Computers and Electrical Engineering 000 (2018) 1–12

Contents lists available at ScienceDirect

Computers and Electrical Engineering

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

Neural network and multi-fractal dimension features for breast cancer classification from ultrasound images☆

Mazin Abed Mohammed[a,c,*], Belal Al-Khateeb[b], Ahmed Noori Rashid[b], Dheyaa Ahmed Ibrahim[b], Mohd Khanapi Abd Ghani[a], Salama A. Mostafa[d]

[a] Biomedical Computing and Engineering Technologies (BIOCORE) Applied Research Group, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia
[b] Computer Science Department, College of Computer Science & Information Technology, University of Anbar, Anbar, Iraq
[c] Planning and Follow Up Department, University Headquarter, University of Anbar, Anbar, Iraq
[d] Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Johor, Malaysia

ARTICLE INFO

Article history:
Received 25 September 2017
Revised 20 January 2018
Accepted 22 January 2018
Available online xxx

Keywords:
Breast cancer
Multi-fractal dimension
Neural network approach
Breast cancer classification
Ultrasound images

ABSTRACT

Breast cancer is considered to be one of the most threatening issues in clinical practice. However, existing breast cancer diagnosis methods face questions of complexity, cost, human-dependency, and inaccuracy. Recently, many computerized and interdisciplinary systems have been developed to avoid human errors in both quantification and diagnosis. A computerized system can be further improved to optimize the efficiency of breast tumour identification. The current paper presents an effort to automate characterization of breast cancer from ultrasound images using multi-fractal dimensions and backpropagation neural networks. In this study, a total of 184 breast ultrasound images (72 abnormal (tumour cases) and 112 normal cases) were examined. Various setups were employed to achieve a decent balance between positive and negative rates of the diagnosed cases. The obtained results manifested in high rates of precision (82.04%), sensitivity (79.39%), and specificity (84.75%).

1. Introduction

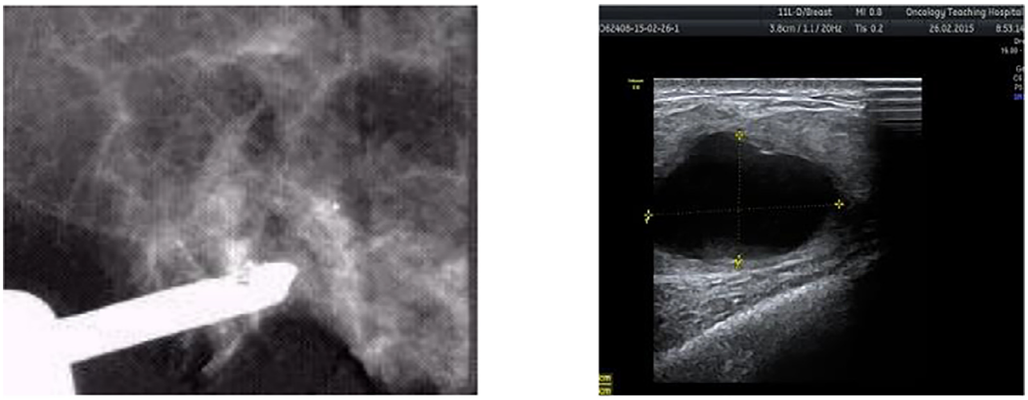
Breast cancer, first identified in Egypt in approximately 1600 BC, is one of the oldest known types of cancer [1]. However, even after intensive research over the last few decades, no solution is available to eradicate this fatal disease. Therefore, innovative methods are needed to find an optimal solution. In current medical practices, the quantification of the region of a tumour and discrimination between positive and negative cases are done manually from real-time scanned images. Thus, the diagnostic procedure entirely depends on the experience of the operator and involves multiple subjective decisions. Subjective decision-making can result in inter- and intra-observer variations. The inter-observer variation is the measure of the difference between the outcomes acquired by at least two observers while looking at similar material. The intra-observer variation occurs when one observer evaluates a similar material more than once. Such inconsistencies cause difficulties and

☆ Reviews processed and recommended for publication to the Editor-in-Chief by Associate Editor Dr. G. R. Gonzalez.

* Corresponding author at: Biomedical Computing and Engineering Technologies (BIOCORE) Applied Research Group, Faculty of Information and Communication Technology, Universiti Teknikal Malaysia Melaka, Melaka, Malaysia.

E-mail addresses: mazin_top_86@yahoo.com (M.A. Mohammed), belal@computer-college.org (B. Al-Khateeb), rashidisgr@gmail.com (A.N. Rashid), dheyaa.ibrahim88@gmail.com (D.A. Ibrahim), khanapi@utem.edu.my (M.K. Abd Ghani), salama@uthm.edu.my (S.A. Mostafa).

https://doi.org/10.1016/j.compeleceng.2018.01.033
0045-7906/



a) Mammogram image b) Ultrasound image

Fig. 1. The images of a breast tumour taken from two different screening methods.

errors in the diagnosis stage and consequently increase patient anxiety. To avoid human errors in both quantification and diagnosis stages, computer-based image processing and analysis tools need to be developed.

From the recent studies on cancer conducted in Iraq, it was observed that approximately one-fourth of the enrolled malignancy cases were affected by breast cancer, and it is the primary factor of death among Iraqi women [1]. As Iraq is a low-income country, the establishment of full-featured detection systems, such as mammography machines and ultrasounds, are difficult. A 2010 study noted that 143 out of 721 (19.8%) Iraqi females, who presented with substantial breast masses in a screening for early discovery of breast malignancy, had abnormal breast growth. Among 90.6% of the patients with recognized lumps, only 32% were therapeutically treated in the first month, while 16% consulted a specialist 1 year later [1]. Mammography and ultrasound machines are accessible in the main hospitals in every city in Iraq. Clearly, due to financial difficulties, it is impossible for specialists to use mammography and ultrasound for the screening of every Iraqi woman.

Breast cancer can be classified into two categories—normal and abnormal, and the abnormal (tumour) category can be divided into two classes—benign (non-harmful) and cancerous (malignant). Benign tumours are not injurious to health; their cells have a close resemblance to normal cells. Benign tumours grow relatively slowly and do not attack the adjacent tissues or spread to different parts of the body [2]. The images of a breast tumour provide oncologists with a full description of the disease, such as the size, shape, and area of spread along with the location of the tumour. Formerly, mammography, which uses X-rays to capture images of the breast, was the most commonly used screening method. This method is hazardous because the patients face a high amount of radiation during each screening, thus leading to leukaemia or other long-term diseases [3]. Ultrasound scanning is a safer screening alternative. Ultrasound scanning involves the exposure of body parts to high-frequency sound waves to create images of the internal structure of the human body. Ultrasound scanning does not use ionizing radiation as in mammography, making it a safer alternative. Ultrasound scanning also produces images with a relatively high resolution [4]. Mammogram images are less clear and the tumour is often poorly defined, whereas an ultrasound image clearly shows the tumour and surrounding tissues (Fig. 1).

A biopsy is a traditional way of detecting breast cancer. While the ultrasound method detects breast abnormalities, biopsy uses pathological investigation (by cutting breast tissue) to detect whether the tumour is benign or malignant. In modern medical practices, many biopsy methods are available based on the location, size, appearance, and characteristics of the abnormalities—fine and core needle biopsy, vacuum helped biopsy, expansive enter biopsy, and open surgical biopsy.

A hazard factor can be anything that transforms a disease into malignancy [5,6]. Different types of cancers have their own hazard factors. For instance, sunlight is a hazard factor for skin disease and smoking is a hazard factor for various types of malignancies. Having a hazard factor does not imply that one will be affected by cancer. Many individuals with hazard factors never encounter the stage of malignancies. Scientific studies have discovered several hazard factors that make an individual more prone to breast cancer. Well-known Machine Learning algorithms, such as genetic algorithms, neural networks, support vector machines, clustering analysis, are the most popular approaches for finding hazard factors. Since many studies are available in this domain, the current study does not aim to present the technical aspects of these algorithms. The majority of IDSs consider only a single algorithm and solve a specific problem. However, recent studies reported the use of structured frameworks involving ML algorithms, such as committees of machines, where more than one algorithm work together to solve a given problem [7,8].

Over recent decades, many researchers have been trying to find the optimal solution to enhance breast cancer diagnosis; however, they have only determined approximate solutions that differ in efficiency depending on the search space. In the literature [9], many enhanced NN trainings were utilized for the classification of breast tumour datasets. In addition,