

Intelligent Breast Cancer Prediction Empowered with Fusion and Deep Learning

Shahan Yamin Siddiqui^{1,2}, Iftikhar Naseer³, Muhammad Adnan Khan⁴, Muhammad Faheem Mushtaq⁵, Rizwan Ali Naqvi^{6,*}, Dildar Hussain⁷ and Amir Haider⁸

¹School of Computer Science, National College of Business Administration and Economics, Lahore, 54000, Pakistan

²School of Computer Science, Minhaj University Lahore, Lahore, 54000, Pakistan

³Department of Computer Science & Information Technology, Superior University, Lahore, 54000, Pakistan

⁴Department of Computer Science, Faculty of Computing, Riphah International University, Lahore Campus, Lahore, 54000, Pakistan

⁵Department of Information Technology, Khwaja Fareed University of Engineering and Information Technology, Rahim Yar Khan, 64200, Pakistan

⁶Department of Unmanned Vehicle Engineering, Sejong University, Seoul, 05006, Korea

⁷School of Computational Sciences, Korea Institute for Advanced Study (KIAS), 85 HoegiroDongdaemungu, Seoul, 02455, Korea

⁸Department of Intelligent Mechatronics Engineering, Sejong University, Seoul, 05006, Korea

*Corresponding Author: Rizwan Ali Naqvi. Email: rizwanali@sejong.ac.kr

Received: 26 August 2020; Accepted: 17 November 2020

Abstract: Breast cancer is the most frequently detected tumor that eventually could result in a significant increase in female mortality globally. According to clinical statistics, one woman out of eight is under the threat of breast cancer. Lifestyle and inheritance patterns may be a reason behind its spread among women. However, some preventive measures, such as tests and periodic clinical checks can mitigate its risk thereby, improving its survival chances substantially. Early diagnosis and initial stage treatment can help increase the survival rate. For that purpose, pathologists can gather support from nondestructive and efficient computer-aided diagnosis (CAD) systems. This study explores the breast cancer CAD method relying on multimodal medical imaging and decision-based fusion. In multimodal medical imaging fusion, a deep learning approach is applied, obtaining 97.5% accuracy with a 2.5% miss rate for breast cancer prediction. A deep extreme learning machine technique applied on feature-based data provided a 97.41% accuracy. Finally, decision-based fusion applied to both breast cancer prediction models to diagnose its stages, resulted in an overall accuracy of 97.97%. The proposed system model provides more accurate results compared with other state-of-the-art approaches, rapidly diagnosing breast cancer to decrease its mortality rate.

Keywords: Fusion feature; breast cancer prediction; deep learning; convolutional neural network; multi-modal medical image fusion; decision-based fusion



This work is licensed under a Creative Commons Attribution 4.0 International License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

1 Introduction

Breast cancer is a great health threat and a significant factor in female mortality. The occurrence of breast cancer is increasing every day. It has become the second-leading disease due to its rapid spread among women worldwide [1]. The early diagnosis of breast cancer can effectively mitigate its risk of mortality. It increases the ratio of life survival through proper treatment because it is one of the most curable malignancies if detected earlier [2]. The detection procedure for breast cancer is expensive and time-consuming. The diagnosing process depends on the consistency and knowledge of medical examiners [3]. The two primary types of tumors are benign (noncancerous) and malignant (cancerous). These two types are further divided into subdivisions and properties. Malignant is considered life-threatening, whereas benign is not typically harmful [4]. Humans are often prone to making omissions so misdiagnosis can take the patient to a noncurable stage. Hence, a fast and efficient computer-aided diagnosing technique can become an assistive tool for the early and accurate detection of breast cancer. The mammogram screening method assists the early detection of breast cancer and enables physicians to make accurate decisions regarding breast cancer treatment [5].

In the process of breast cancer detection, computer-aided diagnosis (CAD) is an assistive tool for early detection. The CAD techniques are effectively used in mammograms to decrease the burden on medical experts and reduce the misdiagnosis of breast cancer [6]. Recently, artificial intelligence technology has been applied to various machine learning and computer vision problems. The deep learning (DL) approach has been used in many scientific and engineering applications, which have increased their performance using DL technology [7,8]. Recently a variety of DL technologies have been successfully adopted in the medical domain in the prediction of heart disease, infant brain [9], lung detection [10], and breast cancer classification [11].

Moreover, DL has been used for many types of cancer and has scored significant success in breast cancer screening [12]. With the emergence of DL, various research has been conducted by considering deep architectures, but a significant type of DL technique is the convolutional neural network (CNN) [13]. Araujo applied a CNN to categorize breast biopsy images to diagnose breast cancer. The accuracy of the proposed model was 83.3% for cancerous and noncancerous detection, whereas the accuracy was 77.8% for invasiveness, carcinoma in situ, benign, and normal tissues [14].

In another study, Yao et al. [15] presented a novel DL model to classify histological images into four classes. This model extracted images featured by an apparent combination of a CNN and a recurrent neural network (RNN). Afterward, the extracted features were used as input in the RNN. The fusion method was applied to three datasets of histological images in the proposed model. Similarly, Wang et al. opted for a CNN and hybrid CNN with a support vector machine (SVM) model for the classification of the breast cancer histological image dataset. Considering the above, the proposed study applied a multi-model approach and achieved 92.5% accuracy [16].

2 Literature Review

The focus of research efforts related to the prevalence of breast cancer is primarily on the diagnosis and detection of tumors. This section of the research briefly summarizes the existing research on breast cancer. During the last decade, research related to this topic has increased, and various computer-based systems have been developed to overcome this challenge.

Wang et al. [17] suggested the mammogram as an essential element in CAD for early breast cancer diagnosis and treatment. Wang et al. designed a model based on feature fusion with CNN

deep features for the detection of breast cancer. The model consists of three phases. The first phase is the unsupervised extreme learning machine (ELM) and CNN deep features used for mass detection. In the second phase, deep, density, texture, and morphological features are used to create a feature set. Third, ELM classifiers are used to classify malignant and benign breast masses using a fused feature set. The experimental results have shown that an ELM classifier has a good capability to identify, classify, and handle multidimensional feature classification.

Chiang et al. [18] suggested a computer-aided detection system to detect tumors using a breast ultrasound. The proposed system is based on a 3D CNN and prioritized candidate aggregation. First, the volumes of interest are extracted using a sliding window method. Then, a 3D CNN was used to estimate the tumor probability for each volume of interest. Those with a high estimated probability were marked as tumor candidates, and the situation of each candidate may overlap. For the cumulative overlap of candidates, an innovative system was designed. In the aggregation process, tumor probability was used to form the candidate prioritization. The experimental results for 171 tumors using the proposed model obtained sensitivities of 95% with an execution time of less than 22 s, which demonstrates that the proposed method is faster than the existing approaches.

Agnes et al. [19] proposed a model called multiscale all CNN (MA-CNN) for the detection of breast cancer. The CNN approach was used in the MA-CNN model to classify mammogram images. The CNN classifier improves the performance for multiple scale features obtained from mammogram images. The proposed MA-CNN model identifies mammographic images into benign and malignant classes, and the experimental results proved that the proposed model is a powerful tool for the detection of breast cancer using mammogram images.

Shen et al. [20] developed a model based on a DL algorithm for the detection of breast cancer. The proposed model uses end-to-end training techniques to screen mammograms. Using this approach, the system requires annotations in the initial training stage and rest stages to require image-level labels only. The CNN was used to classify mammogram screens, and in the breast, the dataset was used to train and evaluate the proposed model. The results reveal that the proposed model achieved high accuracy compared with other existing work on heterogeneous mammography. Zhou et al. [21] presented a CNN-based radionics approach to detect breast cancer. The proposed model applies shear-wave elastography data to obtain important morphology information and feature extraction procedures through a CNN. In the training phase, 540 images in which 318 images were malignant and 222 were benign were used to train this model. The experimental results indicated 95.8% accuracy, 95.7% specificity, and 96.2% sensitivity, respectively.

Qiyuan et al. [22] stated that multiparametric magnetic resonance imaging increased the performance of radiologists in analyzing breast cancer. The proposed study used 927 images, and a pretrained CNN model was applied for feature extraction. The SVM classifier was trained to obtain benign and malignant images from CNN extracted features. The sequence of various levels of feature fusion, image fusion, and classifier fusion was examined, achieving an accuracy of 95%.

Chaves et al. [23] stated that the early diagnosis of breast cancer can increase the chance of treatment and cure patients. According to the researchers, infrared thermography is an essential and promising technique to detect breast cancer. It was found to be low cost and less harmful than radiation when it was applied in women of a young age. The authors applied pretrained transfer learning CNNs to detect breast cancer on infrared images, and the dataset consisted of 440 infrared images that were further divided into two classes, normal and pathology. The experimental results reveal that CNN with infrared images can play a pivotal role in the early detection of breast cancer. George et al. [24] proposed a model for the detection of breast cancer

based on nucleus feature extraction by the CNN. The CNN approach was used to extract features from images, and SVMs with the feature fusion technique were applied to extracted CNN features to categorize breast cancer images. With the help of the feature fusion method, the local nucleus features were transformed into compact image features that improve performance. The proposed model achieved promising results with 96.66% accuracy compared with other existing approaches.

Kadam et al. [25] proposed feature ensemble learning consisting of softmax regression and sparse autoencoders to detect breast cancer. Softmax regression and sparse autoencoders classify tumors into benign and malignant classes. The Wisconsin breast cancer dataset in this study and the prediction results outperform with 98.60% accuracy. Additionally, the findings were also compared with the previous work. The statistical analysis is a useful and beneficial model for breast tumor classification. Singh et al. [26] used an SVM classifier technique for detecting breast cancer. The proposed method was tested on the database and achieved 92.3% accuracy with a cubic SVM classifier.

Motivated by the previous research, the researchers focused on cloud and decision-based fusion for an intelligent breast cancer prediction system using a hierarchical DL approach. Furthermore, DL approaches have been implemented to train and evaluate the proposed model. Moreover, this research is significant because of its approach to fusing different datasets and concludes the results based on a fused dataset. The proposed research is also generalizable to other datasets and will be beneficial to future researchers in the medical field, especially for the early prediction of breast cancer.

3 Proposed CF-BCP System Model

The cloud and decision-based fusion model for an intelligent breast cancer prediction system using hierarchical DL (CF-BCP) is proposed to provide diagnoses. The proposed CF-BCP training model comprised two types of datasets such as image and feature based. The multimodal medical image fusion technique was applied to the image dataset, and the preprocessing technique was used to remove noise from the image data acquisition layer. The moving average technique was used for handling missing values in the electronic medical record (EMR) dataset.

In the application layer, the CNN is applied for breast cancer prediction in the image dataset. In the evaluation layer, the accuracy and miss rate of the proposed CF-BCP model was investigated. For the EMR data, deep ELM (DELM) was applied for breast cancer prediction. If the learning criteria are not met in both conditions, then the system must retrain, whereas if the learning criteria are met, then the data are stored on the cloud, and the next step is the decision-based fusion empowered with fuzzy logic activation. Decision-based fusion empowered with fuzzy logic determines whether the fused images are benign or malignant. In the second training layer, the fused malignant data are used to detect breast cancer types, such as ductal, lobular, mucinous, and papillary carcinoma, and are also stored on the cloud, which is illustrated in Fig. 1.

After the preprocessing step, the proposed CF-BCP model imports fused data from the cloud to predict breast cancer. If breast cancer is not detected, then the model discards the prior part. If breast cancer is detected, then a fused database for intelligent breast cancer prediction (activation layer) is created, and a breast cancer stage prediction model is imported from the cloud. After the detection of breast cancer stages, the patient is referred to the hospital for further treatment, which is presented in Fig. 2.

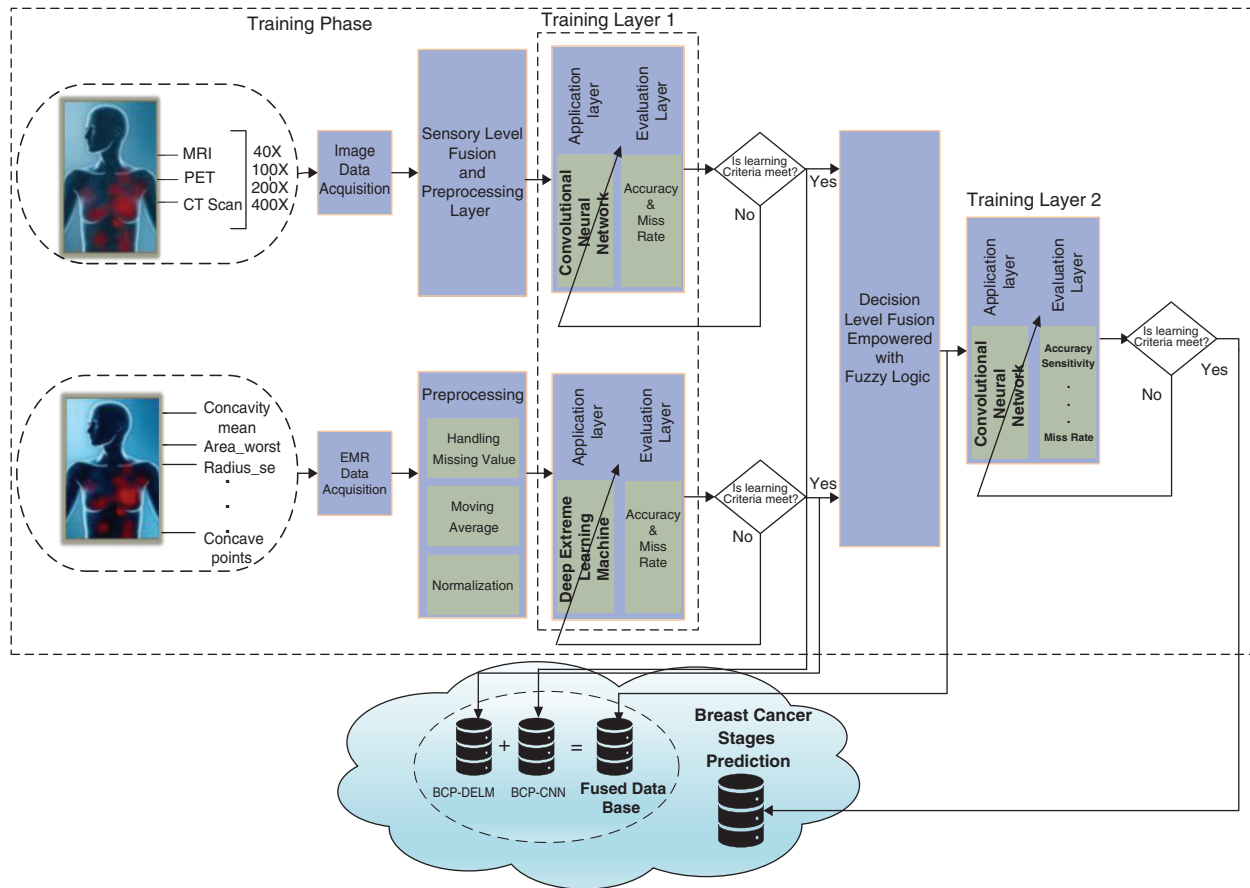


Figure 1: Proposed CF-BCP training model

3.1 Convolutional Neural Network

Deep learning (DL) is a widespread technique applied in various fields ranging from lifespan to forecasting transport, diseases, agriculture, stock markets, and so on. Moreover, DL is very helpful in different areas due to its fast learning procedure.

The CNN involves two segments: Convolutional and pooling layers. The proposed CF-BCP uses two CNN layers. The leading layer is used for diagnosing breast cancer, and the other layer is used for predicting breast cancer types. Moreover, the CNN comprises three layers: The input, hidden, and output layers. The size of the input images is transformed into $700 \times 460 \times 3$, where 700×460 represents the width and height of the fused input images, and 3 represents the number of channels. The convolutional layer can resolve more computational tasks. The purpose of this layer is to recover features by applying filters, preserving the spatial relationships among pixels. The pooling layer reduces the dimensions of the fused images and uses less computational time. Two pooling layers are often used: max pooling and average pooling.

The average pooling layer is used in the proposed CF-BCP model in Fig. 3. This layer is used to preserve the specific image features captured in the CNN procedure. Additionally, all inputs relate to the rectified linear unit activation function that represents a fully connected layer. Extracted data from the last layer are compiled into a fully connected layer to obtain the final output. Finally, the convolutional layer converts it into a single flattened length and becomes a

fully connected layer. The softmax layer is applied to transform logits into probabilities. In the last layer of the CNN model, the accuracy values are marked.

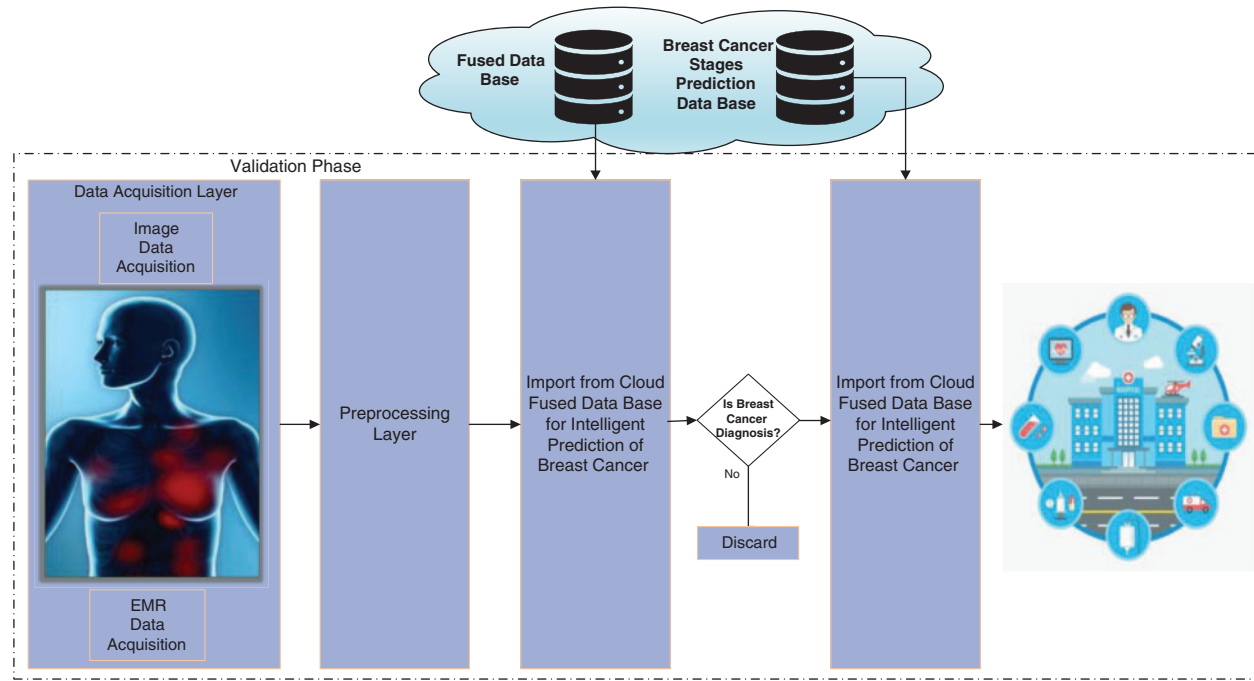


Figure 2: Proposed CF-BCP validation model

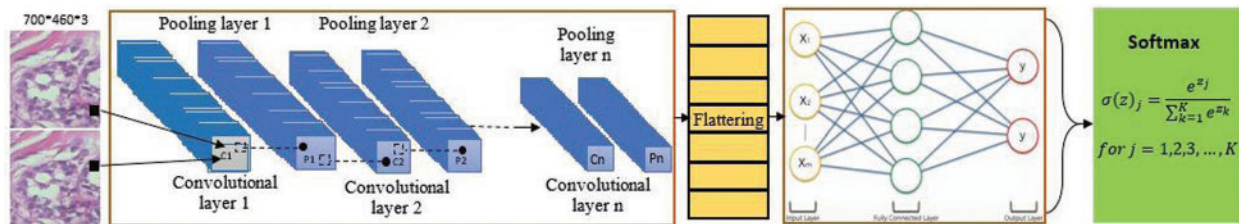


Figure 3: Convolutional neural network model for the proposed CF-BCP system model

In the mathematical model, the target of back prorogates using the derivative of Eq. (1) with respect to (w.r.t.) the weights $\frac{\partial A}{\partial t}$ and bias $\frac{\partial A}{\partial x}$:

$$A = - \sum_{m=1}^z (W_m \log(w_m)), \tag{1}$$

where m denotes the number of classes depending on applications. The softmax transformation is shown in Eq. (2):

$$w_m = \frac{e^{C_m}}{\sum_{n=1}^k e^{C_U}}, \tag{2}$$

where C_m denotes the logits, which are converted into probabilities using softmax $C_a = \sum_{n=1}^{\eta} (T_{na} * B_n^l)$,

where C_a is obtained using interconnected weights with B_n^l .

Next, we find the loss w.r.t the weights that consist of the two summations shown in Eq. (3):

$$\frac{\partial A}{\partial T_{n,a}^l} = \sum_{n=1}^{\eta'} \sum_{a=1}^z \left(\frac{\partial A}{\partial C_a^l} \frac{\partial C_a^l}{\partial T_{n,a}^l} \right), \tag{3}$$

where $n = 1$ to k , $l = 1$ to z , and $\frac{\partial w_m^l}{\partial C_a^l}$ = softmax derivative (*Sd*). In Eq. (1), loss has w_m as its parameter, which is indirectly related to C_m in terms of the following expression: $w_m = \frac{e^{C_a}}{\sum_{u=1}^z e^{C_u}}$.

In addition, $C_a = \sum_{n=1}^{\eta} (T_{na} * B_n^l)$ is given as $C_m = C_a$.

Condition 1: $m = a$

Taking the derivative of Eq. (2), by the quotient rules, results in the following:

$$\frac{\partial w_m^l}{\partial C_{(m=a)}^l} = \frac{e^{C_m} \sum_{u=1}^z e^{C_u} - e^{C_m} e^{C_a}}{\sum_{u=1}^z e^{C_u} * \sum_{u=1}^z e^{C_u}}. \tag{4}$$

Taking the common $\frac{e^{C_a}}{\sum_{u=1}^z e^{C_u}}$ from Eq. (4), we obtain the following:

$$\frac{\partial w_m^l}{\partial C_a^l} = \frac{e^{C_a}}{\sum_{u=1}^z e^{C_u}} \left[\frac{\sum_{u=1}^z e^{C_u} - e^{C_a}}{\sum_{u=1}^z e^{C_u}} \right].$$

By dividing, we obtain the following:

$$\frac{\partial w_m^l}{\partial C_a^l} = \frac{e^{C_a}}{\sum_{u=1}^z e^{C_u}} \left[1 - \frac{e^{C_a}}{\sum_{u=1}^z e^{C_u}} \right] \{ : m = a \}.$$

As we know, $w_m = \frac{e^{C_a}}{\sum_{n=1}^k e^{C_u}}$; thus, the above equation can be further written as follows:

$$\frac{\partial w_m^l}{\partial C_a^l} = w_m^l (1 - w_m^l) = w_m^l (1 - w_m^l) \text{ for } (m = a). \tag{5}$$

Moreover, when $m \neq l$ th unit, this indicates a low probability.

Condition 2: $m \neq a$

Taking the derivative of Eq. (3), through quotient rules w.r.t. C_a^l , we have the following:

$$\frac{\partial w_m^l}{\partial C_a^l} = \frac{\frac{\partial}{\partial C_a^l} e^{C_m} * \sum_{u=1}^z e^{C_u} - e^{C_m} \frac{\partial}{\partial C_a^l} [\sum_{u=1}^z e^{C_u}]}{\sum_{u=1}^z e^{C_u} * \sum_{u=1}^z e^{C_u}}.$$

This can be written as follows:

$$\frac{\partial w_m^l}{\partial C_a^l} = 0 - \frac{e^{C_m} * e^{C_a}}{\sum_{u=1}^z e^{C_U} * \sum_{u=1}^z e^{C_U}} = - \frac{e^{C_m}}{\sum_{u=1}^z e^{C_U}} * \frac{e^{C_a}}{\sum_{U=1}^z e^{C_U}}.$$

As we know, $w_m^l = \frac{e^{C_m}}{\sum_{U=1}^z e^{C_U}}$ and $w_a^l = \frac{e^{C_a}}{\sum_{U=1}^z e^{C_U}}$; thus, we can derive this equation as follows:

$$\frac{\partial w_m^l}{\partial C_a^l} = -w_m^l w_a^l \quad \text{for } (m \neq a). \quad (6)$$

We can summarize Eqs. (5) and (6) as follows:

$$\frac{\partial w_m^l}{\partial C_a^l} = \begin{cases} w_m^l (1 - w_m^l) & \text{for } (m = a) \\ -w_m^l w_a^l & \text{for } (m \neq a) \end{cases}. \quad (7)$$

The cross-entropy loss does not have any module of C_a^l ; therefore, taking the partial derivative of C_a^l w.r.t. $\log(w_U)$ results in the following:

$$A = - \sum_{m=1}^z (W_m * \log(w_m)).$$

Taking the derivative, the equation becomes

$$\begin{aligned} \frac{\partial A}{\partial C_a^l} &= - \sum_{m=1}^z \left(W_U * \frac{\partial}{\partial C_a^l} \log(w_U) \right), \\ \frac{\partial A}{\partial C_a^l} &= - \sum_{m=1}^z W_U \left(\frac{\partial}{\partial w_u^l} \log(w_U) \right) \frac{\partial w_U^l}{\partial C_a^l}, \\ \frac{\partial A}{\partial C_a^l} &= - \sum_{m=1}^z \frac{W_U}{w_u^l} \frac{\partial w_U^l}{\partial C_a^l}, \end{aligned} \quad (8)$$

where $\frac{\partial w_u^l}{\partial C_a^l}$ has already been calculated as shown in Eq. (7), and Eq. (8) is divided into two parts:

$$\frac{\partial A}{\partial C_a^l} = - \frac{W_U^l}{w_u^l} * w_u^l (1 - w_a^l) - \sum_{u \neq 1}^z \left(- \frac{W_u^l}{w_u^l} * w_u^l w_a^l \right),$$

where $\sum_{u \neq 1}^z \left(- \frac{W_u^l}{w_u^l} * w_u^l w_a^l \right)$ for $u \neq a$, and $\frac{W_u^l}{w_u^l} * w_u^l (1 - w_a^l)$ for $u = a$.

We can simplify this as follows:

$$\frac{\partial A}{\partial C_a^l} = -w_u^l (1 - w_a^l) + \sum_{u \neq 1}^z w_u^l w_a^l.$$

We can further simplify this:

$$\frac{\partial A}{\partial C_a^l} = -w_u^l + W_u^l w_a^l + \sum_{u \neq l}^z w_u^l w_a^l,$$

$$\frac{\partial A}{\partial C_a^l} = w_a^l \left(w_u^l + \sum_{u \neq l} w_u^l \right) - w_u^l,$$

where $(w_u^l + \sum_{u \neq l} w_u^l)$ represents I , $\frac{\partial A}{\partial C_a^l} = (w_a^l - w_u^l)$, and $\frac{\partial A}{\partial C_a^l} = (w_a^l - w_u^l), \{ \cdot : u = a \}$.

Then, we enter the value of $\frac{\partial A}{\partial C_a^l}$ into Eq. (3):

$$\frac{\partial A}{\partial T_{n,a}^l} = \sum_{n=1}^{\eta} \sum_{a=1}^z \left(\frac{\partial A}{\partial C_a^l} \frac{\partial C_a^l}{\partial T_{n,a}^l} \right),$$

$$\frac{\partial A}{\partial T_{n,a}^l} = \sum_{n=1}^{\eta} \sum_{a=1}^z (w_a^l - w_u^l) b_n^l, \tag{9}$$

where $\frac{\partial C_a^l}{\partial T_{n,a}^l} = b_n^l$ as input weights. Eq. (9) presents the derivative of the loss w.r.t. the weights for the fully connected layer.

3.2 Deep Extreme Learning Machine

The DELM is a significant method that is primarily used for prediction. Fig. 4 represents the DELM architecture that consists of an input layer, hidden layers, and output layer. In the input layer, the various features are used as input. Six hidden layers are used in the proposed CF-BCP model.

For the mathematical DELM, Eq. (10) presents the input layer, and Eq. (11) represents the output of the first layer:

$$\alpha_i = s_1 + \sum_{j=1}^n (\beta_{ji} * z_j) \tag{10}$$

$$\gamma_i = \frac{1}{1 + e^{-\alpha_i}}, \quad \text{where } i = 1, 2, 3, \dots, z. \tag{11}$$

The following is the feedforward propagation for the second layer to the output layer in Eq. (12):

$$\alpha_{1k} = s^k + \sum_{i=1}^m (x_{i1k=1} * \gamma_i^{k=1}) \tag{12}$$

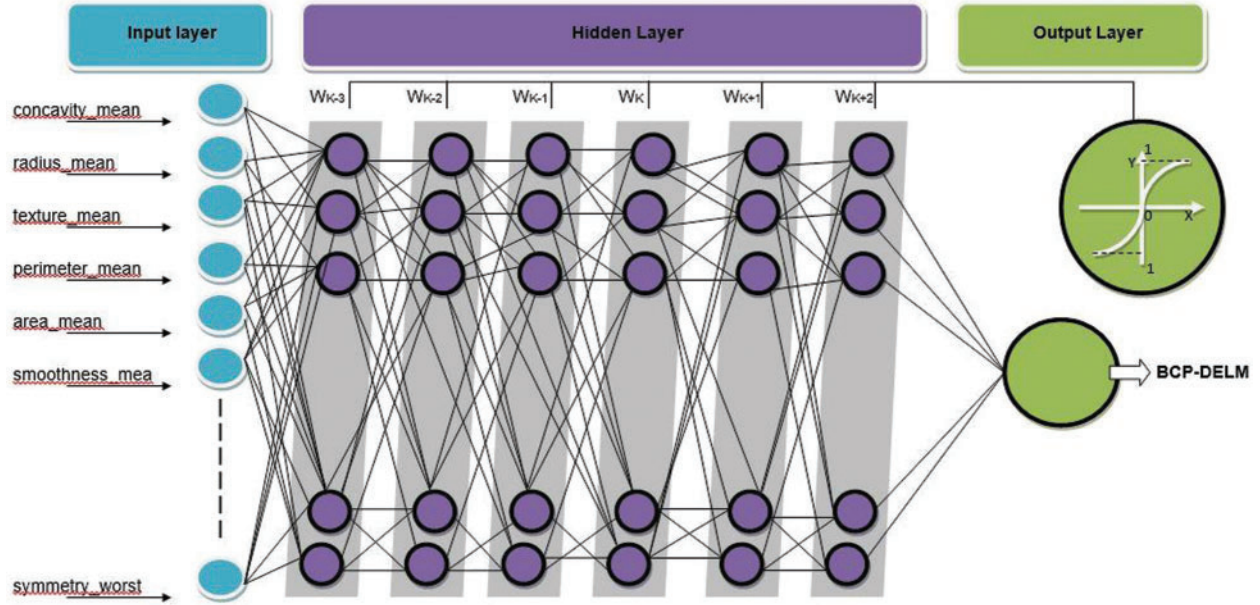


Figure 4: Deep extreme learning machine (DELM) model for the proposed CF-BCP system model

The activation feature of the output layer is indicated in Eq. (13):

$$\gamma_{1^k} = \frac{1}{1 + e^{-\alpha_1^{k-1}}} \quad \text{where } l = 2, 3, \dots, z, \quad (13)$$

$$\alpha'_{1^k} = s^k + \sum_{i=1}^m (x_{i1^k} * \gamma_i^k) \quad \text{where } k = 1, 2, 3, \dots, 6. \quad (14)$$

The error in backpropagation is written as follows in Eq. (15):

$$\epsilon = \frac{1}{2} \sum_1 (\text{Target}_1 - \gamma_1^{k=6})^2, \quad (15)$$

where Target_1 , and $\gamma_1^{k=6}$ represent the desired and calculated outputs, respectively.

Eq. (16) reflects the rate of the weight shift that is written for the output layer:

$$\Delta \beta \propto -\frac{\partial \epsilon}{\partial \beta}$$

$$\Delta \gamma_{i,1^k=6} = -\epsilon \frac{\partial \epsilon}{\partial \gamma_{i,1^k=6}} \quad (16)$$

It is written by adding the chain rule, as follows in Eq. (17):

$$\Delta \gamma_{i,1^k=6} = -\epsilon \frac{\partial \epsilon}{\partial \gamma_1^k} \times \frac{\partial \gamma_1^k}{\partial \alpha_{1^k}} \times \frac{\partial \alpha_{1^k}}{\partial \gamma_{i,1^k}} \quad (17)$$

After implementing the chain rule (substituting Eq. (17)), it is possible to obtain the weight value modified as shown in Eq. (18):

$$\begin{aligned} \Delta\gamma_{i,lk=6} &= \epsilon \left(\text{Target}_l - \gamma_1^k \right) \times \gamma_{1k} (1 - \gamma_{1k}) \times \left(\gamma_i^k \right), \\ \Delta\gamma_{i,l}^k &= \epsilon \mu_{1k} \gamma_1^k, \end{aligned} \tag{18}$$

where $\mu_{1k} = (\text{Target}_l - \gamma_{1k}) \times \gamma_{1k} (1 - \gamma_{1k})$, and so on.

$$\begin{aligned} \Delta\beta_{j,ik} &\propto - \left[\sum_1 \frac{\partial \epsilon}{\partial \gamma_{1k}} \times \frac{\partial \gamma_{1k}}{\partial \alpha_{1k}} \times \frac{\partial \alpha_{1k}}{\partial \gamma_{jk}} \right] \times \frac{\partial \gamma_{jk}}{\partial \alpha_{jk}} \times \frac{\partial \alpha_{jk}}{\partial \beta_{j,ik}}, \\ \Delta\beta_{j,ik} &= -\epsilon \left[\sum_1 \frac{\partial \epsilon}{\partial \gamma_{1k}} \times \frac{\partial \gamma_{1k}}{\partial \alpha_{1k}} \times \frac{\partial \alpha_{1k}}{\partial \gamma_{jk}} \right] \times \frac{\partial \gamma_{jk}}{\partial \alpha_{jk}} \times \frac{\partial \alpha_{jk}}{\partial \beta_{j,ik}}, \\ \Delta\beta_{j,ik} &= \epsilon \left[\sum_1 \left(\text{Target}_l - \gamma_1^k \right) \times \gamma_{1k} (1 - \gamma_{1k}) \times \left(\gamma_i^k \right) \right] \times \gamma_{1k} (1 - \gamma_{1k}) \times \eta_j, \\ \Delta\beta_{j,ik} &= \epsilon \left[\sum_1 \mu_{1k} (x_{i,1k}) \right] \times \gamma_{1k} (1 - \gamma_{1k}) \times \eta_j, \\ \Delta\beta_{j,ik} &= \epsilon \mu_{1k} \eta_j, \end{aligned}$$

where $\mu_{1k} = \left[\sum_1 \mu_{1k} (x_{i,1k}) \right] \times \gamma_{1k} (1 - \gamma_{1k})$.

The output and hidden layers in Eq. (19) in weights are the updating and biases between them:

$$x_{i,lk=6}^+ = x_{i,lk=6} + \delta_{e^{k=6}} \Delta\gamma_{i,lk=6} \tag{19}$$

In Eq. (20), the weight and bias changes between the input and hidden layers are represented:

$$\beta_{j,ik}^+ = \beta_{j,ik} + \delta_{e^k} \Delta\beta_{j,ik} \tag{20}$$

where δ_e is the learning rate of the BCP-DELM, and the value of δ_e is between 0 and 1. The convergence of BCP-DELM depends upon the careful selection of the value of δ_e .

3.3 Decision-Based Fusion Empowered with Fuzzy Logic

The proposed decision-based fusion model empowered with fuzzy logic is based on knowledge, expertise, and logical reasoning ability. The fuzzy logic model has the capacity to manage the uncertainty and imprecision of the data using a proper method. The proposed cloud and decision-based fusion for an intelligent breast cancer prediction system using the hierarchical DL (CF-BCP) model mathematically is written as follows:

$$\mu_{CNN-Layer-1 \cap DELM-Layer-1} (cnn, delm) = \min \left[\mu_{CNN-Layer-1} (cnn), \mu_{DELM-Layer-1} (delm) \right]$$

Fig. 5 represents the decision-based fusion lookup diagram for the detection of breast cancer and its types:

R_{dbf}^1 = If CNN layer 1 is benign and DELM layer 1 is benign, then the breast cancer diagnosis is benign.

R_{dbf}^2 = If CNN layer 1 is benign and DELM layer 1 is malignant, then the breast cancer diagnosis is malignant.

R_{dbf}^3 = If CNN layer 1 is malignant and DELM layer 1 is benign, then the breast cancer diagnosis is malignant.

R_{dbf}^4 = If CNN layer 1 is malignant and DELM layer 1 is malignant, then the breast cancer diagnosis is malignant.

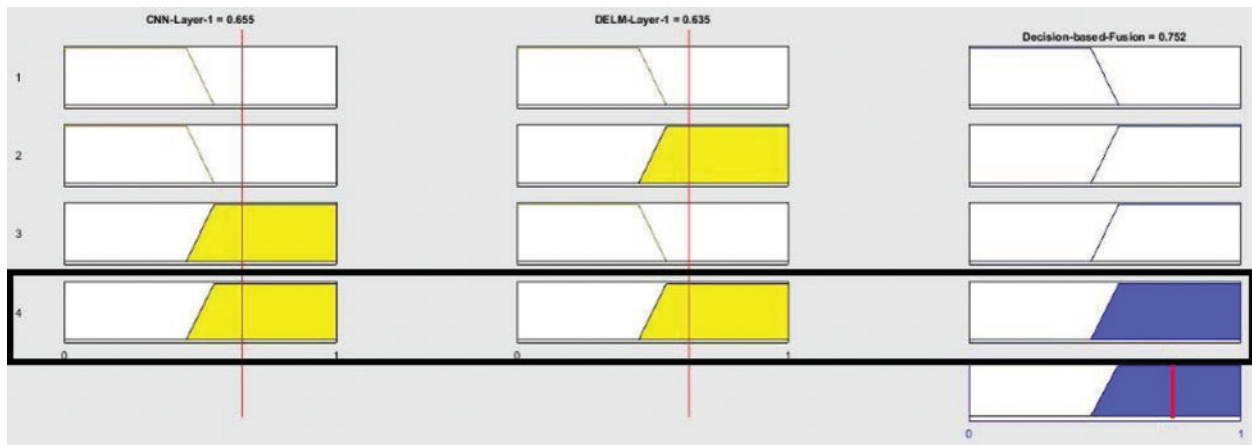


Figure 5: Proposed CF-BCP lookup diagram for decision-based fusion

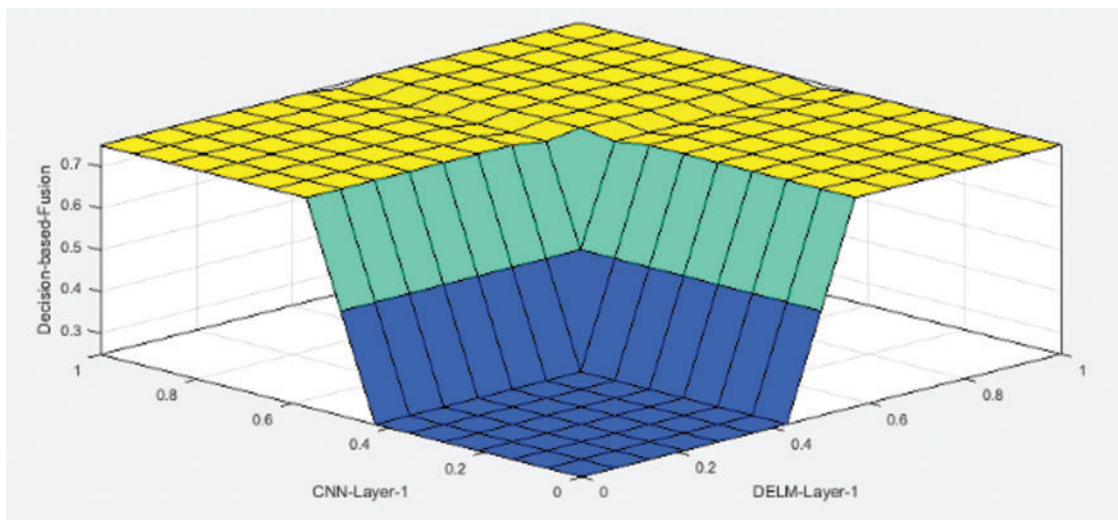


Figure 6: Proposed CF-BCP model rule surface for decision-based fusion

Fig. 6 displays the output of the detection for breast cancer and its types, consisting of CNN Layer 1 and DELM Layer 1. If CNN layer 1 is 0 to 0.45 and DELM Layer 1 is 0 to 0.45, then

there is no chance of breast cancer, which thus presents as benign. If CNN layer 1 is benign and DELM layer 1 is malignant, then the breast cancer is diagnosed as malignant, and in all other conditions, malignancy is detected.

4 Proposed CF-BCP Results

The proposed cloud and decision-based fusion for an intelligent breast cancer prediction system using a hierarchical DL (CF-BCP) model was developed for the earliest prediction of breast cancer and its severity. Using MATLAB (2019a) simulations, the results are obtained for detection. The proposed CF-BCP model consists of two DL approaches: The CNN and DELM. In Layer 1, the DL CNN and DELM approaches were used on 7909 and 569 fused samples, respectively. For both approaches, 80% of the fused samples were used for training purposes, and 20% were used for validation. The accuracy and miss rates of the proposed CF-BCP model were compared with the other existing state-of-the-art techniques:

$$\text{Accuracy} = \frac{\frac{\alpha_{Ti}}{F_{Ti}} + \frac{\alpha_{Tk}}{F_{Tk}}}{\frac{\alpha_{Ti}}{F_{Ti}} + \frac{\sum_{j=1}^z (\alpha_{Tj, j \neq i})}{F_{Tj}} + \frac{\alpha_{Tk}}{F_{Tk}} + \frac{\sum_{l=1}^n (\alpha_{Tl, l \neq k})}{F_{Tk}}}, \quad \text{where } i/j/k/l = 1, 2, 3, \dots, z; \tag{21}$$

$$\text{Miss Rate} = \frac{\frac{\sum_{l=1}^z (\alpha_{Tl, l \neq k})}{F_{Tk}}}{\frac{\sum_{l=1}^z (\alpha_{Tl, l \neq k})}{F_{Tk}} + \frac{\alpha_{Ti}}{F_{Ti}}}, \quad \text{where } i/k/l = 1, 2, 3, \dots, z; \tag{22}$$

$$\text{Average Percentage} = \frac{(\text{Percentage of CNNLayer} - 1) + (\text{Percentage of DELMLayer} - 1)}{(\text{sample size 1} + \text{sample size 2})} \times 100. \tag{23}$$

The proposed CF-BCP model diagnoses breast cancer as benign or malignant, where benign represents no breast cancer, and malignant represents breast cancer.

Tab. 1 represents the detection of the proposed CF-BCP model for training and used 80% fused samples. In Layer 1, the DL CNN model attained 97.81% accuracy and a 2.91% miss rate. Moreover, DELM Layer 1 obtained 98.23% accuracy and a 1.77% miss rate. The average performance of the proposed CF-BCP Layer 1 model achieved 97.84% accuracy and a 2.16% miss rate.

Tab. 2 represents the detection of the proposed CF-BCP model for validation and used 20% fused samples. In Layer 1, the DL CNN model attained 97.24% accuracy and a 2.76% miss rate. Further, DELM Layer 1 obtained 97.18% accuracy and a 2.82% miss rate. The average performance of the proposed CF-BCP Layer 1 model achieved 97.41% accuracy and a 2.59% miss rate.

Tab. 3 lists the proposed CF-BCP Layer 2 for predicting breast cancer types during the training phase. In total, 4344 fused samples (80%) were used during the training phase and were further divided into 2761, 501, 634, and 448 fused samples of malignant T1, T2, T3, and T4, respectively. In malignant T1, a total of 2761 fused samples were taken, in which 2747 samples were predicted correctly as a malignant T1, and 14 fused samples were predicted incorrectly. In malignant T2, 501 fused samples were taken, in which 485 samples were predicted correctly as a malignant T2, and 16 fused samples were wrongly predicted. For malignant T3, 634 fused samples were taken, in which 609 fused samples were validly predicted as a malignant T3, and 25 samples were invalidly predicted. For malignant T4, 448 fused samples were taken, in which 432 samples

were validly predicted as malignant T4, and 16 samples were wrongly predicted by the proposed CF-BCP model.

Table 1: Layer 1 performance for the proposed CF-BCP (training)

(80% fused samples for training)		
Approaches	Accuracy (%)	Miss rate (%)
CNN <i>Layer 1</i>	97.81	2.19
DELM <i>Layer 1</i>	98.23	1.77
Average	97.84	2.16
performance proposed CF-BCP Layer 1		

Table 2: Layer 1 performance for proposed CF-BCP (validation)

(80% of fused samples for validation)		
Approaches	Accuracy (%)	Miss rate (%)
CNN <i>Layer 1</i>	97.24	2.76
DELM <i>Layer 1</i>	97.18	2.82
Average	97.41	2.59
performance proposed CF-BCP Layer 1		

Table 3: Layer 2 Decision matrix for the proposed CF-BCP (training)

(80% fused samples for training)				
$F_{\text{malignant}} = 4344$	$\alpha_{\text{malignant,T1}}$	$\alpha_{\text{malignant,T2}}$	$\alpha_{\text{malignant,T3}}$	$\alpha_{\text{malignant,T4}}$
$F_{\text{malignant,T1}} = 2761$	2747	14	0	0
$F_{\text{malignant,T2}} = 501$	12	485	4	0
$F_{\text{malignant,T3}} = 634$	0	13	609	12
$F_{\text{malignant,T4}} = 448$	0	0	16	432

Tab. 4 presents the proposed CF-BCP Layer 2 for the prediction of breast cancer types during the validation phase. In total, 1085 fused samples (20%) were used during the validation phase and were further divided into 690, 125, 158, and 112 fused samples of malignant T1, T2, T3, and T4, respectively. In malignant T1, 690 fused samples were taken, in which 686 samples were predicted correctly as malignant T1, and 4 fused samples were wrongly predicted. In malignant T2, 125 fused samples were taken, in which 121 samples were predicted correctly as malignant T2, and 4 fused samples were wrongly predicted. For malignant T3, 158 fused samples were taken, in which

149 fused samples were validly predicted as malignant T3, and 9 samples were invalidly predicted. For malignant T4, 112 fused samples were taken, in which 107 samples were validly predicted as malignant T4, and 5 samples were wrongly predicted by the proposed CF-BCP model.

Table 4: Layer 2 decision matrix for the proposed CF-BCP (validation)

(80% fused samples for validation)

$F_{\text{malignant}} = 1085$	$\alpha_{\text{malignant,T1}}$	$\alpha_{\text{malignant,T2}}$	$\alpha_{\text{malignant,T3}}$	$\alpha_{\text{malignant,T4}}$
$F_{\text{malignant,T1}} = 690$	686	4	0	0
$F_{\text{malignant,T2}} = 125$	3	121	1	0
$F_{\text{malignant,T3}} = 158$	0	5	149	4
$F_{\text{malignant,T4}} = 112$	0	0	5	107

Tab. 5 lists the overall performance of the proposed CF-BCP model for the training and validation phases. The proposed CF-BCP model achieved 98.37% overall accuracy and a 1.63% miss rate in the training phase. For the validation phase, the proposed CF-BCP model obtained 97.97% overall accuracy and a 2.03% miss rate.

Table 5: Layer 2 overall performance for the proposed CF-BCP

Overall Performance	Accuracy (%)	Miss rate (%)
Proposed CF-BCP training	98.37	1.63
Proposed CF-BCP validation	97.97	2.03

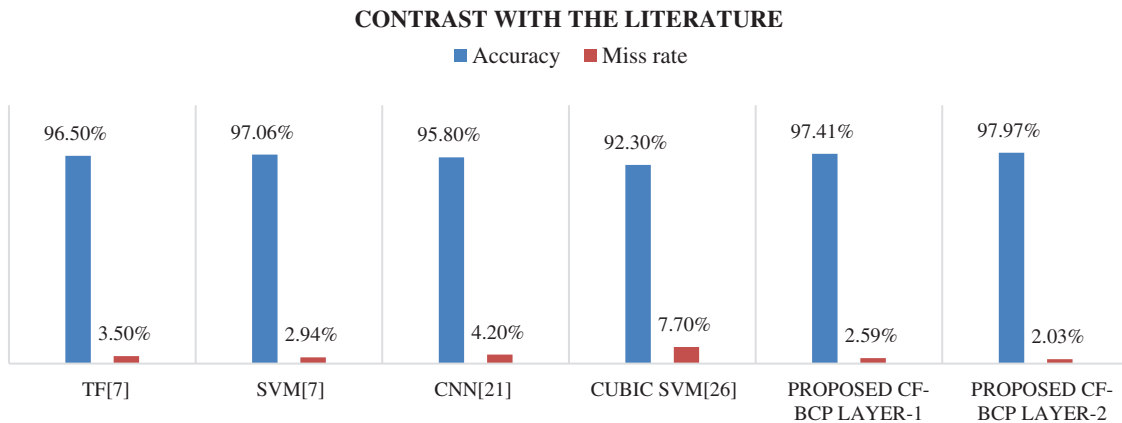


Figure 7: Accuracy chart contrasted with state-of-the-art approaches for the proposed CF-BCP model

Fig. 7 represents a contrast between state-of-the-art methods and the proposed CF-BCP model. The proposed CF-BCP Layer 1 model achieved 97.41% accuracy for the detection of

breast cancer, which is better than the existing approaches. The proposed CF-BCP Layer 2 model also detects breast cancer types, which achieves 97.97% accuracy for the validation phase.

5 Conclusion

Rapidly spreading breast cancer has widely affected women's lives. Reliable and early detection leads to a reduction in the breast cancer death ratio. Moreover, CAD systems are highly assistive for medical practitioners in diagnosing breast tumors. Therefore, researchers have focused on early detection and proper treatment to increase the chances of survival. The significant contribution of the current study is that it presents a novel detection model consisting of cloud and decision-based fusion for an intelligent breast cancer prediction system using a hierarchical DL approach to diagnose breast cancer. Another important contribution of the proposed CF-BCP model is that it has great potential to diagnose different types of breast cancer. The CF-BCP model accomplishes an accuracy of 97.41% for multimodal medical imaging fusion in detecting breast cancer phases and 97.97% accuracy in detecting breast cancer types after decision-based fusion empowered with fuzzy logic. The results of the proposed CF-BCP model are compared with the state-of-the-art approaches, which demonstrates that the proposed model could greatly increase the efficiency and productivity of medical practitioners. The suggested model is generalizable for new datasets because of its flexible and extendable nature.

Acknowledgement: Thanks to our families & colleagues, who supported us morally.

Funding Statement: This work is supported by the KIAS (Research No. CG076601) and in part by Sejong University Faculty Research Fund.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

References

- [1] Z. Wang, M. Li, H. Wang, H. Jiang, Y. Yao *et al.*, "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features," *IEEE Access*, vol. 7, no. 7, pp. 105146–105158, 2019.
- [2] Y. Wang, L. Sun, K. Ma and J. Fang, "Breast cancer microscope image classification based on cnn with image deformation," in *Int. Conf. Image Analysis and Recognition*, Póvoa de Varzim, Portugal, pp. 845–852, 2018.
- [3] Y. Jiang, L. Chen, H. Zhang and X. Xiao, "Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module," *PLOS ONE*, vol. 14, no. 3, pp. 1–21, 2019.
- [4] I. Fondón, A. Sarmiento, A. I. García, M. Silvestre, C. Eloy *et al.*, "Automatic classification of tissue malignancy for breast carcinoma diagnosis," *Computers in Biology and Medicine*, vol. 96, pp. 41–51, 2018.
- [5] J. S. Whang, S. R. Baker, R. Patel, L. Luk and A. Castro, "The causes of medical malpractice suits against radiologists in the United States," *Radiology*, vol. 266, no. 2, pp. 548–554, 2013.
- [6] M. Moghbel and S. Mashohor, "A review of computer assisted detection/diagnosis (cad) in breast thermography for breast cancer detection," *Artificial Intelligence Review*, vol. 39, no. 4, pp. 305–313, 2013.
- [7] F. Khan, M. A. Khan, S. Abbas, A. Athar, S. Y. Siddiqui *et al.*, "Cloud-based breast cancer prediction empowered with soft computing approaches," *Journal of Healthcare Engineering*, vol. 2020, pp. 1–11, 2020.

- [8] W. X. Ju and X. M. Zhao, "Cirrhosis recognition by deep learning model google NET-PNN," *Computer Engineering and Applications*, vol. 55, no. 5, pp. 112–117, 2019.
- [9] W. L. Zhang, R. J. Li and H. T. Deng, "Deep convolutional neural networks for multi-modality iso-intense infant brain image segmentation," *IEEE International Symposium Biomedical Imaging*, vol. 108, pp. 1342–1345, 2015.
- [10] J. Tan, Y. Huo, Z. Liang and L. Li, "Expert knowledge-infused deep learning for automatic lung nodule detection," *Journal of X-ray Science and Technology*, vol. 27, no. 1, pp. 17–35, 2019.
- [11] C. Tao, K. Chen and L. Han, "New one-step model of breast tumor locating based on deep learning," *Journal of X-ray Science and Technology*, vol. 27, no. 5, pp. 839–856, 2019.
- [12] R. A. Hubbard, K. Kerlikowske, C. I. Flowers, B. C. Yankaskas, W. Zhu *et al.*, "Cumulative probability of false-positive recall or biopsy recommendation after 10 years of screening mammography: A cohort study," *Annals of Internal Medicine*, vol. 55, no. 155, pp. 481–492, 2011.
- [13] A. Hamidinekoo, E. Denton, A. Rampun, K. Honnor and R. Zwigelaar, "Deep learning in mammography and breast histology, an overview and future trends," *Medical Image Analysis*, vol. 47, pp. 45–67, 2018.
- [14] T. Araújo, G. Aresta and E. Castro, "Classification of breast cancer histology images using convolutional neural networks," *PLoS One*, vol. 12, no. 6, pp. 10–18, 2017.
- [15] H. Yao, X. Zhang, X. Zhou and S. Liu, "Parallel structure deep neural network using cnn and rnn with an attention mechanism for breast cancer histology image classification," *Cancers*, vol. 11, no. 12, pp. 1–14, 2019.
- [16] Y. Wang, L. Sun, K. Ma and J. Fang, "Breast cancer microscope image classification based on cnn with image deformation," in *Int. Conf. Image Analysis and Recognition*, Canada, pp. 845–852, 2018.
- [17] Z. Wang, M. Li, H. Wang, H. Jiang, Y. Yao *et al.*, "Breast cancer detection using extreme learning machine based on feature fusion with CNN deep features," *IEEE Access*, vol. 7, pp. 105146–105158, 2019.
- [18] T. C. Chiang, Y. S. Huang, R. T. Chen, C. S. Huang and R. F. Chang, "Tumor detection in automated breast ultrasound using 3-D CNN and prioritized candidate aggregation," *IEEE Transactions on Medical Imaging*, vol. 38, no. 1, pp. 240–249, 2018.
- [19] S. A. Agnes, J. Anitha, S. I. A. Pandian and J. D. Peter, "Classification of mammogram images using multiscale all convolutional neural network," *Journal of Medical Systems*, vol. 44, no. 1, pp. 1–13, 2020.
- [20] L. Shen, L. R. Margolies, J. H. Rothstein, E. Fluder, R. M. Bride *et al.*, "Deep learning to improve breast cancer detection on screening mammography," *Scientific Reports*, vol. 9, no. 1, pp. 1–12, 2019.
- [21] Y. Zhou, J. Xu, Q. Liu, C. Li, Z. Liu *et al.*, "A radiomics approach with cnn for shear-wave elastography breast tumor classification," *IEEE Transactions on Biomedical Engineering*, vol. 65, no. 9, pp. 1935–1942, 2018.
- [22] H. Qiyuan, H. M. Whitney and M. L. Giger, "A deep learning methodology for improved breast cancer diagnosis using multiparametric MRI," *Scientific Reports*, vol. 10, no. 1, pp. 1–17, 2020.
- [23] E. Chaves, C. B. Gonçalves, M. K. Albertini, S. Lee, G. Jeon *et al.*, "Evaluation of transfer learning of pre-trained CNNs applied to breast cancer detection on infrared images," *Applied Optics*, vol. 59, no. 17, pp. 1–7, 2020.
- [24] K. George, P. Sankaran and K. P. Joseph, "Computer assisted recognition of breast cancer in biopsy images via fusion of nucleus-guided deep convolutional features," *Computer Methods and Programs in Biomedicine*, vol. 2020, pp. 105531–105532, 2020.
- [25] V. J. Kadam, S. M. Jadhav and K. Vijayakumar, "Breast cancer diagnosis using feature ensemble learning based on stacked sparse autoencoders and softmax regression," *Journal of Medical Systems*, vol. 43, no. 8, pp. 263–264, 2019.
- [26] S. Singh and R. Kumar, "Histopathological image analysis for breast cancer detection using cubic svm," in *7th IEEE Int. Conf. on Signal Processing and Integrated Networks*, Noida, India, pp. 498–503, 2020.