Iterative randomized irregular circular algorithm for proliferation rate estimation in brain tumor Ki-67 histology images

Yazan M. Alomari [a,*], Siti Norul Huda Sheikh Abdullah [b], Reena Rahayu Md Zin [c], Khairuddin Omar [b]

[a] College of Applied Studies and Community Service, Faculty of Sciences and Management, University of Dammam, Saudi Arabia
[b] Pattern Recognition Research Group, Center for Artificial Intelligence Technology, Faculty of Information Science and Technology, University Kebangsaan Malaysia, 43600 Bangi, Malaysia
[c] Department of Pathology, UKM Medical Center, University Kebangsaan Malaysia, 56000 Cheras, Kuala Lumpur, Malaysia

ARTICLE INFO

ABSTRACT

Proliferation rate estimation (PRE) is clinically performed from Ki-67 histopathology images. As brain tumor tissues are very complex, accurate PRE determination requires manual cell counting that is tedious, time consuming and inherently inaccurate due to inter-personal variations. Therefore, pathologists usually determine the PRE based on their experience and visualization without actual counting. Automating PRE can substantially increase the efficiency and accuracy of pathologists' determination of PRE. In addition, developing a deterministic and reproducible proliferation rate value is crucial to reduce inter-observer variability. In this paper, a PRE Computer Aided Diagnosis (PRECAD) system has been developed to automate the determination of PRE from Ki-67 histopathology microscopic images for brain tumors. The process involves six steps: color space transformation, customized color modification, nuclei segmentation based on K-Means clustering, preprocessing the extracted cells, counting based on an iterative structured circle detection (IRIC) algorithm, and finally, calculating the PRE value. The proposed IRIC algorithm is able to detect irregular and overlapping cells by introducing dynamic initialization to the basic RCD method, dividing the entire image into partitions based on 8-neighbor connected components. We initiated a new selection method for determining a best circle candidate that yields a reduced probability of incorrectly detecting circles, and proposed a new technique for detecting irregular cells via a dynamic number of iterations that guarantees finding all the cells in a selected partition. Using the same innovations mentioned above, our proposed IRIC algorithm can also be used to detect irregular and two or more overlapping cells. The proposed PRECAD system achieved high accuracy, as determined by quantitative analysis of precision, recall and F-measurement test values of 97.8%, 98.3% and 98% for blue cells and 98.7%, 98% and 98.4% for brown cells, respectively. Thus, our proposed PRECAD system is as reliable as a pathologist for estimating the proliferation rate, while also featuring inherent reproducibility.

© 2015 Elsevier Ltd. All rights reserved.

1. Introduction

Cancer is a growing problem in the world, and in Malaysia it is one of the major causes of death. More than 30,000 incidences of cancer annually, and unfortunately most of these cases are found in the late stage of the disease (Chin & Lim, 2002).

Fig. 1 shows the age-specific cancer incidence per 100,000 people in Peninsular Malaysia in 2006 (Zainal Ariffin, Zainudin Mohd., & Nor Saleha Ibrahim, 2006). In the USA, Brain tumors are the second-leading cause of cancer death in Children and teens under age 20,

Females under 20 years of age and Males less than 40 years of age. The overall, age-adjusted rate of death from a tumor of the brain or spinal cord was 4.4 deaths for every 100,000 people between 2001 and 2005 (Rothman, 2009).

A brain tumor is an abnormal growth of the brain or central spinal tissue, which disrupts proper brain function. Doctors classify tumors based on where the tumor cell originated and whether they are cancerous (Malignant) or not (Benign). Benign tumors are the least aggressive type and usually do not have cancer cells. Malignant tumors contain cancer cells that grow rapidly and are life-threatening to the patient (Glass-Macenka, Hays, Varner, & Weiss, 2013). The diagnostic evaluation of a patient with suspected a brain tumor is very complicated process and involves number of specialists. It typically includes a brain scan, often an MRI, as a first step. If imaging raises Suspicion of brain tumor, a brain biopsy is usually performed. A biopsy is the
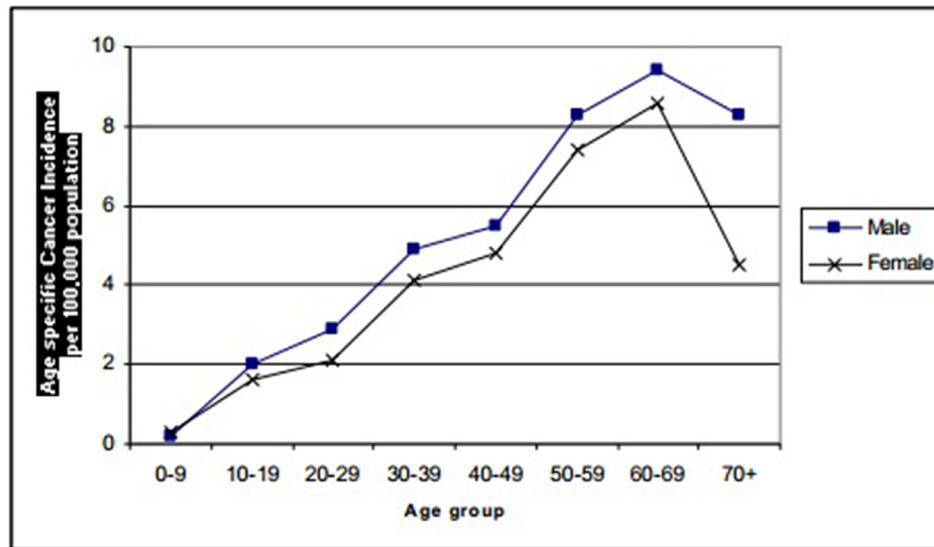
Fig. 1. Age-specific CNS cancer incidence per 100,000 (CR) population by sex in Peninsular Malaysia 2006(Dr. Zainal Ariffin, Dr. Zainudin Mohd., & Dr. Nor Saleha Ibrahim, 2006).

removal of small piece of a tumor tissue for diagnostic purpose (Glass-Macenka et al., 2013). This biopsy is treated and sliced on the lab slides, and the diagnostic process for each case always starts with staining the specimen with dyes (Veta et al., 2013), followed by analysis under a microscope by a pathologist. Under the microscope, the pathologist can examine the biopsy slides with high magnification (40X and 60X) and capture images. Diagnosis is based on these images and the slides. This analysis determines the brain tumor type and grade, based on the appearance of the tumor cells and enables estimation of how quickly the tumor is likely to grow and spread (He, Long, Antani, & Thoma, 2009; "National Cancer Institute," 2013).

Proliferation rate is a measurement that can be used in conjunction with other measurements as prognostic indicators, which helps guide treatment protocols in clinical practice (al-Lahham, Alomari, Hiary, & Chaudhary, 2012; Chassevent, 2001). Cancer cells are always in rapidly growing and dividing phases, and proliferation rate is a *percentage of cancer cells that are actively dividing*. It is measured by determining the number of cancer (positive) cells count over the whole cells count (Klauber-DeMore, Van Zee, Linkov, Borgen, & Gerald, 2001; Urruticoechea, Smith, & Dowsett, 2005). Many techniques have been developed to estimate proliferation rate: Mitotic index, S-phase fractions, nuclear antigen immunohistochemistry (IHC) of markers that that include Ki-67 and PCNA-Staining Cyclins and PET. The Ki-67 labeling index is now widely used to measure the proliferation rate. It is an excellent marker to determine the growth fraction in a population of cells. The fraction of Ki-67-positive tumor cells (the Ki-67 labeling index) is often an indicator of cancer cells. It is usually used in prostate, brain and breast cancers as well as nephroblastoma. For these types of tumors, the prognostic value of survival and tumor recurrence has repeatedly been proven in uni- and multivariate analysis (Ki-67 (protein), 2009). Based on experimental lab work, Ki-67 has many advantages in that it can performed easily, it has a high sensitivity to antibodies and uses a small amount of tissue (Beresford, Wilson, & Makris, 2006). Because the work of cell counting is very tedious, time consuming and inaccurate due to interpersonal variation, the process of estimation proliferation rate can be viewed as highly subjective. Often, the pathologist looks quickly at the image and estimates the nuclei count and the proliferation rate as well.

In determining the Ki-67 labeling index, cells that have the brown color are considered as positive cells and blue cells are considered

negative as shown in Fig. 2. Proliferation rate using Ki-67 is calculated as a percentage of the total number of positive tumor cells, which indicates the growth fraction of the tumor (Klauber-DeMore et al., 2001; Urruticoechea et al., 2005).

Our objective in this paper is to propose a PRECAD system to analyze Ki-67 histopathologic images for brain tumor tissue and to automate nuclei segmentation, counting, and proliferation rate estimation. Automating this process can (1) increase the accuracy of pathologists, (2) save the time of pathologists, (3) generate a second opinion that helps in diagnosis and (4) aid in producing a more deterministic decision for estimating patient lifespan (He et al., 2009). Gold standard analysis of the images is performed by two experts. Evaluation and quantitative analysis have been performed for segmentation and counting results. Based on expert gold standard, sensitivity, specificity, F-measure and the paired T-Test for both automated and manual counting methods was calculated to determine accuracy. In the following sections of this paper, we discuss the related work Ki-67 histopathology image analysis (Section 2); the Methodology used (Section 3); the results and experiments (Section 4); and present the conclusion in the final section.

2. Related work

Recently, many studies have been conducted that focus on using automated CAD systems for analyzing various cancer types. These can help radiologists or pathologists in determining diagnosis and prognosis. These automated systems support them as a second opinion. "Second opinion" means that the analysis results from these systems is considered for its advantages of speed, accuracy, reproducibility and effort conservation. Such analyses will not replace the radiologist or pathologist, however, as the final decision and recommendations will be human-driven (Gurcan et al., 2009; He et al., 2009). Brain tumor detection (Akram & Usman, 2011; Savitha, Prajna, & Ujwal, 2013), grading (Patil & Udupi, 2013), and growth-rate estimation (Anari, Mahzouni, & Amirfattahi, 2010; Gurcan et al., 2009; Yousefi, Ahmadian, Khodadad, Saberi, & Daneshmehr, 2013) are some clinical practices in which automated Computer-aided systems are used to help the clinician in diagnosis, definition of disease severity and to guide the treatment protocols.

Many types of imaging modalities are used in cancer clinics including mammography, ultrasound, Magnetic Resonance Imaging (MRI), Computed Tomography (CT), digital mammography