Guest Editorial
Computer-Aided Diagnosis in Medical Imaging

I. Introduction

Recent developments in the computerized analysis of medical images are expected to aid radiologists and other healthcare professionals in various diagnostic tasks of medical image interpretation. This special issue of IEEE-TRANSACTIONS ON MEDICAL IMAGING (IEEE-TMI) on computer-aided diagnosis (CAD) focuses on progress in this new era of image interpretation.

In medical imaging, the accurate diagnosis and/or assessment of a disease depends on both image acquisition and image interpretation. The role and contribution of radiology to medical diagnosis has expanded tremendously due to advances in image quality compliance regulations, image detector systems, and computer technology. For example, a major contributor to improvement in medical imaging has been cross-sectional imaging [e.g., X-ray computed tomography (CT) and magnetic resonance imaging (MRI)], which depends greatly on computer power and data storage capabilities, and produces many three-dimensional (3-D), high-quality images for interpretation. The image interpretation process, however, has only recently begun to benefit from computer technology. Most interpretations of medical images are performed by radiologists; however, image interpretation by humans is limited due to the nonsystematic search patterns of humans, the presence of structure noise (camouflaging normal anatomical background) in the image, and the presentation of complex disease states requiring the integration of vast amounts of image data and clinical information. CAD, defined as a diagnosis made by a radiologist who uses the output from a computerized analysis of medical images as a "second opinion" in detecting lesions, assessing extent of disease, and making diagnostic decisions, is expected to improve the interpretation component of medical imaging. With CAD, the final diagnosis is made by the radiologist.

This special issue of IEEE-TMI on CAD contains 11 papers by various investigators from around the world (five from the U.S., one from Canada, two from the Netherlands, one from Norway, and two from Japan). Research in CAD is a rapidly growing, dynamic field with new computer techniques, new imaging modalities, and new interpretation tasks. The CAD research presented in this special issue ranges from cancer (breast, lung, and colon) detection in single-projection images and CT, to scoliosis screening, to multimodality assessment of myocardial viability.

Publisher Item Identifier S 0278-0062(01)11275-9.

II. Lesion Detection

The screening of asymptomatic people involves radiologists visually scanning the images of mostly healthy subjects for a specific abnormality. The interpretation of screening images lends itself to CAD since it is a repetitive, burdensome task involving mostly normal images—a situation prime for oversight errors. Many computerized analysis methods have been developed for screening mammography, one of which has led to an FDA-approved commercial system in 1998 and, since early 2001, a reimbursement program for the technique in the U.S. The role of CT in screening programs is also rapidly growing especially in the thorax for lung cancer screening and in the colon (colonography) for the detection of suspect polyps for colon cancer.

In this issue, Hatanaka *et al.* (pp. 1209–1214) extend methods for the computerized detection of masses on mammograms by focussing on the problem of partial lesions at the edge of the film. Use of a sector-form model in the template matching process yielded improved mass detection. Such step-by-step improvements with each focussed on different "problems" ultimately will improve overall detection. Mudigonda *et al.*(pp. 1215–1227) propose a new method for the computerized detection of masses on mammograms by analyzing oriented flow-like textural information along with features in adaptive ribbons of pixels along the margins.

Many CAD papers during the last two decades have involved either mammograms or chest radiographs. This "early" research was performed on digitized radiographs. While the computerized analysis of mammograms is mainly focussed on one disease, breast cancer, the computerized analysis of chest radiographs ultimately requires the diagnosis of a multitude of diseases (e.g., lung cancer, pneumothorax, interstitial diseases). Various reviews have reported on the status and methods for CAD in mammography [1]–[3]. In this special IEEE-TMI issue on CAD, van Ginneken *et al.* (pp. 1228–1241) present a comprehensive overview of computerized analysis methods for the chest radiograph. We expect that the development and implementation of computer techniques for projection radiography of the chest (posteroanterior or lateral) will advance rapidly with the advent and acceptance of digital chest imaging units—taking CAD to a "push button" stage in this electronic imaging environment.

For the screening of lung cancer, single-projection chest radiographs and thoracic CT scans have been considered. The existence of 3-D image data from CT removes much structure noise; however, neighboring pulmonary vessels or nodules near the chest wall can still contribute to oversight errors. These 3-D data sets greatly increase the number of images that must be reviewed by a radiologist in a screening program - leading to an overwhelming task for the human search process. Accordingly, image interpretation may greatly benefit from a computer search aid. Brown *et al.* (pp. 1242–1250) in this issue report on a patient-specific model for automatically detecting lung nodules in CT images. This patient-specific technique uses a patient's baseline image to assist in the segmentation of subsequent images so that changes in size and/or shape of nodules can be measured automatically.

Gokturk *et al.* (pp. 1251–1260) and Yoshida and Nappi (pp. 1261–1274) report on recent developments in their techniques for the computerized detection of polyps in colonography. Using a method of random orthogonal shape sections, Gokturk *et al.* (pp. 1251–1260) report improvement in specificity at a high sensitivity. Yoshida and Nappi (pp. 1261–1274) describe their 3-D geometric features that characterize polyps and folds in the colonic walls. The combination of colonography and CAD challenges the position of conventional colonoscopy as the gold standard for colon cancer screening.

III. Lesion Characterization

Once a lesion is detected, characterization is necessary to determine the status of the lesion, e.g., the likelihood that the lesion is cancerous. Complex anatomy, variation in the presentation of malignant and benign states, and varying abilities of the radiologist can lead to interpretation errors. Methods for the computerized analysis of a medical image contain many stages. Improvements at one stage may influence performance at later stages and, subsequently, the overall performance. Likewise, one can use the performance at later stages to examine the accuracy of earlier stages. Sahiner *et al.*(pp. 1275–1284) investigate the effect of lesion segmentation on computerized lesion characterization and, encouragingly, were unable to show a significant difference in the performance of their mammographic mass classification technique when either radiologist segmentation or computer segmentation was employed—thus, suggesting similar usefulness of segmentation by either computer or human.

It is interesting to note that many changes in an image that lead to improvements in human interpretation performance also lead to improvements in computer image interpretation performance. Huo *et al.* (pp. 1285–1292) investigate the usefulness of applying computerized classification methods, which were previously developed for conventional mammographic views, on an independent set of cases that contained the conventional views as well as special compression views of mammographic mass lesions. The improvement in computer performance with the special views is related to the improved segmentation of the mass margin on the special views over that with the conventional views—a situation similarly encountered by radiologists and, thus, a major reason for requesting special views of a suspect lesion.

Research in the computerized analysis of breast lesions includes mammography, sonography, and MRI. Torheim *et al.* (pp. 1293–1301) report on a technique for breast MRI involving 1) semi-automatic region of interest selection; 2) noise reduction; and 3) classification to distinguish between malignant and benign breast lesions. Their method employs temporal parameters obtained from the time-intensity curves.

IV. Disease Assessment

Various computerized image analysis methods are based on a single task for a single imaging modality. Behloul *et al.* (pp. 1302–1313), however, describe a multimodality framework based on neuro-fuzzy techniques for the task of myocardial viability assessment in a positron emission tomography PET-MRI data fusion application. The investigators use two levels: a modality-independent inference level and a modality-dependent application level.

Computerized image analysis has been applied mainly to medical imaging techniques such as X-ray, sonography, and MRI. A role for computer-aided diagnosis is emerging for applications involving less well-known modalities such as thermography and Moire imaging. Kim *et al.* (pp. 1314–1320) are developing a method for the computerized analysis of Moiré topographic images of the back to aid orthopedists in scoliosis inspection of Japanese school children. Moiré topographic imaging is useful in assessing shape deformation.

V. Issues in CAD Research and in the Reporting of Progress

With each new development in computerized medical image analysis, investigators, journal readers, and end-users are anxious to ask the question, "How good is the new technique compared with other techniques?" It is difficult to compare the various computerized methods under development due to the use of different databases and the varying criteria for reporting and evaluating computer results. While an independent test site with an independent database for the evaluation of each new development would be quite beneficial, it is not yet always practical. However, some guidelines may assist the communication of the merits of a new technique through publication in the scientific literature. We leave it as a future challenge to all investigators who aim to publish to incorporate these suggestions.

Application of the same computer analysis method to different databaseswill potentially yield different performance levels. It is possible, for example, that a computerized detection scheme could achieve a sensitivity of 70% at two false-positives per image with one database and a sensitivity of 100% at two false-positives per image with another database. The characteristics of a database will influence the training (e.g., feature selection and classifier training) of a computer method as well as its reported performance level. Databases can be described by objective measures, such as lesion size and contrast, and by subjective measures, such as lesion subtlety, which depends on the observer who gives the subtlety rating, the specific task, and the presence or absence of other images and/or information.

How a database is used will also influence the development and reported performance of a computerized method. For the training and testing of artificial neural networks, it is important that multiple images of the same lesion be grouped together and not separated between the training and testing sets. For example, a computerized scheme for mammography that is trained on the medio-lateral-oblique view of a lesion will yield biased results if the craniocaudal view of the lesion is used in the testing. The reported performance of a computerized scheme will also vary depending on whether sensitivity is given in terms of the percentage of detected lesions per image or per case.

Different scoring methods will affect the "performance" of a computerized method. For example, in the detection of masses, some investigators use the "percent of overlap" between the actual lesion and the computer-detected region as a means of determining a true detection. It should be noted that there are different definitions of percent overlap. For example, some investigators define overlap as the intersection of the two regions, whereas other investigators define overlap as the intersection of the two regions divided by the union of the two regions (a more strict criterion). In addition, different investigators may use different "truths" in developing and evaluating their computerized detection algorithms. For example, in the detection of lung nodules in CT scans, the results should clearly state whether "truth" is defined as a malignent lesion any nodular opacity, or any "abnormality"

The ultimate acceptance of CAD will depend not only on the performance of the computerized method alone, but also on how well the human performs the task when the computer output is used as an aid and on the ability to integrate the computerized analysis method into routine clinical practice. Observer studies have shown that radiologists' performance increased when using a computer output as an aid. It is important to note, however, that observer experience may influence the reported effect of CAD. Investigators have demonstrated the effect of CAD on radiologists' performance in the detection of lung nodules on chest radiographs and found that the gain in performance with the use of CAD output was different for thoracic radiologists, general radiologists, and radiology residents. Thus, it is important to determine and report the amount of experience of the observers and their current reading load in the relevant diagnostic task. It should be noted that a computerized method will be useful even at a less-than-perfect sensitivity, especially if the lesions detected by the computer do not overlap completely with those detected by a radiologist. In addition, it is important to consider the effect of disease prevalence on the observer study and whether or not the prevalence in the study matches that in the targeted population. For example, in screening mammography, typically less than 1% of the cases will demonstrate.

Although it is too early to have well-defined guidelines for evaluation of CAD methods in general, more developed areas of CAD research exist where higher evaluation standards are applicable. For instance many methods are described in the literature for the detection of abnormalities in mammograms, and large public databases are available (e.g., marathon.csee.usf.edu/Mammography/Database.html). We expect that soon editorial boards will pay much more attention to the details of evaluation methods as part of their review process. The inability to obtain graound truth complicates the evaluation process for many medical image analysis fields (e.g., image registration and segmentation); in CAD, however, we have opportunities to establish gidelines for evaluation, since we are able to obtain truth for detection and classification tasks. In the future, investigators who work in such areas of CAD, which have much prior art, should address evaluation issues properly in order to have their work published in IEEE-TMI, since readers will want to know if and why new proposed methods really are better than existing ones.

VI. Summary and Discussion

Integration of CAD into clinical practice has been shown for screening mammography, and we anxiously await the introduction of other CAD methods to the clinical arena. In the clinical setting, CAD methods might be used routinely as part of a screening protocol or used only when requested by a radiologist interpreting a particular case. Ultimately, a CAD workstation would be configured for each radiologist to allow individualized control over the sensitivity and specificity of the computer output with adjustments depending on the nature of the case material and personal preference. For example, a radiologist might prefer a computer output with high sensitivity for examining high-risk patients being screened for cancer, whereas a lower computer sensitivity and potentially a correspondingly higher specificity might be desired for patients at low risk for cancer.

The practice of interpreting medical images is being modified by the information technology revolution and it appears that both the medical profession and the public (patients) welcome enthusiastically such advances. In the future, it is quite likely that all medical images will undergo some form of computer analysis in order to benefit the diagnosis.

Maryellen L. Giger, *Guest Editor*
The University of Chicago
Dept. Radiology
Chicago, IL 60637 USA

Nico Karssemeijer, *Guest Editor*
University Medical Center Nijmegen
Dept. Radiology
Nijmegen, 6500 HB The Netherlands

Samuel G. Armato, III, *Guest Editor*
The University of Chicago
Dept. Radiology
Chicago, IL 60637 USA

References

[1] M. L. Giger, "Computer-aided diagnosis," in *AAPM/RSNA Categorical Course in Diagnostic Radiology Physics: Physical Aspects of Breast Imaging—Current and Future Considerations*, A. Haus and M. Yaffe, Eds. Oakbrook, IL: Radiological Soc. N. Amer., 1999, pp. 249–272.

[2] M. L. Giger, Z. Huo, M. A. Kupinski, and C. J. Vyborny, "Computer-aided diagnosis in mammography," in *Handbook of Medical Imaging*, M. Sonka and M. J. Fitzpatrick, Eds. Philadelphia, PA: SPIE, 1999, vol. 2, medical imaging processing and analysis, pp. 249–272.

[3] C. J. Vyborny, M. L. Giger, and R. M. Nishikawa, "Computer-aided detection and diagnosis," in *Radiologic Clinics of North America*, S. A. Feig, Ed. Philadelphia, PA: Saunders, 2000, vol. 38, Breast imaging, pp. 725–740.

Maryellen L. Giger (M'90) received the B.S. degree in mathematics and physics in 1978 from Illinois Benedictine College, Lisle, IL, the M.Sc. degree in physics in 1979 from the University of Exeter, Exeter, U.K., and the Ph.D. degree in medical physics in 1985 from The University of Chicago, Chicago, IL.

She is presently Professor of Radiology and the College at the University of Chicago and is Director of the Graduate Programs in Medical Physics at the University. She has authored or co-authored over 100 peer-reviewed manuscripts and 100 proceeding papers, and serves as a reviewer for various granting agencies, including the National Institutes of Health (NIH) and the United States Army. Her research in CAD spans mammography, sonography and MR of the breast for breast cancer detection, diagnosis, and cancer risk assessment; chest radiography and CT for lung cancer detection; and skeletal radiography for assessment of osteoporosis and risk of future fracture.

Dr. Giger is an Associate Editor for *Medical Physics* and IEEE TRANSACTIONS ON MEDICAL IMAGING, and she is a fellow of the American Association of Physicists in Medicine (AAPM) and the American Institute for Medical and Biological Engineering (AIMBE).

Nico Karssemeijer (M'98) received the MS degree in physics from Delft University of Technology, Delft, The Netherlands in 1980. He received the Ph.D. degree in 1989 for his research on model guided medical image analysis, in which he developed a method for segmentation of CT images of the abdomen based on statistical atlasses representing organ shape variability.

He became interested in biomedical data analysis when he worked at the Laboratory of Kinesiology of Leiden University, Leiden, The Netherlands. He became involved in medical imaging when he joined the laboratory of Medical Physics and Biophysics of Nijmegen University, Nijmegen, The Netherlands. Since earning the Ph.D. degree, he has held a research position at the Department of Radiology of the University Medical Center Nijmegen, where his primary interest is development of computer-aided diagnosis. The main focus of his research is improvement of the detection and diagnosis of breast cancer. His current research activities include projects aimed at digitization of the Dutch breast cancer screening program in cooperation with the "National Expert and Training Centre for Breast Cancer Screening" in Nijmegen.

Dr. Karssemeijer has been Associate Editor of IEEE TRANSACTIONS ON MEDICAL IMAGING since 1996.

Samuel G. Armato, III, received the B.A degree in physics in 1987 and the Ph.D. degree in medical physics in 1997, both from The University of Chicago, Chicago, IL. His doctoral research involved automated lung segmentation in posteroanterior and lateral chest radiographs and applications of segmentation to the quantitative analysis of lung fields and registration with radionuclide lung scan images.

He is an Assistant Professor of Radiology at The University of Chicago. His current research involves the investigation of computer-aided diagnostic methods in thoracic CT scans, including lung segmentation, lung nodule detection and classification, and pleural analysis.

Dr. Armato serves as a reviewer for a number of granting agencies and scientific journals.