Determination of Similarity Measures for Pairs of Mass Lesions on Mammograms by Use of BI-RADS Lesion Descriptors and Image Features[1]

Chisako Muramatsu, PhD, Qiang Li, PhD, Robert A. Schmidt, MD, Junji Shiraishi, PhD, Kunio Doi, PhD

Rationale and Objectives. To determine similarity measures for selection of pathology-known similar images that would be useful for radiologists as a reference guide in the diagnosis of new breast lesions on mammograms.

Materials and Methods. The images were obtained from the Digital Database for Screening Mammography developed by the University of South Florida. For determination and evaluation of similarity measures, the ''gold standard'' of similarities for 300 pairs of masses was determined by 10 breast radiologists. For determining similarity measures that would agree with radiologists' similarity determination, an artificial neural network (ANN) was trained with the radiologists' subjective similarity ratings and the image features. The image features were determined subjectively using the Breast Imaging Reporting and Data System (BI-RADS) lesion descriptors and objectively by computerized image analysis. The similarity measures determined by the ANN were compared to the gold standard and evaluated in terms of the correlation coefficient.

Results. The similarity measures determined using the BI-RADS descriptors only were not as useful as those determined by use of the image features only. When the BI-RADS margin ratings were combined with the image features, the correlation coefficient between the subjective ratings and the objective measures improved slightly ($r = 0.76$) compared to those based on the image features alone ($r = 0.74$).

Conclusions. The inclusion of the BI-RADS margin descriptors may be useful for determination of similarity measures, especially when it is difficult to obtain the manual outlines of the masses and if the BI-RADS descriptors were provided consistently by radiologists.

Key Words. Similar images; computer-aided diagnosis; breast mass; mammograms.

© AUR, 2009

Breast cancer is the most frequently diagnosed cancer in women in the United States (1). To achieve a decrease in the cancer mortality rate, early detection of cancer lesions is important. Mammography is considered a useful screening method for early detection in the general population. However, it can be difficult to diagnose breast cancer on mammograms; sometimes cancers might be missed, and usually many patients with benign lesions are sent for biopsy (2–4). Computer-aided diagnosis, in which radiologists make a diagnosis by taking into consideration the outputs from the computer analysis of medical images, has been demonstrated to improve radiologists' diagnostic accuracy in the classification of breast lesions (5–7). With such a computer-aided diagnosis, the likelihood of malignancy of the lesion is presented to radiologists.

It has been suggested that the presentation of images similar to that of a new unknown lesion would be helpful for distinction between benign and malignant lesions because radiologists learn their diagnostic skills by observing many cases in clinical practice, review courses, and textbooks (8–10). Studies have shown the potential usefulness of

Acad Radiol 2009; 16:443–449

[1] From the Kurt Rossmann Laboratories for Radiologic Image Research, Department of Radiology, The University of Chicago, Chicago, IL (C.M., Q.L., R.A.S., J.S., K.D.); Department of Intelligence Image Information, Division of Regeneration and Advanced Medical Sciences, Gifu University, 1-1 Yanagido, Gifu, Japan (C.M.); Duke Advanced Imaging Labs, Department of Radiology, Duke University, Durham, NC (Q.L.). Received August 28, 2008; accepted October 15, 2008. Supported by the Department of Defense Breast Cancer Research Program Predoctoral traineeship award W81XWH-05-1-0249.
Address correspondence to: C.M. e-mail: chisa@fjt.info.gifu-u.ac.jp

© AUR, 2009
doi:10.1016/j.acra.2008.10.012

similar images for classification of breast masses (11) and lung diseases (12). For similar images to be useful, we believe that images must be similar from the point of view of the diagnosing radiologists. However, in most previous studies, the similar images were selected based only on objective image features (13–17) and, to our knowledge, there were only a few studies (12,18) in which computed similarities were evaluated based on the radiologists' subjective similarities. Moreover, the number of studies on investigation of subjective similarity for image pairs by radiologists is limited (19–23). In this study, subjective similarity ratings for pairs of masses were determined by breast radiologists, and the average similarity ratings were employed as the ''gold standard'' for determination and evaluation of objective similarity measures. By training an artificial neural network (ANN) with radiologists' subjective similarity ratings as teacher and the image features as input data, it may be possible to determine similarity measures that would agree with radiologists' similarity ratings for unknown pairs.

When radiologists determine the similarity of a pair of masses, most, if not all, of them consider the margin characteristics of the masses. It is common practice that, when a radiologist finds a mass lesion, he or she reports the shape and margin descriptions in accordance with the Breast Imaging Reporting and Data System (BI-RADS) (24). These shape and margin descriptors suggest the likelihood of malignancy of the lesion. Baker et al (25) employed the BI-RADS descriptors in their computerized scheme for classification of benign and malignant lesions with some success. The BI-RADS descriptors assigned by radiologists can indicate subjective features of a lesion independent of the background structure. However, the subjective judgments would include variations. On the other hand, the computed image features are objective; however, these can be influenced by the image background and are sometimes incorrect. If both subjective and objective information can be used as input data for the ANN, the similarity measures may be determined more accurately. In this study, we investigated the similarity measures for pairs of masses using the BI-RADS lesion descriptors and computed image features for selection of similar images.

MATERIALS AND METHODS

Image Database

The images of breast masses used in this study were obtained from the Digital Database for Screening Mammography (DDSM) (26), which was developed by the University of South Florida and available to the public via website. The patients' health information was not included in the database; therefore, the consent form was waived. Regions of interest around the masses were obtained from craniocaudal and mediolateral-oblique views. In our study, lesions considered as architectural distortion, asymmetric density, and lymph node by a breast radiologist in our institution were not included. Lesions with findings of architectural distortion may be considered in the future when the number of such cases has increased. Some images were excluded for the following reasons: image quality was poor; a mass was very large and did not fit in a 5×5 cm square; a mass was partially visible at the edge of the film; and the pathology of a mass was not proved by biopsy. With these criteria, this study included 1568 regions of interest with 728 malignant and 840 benign masses, with pathologies confirmed by biopsy. This study was approved by the institutional review board.

For each mass, an outline of the mass was determined manually by one of two radiologists for accurate determination of the computerized image features. The contrast and density level of each region of interest were adjusted visually by a breast radiologist to the appropriate level for viewing and image feature determination. The BI-RADS lesion descriptors were provided in the DDSM. The shape descriptors for the 1568 masses included ''round,'' ''oval,'' ''lobular,'' ''irregular,'' and other (''tubular,'' ''n/a,'' ''architectural distortion,'' and ''asymmetry'' as determined by the original radiologists in the DDSM). The margin descriptors included ''circumscribed,'' ''obscured,'' ''microlobulated,'' ''ill-defined,'' and ''spiculated.'' Figure 1 shows the distributions of the numbers of lesions in each category.

Subjective Similarity Ratings by Breast Radiologists

For establishing the ''gold standard'' of the subjective similarity for pairs of masses, 300 pairs of masses were selected from the 1568 masses in the database. For inclusion of various types of mass pairs, it is preferable to include many pairs; however, because radiologists' time is valuable and limited, we believed that the 300 pairs were adequate. First, 50 masses including 25 that were benign and 25 that were malignant were selected as ''unknown'' masses by a breast radiologist to include masses with various sizes and characteristics. Subsequently, six masses, including three benign and three malignant masses, were selected as ''known'' masses, each of which was paired with the unknown mass; therefore, the 300 pairs included 75 malignant-malignant pairs, 75 benign-benign pairs, and 150 malignant-benign pairs. For understanding the difference between similar and dissimilar pairs, the subjective similarity data must include a wide range. If the cases were selected randomly, most pairs would be dissimilar and would not be useful for this study. Therefore, the known images were selected manually by a consensus of three physicists with the experience in medical imaging and mammography research to include pairs with a wide range of expected similarity. In addition, the known

Download English Version:

https://daneshyari.com/en/article/4219860

Download Persian Version:

https://daneshyari.com/article/4219860

Daneshyari.com