ARTIFICIAL NEURAL NETWORK: AS EMERGING DIAGNOSTIC TOOL FOR BREAST CANCER

Ishwinder K. Sandhu¹, M. Nair², H. Shukla³ & S.S. Sandhu³*

¹Department of Computer Science and Engineering, ITM University, Gurgaon, Haryana.
²Center for Scientific Research & Development, Peoples University, Bhopal, M.P.
³*Department of Biological Sciences, R.D.University, Jabalpur, M.P.

*Corresponding Author Email: ssandhu@rediffmail.com

ABSTRACT

In the present scenario, cases of breast cancer disease are on the rise as a result of which it has become one of the most deadly diseases around the world. Classical methods required cytopathologists or oncologists to examine the breast lesions for detection and classification of various stages of the cancer; which was not only labour intensive but inaccurate. Hence an efficient method using artificial neural networks as decision making tool is utilized for the diagnosis of breast cancer. Early detection and diagnosis of breast cancer is necessary for reducing the premature deaths caused by it. But the conventional biological techniques used for breast cancer diagnosis have not proved to be very efficient thereby increasing the mortality rate among cancer patients. These conventional therapies though improve the condition of patient but put them at the stake of enormous side effects. In the recent times, usage of Artificial Neural Networks (ANN) has emerged as a reliable technique for the efficient and quick diagnosis of breast cancer. Several research groups are working world wide on the development of neural networks in medical diagnosis. This review describes neural network approaches to breast cancer diagnosis including different architectures of neural networks.

KEY WORDS
Neural Networks; Artificial Neural Networks; UIMS, diagnosis.

INTRODUCTION

Breast Cancer is one of the most common diseases among women which lead to mortality. It is the most frequently encountered cancer type in women and is the second most terminal one after the lung cancer. It is also seen in men but not much frequently. In India, an average age for developing a breast cancer has undergone a shift. Preoperative or neo-adjuvant chemotherapy is an option with large operable breast cancer so to facilitate breast conservation and to downstage the disease. There has been a number of options for the prognosis and diagnosis of breast cancer. These markers provide a window of opportunity for the detection of breast cancer. Globally, the five most common cancers considered in both sexes were cancers of the lung (1,824,701; 13%), breast (1,676,633; 11.9%), colorectum (1,360,602; 9.7%), prostate (1,111,689; 7.9%), and cervix uteri (527,624; 3.7%), comprising 46.2% of the 28 cancers reported [1]. There have been various strategies utilized for the early detection of breast cancer so to reduce mortality. Newer diagnosis techniques in imaging, tissue diagnosis and cytobiological assessments have held a great promise for the
early detection and identification of risk of the disease. Newer and more specific screening and diagnostic tests are required for early detection of breast cancer, especially in high-risk groups. Therefore, information technology has developed different computerized methods for easy and quick diagnosis of breast cancer, out of which ANN is the most significant method.

CANCER
Cancer is characterized as an abnormal cell growth with the potential to reach to other body parts. It is divided into 100 types depending upon the type of cell from where it started [2]. Cancer generally accounts for 7.6 million deaths (around 13% of all deaths) in 2008, with an estimated 12 million deaths in 2030 [3]. A normal cell in a human being grows, divides and multiplies in a controlled fashion. Every cell in the human body follows a cell cycle comprising of different phases i.e. G1, S, G2 and M (fig. 1A), where “S” stands for “Synthesis”, during this phase DNA synthesis and chromosome duplication takes place, “M” stands for “Mitosis”, during this phase the cell divides and forms two daughter cells, “G1” and “G2”, stand for “gap”, during these phases there is a regulation of cell cycle progression on the basis of different external and internal signals (checkpoints) [4,5,6]. In order to facilitate smooth transition of cell from one phase to another these checkpoints are necessary thus, to ensure that a cell enters the next phase only when all the tasks which were to be performed in the prior phase have been completed [7].

A cancerous cell disregards these checkpoints (fig. 1B) and continues to grow in an uncontrolled manner. Cancer cells thus have two heritable properties: they and their progeny (1) reproduce in defiance of the normal restraints on cell division and (2) invade and colonize territories normally reserved for other cells. Apart from this a cancer cell also does not undergo programmed cell death or apoptosis. In Cancer cells, the length of the telomere is not shortened over several cell divisions and hence the cells continue to divide and bypass senescence (fig. 2).

Fig 1 A: Cell Cycle - Cell grows in interphase, which consist of three phases: S phase; G1 phase; G2 phase whereas in the M phase the nucleus and the cytoplasm divides.
B: Checkpoints of the cell cycle - Signals from the environment causes the control systems to arrest at specific checkpoints.

Fig 2: Telomeres and their role in cellular integrity - Telomeres are non-coding repetitive DNA sequences (TTAGGG) that protect the ends of linear eukaryotic chromosomes.
**Fig 3:** Schematic of ANN

**Fig 4:** Recurrent Network
Breast Cancer

Breast cancer is the most common disease among the women all around the world since last few years [8,9]. In fact, it has become one of the leading causes of death in women [10]. In this disease formation of malignant cells (cancer) takes place in the tissues of the breast. It occurs when a mutation takes place in the cells that line the lobules (lobular carcinoma) or the ducts (ductal carcinoma) that supply milk. The most common type of breast cancer is ductal carcinoma. There are many factors which causes breast cancer these includes certain inherited genes (BRCA1 (Breast cancer 1), BRCA2 (Breast cancer 2), TP53 (tumor suppressor protein p53) and ATM. Variations of the BRCA1, BRCA2, CDH1, STK11, and TP53 genes increase the risk of developing breast cancer whereas AR, ATM, BARD1, BRIP1, CHEK2, DIRAS3, ERBB2, NBN, PALB2, RAD50, and RAD51 genes are associated with breast cancer. These genes accounts for around 40-50% of all cases of inherited breast cancer. Inherited changes in several other genes, including CDH1, STK11, and TP53, have also been found to increase the risk of developing breast cancer. In addition to these specific genetic changes, researchers have also identified many personal and environmental factors (estrogen, radiation, electromagnetic fields, xenoestrogens and exposure to chemicals) that influence the risk of developing breast cancer. These factors include gender, age, ethnic background, a history of previous breast cancer in closely related family members,
certain changes in breast tissue and hormonal factors [11, 12].

**Types of Breast Cancer**

There are mainly two types of breast cancer as mentioned below [13]:

**In Situ Breast Cancer:** This cancer remains within the ducts or lobules of the breasts, where the cancer occurring within the ducts is called as Ductal Carcinoma in situ whereas when the cancer occurs in the lobule of the breast it is called Lobular carcinoma in situ. Lobular carcinoma in situ, are the most common cancer among premenopausal women and cannot be detected via physical examination but also requires mammograms.

**Infiltrating Breast Cancer:** In this type of breast cancer the cancer cells penetrate and cross the membrane surrounding the milk duct or lobule, and thereby invade the adjacent tissues. Thereby resulting in the formation of a lump which can be easily detected via physical examination. Infiltrating Ductal Carcinoma occurs across the ducts while infiltrating lobular carcinoma occurs across the lobules.

**Other types of breast cancer:** Besides above two there are other cancers (cystosarcoma phyllodes, inflammatory cancer, breast cancer during pregnancy, paget’s disease) which occur comparatively less.

**Clinical Stages of Breast Cancer**

Cancer is divided into different stages on the basis of three main features viz., its invasive or non-invasive character, the widespread of cancer in lymph nodes and size of the tumor. Staging is used basically to describe the extent of the cancer spread. American Joint Committee on Cancer (AJCC) utilizes the results of physical exam, biopsy and imaging tests (called the clinical stage), or the results of surgery (called the pathologic stage) for describing the stages of the breast cancer. Pathologic staging is more accurate than the clinical stage, as it allows the clinicians to have an immediate idea of the extent of the cancer.

On the basis of the TNM staging system [14] cancers are classified according to their T, N, and M stages. Stage T (0 to 4) depicts the size of the tumor and its invasion in skin or wall of the chest, Stage N (0 to 3) describes the number of lymph nodes affected by the spreading of cancer to lymph nodes near the breast and Stage M (0 to 1) indicates the spread to distant organs (bones or lungs).

Stage 0 breast cancer is also called as Ductal carcinoma in situ, a pre-cancer of the breast

Stage I: In this stage, the size of tumor is so small that it does not extend outside the breast (Stage I a), small clusters of cancer cells reside in the axillary lymph nodes, but do not invade distant sites.

Stage II: It comprises of stage IIA and stage IIB. In stage IIA the size of tumor is 2 cm or less and cancer has spread to 1 to 3 axillary lymph nodes, 1 to 3 lymph nodes under the arm or small quantity of cancer are found in internal mammary lymph nodes. In the stage IIB, the cancer spreads to axillary lymph nodes, internal mammary lymph nodes but not to other sites and its size remains between 2 cm to 5 cm.

Stage III: It is also divided into three phases viz., stage IIIA, stage IIIB and stage IIIC. In stage IIIA the tumor size is 5 cm or less spreading upto 4-9 axillary lymph nodes, or enlarges into internal mammary lymph nodes but does not extends to distant sites. On the other hand in stage IIIB, besides spreading into axillary lymph nodes and internal mammary lymph nodes, the tumor also
extends towards to the chest wall causing swelling or an ulcer. In Stage III, cancer stretches from axillary lymph nodes and mammary lymph nodes to lymph nodes above and below the clavicle but still does not spreads to distant sites.

Stage IV: By this stage the cancer may be of any size and may or may not have spread to nearby lymph nodes but it finally invades other organs (bones, lungs, liver, or brain) or lymph nodes far from the breast.

Diagnosis of Breast Cancer

Early diagnosis is important to reduce the deadly impact of breast cancer. If breast cancers are detected when they are very small, the large majority of patients can be cured of their disease. Various medical diagnostic methods have been developed for diagnosis of breast cancer. Mamogram detection is the most commonly used method of them. It is an X-ray of the breast that takes pictures of the fat, fibrous tissues, ducts, lobes, and blood vessels. But in spite of the development of medical diagnostic techniques maximum of the breast cancer cases (about >90%) are diagnosed in advance stages i.e. stage II, III and stage IV [15]. This delay in diagnosis not only increases the cost of treatment but also decreases the chances of survival of the patients. As a result, the problem of Breast Cancer Diagnosis (BCD) has attracted many researchers in the area of computational intelligence, data mining and statistical fields.

Recently, Artificial Neural Networks (ANNs) has emerged as an effective method for pattern recognition, machine learning and data mining. It is inspired from the biological neural network of mammalian brain, capable of complex decision making and pattern recognition. Neural networks are used to increase the accuracy and objectivity of medical diagnosis as they allow physicians to distinguish benign breast tumors from malignant ones [16]. Therefore, several research groups are working world wide on the development of neural networks in medical diagnosis.

Artificial Neural Networks: A New Diagnostic Approach

Artificial Neural Network (ANN) is an intelligent system which is inspired by the biological nervous systems which utilizes and processes information. The information processing system which largely comprises of highly interconnected processing neurons working in unison is described as the key element of this prototype.

Neural Network (NN), an imitator of BNN (Biological Neural Network), is a highly interconnected neuron system and is processed using parallel distributed processing system. This system therefore acquires the ability to learn and thereby access knowledge and make it available for use [17]. Simplified versions of our central nervous system known as Neural Networks are hence provoked by the functions performed by a human brain. Neural Networks (NN) are just simplified models of human nervous system which performs functions such as logical inference, cognition, pattern recognition, etc [18]. Neurons are hence the structural entities of a human brain. Hence this simplified imitation of neurons is known as to be Artificial Neural Networks (ANN). It is also termed as Artificial Neural System (ANS).

There exists a learning process which configures an ANN for certain applications involving pattern recognition or classification of data. The learning process involves adjustments to the synaptic connections between the neurons. This methodology is explained with the help of a schematic model (fig. 3) showing the behavior of a neuron with each component depicting the
similarities to the actual elements of a biological neuron and provides a basis to Artificial Neural Networks.

Different Neural network architectures like Feed forward Network, Recurrent Network, and Multilayer Feed forward Network are broadly specified in the literature [17]. Feed-forward ANNs are straight forward networks that allow signals to travel one way only. Recurrent network consists of only single feedback loop, such that activation can flow round the loop (fig. 4). Multilayer Feed-forward network consists of several layers irrespective of an input and output layer (fig.5). In medical research, the most commonly used ANN is the multilayer perceptrons which use back propagation training. This back propagation consists of fitting the parameters (weights) of the model by a criterion function, usually squared error or maximum likelihood, using a gradient optimization method. In recent times, staging systems for cancers have evolved for better cancer progression studies.

The TNM system is one of the most widely used cancer staging systems based on the size and/or extent (reach) of the primary tumor (T), the amount of spread to nearby lymph nodes (N), and the presence of metastasis (M) or secondary tumors formed by the spread of cancer cells to other parts of the body [19]. Therefore, to reduce the computation time of diagnosing the breast cancer with reduced death rate, this computerized breast cancer diagnosis has been developed.

Artificial Neural Network (ANN) models are nowadays utilized for breast cancer prognosis, thereby predicting its prognosis and recurrence so to aid post-operative treatments. As a prediction model these methods have been used in censored survival data. [20, 21, 22, 23, 24].

To assist prognosis, an ANN model learns to predict the time to recur (recurrence time) from previous patients and then predicts outcome for the new patient [25]. An ANN tool reduces the work loads on clinicians by detecting artifact and thereby helps them in decision making. There are various neural network algorithms studied by various researchers to detect and classify the Breast cancer, skin cancer etc. Table 1 provides a preview on the various ANN used for classification, detection and of cancers.

<table>
<thead>
<tr>
<th>S.No.</th>
<th>Algorithms</th>
<th>Utilized in</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.</td>
<td>Back propagation Algorithm</td>
<td>Classification of Breast cancer into malignant or benign with the accuracies of 94.11% and 100%</td>
</tr>
<tr>
<td>2.</td>
<td>Multilayer Perceptron Training Algorithm</td>
<td>Classification of mammographic images of breast cancer. Accuracy obtained is 95.49%.</td>
</tr>
<tr>
<td>3.</td>
<td>Back propagation Algorithm</td>
<td>Breast cancer detection and classification using ANN. provided an accuracy of 94%</td>
</tr>
<tr>
<td>4.</td>
<td>Hierarchical Radial Basis Function</td>
<td>Breast cancer detection using hierarchical RBF with the accuracy of 97.09%.</td>
</tr>
</tbody>
</table>

Over the hindsight, ANN has also been applied in different survival analysis studies of circulating tumor cells in metastatic breast cancer patients [26], classification of micro-

---

International Journal of Pharmacy and Biological Sciences (e-ISSN: 2230-7605)

S.S. Sandhu* et al

www.ijpbs.com or www.ijpbsonline.com

Int J Pharm Bio Sci
calcification in mammograms [27], classification of breast cancer [28], prediction and classification of cancer patients based upon their gene expression profiles [29]. Previous research in this arena has been undertaken by various researchers. Kim et al. [30] compared the performance of logistic regression and ANN in differentiation of benign and malignant breast masses. In recent times, Faradmal et al. [31] compared log-logistic regression and artificial neural network models in prediction of breast cancer (BC) survival. The study demonstrated the use of ANN method for prediction of survival in field of breast cancer in comparison to LLM model. Bhooshan et al. [32] classified breast lesions using computer-aided diagnosis (CADx) of precontrast high spectral and spatial resolution (HiSS) MRI to that of clinical dynamic contrast-enhanced MRI (DCE-MRI). The study utilized Bayesian artificial neural networks and receiver operating characteristic (ROC) analysis which evaluated the performance with leave-one-lesion-out validation.

Subbbaiah et al. [33] proposed a complete robust back propagation ANN model (using Neurointelligence software) of breast carcinomas diagnosed on fine-needle aspiration cytology (FNAC) based on cytomorphological, morphometric, nuclear densitometric data and gray level co-occurrence matrix (GLCM) of ductal carcinoma and fibroadenomas. Wang et al. [34] utilized ANN (STATISTICA(*) software) to predict the five-year survivability of breast cancer patients who were diagnosed and received radiotherapy treatment. The accuracy rate was found to be 85% with receiver operating characteristic (ROC) curve of 0.79 thereby showing ANN as a good tool to predict the five-year survivability of breast cancer patients.

Mojarad et al. [35] corroborated the prognostic significance of a group of well-established prognostic markers (tumour size, oestrogen receptor status (ER), progesterone receptor status (PR), Ki-67 and p53 expression) for breast cancer outcome prediction in terms of nodal status hence obviating the necessity to carry out axillary dissection.

Shi et al. [36] validated the use of artificial neural network (ANN) models for predicting quality of life (QOL) after breast cancer surgery and compared the predictive capability of ANNs with that of linear regression (LR) models.

Burke et al. [37] employed an ANN composed of three interconnected layers of nodes: an input layer, with each input node corresponding to a patient variable; a hidden layer and an output layer. A transfer function (known as activation function) to send the information to the adjacent layer nodes. ANN discriminates breast cancer and also helps in the risk estimation via discriminating malignant breast lesions from benign ones. In the past, several Artificial Neural Network (ANN) models have been developed for breast cancer risk prediction.

A successful breast cancer diagnosis requires a systematic approach utilizing image analysis, characterization, and integration of clinical and mammographic variables. Thus various new prognostic factors have been identified and novel methods have been introduced using artificial neural network statistical models to discriminate benign and malignant cancers. But irrespective of the novel methods utilized, the accuracy of predicting the probability of breast cancer so to aid decision making is still uncertain. Chuaqui et al. [38] describes the measurement of gene-expression profiles using microarray technology which is
increasingly becoming popular among the biomedical research community. High-end computational analysis is also used to identify gene expression patterns, such as clustering, multidimensional scaling, or pattern identification, including neural networks and heuristic algorithms. These data pre-processing and numerical management with a statistical approach accurately identify the differences in the gene expression between sample sets. Likewise, neural networks have been used to examine colon cancer and a range of cancers where the model was found to be 83% accurate at predicting which patients had tumors, based on the blood-plasma lipid profile, and only 8% of patients were identified as false positives [39]. The integrative approach between histological, biochemical and clinical information helps in predicting and estimating the behavior and outcome of the disease.

Khan et al. [40] developed a method for classifying cancers in to specific diagnostic categories based on their gene expression signatures using Artificial Neural Networks (ANNs) using small, round blue-cell tumors (SRBCTs) as a model. ANN research mainly aims to provide a filter between the cases and thereby distinguish the cancers hence reducing the cost of medication and helping clinicians in focusing on cancer prone patients. Hence, an Artificial Neural Network (ANN) system is utilized for diagnosis, prognosis and prediction of cancer [41].

In ANN, the networks are firstly structured as per the particular application. To start this process initial weights are randomly chosen where the ANN is trained by exposing it to a set of existing data (based on the follow-up history of cancer patients) and the outcomes is known. Back-propagation algorithm, a learning technique used in multilayer networks is described as one of the most effective approaches to machine learning algorithm where the information flows from the direction of the input layer towards the output layer [42].

Training in ANN’s is achieved via examples adjusting the connection weights in ANN’s iteratively. The number of iterations of the training algorithm and the convergence time varies depending on the weight initialization. After the repetition of the processes, for a sufficiently large number of training cycles, the network usually converges to a state where the error in the calculations is small thus implying the network to be learned to a certain target function.

CONCLUSION
Over the hindsight, different methods and softwares have been utilized to improve the diagnosis and prognosis of breast cancer. An integrative approach utilizing all the advanced techniques is the best answer to the unanswered questions. In this context, Artificial Neural Networks certainly arose as a valuable tool for the diagnosis of breast cancer as they are automated with intelligent decision making strategy which is not affected from human error factors like emotion, lack of attention or experience. Besides neural networks also provides an inputs in different layered perceptrons without maintaining any software. Thereby, its utilization in diagnosis diminishes the ample time which is required for diagnosis as patients would be checked for cancer quickly and painlessly thereby detecting the disease at an early stage. Thus, ANN is an effective option for cancer diagnosis so to help clinicians and oncologists in the prediction and prognosis of cancer. Hence, neural network will not only assist radiologists by helping in
diagnoses but also it will reduce the economic and mental burden on patients caused due to prolong treatment of breast cancer.

ACKNOWLEDGEMENTS

The authors of this review would like to thank Dr. K. N. Singh Yadav, Vice-Chancellor, R.D. University Jabalpur for his kind support and help. We would also like to thank the Head, Dept. of Biological Sciences, R. D. University, Jabalpur, India for technical and linguistic assistance.

REFERENCES

*Corresponding Author: ssandhu@rediffmail.com