrow) indicating the correct detection of the nodule by our CAD scheme. When we used our database of 100 chest images with lung nodules

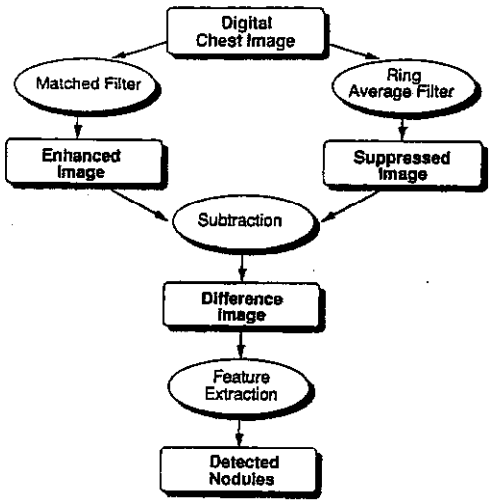

Fig. 1 Basic scheme for computerized detection of lung nodules in chest images.

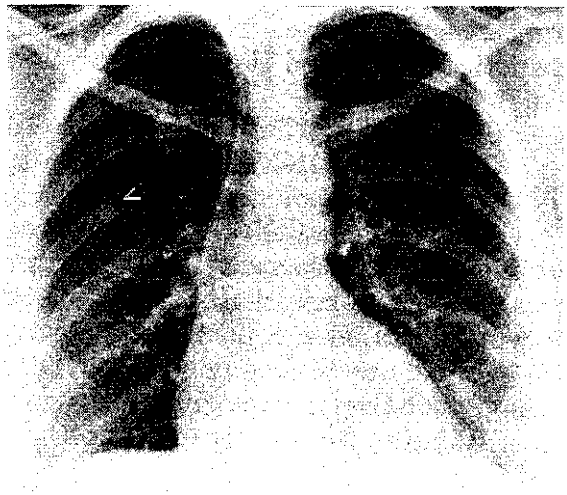

Fig. 2 Chest image with a lung nodule and the computer output (arrow).

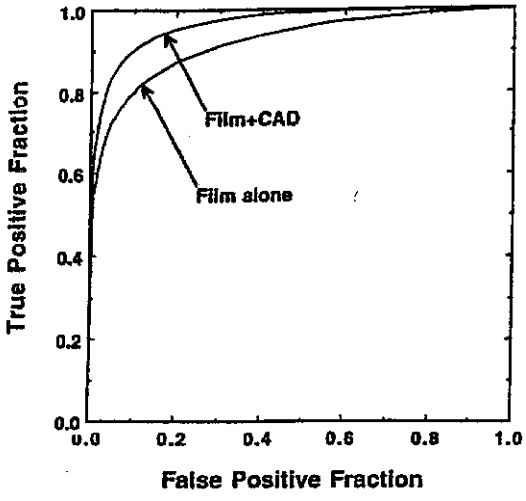

Fig. 3 ROC curves for radiologists' detection of nodules without and with the computer output.

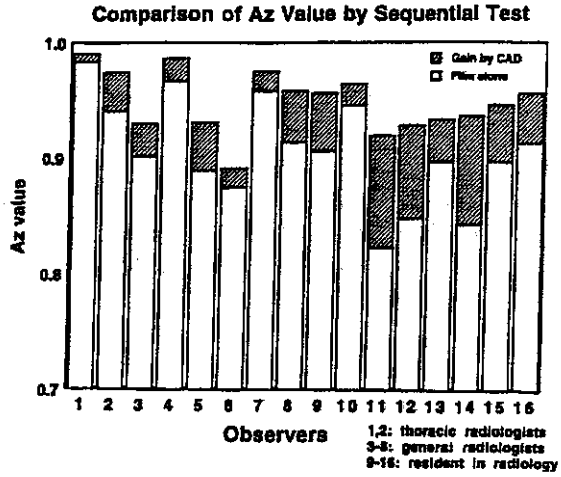

Fig. 4 A_Z values for 16 radiologists without and with the computer output. Gain by CAD is indicated by shaded area.

and 100 normal cases, the perfor-mance of our scheme provided 70% sensitivity at a false positive rate of 1.7 per image. One of the important questions concerning the basic concept of CAD is, "Can the radiologists' performance be improved if the computer output such as that in Fig. 2 is presented to them?" A positive answer is clearly indicated in Figs. 3 and 4. Figure 3 shows ROC curves for radiologists' detection of nodules without and with the computer output. It is apparent that radiologists' performance in detecting lung nodules is improved significantly by use of the computer output. Figure 4 shows A_Z values for sixteen radiologists, including chest radiologists, general radiologists, and residents, without and with the computer output. It is evident that all of the radiologists who participated in this study indicated an increase in A_Z values and thus benefited from the computer output. It should be noted also that, without the computer output, the average A_Z value for residents was lower than the average A_Z value for chest and general radiologists. However, the average A_Z value for residents with the computer output was comparable to the average A_Z value for chest and general radiologists without the computer output. This result seems to indicate that CAD can improve the accuracy and consistency of radiologic diagnosis.

4. DISCUSSION

Although a considerable effort has been made to develop CAD schemes in the past, CAD is still far from being in wide-spread routine clinical use in various diagnostic examinations. The question then arises "What is needed for CAD in the future?" At least three topics may be identified. First, clinical applications for prospective studies with a prototype CAD system need to be initiated. Second, the performance of CAD schemes needs to be improved in terms of the sensitivity in detecting lesions, the number of false positives, and CPU time. Third, new research projects for CAD should be undertaken.

Some of the CAD schemes developed up to now appear to be ready for initial clinical trials, because of the relatively high performance levels already achieved by these schemes. At present, there are two areas that may be identified for potential clinical applications of CAD schemes. One of these includes mass screening programs such as mammography in the United States and some countries in Europe, and also chest radiography in Japan and some other countries. Because the majority of cases in mass screening examinations are normals and because it is difficult for radiologists to maintain high alertness at all times, the computer output indicating potential sites of lesions would be helpful. The second area for the potential clinical application of CAD schemes includes the existing digital imaging systems such as computed radiography (CR), digital subtraction angiography (DSA), CT, and MRI. Once a proper CAD algorithm has been developed, it would be relatively straightforward to implement the algorithm, which will result in added value to the existing digital system.

The performance of the current CAD schemes needs to be improved further. It may be useful to have some target levels for improved algorithms, at least for the near future. The sensitivity in detecting lesions by computer should be comparable to or greater than that by radiologists, which would be in

the range of 70-90%. It should be noted that the computer's sensitivity does not have to be substantially greater than the radiologists' sensitivity. This is because the benefit of CAD is largely due to the difference in performance between radiologists and the computer. In some cases, radiologists may be correct and the computer may be incorrect. However, in other cases, the computer may be correct and the radiologist may be incorrect. In such cases, the radiologist may be able to correct his/her initial mistakes by observing the correct computer output, thus improving the radiologist's performance.

The number of false positives should be as low as possible. One useful target level for the number of false positives for CAD schemes is less than one per image. The majority of current CAD schemes include, on average, more than one false positive per image. This implies that, on average, all of the images subjected to computer analysis contain at least one false positive, which can be disturbing to some radiologists, particularly when they read many normal cases in a screening examination. Of course, if the average number of false positives per image is less than one, some images are expected to have no false positives.

The CPU time required for CAD algorithms needs to be as short as possible. However, many of the computer algorithms do not need to run on a real-time basis, because interpretations of radiologic images by radiologists are not commonly made immediately after the image acquisition following radiological examinations, and thus there usually is adequate time to run the CAD algorithm and to obtain the computer output prior to the reading of the image. This situation is generally applicable to screening programs. It is desirable, however, for the CPU time to be short enough so that the computer program can be run on a new case which may be required to provide a diagnosis immediately after the image becomes available to the radiologist. If the CPU time is less than 10 seconds, the computer program for CAD can be used on a real-time basis for many routine clinical situations, at least in the initial phase of CAD applications.

For those investigators who may wish to do research in CAD, it would be useful to identify some areas for new research subjects. One is to investigate CAD schemes for detection and characterization of lesions that have not been investigated before. There are numerous kinds of known abnormalities which have been subjected to radiologists' image interpretations, and CAD schemes for these abnormalities may be useful to radiologists. Only a small fraction of abnormal patterns in radiologic images have been studied for CAD schemes so far. Another important area for CAD research is the understanding of normal patterns. Previous studies have largely been focused on specific abnormal patterns such as lung nodules and mammographic masses. The understanding of normal anatomic patterns in chest radiographs and mammograms would be very helpful in improving the overall performance of CAD schemes by eliminating some false positives, for example. Another area for new research is the application of CAD to 3D images. At present, considerable effort is focused on the visualization of 3D images by use of CT image data, which would assist radiologists in the search for abnormalities.

However, radiologists would be benefited more if the computer could find subtle lesions such as polyps in 3D images of the colon.

Although the goal of current CAD research is generally to develop algorithms and prototype systems that will achieve relatively simplified and limited tasks, it is probably more important to have a long-range target to establish a new field of imaging science that is based on a technical understanding of the contents of medical images. In essence, the major task for this new field of medical imaging science is to translate the knowledge of image interpretation accumulated in the radiologists' brains into concepts and terminologies that are understandable by physicists, computer scientists, and engineers. Therefore, it is absolutely necessary to have close collaboration between radiologists and physicists.

However, one should not underestimate the very difficult task of understanding the contents of medical images in physical and technical terms. There are many imaging modalities, many different body parts to be examined, many different types of abnormalities and their variations, as well as variations in normal patterns. Theoretically, when the contents of medical images as well as the highly sophisticated information-processing steps that occur in radiologists' brains are understood from a technical standpoint, it will be possible to develop algorithms that can perform detection and recognition tasks at a level comparable to that which radiologists can achieve. In practice, it is likely to take a very long time to understand most of the complex processes involved in radiologists' image interpretation tasks because, even for a given single image, it is common for a radiologist to perform multiple, parallel operations in the detection, recognition, and classification of many different kinds of abnormalities.

An interesting and commonly asked question is, "Can computer vision, as represented by CAD schemes, perform better than human vision?" Before we attempt to answer this question, it may be useful to consider the differences between human vision and computer vision, and also to review some historical background related to the development of CAD schemes. First, we must recognize that computer vision is still in a very early stage of development. However, the capability of human vision as applied to radiologists' interpretation of images is at a highly sophisticated level. Despite all-too-common misses of lesions such as lung nodules in chest images, the overall capability of radiologists in image interpretation is far beyond what current CAD schemes will be able to accomplish in the foreseeable future. It may be a very long time before the computer can challenge human vision in a broad variety of realistic tasks. Near-term success seems much more likely in specific tasks with narrowly focused goals, such as the detection of mammographic microcalcifications by use of computers, to assist radiologists in ways that compensate for their weaknesses as human observers.

Second, we must not forget that there were some attempts in the 1960s and 1970s to develop computer vision for "automated diagnosis" in radiology. These early attempts were not successful, probably in part because computers then in use were not as powerful as those used today, because imaging science

was in an earlier phase, and because digital images were not commonly available. However, it seems that a more serious flaw in these early attempts was the unrealistic expectation of higher performance from computer vision than from human vision. Investigators appear to have assumed and believed that machines and computers would be better than human observers and, therefore, that radiologic diagnoses could be made by computers in place of human beings.

Because of the unfortunate historical failure of early computer-vision approaches, there is still a strong and perhaps irrational negative bias among some experienced radiologists, and even among some computer scientists, against current research on computerized analysis of radiographic images. Current research has a more modest goal which is to provide computer output that will assist, rather than replace, radiologists by offering a "second opinion." This is an important lesson. Extreme views on CAD -- either too critical or too optimistic -- may impede the successful development of CAD schemes. We should not underestimate the complexity and difficulty of computer-vision research. We must be cautious to avoid making another mistake in formulating strategies for computer vision in diagnostic radiology. In the long run, it may be possible to develop a computer algorithm that will allow automated screening of diagnoses. However, it would be prudent at present to demonstrate convincingly and unquestionably that some limited applications of computer vision can be practical and useful in clinical practice by helping radiologists to improve their overall diagnostic performance. Once such a solid foundation has been established, a wider range of computer-vision approaches can be explored in an attempt to exceed the capabilities of human vision.

5. CONCLUSION

It is expected that CAD will have a major impact on diagnostic radiology in the near future. However, it is necessary to improve the performance of current CAD schemes further. In the long run, it may be desirable to establish a new science related to a technical understanding of the contents of radiologic images and also to radiologists' image interpretation process in diagnostic radiology.

ACKNOWLEDGMENT

This work is supported by USPHS Grants CA24806 and CA62625. The author is grateful to his colleagues in the Rossmann Laboratories for their numerous discussions and suggestions; Mrs. E. Lanzl for improving the manuscript; and Ms. E. Ruzich for her secretarial assistance. K. Doi is a shareholder of R2 Technology, Inc., Los Altos, CA. It is the policy of the University of Chicago that investigators disclose publicly actual or potential significant financial interests that may appear to be affected by the research activities.

REFERENCES

(1) MacMahon H, Doi K, Chan HP, Giger ML, Katsuragawa S, Nakamori N: Computer-aided diagnosis in chest radiology. J Thoracic Imaging 5: 67-76, 1990.

(2) Doi K, Giger ML, Nishikawa RM, Hoffmann KR, MacMahon H, Schmidt RA, Chua KG: Digital radiography: A useful clinical tool for computer-aided diagnosis by quantitative analysis of radiographic images. Acta Radiologica 34: 426-439, 1993.

(3) Giger ML, Doi K, MacMahon H, Nishikawa RM, Hoffmann KR, et al.: An "intelligent" workstation for computer-aided diagnosis. RadioGraphics 13: 647-656, 1993.

(4) Nishikawa RM, Haldemann RC, Papaioannou J, Giger ML, Lu P, Schmidt RA, Wolverton DE, Bick U, Doi K: Initial experience with a prototype clinical "intelligent" mammography workstation for computer-aided diagnosis. Proc SPIE 2434: 65-71, 1995.

(5) Giger ML, MacMahon H: "Image processing and computer-aided diagnosis". In: Radiologic Clinics of North America, (eds. R. A. Greenes, R. Bauman). Saunders Publishing Co., vol. 34, pgs. 565-596, 1996.

(6) Doi K, Giger ML, Nishikawa RM, Schmidt RA: Digital Mammography '96 (Elsevier Science, Amsterdam), 1996.

(7) Yin FF, Giger ML, Doi K, Metz CE, Vyborny CJ, Schmidt RA: Computerized detection of masses in digital mammograms: Analysis of bilateral subtraction images. Med Phys 18: 955-963, 1991.

(8) Chan HP, Doi K, Galhotra S, Vyborny CJ, MacMahon H, Jokich PM: Image feature analysis and computer-aided diagnosis in digital radiography. 1. Automated detection of microcalcifications in mammography. Med Phys 14: 538-548, 1987.

(9) Chan HP, Doi K, Vyborny CJ, Schmidt RA, Metz CE, Lam KL, Ogura T, Wu Y, MacMahon H: Improvement in radiologists' detection of clustered microcalcifications on mammograms: The potential of computer-aided diagnosis. Invest Radiol 25: 1102-1110, 1990.

(10) Nishikawa RM, Giger ML, Doi K, Vyborny CJ, Schmidt RA: Computer-aided detection of clustered microcalcifications: An improved method for grouping detected signals. Med Phys 20: 1661-1666, 1993.

(11) Yoshida H, Doi K, Nishikawa RM, Giger ML, Schmidt RA: An improved computer-assisted diagnostic scheme using wavelet transform for detecting clustered microcalcifications in digital mammograms. Acad Radiol 3: 621-627, 1996.

(12) Zhang W, Doi K, Giger ML, Nishikawa RM, Schmidt RA: An improved shift-invariant artificial neural network for computerized detection of clustered microcalcifications in digital mammograms. Med Phys 23: 595-601, 1996.

(13) Jiang Y, Nishikawa RM, Wolverton DE, Metz CE, Giger ML, Schmidt RA, Vyborny CJ, Doi K: Malignant and benign clustered microcalcifications: Automated feature analysis and classification. Radiology 198: 671-678, 1996.

(14) Huo Z, Giger ML, Vyborny CJ, Wolverton DE, Schmidt RA, Doi K: Automated computerized classification of malignant and benign mass lesions on digitized mammograms. Acad Radiol 5: 155-168, 1998.

(15) Giger ML, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography. 3. Automated detection of nodules in peripheral lung fields. Med Phys 15: 158-166, 1988.

(16) Kobayashi T, Xu X-W, MacMahon H, Metz CE, Doi K: Effect of a computer-aided diagnosis scheme on radiologists' performance in detection of lung nodules on radiographs. Radiology 199: 843-848, 1996.
(17) Xu XW, Doi K, Kobayashi T, MacMahon H, Giger ML: Development of an improved CAD scheme for automated detection of lung nodules in digital chest images. Med Phys 24: 1395-1403, 1997.
(18) Katsuragawa S, Doi K, MacMahon H: Image feature analysis and computer-aided diagnosis in digital radiography: Detection and characterization of interstitial lung disease in digital chest radiographs. Med Phys 15: 311-319, 1988.
(19) Ishida T, Katsuragawa S, Kobayashi T, MacMahon H, Doi K: Computerized analysis of interstitial disease in chest radiographs: Improvement of geometric-pattern feature analysis. Med Phys 24: 915-924, 1997.
(20) Nakamori N, Doi K, MacMahon H, Sasaki Y, Montner S: Effect of heart size parameters computed from digital chest radiographs on detection of cardiomegaly: Potential usefulness for computer-aided diagnosis. Invest Radiol 26: 546-550, 1991.
(21) Sanada S, Doi K, MacMahon H: Image feature analysis and computer-aiuded diagnosis in digital radiography: Automated detection of pneumothorax in chest images. Med Phys 19: 1153-1160, 1992.
(22) Armato SG, Giger ML, MacMahon H: Computerized delineation and analysis of costophrenic angles in digital chest radiographs. Acad Radiol 5: 329-335, 1998.
(23) Fujita H, Doi K, Fencil LE, Chua KG: Image feature analysis and computer-aided diagnosis in digital radiography. 2. Computerized determination of vessel sizes in digital subtraction angiographic images. Med Phys 14: 549-556, 1987.
(24) Hoffmann KR, Doi K, Chen SH, Chan HP: Automated tracking and computer reproduction of vessels in DSA images. Invest Radiol 25: 1069-1075, 1990.
(25) Hoffmann KR, Doi K, Fencil LE: Determination of instantaneous and average blood flow rates from digital angiograms using distance-density curves. Invest Radiol 26: 207-212, 1991.

Computer-Aided Diagnosis in Medical Imaging
K. Doi, H. MacMahon, M.L. Giger and K.R. Hoffmann, editors

Development of a mammogram CAD system: Performance studies with large databases

H. Fujita [a], T. Hara [a], T. Suzuki [b], D. Fukuoka [a], T. Endo [c], T. Iwase [d]
N. Ohuchi [e], T. Ikeda [b], and H. Yoshimura [f]

[a] Department of Information Science, Faculty of Engineering,
Gifu University, Gifu-shi 501-1193, Japan
e-mail: fujita@fjt.info.gifu-u.ac.jp
[b] Cancer Detection Center, Miyagi Cancer Association, Sendai 980-0011, Japan
[c] Department of Radiology, Nagoya National Hospital,
Nagoya-shi 460-0001, Japan
[d] Department of Breast Surgery, Aichi Cancer Center Hospital,
Nagoya 464-8681, Japan
[e] Second Department of Surgery, Tohoku University School of Medicine,
Sendai 980-8574, Japan
[f] Central Research Laboratory, Konica Corporation, Hino-shi 191-8511, Japan

The purpose of this work is to evaluate our computer-aided diagnosis (CAD) system for mammograms by using some large databases consisting of more than 2,000 cases of Japanese women, which were obtained from a cancer center hospital and a breast screening program in Japan. The system performance for two datasets from the hospital (all outpatients' mammograms during two one-month periods) was 91% (80%) true-positive rate for detection of clustered microcalcifications (masses) with 0.7 (1.0) false-positive (FP) findings per image. An initial result for the screening image data indicated 88% of cancer detection but with FPs of 0.7/image for clustered microcalcifications and 1.6/image for masses; the slightly higher value of FPs for masses was considered to be due to the use of different data for the system tuning. These preliminary results imply that our CAD system may be effective in a clinical site for reading mammograms. Furthermore, an initial promising result in a CAD system for screening 3-D breast ultrasonograms is described.

Key words: CAD, Breast cancer, Mass screening, Mammography, Ultrasonogram

1. INTRODUCTION

The incidence of breast cancer has been increasing in Japan, and it is said that this cancer may become the first mortality in women by the year 2000. Some decades ago, mass screening for breast cancer was conducted in only a few hospitals, but today it is nationwide. In Japan, however, mass screening for breast cancer basically consisted of physical examination of the breasts and the regional lymph nodes. Since 1987, several research groups, as one of the cancer research projects sponsored by the

Ministry of Health and Welfare in Japanese government, have been working on breast cancer screening [1], recently on the necessity and introduction of mammography [2-4]. Based on the recent Ohuchi's group research [4], Japan Association of Breast Cancer Screening has proposed a guideline on breast cancer screening system with mammography: physical examination combined with mammography every two years for women aged over 50 and only physical examination every year for women aged 40-49. This proposal will be officially accepted by the spring in 2000. However, the number of physicians for reading mammograms will be definitely insufficient, and one of the solutions is to develop a computer-aided diagnosis system for mammography.

We have been developing a computer-aided diagnosis system (CAD) for the detection of masses and clustered microcalcifications on digitized mammograms [5-15]. It is considered that the CAD analysis may not be easier because Japanese women's breasts are typically dense. A CAD system for breast cancer detection on ultrasonograms used in a mass screening program has been also developed [16,17]. Thousands of cases have been examined by our systems.

The purpose of this work is to evaluate our CAD system on mammograms by using over 2,000 cases of Japanese women and to discuss the potential utilities of the CAD in clinical situations. Two "one-month" cases from a hospital and two thousands cases of screening mammograms from an annual mass screening program were employed to estimate our system performance. An initial result in a CAD system for breast ultrasonograms is also described.

2. A CAD SYSTEM FOR MAMMOGRAPHY

2.1 Our CAD system for mammography

The CAD system consists of a laser scanner (Konica LD5500), a workstation (Sun Ultra-1 170E) with a color CRT monitor (app. 1k x 1k pixels), an 18GB hard drive, and a high-brightness B/W monitor (app.2k x 2.5k). All mammograms were digitized at a 100- or 50-micron pixel size, 12-bit grey level and optical density range from 0.0 to 4.0. Then, the detection and analysis programs were run to detect the masses and clustered microcalcifications, based on the methods of adaptive thresholding technique [5,6,10-12] and tripple-ring filter analysis in density gradient [5,7-9,13], respectively, and to analyze these candidates in terms of malignancy.

2.2 Image databases

The mammograms which were collected, digitized and analyzed for our CAD experiments are as follows:

1) all outpatients' mammograms (MLO and CC views) taken at Aichi Cancer Center Hospital during two one-month periods: the month of July, 1996 (230 cases consisting of 922 images: Dataset A) and the month of April, 1997 (240 cases consisting of 862 images: Dataset B). Datasets A and B include 23 and 26 malignant cases, respectively.

2) screening mammograms (MLO view only) taken in 59 municipalities in Miyagi Prefecture as a mass screening program. The screening for about 7,000 women per year, aged over 40, has been carried out with a physical examination combined with

mammography. About 0.2% was found as malignant cases each year. Mammograms of a single MLO view of each breast were obtained using the mobile mammography unit in the bus, in which all equipments were included such as an x-ray generator and an automatic developer. The mammographic findings were classified into 5 categories based on malignant potential from class I to V: class I as no finding, II as benign, III as benign but malignancy not ruled out, IV as malignancy suspected, and V as malignancy. There are two datasets. Dataset C consists of about 980 cases with 12 malignant cases, which were taken in April,'96 to March,'97. Dataset D consists of about 870 cases with 10 malignant cases, which were taken in April,'97 to March,'98 for the cases diagnosed as classes of III-V and in April,'97 to July,'98 for the cases diagnosed as normal (class of I or II). If the patient was diagnosed as one of the classes of III-V, another mammograms of CC view were taken, and physical and echo exams were performed at the same place. Then the hospital for detail exams was introduced to the patient. Applicants usually take this screening every two years.

2.3 Performance result

The datasets A and B were firstly analyzed by our another dataset of 150 cases of MLO-view mammograms which were obtained previously from the same hospital; some differences such as film maker and development condition are involed in this procedure. Table 1 shows the detection performance for these datasets. Table 2 indicates the adjusted result of the performance by modifying the some parameters of the detection algorithms because of relatively large number of false positives (FPs) per image in mass detection in both datasets and low true-positive (TP) rate for mass case in the dataset of B. After this parameter tuning, the true-negative (TN) rate was also improved. We have recently started for studying the CAD performance by using the parameters obtained in this experiment for all outpatients' mammograms at the same hospital. The system performance in average was 91% (80%) true-positive rate for detection of clustered microcalcifications (masses) with 0.7 (1.0) false-positive findings per image.

The datasets C and D were analyzed with the turned parameters of the datasets A and B described above. The results of detection performance for datasets C and D are shown in Table 3. TP[B+M] and TP[M] stand for the true-positive rates including both of benign and malignant lesions as "true positive" and that including only malignant lesion, respectively. This initial result for the screening image data indicates 88% of cancer detection but with FPs (0.7/image for clustered microcalcifications and 1.6/image for masses); the slightly higher value of FPs for masses is considered to be due to the use of different data for the system tuning. Three cancers missed were those of a cluster with low-contrast microcalcifications, a low-contrast mass, and an irregular shape mass.

3. A CAD SYSTEM FOR BREAST ULTRASONOGRAMS

A breast mass screening by using cross-sectional ultrasonograms has started in Osaka city in Japan since 1988 (and also in some other cities) and a good performance in the diagnosis has been shown. However, numerous cross-sectional

Table 1 Detection performances for datasets of A and B from outpatients' mammograms.

Dataset		TP	FP/image	TN
A	CMC	87%	0.5	74%
	Mass	91%	2.0	15%
B	CMC	96%	1.5	34%
	Mass	67%	1.8	16%

Table 2 Detection performances for datasets of A and B from outpatients' mammograms (after the self-turning of parameters).

Dataset		TP	FP/image	TN
A	CMC	86%	0.5	74%
	Mass	83%	1.0	43%
B	CMC	95%	0.9	51%
	Mass	77%	0.9	45%

Table 3 Detection performances for datasets of C and D from screening mammograms.

Dataset		TP [B+M]	TP [M]	FP/image	TN
C	CMC	27/39 (69%)	6/6 (100%)	0.9	59%
	Mass	65/98 (66%)	7/8 (88%)	1.6	15%
D	CMC	22/27 (81%)	2/3 (67%)	0.6	68%
	Mass	37/49 (75%)	6/7 (86%)	1.6	18%

images have to be interpreted by physicians. Therefore, the development of CAD system is required to reduce the false-negative case in the screening.

The overview of our ultrasonogram CAD system is described below [16,17].

The ultrasonograms for each breast include 51 sections. The developed CAD system consists of the following four procedures.

(1) Detection: Tumor candidates are detected by using an active contour model. Our method includes the technique to unify the control points and to produce the first contours with its skeleton functions on density gradient vectors. The distribution of the skeleton function is determined by voting a count from gradient vectors.

(2) False-positive elimination: The false-positive candidates are eliminated by using

shape features including the circularity and the area of candidates.

(3) Classification: The candidates are classified by some texture features including the angular second moment, the contrast, the correlation and the entropy, which are obtained from gray-level co-occurrence matrix.

(4) Annotated image display: 3D image constructed from sectional images is displayed to indicate the detection points in the 3D view.

As an initial performance study, we employed 9 cases (459 images). The TP rate to detect tumors was approximately 78% with about four FP candidates per "constructed 3D" breast, in which difficult cases such as a moderate attenuation of posterior echo were detected correctly. We are now developing the classification method of the detected masses.

4. SUMMARY

The system performances for the outpatient databease and screening database for mammograms were obtained, which show a promising result. It is expected that we will be able to better optimize the CAD performance because we used the parameters tuned by the different database for the screening database. These preliminary results imply that our CAD system may be effective in a clinical site for assisting the reading of mammograms. Moreover, the performance result for our ultrasonogram CAD system is promising although the number of image data was not so much.

ACKNOWLEDGMENTS

The authors are grateful to Messrs. Y. Lee, O. Ohtsuka, S. Arai, Y. Hatanaka, N. Shinohara, T. Yamada, and T. Umeda of Fujita Laboratory for their cooperation, and also to Dr. Yasuyuki Kato, Department of First Surgery, Osaka City University Hospital for his cooperation on the ultrasonogram CAD project.

REFERENCES

1. J. Ota, T. Horino, T. Taguchi, et al., Mass screening for breast cancer: Comparison of the clinical stages and prognosis of breast cancer detected by mass screening and in-patient clinics. Jpn. J. Cancer Res., Vol.80 (1989) pp.1028-1034.
2. N. Ohuchi, K. Yoshida, M. Kimura, et al., Improved detection rate of early breast cancer in mass screening combined with mammography. Jpn. J. Cancer Res., Vol.84 (1993) pp.807-812.
3. N. Ohuchi, K. Yoshida, M. Kimura, et al., Comparison of false nagative rates among breast cancer screening modalities with or without mammography: Miyagi trial. Jpn. J. Cancer Res., Vol.86 (1995) pp.501-506.
4. N. Ohuchi, T. Endo, I. Tsuji, et al., Quality control of breast cancer screening with mammography: A report supported by a Grant-in-Aid for cancer research from the Ministry of Health and Welfare. J. Jpn. Assc. Breast Cancer Scr., Vol.6 (1997) pp.137-143.

5. H. Fujita, T. Endo, T. Matsubara, et al., Automated detection of masses and clustered microcalcifications on mammograms. Proc. SPIE, Vol.2434 (1995) pp.682-692.
6. T. Matsubara, H. Fujita, T. Endo, et al., Development of mass detection algorithm based on adaptive thresholding technique in digital mammograms. Digital Mammography '96, Elsevier, Amsterdam (1996) pp.391-396.
7. T. Hara, K. Hirako, H. Fujita, et al., Automated detection algorithm for clustered microcalcifications based on density gradient and triple-ring filter analysis. Digital Mammography '96, Elsevier, Amsterdam (1996) pp.257-262.
8. K. Hirako, H. Fujita, T. Hara, et al., Development of detection filter for mammographic microcalcifications: A method based on density gradient and triple-ring filter analysis. Systems and Computers in Japan, Vol.27, No.13 (1996) pp.36-48.
9. I. Norhayati, H. Fujita, T. Hara, et al., Automated detection of clustered microcalcifications on mammograms: CAD system application to MIAS database. Physics in Medicine and Biology, Vol.42, No.12, (1997) pp.2577-2589.
10. S. Kasai, H. Fujita, T. Hara, et al., Development of a detection algorithm for masses around thick mammary gland on mammograms, Proc. CAR'98 (1998) pp.213-218.
11. T. Matsubara, H. Fujita, T. Hara, et al., New algorithm for mass detection in digital mammograms, Proc. CAR'98 (1998) pp.219-223.
12. T. Matsubara, H. Fujita, S. Kasai, et al., Development of a new algorithm for detection of mammographic masses, Proc. of 4th International Workshop on Digital Mammography, Nijmegen, The Netherlands (1998) in press.
13. D. Fukuoka, S. Kasai, H. Fujita, et al., Automated detection of clustered microcalcifications on digitized mammograms, Proc. of 4th International Workshop on Digital Mammography, Nijmegen, The Netherlands (1998) in press.
14. M. Goto, A. Morikawa, H. Fujita, et al., Detection of spicules on mammograms based on multi-stage pendulum filter, Proc. of 4th International Workshop on Digital Mammography, Nijmegen, The Netherlands (1998) in press.
15. T. Hara, H. Fujita, T. Endo, et al., Performance studies of a computer-aided diagnostic system on mammograms, Proc. of 4th International Workshop on Digital Mammography, Nijmegen, The Netherlands (1998) in press.
16. D. Fukuoka, T. Hara, H. Fujita, et al., Automated detection method of breast tumors on cross-sectional ultrasonograms, Med. Imaging Inform. Sci., Vol.14, No.3 (1997) pp.148-154 (in Japanese).
17. D. Fukuoka, T. Hara, H. Fujita, et al., Dynamic region-contour-extraction method with automated initial-contour production and unification of contours, Trans. Inst. Electron. Inform. Commun. Eng. (D-II), Vol.J81-D-II, No.6 (1998) pp.1448-1451 (in Japanese).