Biomedical Signal Processing and Control 30 (2016) 149–161

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

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

Effective segmentation of orphan annie-eye nuclei from papillary thyroid carcinoma histopathology images using a probabilistic model and region-based active contour

Angel Arul Jothi J., Mary Anita Rajam V.*

Department of Computer Science and Engineering, College of Engineering Guindy, Anna University, Chennai, India

ARTICLE INFO

Article history:
Received 1 July 2015
Received in revised form
14 November 2015
Accepted 22 June 2016
Available online 14 July 2016

Keywords:
Papillary thyroid carcinoma (PTC)
Histopathology
Orphan annie-eye nuclei
Expectation maximization/Maximization of the posterior marginals (EM/MPM)
Chan-Vese active contour model (ACM)
Detection
Segmentation

ABSTRACT

The presence of orphan annie-eye nuclei is a significant feature for the diagnosis of papillary thyroid carcinoma (PTC), a cancer of the thyroid gland. Automated detection and segmentation of orphan annie-eye nuclei from histopathology imagery is an intricate procedure due to traditional and specific challenges. The specific challenges are posed by the biological properties of these nuclei. This paper propositions an automated method to detect and segment orphan annie-eye nuclei from papillary thyroid carcinoma histopathology images. Our proposed method (EM/MPM-CV) initially uses a Markov random field-based segmentation technique to detect the orphan annie-eye nuclei seeds from the given images. A region-based active contour model (ACM) is initialized and evolved over the nuclei seeds to identify the final nuclei contours. The EM/MPM-CV method is evaluated on 149 PTC histopathology images for detection and segmentation performance. This technique gives a detection sensitivity of 87% and positive predictive value of 93%. The directed Hausdorff distance (DHD) and the mean absolute distance (MAD) values for the proposed method are found to be 3.79 and 1.55 pixels respectively.

1. Introduction

Histopathology is a branch of pathology particularly used as a golden rule for cancer diagnosis [1,2]. A histopathology study done by a pathologist involves viewing stained biopsy tissue sections of patients fixed onto glass slides under a microscope. This is a labor-intense work that is subject to intra-observer and inter-observer variations and demands expertise.

Recently, automated histopathology image analysis systems are being developed for region of interest (ROI) detection, segmentation tissue classification and disease detection. These automated systems are called computer aided detection/diagnosis (CAD) systems. The main goal of a CAD system is to increase the sensitivity of detection/diagnosis while not compromising on the specificity of detection/diagnosis [3]. CAD systems can assist pathologists and provide quantitative disease diagnoses [1,2].

Papillary thyroid carcinoma (PTC) is one of the histological subtypes of thyroid cancer that is most common (accounts for 80%) among human beings [4–7]. Pathologists diagnose PTC by examining stained biopsy tissue sections of thyroid under a microscope looking for vital visual cues like orphan annie-eye nuclei, nuclear groves, nuclear pseudo inclusions and crowding of nuclei. Among these, pathologists consider the presence of orphan annie-eye nuclei as a significant feature for the diagnosis of PTC [8,7]. Consequently, there is a need to develop CAD systems to automatically detect and segment orphan annie-eye nuclei from PTC histopathology images.

Automated orphan annie-eye nuclei detection from PTC histopathology imagery suffers traditional challenges like clumping of nuclei, intensity variation within nuclei, weak or missing nuclei boundaries and morphometric variation between nuclei. Apart from these challenges, the peculiar appearance of the orphan annie-eye nuclei sabotages the segmentation process. While observed under a microscope, an orphan annie-eye nucleus appears to be enlarged in size. The center of an orphan annie-eye nucleus lacks chromatin, while its periphery contains a rim of chromatin. Due to lack of chromatin, an orphan annie-eye nucleus looks optically clear and is also called ground glass nucleus. Consequently, the staining characteristics of these nuclei also vary, that is, the center of these nuclei does not stain and appears white in color and the periphery of these nuclei stain blue. Hence, the property of the pixels at the center of these nuclei differs from those at their periphery. Furthermore, the PTC histopathology images also contain open-faced and close-faced nuclei whose appearance varies significantly from that of an orphan annie-eye nucleus.

* Corresponding author.

http://dx.doi.org/10.1016/j.bspc.2016.06.008
1746-8094/

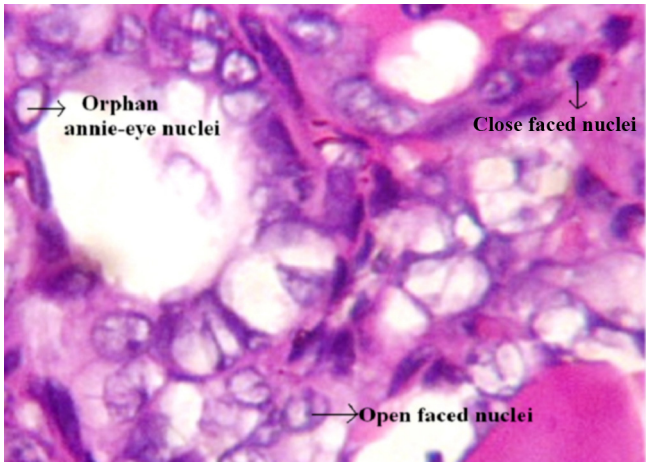

Fig. 1. PTC histopathology image showing orphan annie-eye, open-faced and close-faced nuclei.

The close-faced nucleus possesses dense chromatin and it stains darkly. The open-faced nucleus has loose chromatin and it stains lightly [8,9]. An example image showing the three different types of nuclei is shown in Fig. 1. Therefore, to identify a segmentation method to detect orphan annie-eye nuclei and to capture the size and shape of these nuclei accurately is a challenging task.

This paper describes an automated method (EM/MPM-CV) to detect and separate orphan annie-eye nuclei from other structures such as cytoplasm, close-faced nuclei, open-faced nuclei and the background present in histopathology images of PTC. The steps involved in the proposed model are shown in Fig. 2. The digital images are acquired from tissue biopsy samples and are converted into gray scale. The images are filtered using a Gaussian filter and are subjected to contrast enhancement using linear gray-level intensity transformation. These operations help to emphasize the nuclei areas and also prepare the images for segmentation.

The filtered and enhanced images are then given as input to the expectation maximization/maximization of the posterior marginals (EM/MPM) algorithm [10]. The EM/MPM algorithm partitions the given image into binary scenes where each binary scene contains certain regions of the input image. From the resulting binary scenes, the binary scene containing the orphan annie-eye nuclei is selected automatically based on the highest average pixel intensity. The selected binary scene is given as input to the nuclear seed detection step. In the nuclear seed detection step, orphan annie-eye nuclei seeds are separated from the other objects using a size constraint. A contour is initialized using the detected nuclei seeds and is evolved using the Chan-Vese active contour model (ACM) algorithm [11]. This helps to accentuate the size and shape of the detected orphan annie-eye nuclei and to improve the segmentation accuracy. After segmentation, small non-nuclei objects are removed using morphological operations.

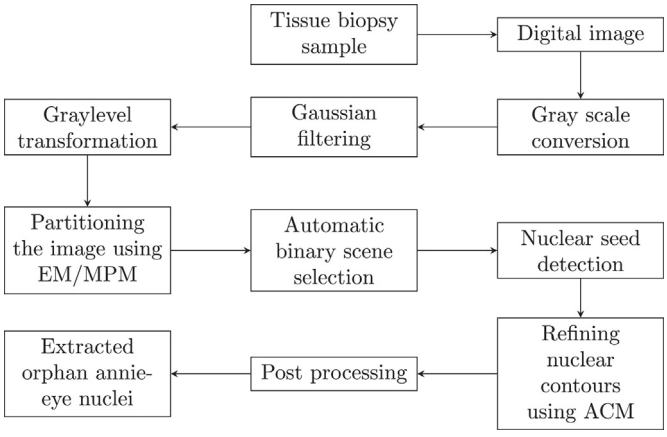

Fig. 2. Diagram illustrating the steps involved in the extraction of orphan annie-eye nuclei from PTC images.

The EM/MPM-CV method can be used as an automatic computer aided detection system that assists pathologists. This enhances the diagnostic reliability and accuracy. Moreover, it can be extended to an automatic diagnosis system to detect cancer in histopathology images by extracting nuclear features and training a classifier.

The novelty of this work is as follows: this work aims to segment orphan annie-eye nuclei from papillary thyroid carcinoma (PTC) histopathology images. To the best of our knowledge, there is no previous work that detects and segments orphan annie-eye nuclei from PTC histopathology images. Our segmentation method is fully automatic and unsupervised. It overcomes the traditional and specific challenges posed by the PTC histopathology images and extracts the orphan annie-eye nuclei by encompassing two algorithms such as the EM/MPM and Chan-Vese ACM. The automatic binary scene selection step and the orphan annie-eye nuclei seed detection step help these algorithms to work in tandem. This synergistic segmentation scheme provides better results when compared to the results of the stand-alone algorithms.

The remainder of the paper is organized as follows: Section 2 describes the related work in the field. Section 3 explains the dataset, ground truth preparation and the proposed model (EM/MPM-CV) for segmenting the orphan annie-eye nuclei. The experiments conducted, metrics used for evaluation and the results obtained are provided in Section 4. The discussion is presented in Section 5. Finally, the conclusion is given in Section 6.

2. Related work

Detecting and segmenting cell nuclei from histopathology images is a primary task in the development of automated image analysis systems [12]. This section summarizes the various works that have concentrated on detecting and segmenting different regions of interest (i.e., nuclei, cells and lymphocytes) from different histopathology images.

Nuclei were automatically segmented from breast histopathology images by applying watershed segmentation at multiple scales and with multiple markers. After a post-processing step that eliminated the false regions, the results of the multiple scales were merged to get the final nuclei areas [12]. An automatic method to detect centroblasts from follicular lymphoma histopathology images was proposed. The method combined the Gaussian mixture model (GMM) and expectation maximization (EM) algorithm along with local adaptive thresholding to segment the cells. After this step, size, shape and texture distribution constraints were imposed on the segmented cells to identify the exact candidate centroblast cells [13]. Watershed transform and gradient vector flow (GVF) snakes were used to segment nuclei from hepatocellular carcinoma images [14]. Thresholding is a traditional method for image segmentation and optimal thresholding was used to segment nuclei from skin histopathology images [15].

A combined region-, boundary- and shape-based active contour model was used to detect and segment nuclei from prostate histopathology images and lymphocytes from breast histopathology images [16]. An automated diagnosis system was developed to segment cells from breast cancer histopathology images, which comprised of GMM and EM to identify the initial nuclear areas. Individual cells were then delineated by applying watershed transform on the resulting image [17]. A supervised method