Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 90 (2016) 169 – 174

Procedia Computer Science

International Conference On Medical Imaging Understanding and Analysis 2016, MIUA 2016, 6-8 July 2016, Loughborough, UK

Spatial Relations of Mammographic Density Regions and their Association with Breast Cancer Risk

Maya Alsheh Ali[a,b,*], Mickaël Garnier[c,d], Keith Humphreys[a,b]

[a] *Department of Medical Epidemiology and Biostatistics, Karolinska Institutet, SE-17177 Stockholm, Sweden.*
[b] *Swedish eScience Research Centre (SeRC), Karolinska Institutet, SE-17177 Stockholm, Sweden.*
[c] *Institut Curie, PSL Research University, CNRS, UMR3664, F-75005 Paris, France*
[d] *Sorbonne Universities, UPMC Univ Paris 06, CNRS, UMR3664, F-75005 Paris, France*

Abstract

We present a new approach for characterising the shape and the spatial relationships of different categories of density in mammograms. Descriptions of regions are encoded using a forces histogram method and across-image variation is captured using functional principal component analysis. We evaluate the association of the features with breast cancer based on a pilot case-control study using logistic regression with percent density, age, and body mass index included as adjustment variables. The spatial relations were significantly associated with breast cancer status (p= 0.009). Our approach can provide insights into the role of different density regions in the development of breast cancer.

Peer-review under responsibility of the Organizing Committee of MIUA 2016
Keywords: Mammography, Spatial organisation, forces histogram, Breast cancer risk

1. Introduction

Breast cancer is the most common cancer diagnosed in women and despite clinical advancements, in the nordic countries, around 20% of breast cancer patients still die of their disease. The most established image-based risk factor for breast cancer is Percent mammographic Density (PD). This is classically evaluated as the ratio of the fibroglandular tissue area over the surface of the entire breast in a mammogram[1]. Volumetric measures of PD, obtained from a 2D image by, for example, placing phantoms between the plates of the mammography machine to act as a reference from which dense tissue thickness can be estimated, have also been developed. It is not clear which type of measure is optimal for risk prediction[2], but, in general, women with high PD ($> 75\%$) have an approximately 6-fold increased risk compared to women with very low PD. Besides PD (or the absolute dense area/volume) there is likely to be additional relevant information which can be extracted from the mammogram. A number of studies have tried to measure heterogeneity of the parenchymal pattern by using texture-based methods ranging from simple (e.g.intensity histogram), to complex (e.g. based on scale space features). Texture features can be extracted from the entire breast region[3] or from specific regions of interest, such as the retroareolar area[4], the central area of the breast or across a lattice covering the breast.

* Corresponding author
E-mail address: maya.alsheh.ali@ki.se (Maya Alsheh Ali).

1877-0509
Peer-review under responsibility of the Organizing Committee of MIUA 2016
doi:10.1016/j.procs.2016.07.019

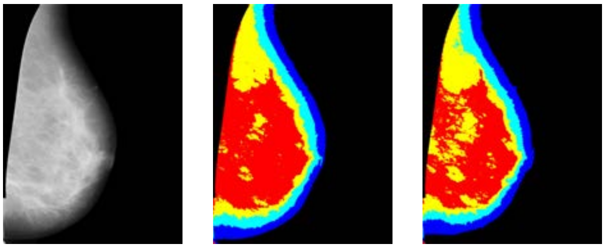

Fig. 1. Illustration of image segmentation. The original image (in (a)) segmented into 4 regions using fuzzy c-means is in (b) and using Otsu's method is in (c), where the colours dark blue, blue, yellow and red correspond to fatty, semi-fatty, semi-dense and dense tissue regions respectively.

It is also interesting to study the spatial distribution/shape of dense tissue in the breast and its relationship with breast cancer risk. It has been suggested that the relative distribution of adipose and fibroglandular tissue is involved in breast cancer development[5]. There are, in fact, multiple ways to describe the spatial organisation of objects inside an image, the most basic relying on qualitative measures such as "to the right of", "close to", or "surround"[6], which are examples of directional, distance and topological descriptions. Keller et al.[7] measured the distance from the centroid of a segmented region of dense tissue to the skin line. Since the relations between two objects is rarely absolute, fuzzy logic may be useful for assessing the degree of truth of a given relation. The noteworthy histogram of angles[8] encapsulates the fuzzy directional relations of all considered directions and several methods stem from this approach[9]. The forces histogram[10] (FH) is a quantitative fuzzy spatial relation description taking into account the directional and the distance relationships as well as the shapes of the objects. The FH approach has been applied to lung CT images to classify lesions according to their relative position with anatomical landmarks[11]. In this article we describe how it can be used to comprehensively assess whether the spatial organisation of different regions of tissue, or the shapes of these regions, is associated with breast cancer risk.

2. Methods

We start our methods description from the point at which the breast region has already been segmented (i.e. tags, background and pectoral muscle have been removed). After segmenting the breast into regions according to the density, we used the FH approach to capture the shape and spatial relations information. We captured the variation inside the set of FHs using functional principal component analysis[12] and assessed associations between principal components and breast cancer status by fitting logistic regression models.

2.1. Segmentation of regions of density

Segmentation of regions within the breast, on a mammogram, is an important and challenging problem and methods are under continual development[13]. The number of regions/density categories used in the literature has varied widely[13] (anything from two to thirteen) and its choice is likely to vary according to the purpose for which the regions are used (e.g. for quantifying density amount, texture or spatial organisation). Here we segment the breast into four regions, which is a common choice and is the same number of regions used in Wolfe's original parenchymal patterns[14] (known as N1, P1, P2 and DY) for categorising according to both the extent of densities and their characteristics (prominence of ducts and dysplasia). Four regions, representing very dense tissue, dense with structures (fibrotic stromal tissue and glandular tissue), fatty tissue (Wolfe's normal breast pattern, N1), and fatty breast edge, have also been used in studies of textural features[15].

We have used two methods to see how sensitive our overall analysis is to the choice of segmentation algorithm, the first being the fuzzy c-means (clustering) method (Fig. 1 (b)) and the second being Otsu's (global thresholding) method (Fig. 1(c)). Both are considered to provide accurate, albeit somewhat different, segmentations[13].

2.2. Shape and spatial relations measurement

To comprehensively measure the relative positions of the different regions of tissue and summarise their shapes, we use the FH method[16]. Although the idea of the FH is elegant, its computation is demanding and complicated. Here