Computers in Biology and Medicine 74 (2016) 91–102

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Detection of lobular structures in normal breast tissue

Grégory Apou [a,*,1], Nadine S. Schaadt [b,1], Benoît Naegel [a], Germain Forestier [c], Ralf Schönmeyer [d], Friedrich Feuerhake [b], Cédric Wemmert [a], Anne Grote [b]

[a] ICube, University of Strasbourg, 300 bvd Sébastien Brant, 67412 Illkirch, France
[b] Institute for Pathology, Hannover Medical School, Carl-Neuberg-Straße 1, 30625 Hannover, Germany
[c] MIPS, University of Haute Alsace, 12 rue des Frères Lumière, 68093 Mulhouse, France
[d] Definiens AG, Bernhard-Wicki-Strasse 5, 80636 Munich, Germany

ARTICLE INFO

Article history:
Received 5 February 2016
Received in revised form
29 April 2016
Accepted 8 May 2016

Keywords:
Whole slide image
Digital histopathology
Normal breast lobule
Image analysis
Convolutional neural network

ABSTRACT

Background: Ongoing research into inflammatory conditions raises an increasing need to evaluate immune cells in histological sections in biologically relevant regions of interest (ROIs). Herein, we compare different approaches to automatically detect lobular structures in human normal breast tissue in digitized whole slide images (WSIs). This automation is required to perform objective and consistent quantitative studies on large data sets.

Methods: In normal breast tissue from nine healthy patients immunohistochemically stained for different markers, we evaluated and compared three different image analysis methods to automatically detect lobular structures in WSIs: (1) a bottom-up approach using the cell-based data for subsequent tissue level classification, (2) a top-down method starting with texture classification at tissue level analysis of cell densities in specific ROIs, and (3) a direct texture classification using deep learning technology.

Results: All three methods result in comparable overall quality allowing automated detection of lobular structures with minor advantage in sensitivity (approach 3), specificity (approach 2), or processing time (approach 1). Combining the outputs of the approaches further improved the precision.

Conclusions: Different approaches of automated ROI detection are feasible and should be selected according to the individual needs of biomarker research. Additionally, detected ROIs could be used as a basis for quantification of immune infiltration in lobular structures.

1. Introduction

Lobular structures are the functional units of the resting mammalian breast that further differentiate into milk-producing glands during lactation. The normal anatomical structures are important because breast cancer and premalignant lesions originate in these epithelial structures and there is evidence that the transition between ductal and lobular structures may be particularly susceptible to oncogenic events [1]. In addition to studies on the origins of cancer, the detection of ducts and lobules is also relevant for an inflammatory condition referred to as *lymphocytic lobulitis* (LLO), which has been observed in the adjacent tissue around breast cancer and in prophylactically removed breast tissue without any evidence for cancer in *BRCA1/2* mutation carriers [2,3]. This phenomenon is not yet well understood and deciphering its possible link with hereditary breast cancer may lead to better disease understanding, new prognostic indicators, or novel treatment options. In order to perform an objective, repeatable, and statistically reliable quantitative study of LLO on large data sets, the ability to automatically detect relevant structures in histological slides is necessary. We refer to ducts and lobules as lobular structure in the following.

Nowadays, such slides can be routinely digitized; the resulting whole slide images (WSIs) can be processed by automated image analysis techniques with the aim to detect lobular structures and to quantify cell numbers [4,5]. Many works are based on detecting and automatically counting cells for cancer diagnosis, grading, and prognosis [6–8]. However, to give more insight to the pathologist, it is necessary to evaluate immune cells beyond estimation of their density, for example by object-based recognition of spatial patterns and interactions at high resolution [9]. Thus, our objective is to pave the way for identifying and classifying cells in those areas in image that are most relevant, like lobules in breast cancer and LLO, finally enabling methods to characterize the spatial distribution of different subtypes of immune cells in relation to these larger image objects.

In general, lobular structures are composed of dense areas of epithelial cells in tubular structures, normally with a clear contrast

* Corresponding author.
E-mail address: gapou@unistra.fr (G. Apou).
[1] Co-first author.

http://dx.doi.org/10.1016/j.compbiomed.2016.05.004
0010-4825/

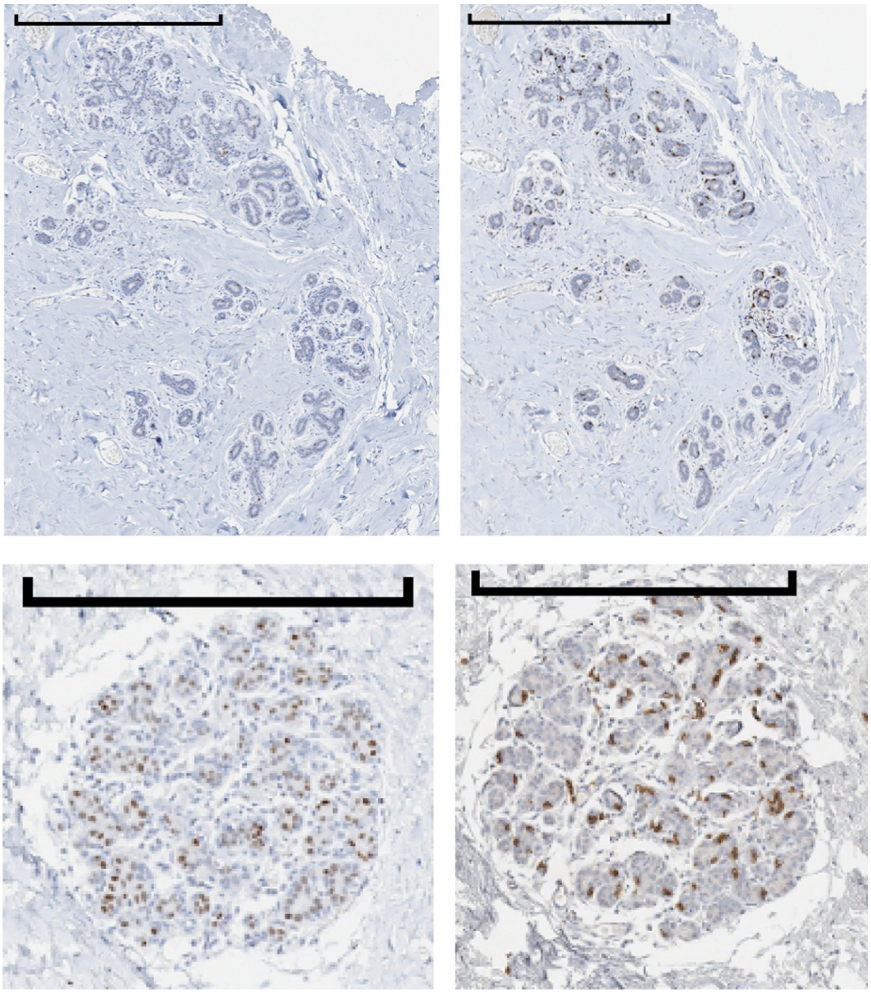

Fig. 1. Top row shows a large lobule with sparse branching in two different stainings (left: ER, right: CD8). Bottom row shows a small lobule with dense branching (left: ER, right: CD8). Scale bars are 0.5 mm.

to the lobular stroma. However, lobules can be very different in size, shape, and texture depending on their functional stage (e.g., phase of the menstrual cycle), the degree of immune cell infiltration, and also appear differently according to the used staining method (see Fig. 1).

Cancer growth adds to this complexity of tissue structures and makes manual or automated lobule detection even more challenging. As about 75% of invasive breast cancer cases are of ductal type and detection of lobular type is more difficult by mammography [10], it is important to distinguish between ducts and lobules in this context, but we do not provide a way to make this distinction. As a preparatory work before delving into the structural complexity of cancer-affected tissue, we focus on normal breast tissue of healthy patients in order to evaluate and compare three different image analysis methods to automatically detect epithelial structures including ducts and lobules in immunohistochemically (IHC) stained WSIs. Even without solving the cancer-related challenges, the work addresses an important demand because the evaluation of LLO may lead to new biomarker patterns with diagnostic and prognostic value. From a technological point of view, the task of defining regions of interest (ROI) for further analysis tackles a problem that occurs frequently in medical image analysis: the limited availability of experts to perform large-scale manual annotation due to restricted resources of trained pathologists remains an important bottleneck for progress. In the context of LLO, this is evident because the statistical base for experimental evidence has been limited by the available time for pathologists who have to manually select every lobule and therefore could so far only annotate small data sets [11,12]. Thus, overcoming the limitations of manual annotation of WSIs by automation is highly desirable. This work builds on previous work [13], where detection of normal lobules in the vicinity of breast cancer was optimized for the purpose of analyzing nuclear expression of estrogen receptor (ER) or progesterone receptor (PR). Grote et al. detected lobules on several segmentation layers using textural, geometric, and relational features, as well as solid tumor using textural features. Other tissue classification techniques have supported the study of pathologies like odontogenic cysts [14] and various cancers [15,16]. The identification of general biological structures has received comparatively little attention, although graph-based approaches exist for unsupervised top-down tissue categorization [17] and bottom-up biological object identification [18].

A machine learning algorithm describes how to identify patterns in existing data (learning) and uses this acquired knowledge to make predictions on new data [19]. A deep learning algorithm is a machine learning algorithm that can learn a hierarchical description of the data with multiple sublevels of nonlinear features [20]. In recent years, a type of deep learning architecture optimized for 2D data called convolutional neural networks (CNNs) [21] have provided state-of-the-art results in various applications of machine learning-based image analysis, from general scene labeling [22] to cancer classification [23] and mitosis detection [24].

After presenting our data set and evaluation criterion, we will describe three methods that were developed in the context of