Expert Systems with Applications 42 (2015) 5356–5365

Contents lists available at ScienceDirect

Expert Systems with Applications

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

Review

Image based computer aided diagnosis system for cancer detection

Howard Lee, Yi-Ping Phoebe Chen *

Department of Computer Science and Computer Engineering, La Trobe University, Melbourne, Australia

ARTICLE INFO

Article history:
Available online 26 February 2015

Keywords:
Segmentation
Computer vision
CAD system
Breast cancer, Lung cancer
Prostate cancer
Skin cancer
Cancer detection
Computed tomography (CT)
Ultrasound
Mammogram
Dermatography
Visual features

ABSTRACT

Cancer is one of the major causes of non-accidental death in human. Early diagnosis of the disease allows clinician to administer suitable treatment, and can improve the patient's survival rate. Traditional diagnosis involves trained clinicians to visually examine the respective medical images for any signs of nodule development in the body. However due to the large scale of the medical image data, this manual diagnosis is often laborious and can be highly subjective due to inter-observer variability. Inspired by the advanced computing technology which is capable of performing complex image processing and machine learning, researches had been carried out in the past few decades to develop computer aided diagnosis (CAD) systems to assist clinicians detecting different forms of cancer. This paper reviews computer vision techniques adopted in medical image analysis, in particular, for cancer detection. The review focused on the detection of the most common form of cancer types, namely breast cancer, prostate cancer, lung cancer and skin cancer. A recent proposed cloud computing frame work has inspired the researchers to utilize the existing works on image based cancer research and develop a more versatile CAD system for detection.

1. Introduction

Cancer is a type of disease in which a group of cells exhibits irregular cell growth cycle. In a normal cell cycle, the cells undergo mitosis process to replicate itself and hence the cell grows (Lee & Chen, 2014a; Nahar, Tickle, Ali, & Chen, 2011); eventually the *programmed cell death* process called apoptosis leads the cells to die in order to regulate its growth. In cancer, the cells lost such balance and grow uncontrollably, to form malignant tumors invading the surrounding tissues. The cancer cell can also migrate to other parts of the body by the bloodstream or lymphatic system, and continue to spread from the new location.

The cause of cancer has not been fully unveiled, however certain habits, such as smoking, exposure to radiations and environmental pollutants are known to cause cancer. Inherited genetic defects are also linked to the cause of some cancers (Avila-Garcia, Trefethen, Brady, Gleeson, & Goodman, 2008; Doi, 2005; Nahar et al., 2011). Early diagnosis of the cancer will allow the clinician to remove the cancer cells via operation or administer suitable treatment plan to eliminate the cancerous cells using chemical or radiation treatment.

Traditionally, cancer can be detected from the presence of certain symptoms, such as irregular markings on the skin, or hard lumps on the body. In clinics, screening tests (Chen & Chen, 2006; Wu, Erwin, & Rosner, 2011) and medical imaging (Huang, Chen, & Lee, 2007; Huang, Hung, Lee, Li, & Wang, 2012; Lee & Chen, 2014b; Nahar, Imam, Tickle, Ali, & Chen, 2012) are the initial stages in cancer detection. Once a suspected cancer has been detected, tissue samples from the suspected region are extracted and examined (Chen & Chen, 2006; Hung, Chen, & Yang, 2011; Nahar, Chen, & Ali, 2007; Nahar et al., 2012).

Non-invasive diagnosis of cancer involves a trained clinician visually examine different types of medical images, and identify the possible locations which resembles signs of malignant tumors. The accuracy of the diagnosis is highly dependent on the experiences of the clinician. Furthermore, with a large volume of medical database, this process is laborious and hardly consistent (Cruz-Roa, Caicedo, & Gonzalez, 2011).

With the advance in digital computing technology, many researchers have combined image processing, pattern recognition, and artificial neural network to develop computer aided diagnosis (CAD) systems to assist the clinicians in the diagnosis process (Lee & Chen, 2014b; Rolim et al., 2010; Verma & Zakos, 2001; Ye, Lin, & Dehmeshki, 2009) Fig. 1 shows a general framework for a CAD system for skin cancer detection (Lee & Chen, 2014b).

In a CAD system for cancer detection, medical images are recorded using the appropriate imaging systems. The image acquisition device can be adjusted to ensure the consistent image quality under the same laboratory environment (Lee & Chen, 2014a; Zheng et al., 2006). The captured images undergo series

* Corresponding author.
E-mail addresses: H17lee@stduents.latrobe.edu.au (H. Lee), phoebe.chen@latrobe.edu.au (Y.-P.P. Chen).

http://dx.doi.org/10.1016/j.eswa.2015.02.005
0957-4174/

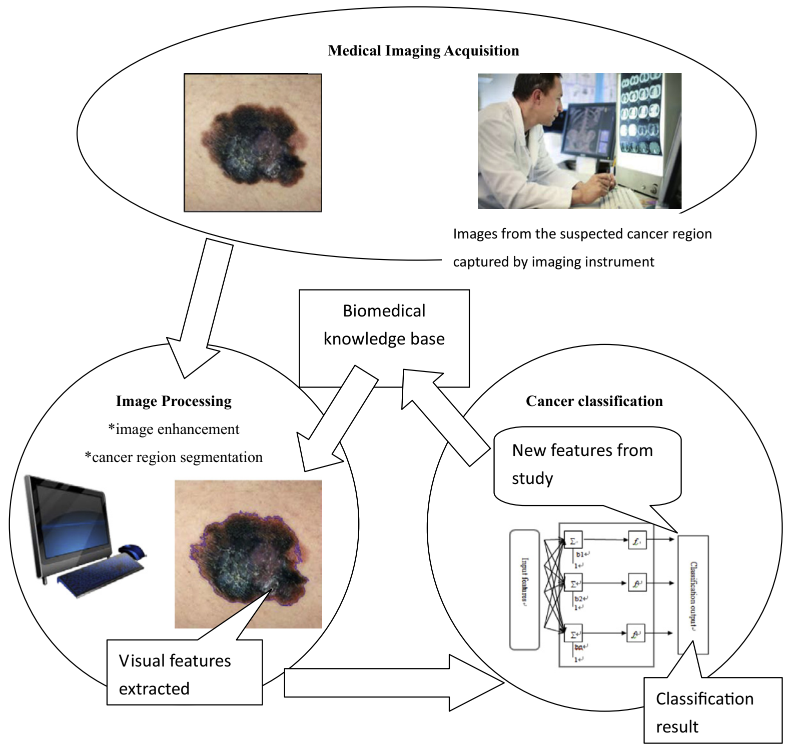

Fig. 1. Framework for the cancer diagnosis system.

of software-based algorithms to isolate the suspected cancerous regions from the rest of the image (Lee & Chen, 2014). A biomedical knowledge based features, such as textures (Tang & Guo, 2011), and shapes (Celebi, Aslandogan, & Bergstresser, 2005; Lee & Chen, 2014a; Soysal & Chen, 2008; Tourassi, Delong, & Floyd, 2006) can be extracted to characterize the extracted segments. This feature space forms a biometric describing the suspected region. A supervised classifier can be implemented and trained using the existing sample images, to learn the distinctive patterns associated with the biometric. The diagnosis can be performed by comparing the feature patterns between the test sample and the trained patterns in the classifier features (Lee & Chen, 2014). Newly developed features from the study can be fed back to the biomedical knowledge database to improve the existing feature sets (Nahar et al., 2007).

Image segmentation plays a crucial role in a CAD system. It aims to isolate the suspected region from the rest of the image (Cheng, Cheng, & Chen, 2013; Lee & Chen, 2014b). It can also incorporate visual features such as color and texture information with other statistical and biological features to distinguish different regions in an image. Hence the unaffected regions can be removed and leaving the suspected regions which resembles certain visual patterns, such as irregular texture, color and intensity (Lee & Chen, 2014b; Umbaugh, Moss, & Stoecker, 1989; Yuksel & Borlu, 2009). An accurate segmentation result will help to determine the location and the size of the tumor, which is important for treatment planning.

The major challenge in this field of research is to build a fully automatic CAD system which can analyze large quantities of images to provide an accurate diagnosis and at the same time, robust enough to handle the biological variations in humans (Lee & Chen, 2014a).

In this paper, we are going to provide an overview of the segmentation process used in these common medical image modalities to detect the most common forms of cancer: breast cancer, prostate cancer, lung cancer and skin cancer. We also investigate the recent work on these cancer detections with medical image processing and analysis. We discuss how image processing techniques have assists the clinicians in IMRT for cancer treatment. To utilize the existing research works on the algorithms to segment and classify for cancer images, a cloud computing frame work for image based cancer research has been proposed. In this paper, we discussed the advantages of this frame work and future works involved in developing a cloud computing based CAD system for cancer research.

2. Breast cancer detection

Breast cancer is a type of cancer originating from breast tissue, and it accounts for 23% of all cancers in women (Berman, 2007). The most effective way to detect breast cancer is through the breast mammogram screening, however the major limitation for mammography diagnosis is sensitivity. This diagnosis is less sensitive in younger women especially cancers in dense breasts become difficult to detect (Baker, 2003; Smith et al., 2004).

Mammography is the most common imaging technique to detect breast cancer. (Boukerroui, Baskurt, Noble, & Basset, 2003; Chang, Wu, Moon, Chou, & Chen, 2003b; Chang et al., 2003a; Chen, Chang, Kuo, Chen, & Huang, 2002; Cheng et al., 2010). In contrary to X-ray mammography, the conventional B-mode ultrasound is used to distinguish benign masses from malignant cancerous masses. Clinical studies have tried to characterize these masses by the texture and geometric properties (Arger et al., 2001; Huang et al., 2012). Fig. 2 illustrates an example of breast cancer detection in mammography and ultrasound.

A suspected benign can be detected by observing the regions showing abnormal high density. The texture and shape features are commonly used to detect the presence of breast cancer (Baker, 2003), however one of the major hurdles in detecting breast cancer in mammography is that the high breast density, usually in young age group, may be difficult to detect the benign in the early stage, when the size is small. In conjunction with mammography, ultrasound technique has also been used to detect the presence of the benign in breast.