ARTICLE

Data Fusion Architecture Empowered with Deep Learning for Breast Cancer Classification

Sahar Arooj[1], Muhammad Farhan Khan[2], Tariq Shahzad[3], Muhammad Adnan Khan[4,5,6], Muhammad Umar Nasir[7], Muhammad Zubair[1], Atta-ur-Rahman[8] and Khmaies Ouahada[3,*]

[1]Faculty of Computing, Riphah International University, Islamabad, 45000, Pakistan

[2]Department of Forensic Sciences, University of Health Sciences, Lahore, 54000, Pakistan

[3]Department of Electrical and Electronic Engineering Science, University of Johannesburg, P.O. Box 524, Johannesburg, 2006, South Africa

[4]School of Computing, Skyline University College, University City Sharjah, Sharjah, 1797, United Arab Emirates

[5]Riphah School of Computing and Innovation, Faculty of Computing, Riphah International University, Lahore Campus, Lahore, 54000, Pakistan

[6]Department of Software, Faculty of Artificial Intelligence & Software, Gachon University, Seongnam, Gyeonggido, 13120, Korea

[7]Department of Computer Science, Bahria University, Lahore Campus, Lahore, 54000, Pakistan

[8]Department of Computer Science, College of Computer Science and Information Technology (CCSIT), Imam Abdulrahman Bin Faisal University (IAU), P.O. Box 1982, Dammam, 31441, Saudi Arabia

*Corresponding Author: Khmaies Ouahada. Email: kouahada@uj.ac.za

ABSTRACT

Breast cancer (BC) is the most widespread tumor in females worldwide and is a severe public health issue. BC is the leading reason of death affecting females between the ages of 20 to 59 around the world. Early detection and therapy can help women receive effective treatment and, as a result, decrease the rate of breast cancer disease. The cancer tumor develops when cells grow improperly and attack the healthy tissue in the human body. Tumors are classified as benign or malignant, and the absence of cancer in the breast is considered normal. Deep learning, machine learning, and transfer learning models are applied to detect and identify cancerous tissue like BC. This research assists in the identification and classification of BC. We implemented the pre-trained model AlexNet and proposed model Breast cancer identification and classification (BCIC), which are machine learning-based models, by evaluating them in the form of comparative research. We used 3 datasets, A, B, and C. We fuzzed these datasets and got 2 datasets, A2C and B3C. Dataset A2C is the fusion of A, B, and C with 2 classes categorized as benign and malignant. Dataset B3C is the fusion of datasets A, B, and C with 3 classes classified as benign, malignant, and normal. We used customized AlexNet according to our datasets and BCIC in our proposed model. We achieved an accuracy of 86.5% on Dataset B3C and 76.8% on Dataset A2C by using AlexNet, and we achieved the optimum accuracy of 94.5% on Dataset B3C and 94.9% on Dataset A2C by using proposed model BCIC at 40 epochs with 0.00008 learning rate. We proposed fuzzed dataset model using transfer learning. We fuzzed three datasets to get more accurate results and the proposed model achieved the highest prediction accuracy using fuzzed dataset transfer learning technique.

1 Introduction

In today's world, cancer is a substantial public health problem. The abnormal development of cells in the breast causes cancer tumors. Breast cancer (BC) can be detected by irregularities such as lumps, micro-calcifications, and areas of asymmetry and distortion in the breast. The use of medical imaging in treating cancer is essential. BC is a significant disease among women worldwide, and it creates a severe public health threat on a worldwide scale. When it comes to decreasing the rate of death, accurate identification and assessment of BC in its initial phases are essential.

BC develops when breast cells divide and grow out of control, forming a tissue mass called a tumor. BC signs include a breast lump, irregularities in the skin of the breasts, and a change in the breast size. BC can occur anywhere in the breast, although the upper outer region is the most common. There are two types of breast cancer: benign and malignant. Benign tumors grow slowly and do not spread. Malignant tumors can spread throughout the body, attack and destroy surrounding normal tissues, and develop quickly. Automatically identifying and locating cancer cells in BC scans is a big challenge because cancerous cells come in different sizes, appearances, and positions.

Clinicians are already using computer-assisted detection and diagnosis (CAD) systems to aid decision-making. In clinical practice, such techniques can dramatically reduce the time and effort required to evaluate cancer. CAD systems aid in the detection and classification of tumors. Breast cancer can be detected and diagnosed using imaging techniques such as diagnostic magnetic resonance imaging (MRI), mammography (x-rays), thermography, and ultrasound (sonography).

Deep learning (DL) is a significant technological development since it has outperformed the condition in various machine learning (ML) tasks [1], such as object detection and categorization. DL approaches adaptively train the proper extraction of feature procedure from the data input concerning the wanted output. Traditional ML methods require a hand-crafted extracting features phase, which is problematic because it depends on a knowledge base. Various publications utilizing DL have been published since its beginning. The convolutional neural network (CNN) is the most common DL framework network [2].

Artificial intelligence, especially DL algorithms and transfer learning (TL), is getting a lot of attention due to its outstanding performance in image identification and classification. TL is a technique of ML that allows you to reuse a model you built for one job for another. It is a typical DL strategy that employs pre-trained models. It is particularly popular in DL because it can train deep neural networks with small data. A pre-trained CNN, such as the ResNet50, VGG16, VGG19, or InceptionV2 ResNet, can be used to classify and identify BC. We will use Customized AlexNet and BCIC in this work. AlexNet [3] was the 1st CNN to achieve performance in the task of object recognition and classification. AlexNet has a significant effect on ML, particularly in DL. AlexNet is a robust model capable of achieving accuracy even with the most challenging data. BCIC is the CNN model proposed in this work. We also achieved the optimum results through BCIC.

We used three datasets, A, B, and C. We fuzzed these datasets and got 2 datasets from Dataset A with 2 classes categorized as benign and malignant (Dataset A2C) and Dataset B with 3 types

categorized as benign, malignant, and normal (Dataset B3C). Rare work has been done on the fusion of BC datasets. We are the earliest to fuse three datasets and compare their results on customized AlexNet and proposed model BCIC. In our research, we used ultrasound and histopathological images.

The main contributions of this work are data fusion on 3 datasets containing different types of images to make 2 merged datasets of them. We have also proposed a CNN-based Model BCIC and compared the performance of our 2 fuzzed datasets among Customized AlexNet and BCIC. The remaining part of the paper is written out as follows. The literature study is given in Section 2, our proposed system model is explained in Section 3, results and simulation are discussed in Section 4, and future work is concluded in Section 5.

2 Related Works

Using the Mammographic Image Analysis Society (MIAS) mammography dataset, this work [4] offered an ML-based 2-level top to bottom categorized strategy for BC classification and detection into 3 classes: benign, malignant, and normal. Before applying feature extraction methods and ML techniques for classification, many data pretreatment techniques are used. Gray level co-occurrence matrix (GLCM) as a method of feature extraction and as a random forest classifier is used in the first stage of classification to distinguish between abnormal and normal cases, followed by Local binary patterns (LBP) as a feature extracting method and as a random forest classifier in the 2nd stage of classification to classify abnormality of the patient into malignant or benign. The 1st stage classification accuracy is 97%, with F1 scores of 97% and 98% for abnormal and normal, respectively. The classification accuracy for the second stage is 75%, with an F1 score of 74% and 76% for malignant and benign classifications, respectively. The whole hierarchy classification method achieves an accuracy of classification of 85%, a Matthews's correlation coefficient (MCC) of 76%, and F1 scores of 98%, 70%, and 74%, accordingly, for normal, benign, and malignant tests.

In this paper [5], two ML algorithms for automated magnification of multi-categorization on a composed break his dataset for BC diagnosis are comprehensively investigated and compared. The first method relies on custom characteristics extracted with the Hu moment, histogram analysis, and Haralick textures. Extracted features are then used to train traditional classifiers, whereas the second method uses TL to use systems before VGG19, VGG16, and ResNet50 as feature extractors and baseline models. The results show that using pre-trained networks as a feature extractor outperformed the baseline and handmade approaches for all magnifications. Furthermore, augmentation has been found to have a critical role in improving classification accuracy. For the categorization of histopathological images, the linear support vector machine (SVM) and VGG16 network deliver the greatest accurateness, which is computed in 2 ways: (a) accuracies based on patch (93.97% for 40, 92.92% for 100, 91.23% for 200, and 91.79% for 400); (b) customer accuracies (93.25% for 40, 91.87% for 100, 91.5% for 200) Furthermore, the (malignant) and (benign) classes were discovered to become the most composite throughout the entire enlargement factor.

They created an ML method with restricted structures to obtain good accuracy in the classification of tumors in this study [6]. They looked at 569 females diagnosed with cancers (212 malignant and 357 benign). Three supervised ML algorithms, SVM, linear regression (LR), and K nearest Neighbour (KNN) were used to create the model. Each model was tested using a 10-fold cross-validation procedure, and evaluation metrics were developed to assess the model's results. LR and SVM models produced 97.66% accuracy with the entire feature evaluation. The SVM performance was enhanced by 98.25% using chosen features. On the other hand, the LR model with limited features

provided 100% correct positive predictions. Selective features in the presented models could enhance the accuracy of a BC diagnostic prediction. This paper proposes two primary deep TL-based models that beat current framework systems in binary classification and multiple classifications by relying on pre-trained Deep CNN and a huge number of ImageNet images. They use pre-trained DesneNet121 and ResNet50 values as initial values in ImageNet, then use a deep learner with data preprocessing to fine-tune these models to distinguish various malignant and benign sample tissues in binary & multiple classifications. The suggested models were tested with optimized hyperparameters in magnification-independent and dependent categorization modes. The proposed method attained a 98% accuracy in multiclass classification. Regarding binary classification, the proposed method achieves an accuracy of up to 100% [7].

In this study [8], a new type of computer-aided detection (CAD) system for breast tumors was proposed for differentiating between tumor and non-tumor by using an MRI dataset. They used the VGG16, Inception V3, and ResNet50 networks to build their CAD system. They have a 91.25% accuracy rate for CNN represented by VGG16, which is higher than that of conventional machine learning models. They come to the conclusion that ResNet50 and Inception V3 is inferior to the VGG16-based breast tumor CAD system in terms of performance.

Breast cancer can be diagnosed using a variety of techniques, including computer-aided diagnosis (CAD). After evaluating numerous recent studies, they [9] concluded a thorough analysis of the convolutional neural network (CNN)-based diagnosis of breast cancer. They introduced a variety of imaging techniques. The second section contains the CNN structure. They also presented various open-access data sets on breast cancer. Then they separate the jobs of classifying, detecting, and segmenting the diagnosis of breast cancer. Even though CNN's diagnostic has had considerable success, they claim that it still has some limits, such as the scarcity of high-quality data sets. The CNN-based model requires a lot of computation and time to perform the diagnosis when the data set is too big.

In medical image analysis, CNN [10] is now the most well-known type of DL architecture. DL is used in clinical decision-making to detect and classify diseases based on medical pictures. This paper [11] provides a new training strategy for developing illness detection and classification DL models with small datasets. Their method is based on a hierarchical classification method in which the healthy/sick information from the first model is efficiently used to develop subsequent models for classifying the disease into subtypes using a TL method. Multiple input datasets were used to improve accuracy, and the final classification was done using a stacking ensemble approach. A classified dataset for 798 glaucoma eyes and 156 normal was used to show the method's performance, with glaucoma eyes being further categorized into 4 subtypes. Their weighted average accuracy was 80% and 83%.

The research [12] has provided the DeTraC technique based on class decay. This CNN architecture uses TL and a class decomposition approach to improve the performance of medical picture categorization. DeTraC allows for separated learning at the subclass level, with the potential for fast computation. They used three separate groups of histological images, chest X-ray images, and digital mammograms to confirm their suggested method. They compared DeTraC against cutting-edge CNN models to show that it outperforms them in terms of specificity, sensitivity, and accuracy. They used a variety of TL models. Their maximum accuracy with AlexNet is 95.7%.

ML-based BC classification and detection algorithms were provided by Hamed et al. [13], who evaluated them in the manner of comparative studies. According to their comparative investigation, the most current top accuracy models based on basic detection and classification architectures are You only look once (YOLO) and Retina Net. According to them, the YOLO model is adjustable. Still,

it has specific constraints on how close things can be. It cannot attain excellent accuracy with little objects and makes a few localization errors. As a result, the Retina Net is introduced to fill in for the imbalances and irregularities of single-shot different classifiers quickly and efficiently like YOLO.

This paper [14] used a method of 10-fold cross-validation to compare the accuracy of classification and confusion matrix of the diverse classifiers Multi-layer perception (MLP), decision tree (J48), Sequential minimal optimization (SMO), Naive bayes (NB), and Instance-Based for K-nearest neighbor (IBK) on 3 datasets of breast cancer Wisconsin diagnosis breast cancer (WDBC), Wisconsin breast cancer (WBC), and Wisconsin prognosis breast cancer (WPBC). They also combine these classifiers at the categorization level to obtain the best multi-classifier strategy for all datasets. On the fusion of all classifiers SMO+J48+NB+IBK, their highest accuracy is 97%. All the trials were carried out using the WEKA data mining tool.

This research [15] established a complete information model to describe a patient's whole clinical history. DL methods are also used to extract ideas and their attributes from clinical breast cancer records by good pre-trained Bidirectional Encoder Representations from Transformers (BERT) linguistic methods. They used a DL approach to fine-tune BERT's method of obtaining breast cancer categories and features. The clinical corpus for their work came from a cancer hospital, 3A in China. It included 100 breast cancer women's operation records, encounter notes, radiology notes, pathology notes, discharge summaries, and progress notes. Their system obtained 93.53% for the named entity recognition (NER) and 96.73% for relation extraction on the F1 score. They also demonstrated that the support vector machines (SVM) and Bi-LSTM algorithms dominated recognition.

The usage of TL-AlexNet was examined in this paper [16]. BUSI and Dataset B were the two datasets. They were also using fusion on datasets and obtained datasets (B+BUSI). With the help of generative adversarial network (GAN), they developed a revolutionary data augmentation process. When combined with DAGAN and classical augmentation, it attained an accuracy of 84% on Dataset (BUSI+B). Because DL approaches are based on ML, and customized models are built for each dataset, they may be adapted to the particular properties of any dataset. Finally, models trained using the suggested GAN augmentation methodology outperform standard models by a significant margin.

The proposed [17] method for extracting and pooling low- to mid-level features using a VGG19 pre-trained CNN and fusing them with handmade radionics data computed with traditional CAD methods. They analyzed three clinical datasets: FFDM, ultrasound, and MRI, all of which were classified as benign or malignant based on pathology or radiology reports. Their extreme achieved accuracy is 81% on the FFDM dataset, 87% on the ultrasound dataset, and 87% on the MRI dataset.

Their framework [18] of CAD of breast cancer CNN is used. On 2 datasets of mammographic containing ROIs displaying malignant or benign mass cancer, state-of-the-art CNNs are trained and evaluated. 2 training scenarios are used to measure the performance of each researched network: the first includes initializing the network with pre-trained weights, while the other involves random initialization of the networks. Extensive experimental results suggest that fine-tuning a pre-trained network outperforms training from scratch in terms of performance. Their maximum achieved accuracy on fine-tuning is 75.3% of AlexNet and 75.5% of ResNet-152.

The suggested approach [19] included benign and malignant breast cancer forms. Identification and categorization of BC using ultrasound images and histopathology are crucial in computer-aided diagnosis systems. Early detection of breast cancer increases the likelihood of women receiving the appropriate treatment and surviving. Many medical problems are solved using DL, TL, and ML approaches. The absence of a dataset hinders their research. The suggested methodology was developed to aid in the rapid detection and diagnosis of breast cancer. Their primary contribution

is their model's use of the TL approach on the three datasets, A, B, C, and A2. Histopathology and ultrasound imaging are both employed in this investigation. The CNN-AlexNet model employed in this study is a customized version trained using the dataset's specifications. Another contribution of this work is this. Results indicate that on Datasets A, B, C, and A2, the proposed system equipped with TL outperformed existing models in terms of accuracy. The maximum accuracy for Dataset A is 99.4%, whereas the maximum accuracy for Dataset B is 96.70%, for Dataset C, it is 99.10%, and for Dataset A2, it is 100%.

In their work [20], a brand-new Hybrid neuro-fuzzy classifier (HNFC) method is put forth to increase the classification accuracy of input data. Ten benchmark datasets are utilized for this unique technique. The mean value is used to fill in the gaps in the data during the preprocessing stage. The key features from the dataset are then chosen using statistical correlation. The values were normalized after utilizing a data transformation technique. Fuzzy logic was first used to analyze the input dataset before the neural network was used to assess performance. The suggested method's output is estimated with supervised classification methods like Adaptive neuro-fuzzy inference systems (ANFIS) and Radial basis function neural networks (RBFNN). Measures like accuracy and error rate are used to assess classifier performance. According to the analysis, the proposed strategy outperformed the other two approaches by a margin of 86.2% in terms of classification accuracy for the breast cancer dataset.

Some advantages and disadvantages of using DL architecture we discussed here. Advantages: DL architectures can automatically detect and learn features from datasets which means DL cannot require hand-engineered features. DL architectures have the capability to handle complex datasets also it has the capability to handle both structured and unstructured datasets. DL architectures can handle missing data and features and it can still make predictions. Disadvantages: DL architectures have high computational costs and require a large number of GPUs. Those DL architectures have complex layers architectures has difficult to interrupt. Difficult to understand how DL models can make predictions.

Table 1 illustrates the previous work comparison regarding the accuracy, types of images, and limits. Earlier work [9,12,14–18,21–23] have limits such as small datasets, handcrafted features required, lack of accuracy, and types of images.

Table 1: Literature of previous studies

Studies	Type of model	Type of cancer	Type of images	Accuracy	Limitations
Ziquan et al. [9]	ResNet50+ DesneNet121	Breast cancer (Benign and Malignant class)	• Histological images	98%, 100%	• Overfitting • The data set is small
Abbas et al. [12]	AlexNet	Chest cancer	• Histological images • Chest X-ray images	95.7%	• Imbalanced dataset • Lack of diverse model use • Digital mammograms.
Salama et al. [14]	Decision tree j48+ MLP+ NB+ SMO+ IBK	Breast Cancer (Benign and Malignant class)	• Ultrasound dataset	95.13%, 95.27%, 95.99%, 96.99%, 94.56%	• Dataset is small • Required hand-crafted features
Zhang et al. [15]	NER SVM	Breast cancer (Benign and Malignant class)	• Clinical information of women	93.53%, 96.73%	• The data set is small • Required handcrafted features

(Continued)

Table 1 (continued)

| Studies	Type of model	Type of cancer	Type of images	Accuracy	Limitations
Dhabyani et al. [16]	AlexNet+ VGG 16+ Inception+ ResNet+Nasnet	Breast cancer (Benign, Malignant, and Normal Class)	• Ultrasound dataset	78%, 80%, 88%. 82%, 85%, 80%, 93%. 90%, 94%, and 92% on datasets BUSI and B, respectively	• Less Accuracy
Antropova et al. [17]	SVM	Breast cancer (Benign and Malignant class)	• Mammography dataset • MRI dataset • Ultrasound dataset	81% on the Mammography dataset, 87% on the MRI dataset, and 87% on the ultrasound dataset	• Less accuracy • Required hand-crafted features
Tsochatzidis et al. [18]	AlexNet ResNet	Breast cancer (Benign and Malignant class)	• Mammography dataset	75.3% 75.5%	• Less accuracy • The data set is small
Gokhan et al. [21]	Deep belief network model	Breast cancer	• Mammography dataset • ROI Images	96.32%	• Limited number of hidden layers
Gokhan et al. [22]	CNN	Breast cancer	• Mammography dataset	92.84%	• Lack of diverse model use
Gokhan et al. [23]	Simple DL architectures	Breast cancer	• Mammography dataset	95.17%	• Lack of diverse model use

Main contributions of our proposed Model:

- We applied data fusion to 3 datasets to make 2 datasets of them.
- We merged 2 different types of images histopathological and ultrasound images.
- We proposed a CNN-based Model BCIC.
- We compared the performance of our 2 fuzzed datasets among Customized AlexNet and BCIC.
- We achieved the optimum results by using the proposed Model BCIC.

3 Materials and Methods

In recent years, various ML and DL algorithms have been commonly used in medical image processing [20] to discover and analyze objects in medical imaging. We used TL and CNN models to detect and categorize breast cancer. We proposed our CNN-based Model BCIC and also used a pre-trained model AlexNet. TL [24] and CNN models are well-known for their usage in medical image processing and disease diagnosis. We applied these 2 models, AlexNet, and BCIC, to our fuzzed datasets, A2C and B3C.

Fig. 1 shows the system model proposed for data fusion for breast cancer detection and classification. Our system contains 2 significant parts: first is data fusion and preprocessing, and second is the training, identification, and classification of BC. We collected these three datasets, A, B, and C, from the Kaggle repository and previously published research [25–30]. After the data collection, we fuzzed these 3 datasets, A, B, and C, and got these 2 fuzzed datasets, A2C and B3C. After that, we preprocessed these 2 datasets, A2C and B3C, because the collected data was in raw form. We converted these data sets into dimensions of $227 * 227$ and split them into 80:20 according to the requirements of models AlexNet and BCIC.
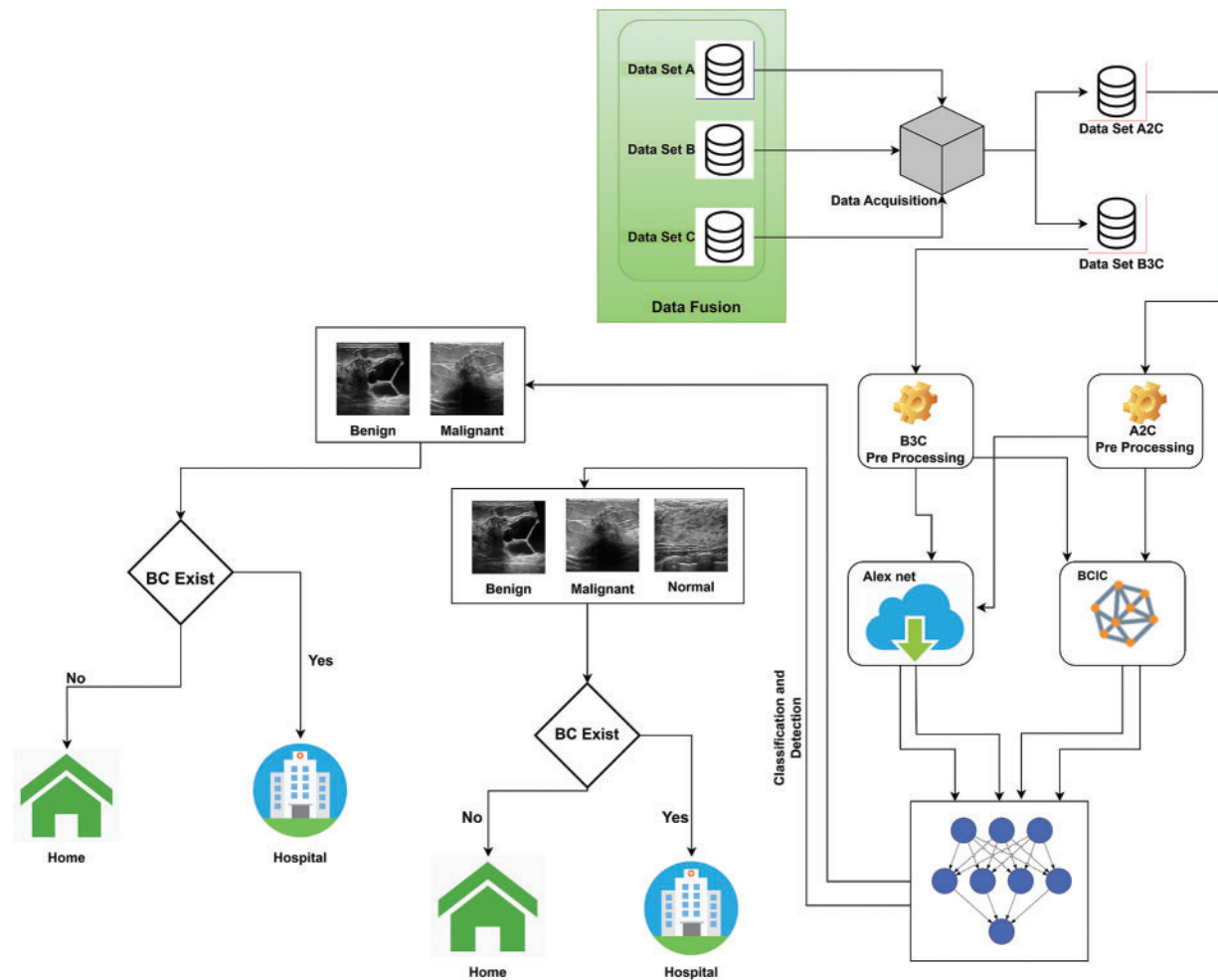
Figure 1: Proposed data fusion empowered with deep learning system model for breast cancer identification

We trained our datasets using AlexNet and BCIC models using preprocessed images in the second part. We trained the models multiple times to get the optimum results. After the training of AlexNet and BCIC, stored these models in the cloud and used these trained models for the detection and classification of Breast cancer. Intellectual models classified Dataset A2C into Benign and Malignant and Dataset A3C Malignant, into Benign and Normal. If the person is healthy, there is no need to go to the hospital; however, if the woman has signs of malignant or benign, she should go to the hospital for tumor therapy.

3.1 Convolutional Neural Network

It is a recognition learning approach, according to DL. CNNs are useful in image processing; this work modifies AlexNet's last three layers according to the 3 classes and sets 3 as input parameters for fully connected layers. The fully connected layer will connect classes of datasets. Training contains multiple options; the most important are epochs, learning rate, and solver. This work is done on 40 epochs, 0.00008 learning rate, and stochastic gradient descent with a momentum (SGDM) solver. This

work also compares results with the proposed model BCIC; that is the reason the same number of epochs and the learning rate is used for both models AlexNet and BCIC.

These modified layers train datasets and models stored in the cloud. This is especially important in identifying and classifying objects, text, medical images, and human tissue. We analyze the performance of DL in detecting and classifying BC. Pooling layers and Convolutional layers make CNNs. CNNs are layered neural networks, and their design determines the accuracy of their layers and training models. Every model has several layers, which lead to structural patterns used in the learning process.

3.2 AlexNet

AlexNet was the primary CNN [11] to achieve performance in object recognition and classification tasks. AlexNet significantly affects ML, particularly in DL and TL. The ultrasound and histological breast pictures in the datasets are grayscale AlexNet was employed to tackle the multi-class classification problem. The images are used as model input containing the dataset's classes. We divided all datasets into 80% and 20% for training and validation, respectively. Breast cancers and normal images were used as input images.

The proposed system used customized AlexNet. This work modifies 1st and last 3 layers of AlexNet. According to the need, the input image layer is changed; we set the dimension $227 * 227$ for used datasets, and the previous 3 layers are adjusted according to the output of datasets. Data sets A2C have 2 classes; we modify AlexNet's last three layers according to the 2 classes and set 2 as input parameters for fully connected layers. Data sets B3C have 3 classes. AlexNet was modified and used for the classification and detection of BC.

3.3 BCIC

CNN is best in processing medical images, especially in the identification and classification; that is why this work proposed the BCIC model based on CNN. BCIC contains 8 layers, including an image input layer, convolution layer, batch normalization layer, RELU layer, max pool layer, fully connected layers of Softmax layer, and classification layer. Fig. 2 shows the layers of the proposed Model BCIC, and Fig. 3 shows the complete architecture of the proposed model BCIC.

3.4 Dataset

Based on ultrasound and histopathological reports, three datasets were used to classify breast cancer. These datasets were known as Datasets A, B, and C in this work. We fuzzed three datasets into Dataset A2C and Dataset B3C. BC was classified as benign, malignant, and normal. Table 2 summarizes BC types and the three datasets total number of images.

3.5 Dataset A

Dataset A contains 780 images categorized as Benign, Malignant, and Normal. 437 Benign, 210 malignant, and 133 normal images are included in this dataset. This dataset contains ultrasound images. Dataset A is collected from [27,29].
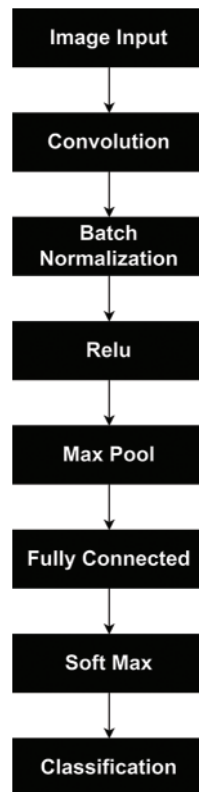
Figure 2: Layers of proposed model BCIC

Layers	Detailed Structure
Image Input	Image Input Size [227*227]
Convolution	Filter Size# 3*3*32 Stride 1,1, Padding Same
Batch Normalization	Epsilon 0.00001
Relu	Same
Max Pool	Pool Size 5,5, Stride 1,1
Fully Connected	Output Size= Num of classes in dataset (2,3)
Soft Max	Same
Soft Max	Classes Output, Output size auto [227*227]

Figure 3: Detailed architecture of proposed model BCIC

Table 2: Summarized Datasets A, B, C, A2C, and B3C

Datasets	Benign	Malignant	Normal	Total No. of images
A	437	210	133	780
B	2479	5304	–	7783
C	547	579	–	1126
A2C	3464	6117	–	9581
B3C	3464	6117	133	9714

3.6 Dataset B

Dataset B contains 7783 images and is categorized as Benign and Malignant. 5304 malignant and 2479 benign images are included in this dataset. This dataset contains histopathological images. Dataset B is collected from [25–29].

3.7 Dataset C

Dataset C contains 1126 images and is categorized as benign and Malignant. 579 malignant and 547 benign images are included in this dataset. This dataset contains histopathological images. Dataset C is collected from [27].

3.8 Dataset A2C

We fuzzed Datasets A, B, and C with 2 classes, benign and malignant, to get this Dataset A2C. This dataset contains 9581 images, 3464 benign, and 6117 malignant images.

3.9 Dataset B3C

We fuzzed datasets A, B, and C with 3 classes malignant, benign, and normal, to get this Dataset B3C. This dataset contains 9714 images, 3464 benign, 6117 malignant, and 133 normal images.

After those preprocessing measures were conducted to facilitate models, this process is necessary to remove the constraints of irregularity detections and set image dimensions [227 ∗ 227} as per the models' requirements. Splitting is a critical element for testing and training models. This work randomly split the dataset into 80% for training and 20% for testing.

4 Simulation and Results

This section contains multiple test results of AlexNet [31] and the proposed Model BCIC on fuzzed Datasets A2C and B3C. Dataset A2C contains binary classes categorized as Malignant and Benign; Dataset B3C contains 3 classes classified as Benign, Malignant, and Normal. Training and testing are done on Matlab version 2020a. This work is done by splitting datasets into 80% and 20%. 80% is for training, and 20% is for testing. Accuracy (ACC), Sensitivity (SEN), Specificity (SPEC), Miss classification rate (MCR), False-negative ratio (FNR), False positive ratio (FPR), Likelihood negative ratio (L −ve R), Likelihood positive ratio (L +ve R), False negative (TN), True negative (TN), False positive (FP), True positive (TP), are the methods of assessment [19,30–31] which are used to evaluate the performance of AlexNet and BCIC. JP shows the output positive, KP shows the input positive,

JN shows the output negative, KN shows the input negative and T shows the total number of images in the dataset.

$$ACC = \frac{J_P/K_P + J_N/K_N}{K_P + K_N} * 100 \tag{1}$$

$$SEN = \frac{J_P/K_P}{J_P/K_P + J_P/K_N} * 100 \tag{2}$$

$$SPEC = \frac{J_N/K_N}{J_N/K_N + J_N/K_P} * 100 \tag{3}$$

$$MCR = \frac{J_N/K_P + J_P/K_N}{T} * 100 \tag{4}$$

$$FPR = \frac{J_N/K_P}{J_N/K_N + J_N/K_P} * 100 \tag{5}$$

$$FNR = \frac{J_P/K_N}{J_P/K_P + J_P/K_N} * 100 \tag{6}$$

$$L + veR = \frac{SEN}{1 - SPEC} \tag{7}$$

$$L + veR = \frac{\dfrac{J_P/K_P}{J_P/K_P + J_P/K_N}}{\dfrac{J_N/K_P}{J_N/K_N + J_N/K_P}} \tag{8}$$

$$L + veR = \frac{(J_P/K_P)(J_N/K_N + J_N/K_P)}{(J_N/K_P)(J_P/K_P + J_P/K_N)} \tag{9}$$

$$L + veR = \frac{TP\,(FN + FP)}{FP(TP + FN)} \tag{10}$$

$$L - veR = \frac{1 - SEN}{SPEC} \tag{11}$$

$$L - veR = \frac{\dfrac{J_P/K_N}{J_P/K_P + J_P/K_N}}{\dfrac{J_N/K_N}{J_N/K_N + J_N/K_P}} \tag{12}$$

$$L - veR = \frac{(J_P/K_N)(J_N/K_N + J_N/K_P)}{(J_N/K_N)(J_P/K_P + J_P/K_N)} \tag{13}$$

$$L - veR = \frac{FN\,(TN + FP)}{TN\,(TP + FN)} \tag{14}$$

This proposed article used three datasets, A, B, and C, and got 2 datasets, A2C and B3C, by applying fusion. This work's results are done using two models, customized AlexNet and BCIC. BCIC is the model proposed in this work. These models were trained multiple times on different parameters epochs, solver, and learning rates. The optimum accuracy of BCIC is 94.9% for Dataset A2C and 94.5% for Dataset B3C, and AlexNet gave an accuracy of 76.8% for Dataset A2C and 86.5% for Dataset B3C on 40 epochs and 0.00008 LR with stochastic gradient descent with momentum (SGDM) solver. Table 3 shows the AlexNet and proposed BCIC model training simulation parameters like number of epochsm learning rate, number of layers, image dimension, pooling method and iterations per epoch. Table 4 shows the AlexNet accuracy and BCIC on 40 epochs and 0.00008 learning rate (LR) with the SGDM solver.

Table 3: AlexNet and proposed BCIC model training simulation parameters

Models	No. of epochs	Learning rate	No. of layers	Image dimension	Pooling method	Iterations
AlexNet	40	0.00008	25	$227 * 227 * 3$	MAX	766 iterations per epoch of Dataset A2C, 766 iterations per epoch of B3C.
BCIC	40	0.00008	8	$227 * 227 * 3$	MAX	

Table 4: Proposed BCIC results comparison with AlexNet results on 40 epochs

Models	Epochs	Learning rate	Solver	Accuracy on Dataset A2C	Accuracy on Dataset B3C
Proposed BCIC model	40	0.00008	SGDM	94.9%	94.5%
AlexNet	40	0.00008	SGDM	76.8%	86.5%

Table 5 shows the accuracy and loss rate of Datasets A2C classes (Benign and Malignant) and the accuracy and loss rate of Datasets B3C classes (Benign, Malignant, and Normal). The accuracy of Dataset A2C classes benign and malignant is 92.8% and 96.1%, and the miss rate is 7.2% and 3.9%, respectively. The accuracy of Dataset B3C classes benign, malignant, and normal is 88.7%, 97.9%, and 92.6%, and the miss rate is 11.3%, 2.1%, and 7.4%, respectively.

Table 5: Accuracy and miss rate class-wise of Datasets A2C and B3C

Datasets	Classes	Miss rate	Accuracy	Epochs	Learning rate
A2C	Benign	7.2%	92.8%	40	0.00008
	Malignant	3.9%	96.1%	40	0.00008
B3C	Benign	11.3%	88.7%	40	0.00008
	Malignant	2.1%	97.9%	40	0.00008
	Normal	7.4%	92.6%	40	0.00008

We split our datasets into 80:20, 80% is for training, and 20% is for testing. Table 6 shows the testing confusion matrix of Dataset A2C. The total number of images of Dataset A2C is 9581. 7665 is

used for training, and 1916 is used for testing. 691 benign and 1225 malignant images were used for testing. The proposed model classified 643 images correctly as benign and 48 wrongly as malignant in the case of benign class and 1175 as malignant, and 50 as benign from 1225 images as malignant. Table 7 shows the testing confusion matrix of Dataset B3C. The total number of images on dataset is 9714. 7771 is used for training, and 1943 is used for testing. During testing, 642, 1275 & 26 images are used for benign, malignant, and normal classes. The proposed model classified 615 images as benign, 25 as malignant, and 2 as normal from 642 images of benign, classified 1197 as malignant, 78 as benign, and 0 as normal from 1197 images of malignant, and classified 25 as normal, 0 as benign and 1 as malignant from 26 images of normal.

Table 6: Confusion matrix of Datasets A2C

Dataset A2C	Benign	Malignant
Benign	643	48
Malignant	50	1175

Table 7: Confusion matrix of Datasets B3C

Dataset B3C	Benign	Malignant	Normal
Benign	615	25	2
Malignant	78	1197	0
Normal	0	1	25

The training accuracy graph of Dataset A2C shows in Fig. 4. This graph shows training done at 40 epochs and a 0.00008 learning rate. Maximum iterations are 30,640 with 766 iterations per epoch. This graph shows training accuracy from epochs 1 to 40 of the BCIC model.
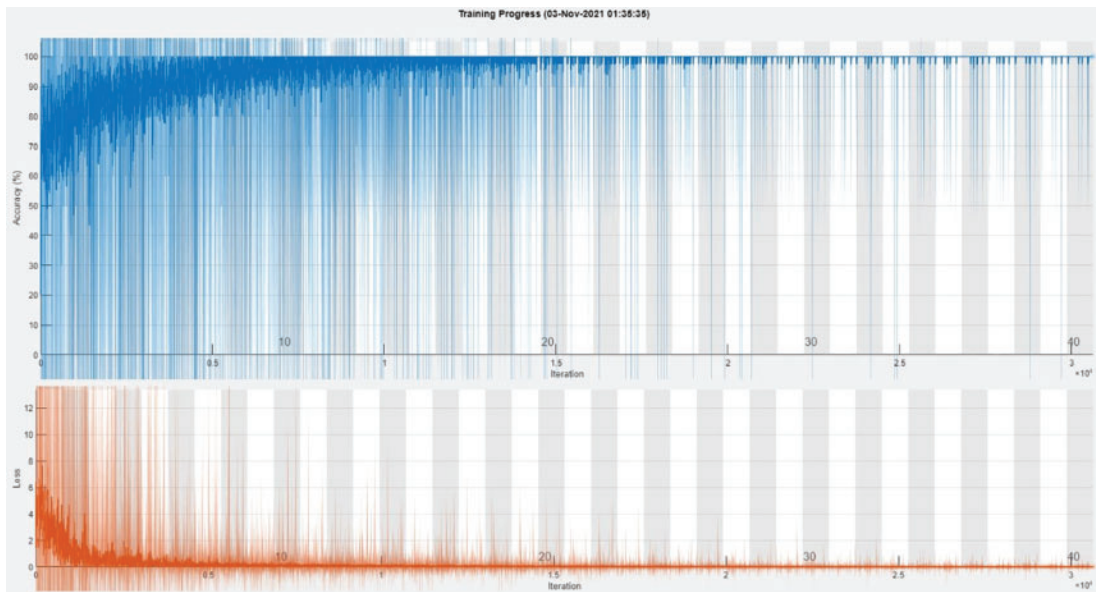
The training accuracy graph of Dataset B3C shows in Fig. 5. This graph shows training done at 40 epochs and a 0.00008 learning rate. Maximum iterations are 31,080 with 777 iterations per epoch. This graph shows training accuracy from epochs 1 to 40 of the BCIC model.

Table 8 shows proposed model gives accurate results on fuzzed Datasets A2C. Dataset A2C is the binary dataset. BCIC gives 94.885% accuracy and a 5.114% miss classification rate.

Table 9 shows proposed model gives accurate results on fuzzed Datasets B3C. Dataset B3C contains 3 classes Benign, Malignant, and Normal. BCIC achieves 94.595% accuracy and 5.404% miss classification rate for the benign class, an accuracy of 94.647% and 5.3525% miss classification rate for the malignant class, and an accuracy of 99.485% and 0.1544% miss classification rate for the normal class.

We examined the accuracy of the proposed method in previous research. Fig. 6 also shows the compared results of BCIC with AlexNet results. In terms of accuracy and miss rate, the proposed Model BCIC performs better than previous research and AlexNet. They [9] used histological images and achieved an accuracy of 100%, which causes overfitting. In this work [12], they achieved 96.7% accuracy with AlexNet, but their dataset is imbalanced. They [14] used a small dataset, but their accuracy is good. In this work [15], they used SVM and required hand-crafted features. They [16] did a fusion of 2 datasets and achieved maximum accuracy of 84% on AlexNet. Their [17] used datasets

are suitable. Still, they have not applied fusion on datasets, and their maximum accurateness is 81% on the FFDM dataset, 87% on the ultrasound dataset, and 87% on the MRI dataset. They [18] have used multiple CNN models, but their maximum achieved accuracy is 75.3% of AlexNet and 75.5% of ResNet-152. This work was done on fuzzed Datasets A2C and B3C. The proposed model BCIC achieves good accuracy, as seen in Fig. 6. The accuracy of the suggested model BCIC is 94.9% for Dataset A2C and 94.5% for Dataset B3C which is greater than previous work.

**Figure 4:** Training accuracy graph of Dataset A2C
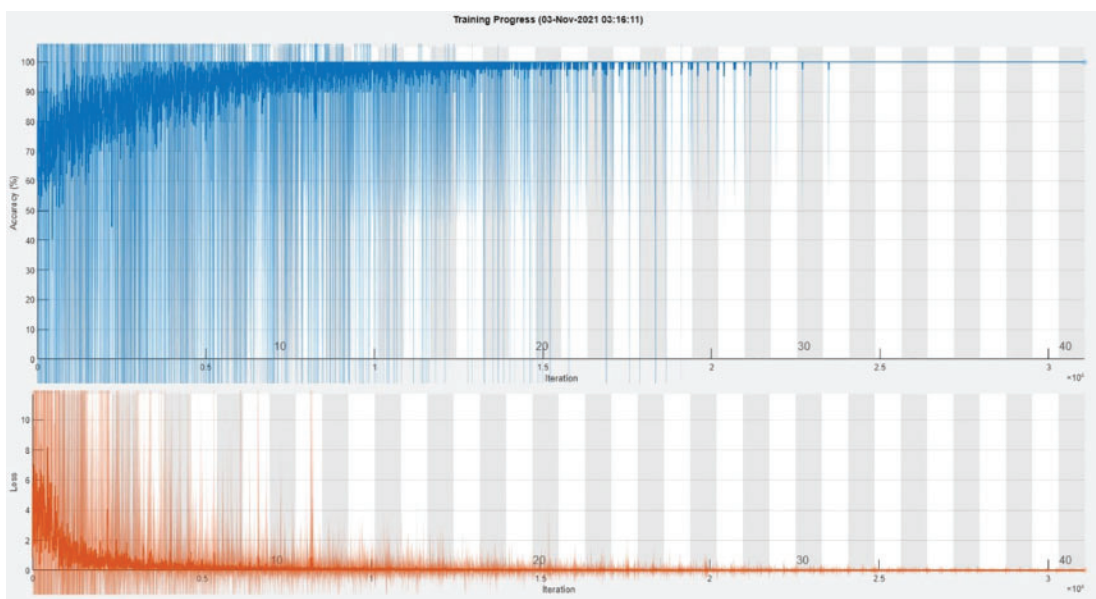
Figure 5: Training accuracy graph of Dataset B3C

Table 8: Quantitative evaluations of Dataset A2C

Statistical measures	Dataset A2C
ACC	94.885%
SEN	92.784%
SPEC	96.075%
MCR	5.114%
FPR	3.924%
FNR	7.215%
L +ve R	23.639
L −ve R	0.0751

Table 9: Quantitative evaluations of Dataset B3C

Statistical measures	Benign class of B3C	Malignant class of B3C	Normal class of B3C
ACC	94.595%	94.647%	99.485%
SEN	88.744%	97.874%	92.592%
SPEC	97.84%	89.166%	99.947%
MCR	5.404%	5.3525%	0.1544%
FPR	2.16%	10.833%	0.0521%
FNR	11.25%	2.1259%	7.4074%
L +ve R	41.085	9.0339	1543.1
L −ve R	0.1150	0.0238	0.0741
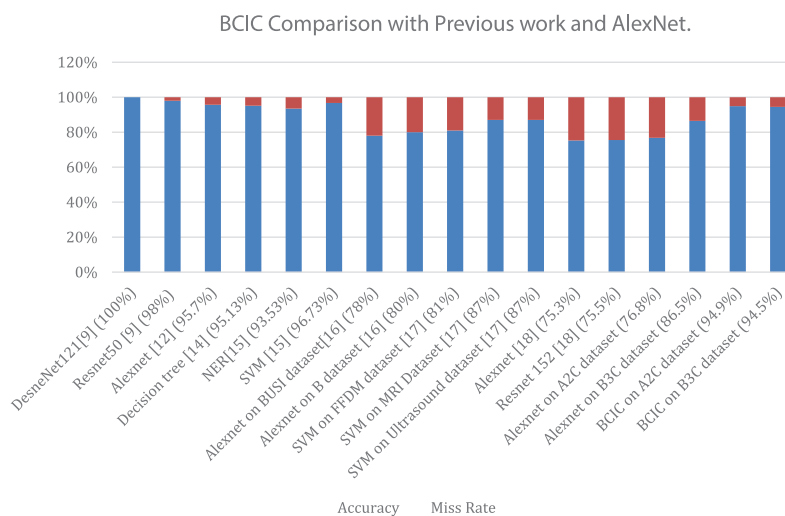
Figure 6: Accuracy and miss rate comparison of the proposed model (BCIC) with literature review and AlexNet

5 Conclusion and Future Work

This research assists in the identification and classification of BC. Early detection and therapy can help women receive adequate treatment and, as a result, decrease the rate of breast cancer disease. Breast cancers are classified as benign or malignant, and the absence of a tumor in the breast is considered normal. In this work, AlexNet and the proposed model BCIC are used. We applied fusion in this work; Datasets A, B, and C were fuzzed into 2 Datasets, A2C and B3C. We achieved 76.8% on Dataset A2C and 86.5% on Dataset B3C by using AlexNet and the optimum accuracy of 94.9% on Dataset A2C and 94.5% on Dataset B3C by using the proposed model BCIC at 40 epochs with 0.00008 learning rate. In this work, we have just compared with AlexNet, but we will extend this work by comparing this model BCIC with various models like ResNet, VGG, etc., and we will also enhance our proposed model BCIC by using different variants. We are satisfied with the findings of the proposed model BCIC and want to expand this research in the future by testing more breast cancer datasets on this model. We have fuzzed two kinds of datasets, ultrasound, and histopathology, which is why accuracy is compromised. In the future, we will fuse the same sort of multiple datasets to enhance accuracy. We will also recommend using the same kind of datasets for fusion to improve accuracy. We will examine this proposed model BCIC in various areas of medical imaging to achieve the best outcomes. In the future will also develop its Computer vision based application which we will use in hospitals for the detection of breast cancer. We have used statistical measures for performance evaluation in this work in the future we will use multiple performance measures like ROC, AUC, etc.

Author Contributions: S.A., M.F.K. and T.S. collected data from different resources. S.A., M.A.K., and M.U.N. performed formal analysis and Simulation, S.A., M.F.K., T.S., and A.R. contributed in writing—original draft preparation, M.A.K., M.Z., and K.O. contributed in writing—review and editing, M.A.K., M.Z. and K.O. performed supervision, A.R., and T.S. drafted pictures and tables, M.Z. and K.O. performed revision and improve the quality of the draft. All authors have read and agreed to the published version of the manuscript.

Availability of Data and Materials: The simulation files/data used to support the findings of this study are available from the corresponding author upon request.

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

References

[1] A. Hamidinekoo, E. Denton, A. Rampun, K. Honnor and R. Zwiggelaar, "Deep learning in mammography and breast histology, an overview, and future trends," *Medical Image Analysis*, vol. 47, pp. 45–67, 2018.

[2] M. M. Ibrahim, D. A. Salem and R. A. Seoud, "Deep learning hybrid with binary dragonfly feature selection for the wisconsin breast cancer dataset," *International Journal of Advanced Computer Science and Applications*, vol. 12, pp. 114–122, 2021.

[3] M. Alkhaleefah, S. C. Ma, Y. L. Chang, B. Huang, P. K. Chittem *et al.,* "Double-shot transfer learning for breast cancer classification from X-ray images," *Applied Sciences*, vol. 10, pp. 3999–4008, 2020.

[4]   M. S. Darweesh, M. Adel, A. Anwar, O. Farag, A. Kotb *et al.,* "Early breast cancer diagnostics based on hierarchical machine learning classification for mammography images," *Cogent Engineering*, vol. 8, pp. 1968324–1968332, 2021.

[5]   S. Sharma and R. Mehra, "Conventional machine learning and deep learning approach for multi-classification of breast cancer histopathology images—a comparative insight," *Journal of Digital Imaging*, vol. 33, pp. 632–654, 2020.

[6]   G. Battineni, N. Chintalapudi and F. Amenta, "Performance analysis of different machine learning algorithms in breast cancer predictions," *EAI Endorsed Transactions on Pervasive Health and Technology*, vol. 6, pp. 41–53, 2020.

[7]   Y. Yari, T. V. Nguyen and H. T. Nguyen, "Deep learning applied for histological diagnosis of breast cancer," *IEEE Access*, vol. 8, pp. 162432–162448, 2020.

[8]   J. Lu, Y. Wu, M. Hu, Y. Xiong, Y. Zhou *et al.,* "Breast tumor computer-aided detection system based on magnetic resonance imaging using convolutional neural network," *Computer Modeling in Engineering & Sciences*, vol. 130, no. 1, pp. 365–377, 2022.

[9]   Z. Q. Zhu, S. H. Wang and Y. D. Zhang, "A survey of convolutional neural network in breast cancer," *Computer Modeling in Engineering & Sciences*, vol. 136, no. 3, pp. 2127–2172, 2023.

[10]   J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks*, vol. 61, pp. 85–117, 2015.

[11]   G. An, M. Akiba, K. Omodakak, T. Nakazawa and H. Yokota, "Hierarchical deep learning models using transfer learning for disease detection and classification based on a small number of medical image," *Scientific Reports*, vol. 11, pp. 1–9, 2021.

[12]   A. Abbas, M. M. Abdelsamea and M. M. Gaber, "DeTrac: Transfer learning of class decomposed medical images in convolutional neural networks," *IEEE Access*, vol. 8, pp. 74901–74913, 2020.

[13]   G. Hamed, M. A. E. Marey, S. E. Amin and M. F. Tolba, "Deep learning in breast cancer detection and classification," in *The Int. Conf. on Artificial Intelligence and Computer Vision*, Cairo, Egypt, pp. 322–333, 2020.

[14]   G. I. Salama, M. Abdelhalim and M. A. Zeid, "Breast cancer diagnosis on three different datasets using multi-classifiers," *International Journal of Computer and Information Technology*, vol. 32, pp. 569–574, 2012.

[15]   X. Zhang, Y. Zhang, Q. Zhang, Y. Ren, T. Qiu *et al.,* "Extracting comprehensive clinical information for breast cancer using deep learning methods," *International Journal of Medical Informatics*, vol. 132, pp. 103985–103996, 2019.

[16]   W. A. Dhabyani, M. Gomaa, H. Khaled and F. Aly, "Deep learning approaches for data augmentation and classification of breast masses using ultrasound images," *International Journal of Advanced Computer Science and Applications*, vol. 10, pp. 1–11, 2019.

[17]   N. Antropova, B. Q. Huynh and M. L. Giger, "A deep feature fusion methodology for breast cancer diagnosis demonstrated on three imaging modality datasets," *Medical Physics*, vol. 44, pp. 5162–5171, 2017.

[18]   L. Tsochatzidis, L. Costaridou and I. Pratikakis, "Deep learning for breast cancer diagnosis from mammograms—A comparative study," *Journal of Imaging*, vol. 5, pp. 37–46, 2019.

[19]   S. Arooj, A. Rahman, M. F. Khan, K. Alissa, M. A. Khan *et al.,* "Breast cancer detection and classification empowered with transfer learning," *Frontiers Public Health*, vol. 10, pp. 924432, 2022.

[20]   S. Chidambaram, S. S. Ganesh, A. Karthick, P. Jayagopal, B. Balachander *et al.,* "Diagnosing breast cancer based on the adaptive neuro-fuzzy inference system," *Computational and Mathematical Methods in Medicine*, vol. 2022, no. 2, pp. 1–11, 2022.

[21]   A. Gökhan, "Breast cancer diagnosis using deep belief networks on ROI images," *Pamukkale Üniversitesi Mühendislik Bilimleri Dergisi*, vol. 28, no. 2, pp. 286–291, 2022.

[22]   A. Gokhan, "Deep learning-based mammogram classification for breast cancer," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 8.4, pp. 171–176, 2020.

[23]   A. Gokhan, "A deep learning architecture for identification of breast cancer on mammography by learning various representations of cancerous mass," in *Deep Learning for Cancer Diagnosis*, vol. 1. Singapore: Springer, pp. 169–187, 2021.

[24] Y. Xue, T. Xu, H. Zhang, L. R. Long and X. Huang, "SegAN: Adversarial network with the multi-scale *L1* loss for medical image segmentation," *Neuroinformatics*, vol. 16, pp. 383–392, 2018.

[25] B. Huynh, K. Drukker and M. Giger, "MO-DE-207B-06: Computer-aided diagnosis of breast ultrasound images using transfer learning from deep convolutional neural networks," *Medical Physics*, vol. 43, pp. 3705, 2016.

[26] M. Said, "Breast cancer," [Online]. Available: https://www.kaggle.com/mostafaeltalawy/brest-cancer (accessed on 21/06/2022)

[27] A. Aabo, "Breast cancer dataset," [Online]. Available: https://www.kaggle.com/anaselmasry/breast-cancer-dataset (accessed on 21/06/2022)

[28] F. A. Spanhol, L. S. Oliveira, S. Petitjean and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Transactions on Biomedical Engineering*, vol. 63, pp. 1455–1462, 2015.

[29] A. Shah, "Breast ultrasound images dataset," [Online]. Available: https://www.kaggle.com/aryashah2k/breast-ultrasound-images-dataset (accessed on 21/06/2022)

[30] M. U. Nasir, T. M. Ghazal, M. A. Khan, M. Zubair, A. U. Rahman *et al.,* "Breast cancer prediction empowered with fine-tuning," *Computational Intelligence and Neuroscience*, vol. 2022, pp. 1–9, 2022.

[31] M. U. Nasir, M. A. Khan, M. Zubair, T. M. Ghazal, R. A. Said *et al.,* "Single and mitochondrial gene inheritance disorder prediction using machine learning," *Computers, Materials & Continua*, vol. 73, no. 1, pp. 953–963, 2022.