Pattern Recognition Letters 36 (2014) 117–124

Contents lists available at ScienceDirect

Pattern Recognition Letters

journal homepage: www.elsevier.com/locate/patrec

Texture and region dependent breast cancer risk assessment from screening mammograms

Xi-Zhao Li *, Simon Williams [1], Murk J. Bottema [2]

School of Computer Science, Engineering and Mathematics, Flinders University, Bedford Park 5042, SA, Australia

ARTICLE INFO

Article history:
Received 31 May 2013
Available online 12 October 2013

Communicated by C. Luengo

Keywords:
Texture analysis
Textons
Screening mammograms
Breast cancer
Risk assessment

ABSTRACT

Breast density is a known risk factor for breast cancer. Here two classes of texture features, one based on textons derived from local pixel intensity variation and one based on oriented tissue structure characteristics are measured on different regions of the breast in an effort to clarify the potential contribution of texture independent of local tissue density to estimate breast cancer risk. The region just behind the nipple is found to be the most significant local region for estimating risk, but estimates based on the entire breast perform better. Texton features are found to perform better than features based on oriented tissue structure.

1. Introduction

1.1. Breast cancer prevalence and early detection

Breast cancer is the second most common cancer diagnosed in women world wide. Around 1.38 million women across the world were diagnosed with breast cancer in 2008, accounting for nearly a quarter (23%) of all cancers diagnosed in women according to a report from Cancer Research UK in 2011 (Cancer Research UK, 2011). Breast cancer incidence in developed countries was generally higher than in developing countries, but breast cancer incidence has increased in most countries over the last few decades, with the most rapid increases occurring in the developing countries. In Australia, 1 in 8 women on average will develop breast cancer at some time in their life but only 1 in 37 will die of breast cancer before the age of 85 according to the report of Australian Institute of Health and Welfare in 2012 (Australian Institute of Health and Welfare, 2012). According to Cancer Australia (Cancer Australia, 2013), the number of newly diagnosed breast cancer cases in women in 2009 increased to 13,668 from 5317 in 1982. By 2020, it is estimated that 17,210 new breast cancer cases will be diagnosed in women in Australia.

Early detection is widely viewed as providing the best opportunity for reducing the morbidity and mortality due to breast cancer. As a result, many countries provide breast cancer screening based on mammography to facilitate early detection. The risk of breast cancer is not equal for all women and so adjusting early detection strategies according to the level of risk of an individual increases the efficiency and reduces the cost of screening programs. Accordingly, a good understanding of risk factors for breast cancer is vital. Known risk factors include age, weight, size of the breast, geographical location, diet, alcohol consumption and many more (McPherson et al., 2000). Of interest in this study are indicators of risk manifest in screening mammograms.

1.2. Texture and breast cancer risk

Wolfe was the first to study the relationship between mammographic appearance and breast cancer risk (Wolfe, 1976a,b). He proposed four breast pattern classes: $N1, P1, P2$, and DY and demonstrated a substantial increase in breast cancer risk progressing from $N1$ patterns to DY. $N1$ denotes a breast comprising mostly fat; $P1$ denotes a breast with a prominent duct pattern but limited in extent; $P2$ denotes an extended and prominent duct pattern; and DY denotes an extremely dense duct pattern.

The observations by Wolfe were reproduced by some studies (Wellings et al., 1975; Brisson et al., 1981; Boyd et al., 1984; Saftlas et al., 1989). However, other studies did not reproduce odds ratios as great as those reported by Wolfe and some even failed to find evidence of a relationship between Wolfe's breast patterns and breast cancer risk (Egan and Mosteller, 1977; Whitehead et al., 1985; Mendell et al., 1977).

The American College of Radiology has since introduced new breast pattern categories known as the BI-RADS classes

* Corresponding author. Tel.: +61 8 82017927.
E-mail addresses: li0691@flinders.edu.au (X.-Z. Li), s.williams@flinders.edu.au (S. Williams), murk.bottema@flinders.edu.au (M.J. Bottema).
[1] Tel.: +61 8 82012037.
[2] Tel.: +61 8 82013652.

0167-8655/$ - see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.patrec.2013.10.001

(Ame, 2003). BI-RADS classes are modified version of Wolfe classes, shifting the focus from structure patterns to the amount and distributions of dense tissue.

As the connection between breast tissue patterns and cancer risk emerged, the computer-aided diagnosis community developed related image analysis methods for determining risk. In an early work, Magnin et al. (1986) used the spatial gray-level dependence method (SGLDM) and gray level difference method (GLDM), which were based on gray-level co-occurrence matrices, to quantify density variations in mammograms to characterize images into Wolfe pattern classes for breast cancer risk evaluation. Caldwell et al. (1990) computed fractal dimensions in mammograms to classify them into the four Wolfe pattern classes. Tahoces et al. (1995) extracted texture features based on the Fourier transform, spatial relationships among grey levels and absolute values of the grey levels from three different regions of interest (ROIs) in craniocaudal (CC) view images to categorize mammogram images into Wolfe pattern classes. Byng et al. (1996) extracted texture features from fractal dimensions and grey-level histograms to classify images according to a six-category classification (SCC). The year after, the same group used texture features calculated from regional skewness and fractal dimension to characterize images into corresponding SCC categories (Byng et al., 1997). Karssemeijer (1998) applied two classifiers with and without pectoral features computed from grey-level histograms to categorize images into four density pattern classes.

In 2003, Petroudi et al. (2003) proposed the application of features extracted from textons generated from filter bank responses to classify mammograms into the four BI-RADS classes. Later, the same group used texton features generated from local $N \times N$ neighborhoods to characterize images into BI-RADS classes (Petroudi and Brady, 2006). Texton features from local $N \times N$ neighborhoods outperformed the filter banks in their previous paper. More recently, He et al., 2012 used textons to generate grey-level histograms to classify images into Tabár and BI-RADS classes. More details can be found in their previous work (He et al., 2009), where the term "cluster center" is used to describe what is now commonly called "texton". Karemore et al. (2012) calculated texture features from Gaussian derivatives at four different scales to classify both cancer vs controls and Estrogen-Receptor (ER) subtype specific classes (ER-positive vs ER-negative).

A key question in computer-aided risk assessment is whether patterns relevant to breast cancer risk are concentrated in a particular region or spread throughout the breast. Huo et al. (2000, 2002) consistently selected a ROI of 256×256 pixels from the central breast region behind the nipple regardless of the breast size to classify images into high or low risk groups. Texture features were extracted from local gray-level variation analysis and an average A_z score of 0.91 was obtained in classifying BRCA1/BRCA2 mutation carriers and non carriers. They also found that high risk images tended to be dense and mammographic patterns appeared as a coarse low contrast texture. Choosing the central region behind the nipple is reasonable since this region is usually the densest part of the breast and density is a significant indicator of breast cancer risk (Huo et al., 2000; Li et al., 2008). In 2004, the same group, studied the effect of ROI size and location on breast cancer risk again using the BRCA1/BRCA2 mutations to assign high and low risk groups (Li et al., 2004). Five ROIs were selected manually from left CC view images: (A) the central breast region immediately behind the nipple, (B) the upper central breast region, (C) the lower central breast region, (D) the center of the central breast regions, and (E) the central left breast region. Their results showed that the size of the ROI was not important but there was a statistically significant decrease in classification performance as the ROI location varied from the central region behind the nipple (A) to other locations (B, C, D, and E). In 2008, they applied power law spectral analysis to mammograms to classify BRCA1 and BRCA2 mutations carriers (Li et al., 2008). They achieved an A_z score of 0.9 in differentiating 30 BRCA1/BRCA2 gene mutation carriers and 60 age-matched low risk women.

1.3. Risk measures and density dependence

There is no gold standard for breast cancer risk. Thus, in the studies above, true breast cancer risk was replaced by schemes for determining Wolfe texture pattern classes, BI-RADS classes, SCC or BRCA1 and BRCA2 mutation carriers. All these are reasonable, but are still surrogates for true breast cancer risk.

The approach we take is based on the co-occurrence of cancer in both breasts. The American Cancer Society (American Cancer Society, 2012), Cancer Australia (Cancer Australia, 2012) and Cancer UK (Cancer Research UK, 2012) report a three- to five-fold increase of risk for cancer in the breast if the contralateral breast is detected with cancer. Accordingly, we take as high risk, images of the cancer-free breast from a screening visit in which cancer was detected in the contralateral breast. Low risk images were taken as either the right or left breast (selected randomly) from a screening visit in which no cancer was detected in either breast.

In the studies cited above, the contribution of texture to risk assessment cannot be separated entirely from the contribution of density. In most imaging systems, the relationship between tissue density and image intensity is non-linear (Wolbarst, 2004, Chapter 29). As a result, texture superimposed on dense tissue does not produce the same response as the same texture superimposed on fatty tissue. Hence, even if the local background is subtracted (a natural by product of many processing steps), the texture measures may still carry density information. Since breast density is known to correlate to breast cancer risk, these texture features may simply be measuring density instead of providing a measure of breast cancer risk independent of density.

In order to better separate the contribution of texture from the contribution of density to risk assessment, image intensities in this study are normalized so as to remove both the local mean intensity and the local variance of intensity (Li et al., 2012).

Two classes of texture features are used; one based on $N \times N$ neighborhood patches and one based on oriented tissue structures. The former was chosen because of the success of these features reported previously (Varma and Zisserman, 2003; Li et al., 2013) and the latter ties texture features directly to biological structure and therefore has the potential to deliver a causal result rather than just an observational one. Although breast density is usually most pronounced in the region just behind the nipple, whether this holds for texture is not known. Accordingly, several regions of the breast as well as the full breast were examined separately.

2. Data and preliminary processing

2.1. Image data

Mammograms for this study were selected from the publicly available Digital Database of Screening Mammography (DDSM) (Heath et al., 1998, 2001). The images selected were acquired by Lumisys and Howtek machines at spatial resolutions ranging from 42 μm to 50 μm per pixel and depth ranging from 12 to 16 bits. The data set used in this study consists of 320 left and right CC view images. Of these, half were images from normal cases (no cancer found in either breast at the current screening round) to serve as the low risk group and half were images of the normal breast from cases found to have cancer in the contralateral breast at the current screening round to serve as the high risk group. Images in the DDSM database are annotated with BI-RADS category ratings and these were used to ensure a wide range of