System for Breast Cancer Diagnosis: A Survey

Rasha, Maizatul Akmar Ismail, Amiruddin Kamsin and Saqib Hakak
Department of Information Systems
Faculty of Computer Science and Information Technology,
University of Malaya, MALAYSIA
e-mail: rashaatallah@siswa.um.edu.my, maizatul@um.edu.my, amir@um.edu.my,
saqibhakak@siswa.um.edu.my

ABSTRACT
One of the most cause of death in the world is breast cancer. This cancer increase the ration of the women death. Cancer happens when uncontrollable cell start appear in the human body and spread. But to reduce breast cancer fatality must detect and diagnose this disease early. Accurate classification of breast tumor is an important task in medical diagnosis. New technical computing starting help medical in diagnosis diseases and to improve the specialists doctors performance. The aim of this survey paper is to define the current state of research in breast cancer and to extract the limitation of the existing systems. There are many software based on neural network, support vector machine, fuzzy logic, deep learning and many others techniques are being used in medical. In this paper two type of technique have been studied deep learning and neural network. The comparison between existing approaches was done based on three main evaluation parameters i.e., accuracy, sensitivity and specificity along with data sets used.

Keywords: Breast Cancer, Deep learning, Neural network

INTRODUCTION

Computer system techniques are used for classify and detect diseases in medical field. Breast cancer is one of the huge and most diagnosed cancers. A report by World Health Organization (WHO) published in 2013 has shown that cancer is one of the leading causes of death. Early detection of cancer saving approximately 37.3% patients (Salama, Abdelhalim et al. 2012). There are two type of cancer malignant and benign. The main issues is to diagnose malignant and benign cancer. Many studies have used computer technology to classify whether the cancer is benign or malignant tumours (Huaqing, et al 2016 ). However, this paper discovers the most recent works starting from 2015 onwards in the area of breast cancer detection. The main contribution of the article can be summarised as:

This survey compare different approaches used for breast cancer involving latest works and discuss open issues for deep learning and neural network.
Approaches used for Cancer Detection

In order to carry out the survey work, several articles related to breast cancer techniques were reviewed and selected from credible sources such as: Web of Science (WoS), Science Direct, IEEE, Springer. The main methods used in these articles were classified as figure 1, Neural Network and Deep Learning.

![Diagram showing different approaches methods used for classifying the types of cancer](image)

Figure 1: The different approaches used for classifying the types of cancer

The current paper integrates different classification techniques which are used at breast cancer. Table 1 provides list of the papers which dealt with single methods.

<table>
<thead>
<tr>
<th>Classifier Types</th>
<th>Author</th>
<th>Aim</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bayesian neural network (BNN)</td>
<td>Breast Cancer Classification Using Extracted Parameters from a Terahertz Dielectric Model of Human Breast Tissue</td>
<td>Increase the classification tumor using BNN</td>
</tr>
<tr>
<td>Cellular neural network (CNN)</td>
<td>Detection of microcalcifications in digitized mammograms with multistable cellular neural networks using a new image enhancement method: automated lesion intensity enhancer (ALIE)</td>
<td>To propose a model using Cellular neural networks (CNNs) to remove pectoral muscle and unnecessary parts from the mammogram images</td>
</tr>
</tbody>
</table>
Cellular neural network (CNN) | Classification of benign and malignant breast tumors based on hybrid level set segmentation | proposed two segmentation approaches for classifying benign and malignant tumours using CNN

Artificial Neural Network (ANN) | Robust mass classification-based local binary pattern variance and shape descriptors | Detect masses on mammographic images based on the Local Binary Pattern Variance (LBPV) and shape descriptors ,also using Artificial Neural Network (ANN) to classify the masses

Deep Learning

Deep learning with Shear wave Elastography | Deep learning based classification of breast tumors with shear-wave elastography | To propose deep learning architecture to classify malignant and benign breast tumours

Convolution Neural Network | A deep feature based framework for breast masses classification | To implement? fine-tuning operation on the trained deep CNN model to acquire the feature extraction

Deep learning with SAE | Discrimination of Breast Cancer with Microcalcifications on Mammography by Deep Learning | To improve the performance of an innovative deep learning model for classifying breast lesions.

Deep Learning based Visual Search | Probabilistic Visual Search for Masses Within Mammography Images using Deep Learning | To work with an entire mammography image as input without the need for image segmentation

The effectiveness of the previously listed techniques that were used at cancer breast classification is evaluated based on how correctly the methods classified cancer. The evaluation is made based on accuracy, specificity and sensitivity criteria. Table 3 indicates the analysis of these three key aspects in medical data using at breast cancer detect.

**Survey on Cancer Detection techniques**

As shown in table 3, this survey presents 2 Network and Deep Learning. The first method was Neural network, Cellular Neural Network (CNN) on of the methods when used it the accuracy increase from 82% to 96% , also used Artificial Neural Network (ANN) which improve the accuracy to 91% , the best become 100% when used Bayesian Neural network . The second method was deep learning, when deep learning combined with visual search, the accuracy was 85% , also when Stacked auto encoder (SAE) used to build deep learning network, the accuracy improved to 87.3%. The accuracy also improved to 91.3% by adding shear wave elastography (SWE), after add different layers at deep learning network the accuracy become 96.7%.
As seen in table 3, the data set which used for evaluation collected from various resources as Digital Database for Screening Mammography (DDSM), Wisconsin Breast Cancer Dataset (WBCD), Mammography Image Analysis Society (MIAS) and Wisconsin Original Breast Cancer (WOBC).

The least number of images were 22 images, the images were collected from Mammography Image Analysis Society (MIAS) dataset, and the result of evaluation was 82% accuracy, sensitivity was 90.9% and specificity was 52.2% by using cellular neural network.

The most number of images were 2803 images. This images were collected from Wisconsin Diagnostic Breast Cancer (WDBC), Wisconsin Original Breast Cancer (WOBC), Ljubljana Breast Cancer Dataset, METABRIC Breast Cancer Dataset and Netherlands Cancer Institute (NKI) Dataset, through this paper the evaluation was 96% for accuracy by using deep learning network.

Table 2: Performance evaluation summary for the surveyed methods

<table>
<thead>
<tr>
<th>Paper</th>
<th>Data Set</th>
<th>Data</th>
<th>Method</th>
<th>References</th>
<th>Accuracy</th>
<th>Sensitivity (TPF)</th>
<th>Specificity (TNF)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast Cancer Classification Using Extracted Parameters from a Terahertz Dielectric Model of Human Breast Tissue</td>
<td>74</td>
<td>Female patients Italy research committee</td>
<td>Bayesian neural network (BNN)</td>
<td>17</td>
<td>97.3%</td>
<td>94.3</td>
<td>94.4</td>
</tr>
<tr>
<td>Detection of microcalcification in digitized mammograms with multistable cellular neural networks using a new image enhancement method: automated lesion intensity enhancer (ALIE)</td>
<td>100</td>
<td>Mammographic Image Analysis Society (MIAS) database</td>
<td>Cellular neural networks-CNN</td>
<td>43</td>
<td>82</td>
<td>90.0</td>
<td>52.2</td>
</tr>
<tr>
<td>Classification of benign and malignant breast tumors based on hybrid level set segmentation</td>
<td>57</td>
<td>Mammographic Image Analysis Society (MIAS) database</td>
<td>Cellular neural network (CNN)</td>
<td>60</td>
<td>97.30%</td>
<td>92.70%, 90.54%</td>
<td></td>
</tr>
<tr>
<td>Robust mass classification-based local binary pattern variance and shape descriptors</td>
<td>600</td>
<td>FARABI database from EL FARABI radiologic centre</td>
<td>Local Binary Pattern Variance (LBPV) with Artificial Neural Network (ANN)</td>
<td>22</td>
<td>91</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>
State-of-the-art for breast cancer techniques

Neural network

(Civcik, Yilmaz et al. 2015) used Cellular neural networks (CNNs) to remove pectoral muscle and unnecessary parts from the mammogram images, supported the model by Automated Lesion intensity Enhancer (ALIE) to enhancing lesion intensities. They also used the multistable CNNs. After applying the combination of these methods on the MIAS database, the accuracy become 82.0%. In 2015 Masmoudi and others researchers (Masmoudi, Ben Ayed et al. 2015) proposed an approach for detect masses on mammographic images based on the Local Binary Pattern Variance (LBPV) and shape descriptors, also using Artificial Neural Network (ANN) to classify the masses. The results indicate that the approach achieves accuracy over 91% which are applied to 600 mammographic, 300 for training and 300 for testing. Also (Rouhi, Jafari et al. 2015) proposed two segmentation approaches for classifying benign and malignant tumours using mammograms. One segmentation method consists of region growing method and other method consists of cellular neural network (CNN) approach. After necessary pre-processing of input mammogram, using region growing method and application of artificial neural network (ANN), necessary features from benign and malignant classification are extracted. In the CNN based approach, existing CNN model is applied for segmentation and classification purposes. The CNN
model is trained using genetic algorithm. In these both methods, tumour boundary information is preserved to detect malign and benign tumours in mammograms. The average of the accuracy is 96%. (Truong, Tuan el at 2015) used Bayesian neural network (BNN) instead of using support vector machine (SVM) to increase the accuracy of tumour classification by using combinations of four model parameters. The dataset is 74 sample, they use 80% of the data (59 samples) for training and 20% (15 samples) for testing. BNN successfully classifies the data using the combinations of four model parameters with an accuracy of 97.3%.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Methods</th>
<th>Classification Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Cicic, Yilmaz et al. 2015)</td>
<td>CNN</td>
<td>82.0%</td>
</tr>
<tr>
<td>(Masmoudi, Ben Ayed et al. 2015)</td>
<td>ANN</td>
<td>91%</td>
</tr>
<tr>
<td>(Rouhi, Jafari et al. 2015)</td>
<td>CNN</td>
<td>96%</td>
</tr>
<tr>
<td>Truong, Tuan el at 2015</td>
<td>BNN</td>
<td>97.3%</td>
</tr>
</tbody>
</table>

Deep Learning
A visual search engine developed based on deep learning to classify if the tissues is mass or not at the mammography images. (Ertosun and Rubin 2015). The model obtained 85% accuracy. Also deep learning used with the stacked auto encoder (SAE) to create a deep network by stacking multiple auto encoders hierarchically to improve the performance of an innovative deep learning model for classifying breast lesions (Wang, Yang et al. 2016). It is applied at SunYat-sen University Cancer Center (Guangzhou, China) and Nanhai Affiliated Hospital of Southern Medical University (Foshan, China). The accuracy increased to 87.3%. When Zhang, Xiao (Zhang, Xiao et al. 2016) proposed deep learning architecture to classify malignant and benign breast tumours. The (Zhang, Xiao et al. 2016) image representations using point-wise gated Boltzmann machine (PGBM) technique. SWE images of 291 patients after all necessary pre-processing were used to evaluate the proposed architecture. (Jiao, Gao et al. 2016) applied fine-tuning operation on the trained deep CNN model to acquire the feature extraction. It is applied at DDSM dataset accuracy 96.7%.

<table>
<thead>
<tr>
<th>Authors</th>
<th>Methods</th>
<th>Classification Accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Ertosun and Rubin 2015)</td>
<td>Deep learning-based visual search</td>
<td>85%</td>
</tr>
<tr>
<td>(Wang, Yang et al. 2016)</td>
<td>Deep learning with SAE</td>
<td>87.3%</td>
</tr>
<tr>
<td>Zhang, Xiao et al. 2016</td>
<td>Deep learning with Shear wave Elastography</td>
<td>91.3%</td>
</tr>
<tr>
<td>(Jiao, Gao et al. 2016)</td>
<td>Deep learning : Deep features from different layers</td>
<td>96.7%</td>
</tr>
</tbody>
</table>

RESEARCH GAPS
Challenges and Open Issues
Accuracy
An accurate classifier is the most important component of any CAD. (Abdel-Zaher and Eldeib 2016). To achieve accurate breast cancer classification is still challenging due to the unknown cause of the disease and the similarities between benign and malignant masse. (Verma, McLeod et al.)
Pattern
Breast tumour SWE images contain artefact, noise, and other irrelevant patterns, such as irregular stiffness distributions. (Zhang, Xiao et al.), also the challenge is how to learn robust representations that can distinguish useful (i.e., task-relevant) patterns from large amounts of distracting (i.e., task-irrelevant) patterns (Nair and Hinton 2009). The important challenge is how to understand and utilize the patterns that may be task-relevant but are difficult to interpret by human observers, such as the black holes absent of colour on SWE, the missing areas with invalid stiffness values. (Zhang, Xiao et al. 2016)

CONCLUSION
In this survey paper, we have discussed various techniques system designed using different soft computing techniques from 2015 to 2017. These expert systems are widely used in medical field for classification and diagnosis of breast cancer tumour. The key concept of soft computing technologies is to build an automated system that learns from the decision parameter of the diseases so that the designed system can be used to diagnose and treatment of patient with unknown disease symptoms.

REFERENCES


