Computers in Biology and Medicine 79 (2016) 80–91

Contents lists available at ScienceDirect

Computers in Biology and Medicine

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

Influence of feature set reduction on breast cancer malignancy classification of fine needle aspiration biopsies

Łukasz Jeleń[a,*], Adam Krzyżak[b], Thomas Fevens[b], Michał Jeleń[c]

[a] Department of Computer Engineering, Wrocław University of Science and Technology, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland
[b] Department of Computer Science and Software Engineering, Concordia University, 1455 De Maisonneuve Blvd. W., Montréal, Québec, Canada H3G 1M8
[c] Department of Pathology and Oncological Cytology, Medical University of Wrocław, Borowska 213, 50-556 Wrocław, Poland

ARTICLE INFO

Keywords:
Computer-aided diagnosis
Breast cancer
Pattern analysis
Cancer classification
Malignancy grading

ABSTRACT

Grading of breast cancer malignancy is a key step in its diagnosis, which in turn helps to determine its prognosis and a course of treatment. In this paper, we consider the application of pattern recognition and image processing techniques to perform computer-assisted automatic breast cancer malignancy grading from cytological slides of fine needle aspiration biopsies. To determine a classification of the malignancy of the slide, a feature set is first determined from imagery of the slides. In this paper we investigated the nature of a wide set of features extracted from biopsy images to determine their discriminatory power and cross-correlation. Feature vector reduction is studied using a correlation map of the features, determining discriminatory power using the Kolmogorov–Smirnov test, significant feature selection, and stepwise feature selection. The reduction of the feature vector simplifies the complexity of classification scheme and does not impair the classification accuracy. In some cases a decrease of the error rate is noted. Based on this analysis, we present an improved classification system for cancer malignancy grading.

1. Introduction

Recent statistics show that breast cancer is the most common cause of death among middle-aged women. In 2015, the Breast Cancer Society of Canada [1] released statistics which shows that each week in Canada about 481 women will be diagnosed with breast cancer. Due to the fact that cancers in their early stages are more vulnerable to treatment, we can assume that most of the diagnosed cases will lead to a successful recovery if they are diagnosed at an early stage. Conversely, most advanced cancers stages are usually almost impossible to treat. The diagnosis process assumes that cancer will be assigned a malignancy grade. The determination of cancer malignancy is also known as a prognostic factor of a cancer because it allows for precise estimation of the cancer behavior, with or without undertaking treatment. According to this factor an appropriate treatment is suggested and therefore it is an important part of the breast cancer diagnosis.

To overcome the problem of late diagnosis, screening mammographic tests were introduced and when a suspicious region in the image is noted, a fine needle aspiration biopsy (FNA) is taken. FNA is a method of extraction of a small tissue sample of the questionable breast tissue. This minimally invasive method allows for the description of the type and malignancy grade of the cancer. In 1957, Bloom and Richardson proposed a grading scale that assists in this difficult task. The proposed scheme was originally derived for assessment of malignancy from histopathological slides and today is very popular among pathologists. They use it for grading not only histological but also cytological tissue.

Here, we are addressing a very important area of research which is the computerized automatic detection of pathologies from cytological images. The process of automatic cancer grading suffers from large variations in cancer imaging and analysis what makes that task very challenging [2]. Recent efforts have applied deep learning to cytological images for breast cancer to deal with the complexity of the data [3]. In this paper, to seek to improve the results presented in our previous work [4], we study the features extracted from biopsy images to determine their relative discriminatory power and cross-correlations with the aim to reduce the size of the final feature vector to simplify the complexity of classification scheme with minimal influence on the error rate while achieving an improvement in classification quality.

The layout of the remainder of the paper is as follows. In Section 2, we review the related literature for computerized classification of breast cancer cytological images, in particular discussing the Bloom–Richardson grading scheme, and the necessary background in the areas of image processing and feature selection. In Section 3, we also

* Corresponding author.
E-mail address: lukasz.jelen@pwr.edu.pl (Ł. Jeleń).

http://dx.doi.org/10.1016/j.compbiomed.2016.10.007
Received 10 March 2016; Received in revised form 21 September 2016; Accepted 11 October 2016

introduce our FNA database used for simulations, as well as the set of extracted features to be able to perform classification to assign a Bloom–Richardson grade. In this section, as well, we introduce the classifiers used to compute the final Bloom–Richardson grade. In Section 4, we present the results of our feature selection methods and the related simulation results for classification. We finish the paper with Section 5 with our conclusions and briefly discuss future work, respectively.

2. Background

In this section, we review related literature and necessary background for the automated classification of cytological imagery of FNAs for breast cancer.

2.1. Cytopathology and breast cancer

The related literature can be viewed as dealing with two main problems arising from the detection of breast cancer pathologies from cytological imagery. The first is *malignancy diagnosis* which deals with the benign/malignant classification of fine needle aspirates. The main result of malignancy diagnosis is to determine whether cancer is present or not, while trying to minimize the number of false positives (the erroneous determination of the presence of cancer). The second problem is *malignancy grading* which involves the classification of the malignancy stage of cancer for all the reasons mentioned above.

A literature review showed that there exists many approaches to computerized analysis of breast cancer imagery with by far the majority dealing with the problem of malignancy diagnosis [5]. To the best of our knowledge, the first description of the computerized breast cytology classification problem was provided by Wolberg et al. in 1990 [6]. In this work, the authors presented an application of a multi-surface pattern separation method applied to cancer diagnosis. The database used in the study consisted of 169 malignant and 201 benign cases and the described algorithm was able to achieve a 6.5% error when 50% of a data was used for training. The authors showed that increasing a training set to 67% allowed them to reduce the error to 4.1%. Apart from the work on the classification problem, Wolberg et al. in [7] described a Wisconsin Breast Cancer Database (WBCD), which is a widely used database of features of breast cancer nuclei. The features are pre-extracted from fine needle aspiration biopsy images. In 1993, a 'snake' active contour algorithm was used by Street et al. [8] for segmentation and precise description of nuclei shapes. Additionally they proposed 10 nuclei features relevant for classification. The described system was based on a multi-surface method and was able to classify the breast cancer cases with an accuracy of 97.3%. In 2005, Xiong et al. [9] described a partial least squares regression as a classification system. They used a WBCD database with 699 (241 malignant, 458 benign) cases and the achieved classification rate was 96.57%. Further review of literature showed that various other researchers have used the WBCD database (see [10] and reference therein) and the achieved results were ranging from 94.74% to 99.54%.

Furthermore, in 2009, Malek et al. [11] described an application of active contours to nuclei segmentation and a fuzzy c-means classifier for classification of 200 (80 malignant, 120 benign) cases achieving a 95% classification rate. In 2010, Niwas et al. [12] described a complex wavelet transform applied to a nuclei Chromatin texture and a k-nearest neighbor classifier. The authors used a data set of 20 malignant and 25 benign cases for classification and the reported classification rate was 93.33%. There are two other approaches that utilize a circular Hough transform to detect cell nuclei. The first approach of Filipczuk et al. [13] used an SVM for subsequent classification of segmented nuclei to decide if the segmentation of the nucleus is correct or not. The authors were able to achieve a classification rate of 98.51% using k-nearest neighbor, naive Bayes, or an SVM classifier on selected features sets of 67 (42 malignant, 25 benign) cases. The second approach was described by George et al. [14]. In the described work, the authors were able to confirm a segmented nuclei with the application of thresholding and fuzzy c-means clustering. Based on such segmentation they were able to extract 12 features that were later introduced to several neural network architectures achieving a sensitivity of 95.49% and specificity of 83.16% for the probabilistic neural network. The data set used by the authors consisted of 92 (47 malignant, 45 benign) cases and the WBCD database as a comparison.

It should be noticed that all mentioned approaches are purely concentrated on malignancy diagnosis of fine needle aspirates for the presence of breast cancer. As mentioned above, the classification of the malignancy stage of cancer is called a malignancy grading, and this is the problem that we are focusing on in this paper. The pre-screening process before taking an FNA results in the situation in which the biopsy slide being classified is nearly always malignant. Henceforth, in this work, we deal with a malignancy grading problem instead of malignancy diagnosis. In previous work, we have explored a framework based on pattern recognition methods along with image processing techniques that was able to extract features from FNA slides which were later used for the automatic assignment of a malignancy grade. In [4], we presented an approach based on the level sets segmentation method. Classification effectiveness was tested on 110 (44 high malignancy, 66 intermediate malignancy) images with results reaching an overall accuracy of 82.7%.

2.2. Bloom–Richardson breast cancer malignancy grading scheme

The diagnosis of breast cancer can be divided into several stages. The first stage is screening which is performed in a non-invasive manner with mammography. With mammography, a physician is able to localize microcalcifications and other indicators within a breast tissue. Depending on the findings in mammography image, a patient can be sent to a Pathologist for a more precise diagnosis. When a suspicious region is found, a fine needle aspiration biopsy (FNA) is taken and a small sample of the questionable breast tissue is extracted. This minimally invasive method allows for the precise description of the type of the cancer. During the FNA examination, the tissue is assigned a malignancy grade, which is also called a prognostic factor because it allows for a precise estimation of cancer behavior with or without undertaken treatment. According to that prognostic factor the appropriate treatment is determined.

Cancer malignancy assessment is an intricate process involving the assessment of various nuclear features. Such an assessment allows for the estimation of a malignancy grade which is determined according to the numerical scale proposed by Bloom and Richardson in 1957 [15]. The Bloom–Richardson grading system (BR) was initially described by Bloom and Richardson [15] and is now widely used by pathologists. In 1989, the originally proposed scheme was modified by Scarff and is now recognized as a modified Scarff–Bloom–Richardson scheme. For the breast cancer diagnosis this scheme is one of the best known prognostic factors [16]. The most widely used system for breast tumor grading is the Elston and Ellis modification of Scarff–Bloom–Richardson histologic grading [17] (also known as the Nottingham histologic grade). All of the above schemes are based on grading three types of features called factors: cells' polymorphy, ability to reform histoformative structures, and mitotic index. The Bloom–Richardson scheme assesses these features in a point-based scale as:

1. **Degree of structural differentiation (SD)** - Originally this factor was defined as tubule formation that described the cells' ability to create tubules. As opposed to histopathological slides, in cytological smears of FNA, tubules are not preserved. Hence, the scoring for this factor is performed based on the evaluation of the cells' ability to form groups within a smear. An example of SD factor is illustrated on Fig. 1a and b. On Fig. 1a one can notice that only a single group is visible, while on Fig. 1b multiple dispersed groups