IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

Breast Cancer Histopathology Classification Using Levenberg–Marquardt Optimised Deep Neural Networks

Milan Srinivas[1], Shri Ranjani S M[2], Vinay Kumar[3], Shravya Bhat[4]

BE, Department of Computer Science and Engineering, VKIT, Bangalore, India[1]

BE, Department of AI and ML, BNM Institute of Technology, Bangalore, India[2]

BE, Department of AI and ML, BNM Institute of Technology, Bangalore, India[3]

BE, Department of AI and ML, Nitte Meenakshi Institute of Technology, Bangalore, India[4]

Abstract: Breast cancer is a dangerous illness that mostly affects women. Increasing survival rates requires early detection of the disease. Artificial intelligence developments, especially in deep neural networks, have greatly improved breast cancer detection. The suggested LMHistNet, a hybrid convolutional neural network model intended to categorise microscopic images of breast tumour tissue acquired during excisional biopsy, is one such development. Levenberg-Marquardt optimization and asymmetric convolutions are used by LMHistNet to categorise breast cancer images into eight subclasses and a binary (benign or malignant) category. To achieve adaptive feature refinement, the model integrates a convolution block attention module. Additionally, batch normalisation is used to normalise input features and speed up convergence. Additionally, this method reduces internal covariance shifts during training. The model's convergence is further improved by the application of a hinge loss function, which makes it a useful instrument for precise breast cancer detection. By extracting various features from histopathological pictures, the LMHistNet model effectively performs binary and eight-class classifications that are both independent and dependent on magnification. LMHistNet's efficacy was shown in a study containing 7,909 histopathological pictures, of which 2,480 were benign and the remainder malignant. Using loss and accuracy curves, the model's performance was assessed at several magnifications (40X, 100X, 200X, and 400X). According to the findings, the model scored 88% accuracy, 89% precision, 88% recall, and 88% F1 score for multiclass categorization into eight subtypes. The model achieved remarkable accuracy, precision, recall, and F1 score of 99% for binary classification, which discerns between benign and malignant tissues. These results highlight the model's strong performance in correctly categorising photos of breast cancer at various magnifications and classes.

Keywords: Breast cancer, convolutional neural networks (CNN), deep learning, histopathological images, Levenberg–Marquardt, transfer learning.

I. INTRODUCTION

One type of cancer that has a major impact on women's death rates is breast cancer [1], [2]. Breast cancer can be caused by both hereditary and non-genetic variables, including alterations in the BRCA1 or BRCA2 genes, age, mammography density, duration of breastfeeding, obesity, alcohol intake, and so forth [3] and [4]. Breast lumps, breast swelling, soreness, pain, non-breast milk nipple discharge, etc. are common signs of breast cancer. the breast early. Early detection of breast cancer can be facilitated by the generation of three-dimensional pictures of the scanned breast through the use of received signals. With little expense and few adverse effects, focused microwave breast hyperthermia (FMBH) [50] is an efficient method for detecting breast cancer.

A useful guiding technique for FMBH is compressive thermoacoustic tomography (CTT), which tracks the distribution of microwave power in the breast.

In the past, manual data interpretation from the various screening techniques was used. Because of the complexity of the data and the need for medical professional's competence, it was a difficult task. Manual analysis takes a lot of time, and experts could become overworked, which could result in incorrect diagnoses. Thanks to advances in technology, breast cancer is now being diagnosed using computer-aided diagnostic, or CAD, systems. Modern approaches based on artificial intelligence (AI), such as deep learning (DL) and machine learning (ML) [7], have made it feasible to rapidly and accurately diagnose breast cancer.

DOI: 10.17148/IJARCCE.2024.13812

The process of applying machine learning algorithms to data entails uncovering hidden patterns that help with disease classification and prediction. Features in machine learning are manually extracted, which is a laborious and time-consuming process. Images of breast cancer are classified using a variety of well-liked machine learning methods, such as support vector machines (SVM), logistic regression, and random forest (RF).

Multiple-layer deep learning algorithms extract features automatically and are faster than machine learning. Convolutional neural networks, or convents for short, are a type of deep learning model that is primarily utilised in computer vision applications [8]. Convolutional, pooling, and fully connected layers make up CNN, which is a very helpful tool for identifying and categorising photos of breast cancer.

A novel convolutional neural network called LMHistNet has been suggested for the binary and multiclass categorization of pictures related to breast histopathology. Differentiating between benign and malignant groups is known as binary classification. The process of multiclass classification entails grouping the histological pictures into eight classes: malignant (such as mucinous carcinoma, lobular carcinoma, ductal carcinoma, and papillary carcinoma) and benign (such as adenosis, fibroadenoma, phyllodes tumour, and tubular adenoma). There have been categories that are both magnification-dependent and independent. The deep convolutional neural network is integrated with the channel block attention module [9] to effectively improve intermediate features.

The Levenberg-Marquardt algorithm is used to update the network's weight. The effectiveness of the suggested LMHistNet has been demonstrated using the Breast Cancer Histopathological Database (BreaKHis) [10], which was established by compiling microscopic pictures from 82 patients at four distinct resolutions.

The principal contributions of the suggested LMHistNet model are as follows.

- Several blocks have been added to the model in order for it to correctly learn the different classification features
- For the efficient classification of breast cancer histopathology pictures, the Convolution Block Attention (CBAM) Module, Convolution Block Normalised block (CGN B), dual Spa-tial Factorization and Grid Reduction (SFGR1 & SFGR2) Blocks, and BN block have been incorporated.
- While the dual Spatial Factorization and Grid Reduction (SFGR1 & SFGR2) Blocks are used to solve the vanishing gradient problem, the position-invariant local features are normalised using the Convolved Group Normalised block (CGN B). Convolution Block Attention is used to create the channel and spatial features (CBAM).
- The Hinge Loss function is used to speed up the model's training.
- Multiclass classification, as well as binary classifications independent of and dependent on magnification, have been carried out.
- Refinement of intermediate features according to channel and attention modules. Spatial attention improves performance and both max and average pooled features are employed for channel attention.
- Using Levenberg Marquardt optimization, the sum of squared differences between the actual and forecasted data was decreased.
- In order to achieve the best accuracy, hyperparameter adjustment was done.

The structure of the paper's remaining portion is as follows: An overview of pertinent studies in the field of breast cancer detection is given in Section II. Section III provides a full description of the LMHistNet architecture; Section IV describes the experimental findings and discussions of the custom LMHistNet; Section V discusses the conclusion and future work.

II. RELATED WORK

MACHINE LEARNING APPROACHES

The Radial-Based Function Kernel Extreme Learning Machine algorithm, whose parameters are modified using the differential evolution algorithm, was employed by Bacha and Taouali [11] to diagnose breast cancer. Using the Kernel Principal Component Analysis (KPCA) methodology, the best qualities are chosen. The Wisconsin Breast Cancer Database (WBCD) and the Mammographic Image Analysis Society (MIAS) dataset are used to validate experimental results.

In order to improve classification accuracy, Dai et al. [12] aggregated the output of several weak classifiers using the ensemble learning capability of RF. Two ensemble learning algorithms—RF and Extreme Gradient Boosting (EGB)—were employed by Kabiraj et al. [13] to forecast breast cancer. RF obtained an accuracy of 74.73%, whereas EGB was less accurate. Linear discriminant analysis was used by Omondiagbe et al. [14] to reduce the dimensionality of the features, and artificial neural networks (ANN), support vector machines (SVM), and Naïve Bayes were used to classify

DOI: 10.17148/IJARCCE.2024.13812

the feature set that was produced. With a sensitivity of 98.41%, accuracy of 98.82%, specificity of 99.07%, and area under the receiver operating characteristic curve of 0.9994, the SVM outperformed the other techniques.

Vijaya Rajeswari et al. [15] retrieved features from mammography using a two-dimensional rough transform. The derived features were then classified using support vector machines (SVM) with a 94% accuracy rate. Based on the collected features at different threshold levels, Wu and Hicks [16] categorised breast cancer into non-triple negative and triple negative categories. Four criteria were used to perform the classification:

In contrast to the other ML models, the SVM classifier achieved higher accuracy, sensitivity, and specificity.

Using five distinct machine learning classifiers, including SVM, KNN, XGBoost, RF, and LightGBM, Michael et al. [17] classified breast ultrasound pictures into benign and malignant categories. The findings were then improved through optimization using the Parzen estimator. With 100% precision, 99.86% accuracy, and 99.6% recall, the Light GBM classifier fared better than the other models. Ara et al. [18] used a variety of machine learning techniques, including RF, KNN, SVM, LR, NB, and DT, to assess the WBCD performance. RF and SVM attained the highest accuracy of 96.5%. Five machine learning algorithms—LR, SVM, RF, KNN, and DT (C4.5)—were examined by Naji et al. [19] with a 97.2% accuracy rate on WBCD.

Breast histopathology images were segmented by Anuranjeeta et al. [20], and the features of the segmented region were retrieved. Several algorithms are used to classify images as malignant or non-cancerous based on the characteristics retrieved. These algorithms include J-Rip, SMO, MLP, LMT, Random forest, Rotation forest, Naïve Bayes, and PART. Rotation forest achieved the highest accuracy of 85.7%.

DEEP LEARNING

Veta et al.'s assessment [45] was conducted to understand breast cancer detection through the widespread usage of images from breast histology. Even with advancements in tissue processing, staining and slide digitization still have inherent variations, which calls for the development of robust analytic techniques to deal with these fluctuations. A deep active learning system was presented by Qi et al. [46] for the binary classification of breast histopathology pictures utilizing entropy-based and confidence-boosting methods. The better categorization was achieved by selecting high-confidence samples more effectively using confidence boosting than through entropy. A collaborative transfer network (CTransNet) with a transfer learning backbone was proposed by Liu et al. [47] to extract and predict from the optimized fused features, achieving a 98.29% accuracy rate for multi-class classification.

A Stacked Sparse Autoencoder was proposed by Xu et al. [48] to extract nuclei differentiating characteristics from the pixel intensities of histopathology pictures. Using high-level characteristics obtained from the auto-encoder, sliding window operation is used to images to represent image patches and successfully categorize as nuclear or non-nuclear.Chowdhury et al. [21] used transfer learning based on ResNet10 to assess images of breast histopathology, and the results showed a 99.58% accuracy rate in cancer early detection. Zhang et al. [22] integrated CNN with a graph convolutional network (GNN), first establishing eight convolution layers. Batch normalization and dropout, which used stochastic pooling rather than max pooling, enhanced the model. The model's accuracy increased when both relation awareness and image level characteristics were extracted. Because the employed dataset was homogeneous and smaller in size, larger datasets will need to be used in future research in order to validate the findings. Gour et al. suggested a CNN based on residual learning with feature learning at various levels of abstraction for the diagnosis of breast tumors [23].The binary classification yielded an accuracy of 84.34%; with more data, the accuracy increased to 92.52%.Alzubaidi et al. [24] used transfer learning to classify biopsy images into four groups. In this method, training is first conducted on a dataset within the same domain and subsequently on the target dataset. The opposite strategy, which involves training first on a dataset from a different domain and then on the target dataset, is also carried out. The accuracy of this method's image classification was 97.4%. CNN was employed by Ting et al. [25] to classify breast mammography images into two groups, with a 90.50% accuracy rate.

A CNN for the classification of breast histology images was created by Bardou et al. [26]. The range of accuracy for the multiclass classification was 83.31% to 88.23%, and the range for the binary classification was 96.15% to 98.33%. Mahmoud [27] achieved an accuracy of 86.8% for binary classification using deep CNN.The Inception Recurrent Residual model was proposed by Alom et al. [28] and achieved testing accuracy of 98.59% for multiclass classification and 99.05% for binary classification. CNN, which was proposed by Gecer et al. [29], achieved a 55% accuracy rate in multiclass image classification.

DOI: 10.17148/IJARCCE.2024.13812

With an accuracy of 77.8% for four class classifications and 83.3% for two-class classifications, Araíujo et al. [30] examined binary and multiclass classifications. Bayramoglu et al. presented single-task and multi-task CNN for magnification-independent breast histopathology pictures [31]. The recognition rate of the single-task model was 83.25%, while the multi-task model yielded an 82.13% recognition rate. Han et al. [32] proposed a structured deep-learning model for multiclass classification that achieved an average accuracy of 93.2%. Li et al. [33] classified four classes with a 95% accuracy rate.

Aljuaid et al.'s [34] use of ResNet, InceptionV3Net, ShuffleNet, and a transfer learning technique allowed for the binary and multiclass classification of breast histopathology images.

ResNet achieved maximum average accuracies of 97.81% to 99.70% for binary and multiclass classification. Majumdar et al. [35] used the results of GoogleNet, MobileNetV3_Small, and VGG11 to suggest an ensemble approach utilizing the gamma function. The BreaKHis dataset was evaluated, yielding an accuracy of 96.95% on the ICIAR-2018 dataset and values of 96.16%, 98.67%, 98.24%, and 99.16% for the 400X, 200X, 100X, and 40X levels of magnification. CNN techniques were employed by Tayel et al. [36] to detect and localize breast tumors.

Jabeen et al. [37] achieved a 99.1% accuracy rate in classifying breast cancer on ultrasound pictures by using optimized reformed differential evaluation (RDE) and reformed grey wolf (RGW) algorithms along with transfer learning on a modified DarkNet-53. Wang et al. [38] suggested a hybrid deep-learning model for feature extraction and binary classification of pictures related to breast histopathology. 86% F1 score, 85% precision, 86% accuracy, and 85% sensitivity were attained with this procedure. Hybrid convolutional neural networks based on ResNet18, AlexNet, ShuffleNet, VGG16, and MobileNetV2 were proposed by Sahu et al. [39]; validation was performed using datasets from mammograms and ultrasounds. Even on tiny datasets, the model performed better.

The final predictions are created by adaptively combining the decision scores of the base models. Altameem et al. [40] proposed an ensemble approach in which fuzzy rankings of the base classification techniques like DenseNet121, VGG-11, ResNet-164, and Inception V4 are built using Gompertz function. Particle swarm optimization and genetic algorithms were employed by Gonçalves et al. [41] to identify optimal hyperparameters and designs for DenseNet-201, ResNet-50, and VGG-16. The F1 scores for all three models were higher than 0.90. For the binary classification of breast histopathology pictures, Singh and Kumar [42] suggested a model based on inception and residual networks, and they were able to achieve a training accuracy of 0.9642. The dataset is expanded by the application of Generative Adversarial Networks (GAN). The model's 97.58% F-score was attained.
A comparison of the several deep learning classification techniques that have been published in the literature may be seen in Table 1.

III. PROPOSED METHODOLOGY

Several innovative features are incorporated into the LMHistNet architecture, which is intended for the binary and multiclass classification of histopathology pictures of breast cancer. It has a Convolved Group Normalised Block (CGN B), which makes use of group normalization to control batch size changes and stabilize the training process, guaranteeing steady feature scaling. Dual Spatial Factorisation and Grid Reduction Blocks (SFGR1 & SFGR2) are another aspect of the architecture that breaks down the spatial dimensions of feature maps to lower computational complexity while maintaining crucial spatial information.

As a result, spatial hierarchies can be effectively extracted from the data. The Convolution Block Attention Module (CBAM), which is further integrated into the model, improves the sensitivity of the model to critical patterns by sharpening the feature maps by emphasizing significant regions and suppressing unimportant features. Furthermore, input feature normalization is achieved by batch normalization, which minimizes the internal covariance shift and speeds up convergence. The Levenberg-Marquardt method, which is renowned for its quick convergence qualities, is used to optimize the entire network. A hinge loss function is used to improve convergence and robustness during training. Together, these components allow LMHistNet to effectively extract a variety of features from histopathological pictures, allowing it to perform both independent and magnification-dependent classifications.

A. LMHistNet Architecture

To maximize the classification of histological images of breast cancer, the LMHistNet design combines a number of sophisticated components. Convolution Block Attention Module (CBAM), in conjunction with dual Spatial Factorisation and Grid Reduction Blocks (SFGR1 & SFGR2), Convolved Group Normalised Blocks (CGN B), and other components, forms the core of the design.

DOI: 10.17148/IJARCCE.2024.13812

- Convolutional Group Normalised Block (CGN B): This block normalizes the features within each group after grouping the feature channels. This method lessens the unpredictability brought on by small batch sizes and improves the stability of the training process by preserving uniform feature scaling across batches.

- Blocks for Grid Reduction and Spatial Factorisation (SFGR1 & SFGR2): By reducing the spatial dimensions of the feature maps, these blocks facilitate learning by performing spatial factorization and grid reduction. This factorization concentrates on important features and eliminates duplicate data, which lowers computing complexity and improves convergence.

- Convolution Block Attention Module (CBAM): This architecture-integrated module applies attention processes to enhance the features. By suppressing less relevant parts and concentrating on the most useful portions of the feature maps, it improves the model's sensitivity to important features and boosts its overall performance.

Figure 1 shows the LMHistNet workflow, which demonstrates how these elements are integrated to produce accurate and efficient histopathological image categorization.

B. CONVOLVED GROUP NORMALISED BLOCK (CGNB)

The LMHistNet architecture processes input histopathological images through a Convolved Group Normalised Block (CGNB), which is enhanced with group normalization and ReLU activation functions. In this block, the input image is initially processed using convolution operations to generate position-invariant local features. These features are then grouped and normalized by group normalization, which helps in stabilizing the training and managing batch size variations.

The convolution operation can be mathematically represented by the following equation:

$$C_l = (\mathrm{g} * \mathrm{k})\,[\mathrm{x}, \mathrm{y}] = \sum_{\mathrm{j}} \sum_{\mathrm{k}} \mathrm{k}\,[\mathrm{j}, \mathrm{k}]\,\mathrm{g}[\mathrm{x} - \mathrm{j}, \mathrm{y} - \mathrm{k}] \quad (1)$$

A 5 × 5 filter size is used for feature extraction, producing a pattern that may be recognized at several scales to produce larger object perceptions and multiscale feature learning. The convolved features Gl[x,y], which were extracted, are put into an average pooling layer. This layer removes spatial information, making the feature more resilient to spatial translations and reducing the risk of overfitting. A decreased spatial feature was obtained using global average pooling. Equation 2 illustrates the gap for a convolved feature map Cl, where sl denotes the size 2 sliding stride and a, b, and c are the feature sizes.

$$G_{ap} = (\mathrm{a} - \mathrm{k} + 1)\,/s_l \times (\mathrm{b} - \mathrm{k} + 1)\,/s_l \times c \qquad (2)$$

The decreased convolved feature that was produced The group normalization layer GN receives the gap, enabling separate distributions to be learned for every group of channels. Features normalized by group Equations 3 and 4 illustrate how GN are calculated using mean and variance, where ε is a tiny constant and si is the set of pixels for which the mean and standard deviation are calculated for n, the set size.

$$Mean\,(\mu_i) = \frac{1}{n} \sum_{k \in s_i} y_k \qquad (3)$$

$$Standard\ deviation\,(\sigma_i) = \sqrt{\frac{1}{n} \sum_{k \in s_i} (y^2_{k-\mu_i}) + \varepsilon} \qquad (4)$$

Convolutional layer augmentation by group normalization, followed by ReLu activation, is incorporated into the Convolved Group Normalisation Block (CGN B), which improves network stability and makes it easier to learn complex spatial information features. Local patterns that distinguish benign from malignant traits are well-represented by hierarchical and spatial patterns, such as edges and textures. Group normalization and activation functions together with non-linearity enable the efficient generation of complex patterns. ReLU activation function and the resulting normalized feature are used to optimize the group normalization GN output. The Spatial Factorisation and Grid Reduction (SFGR) Block applies to CGNO.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

C. SFGR1 BLOCK: SPATIAL FACTORIZATION AND GRID REDUCTION

As seen in Figure 1, the LMHistNet consists of two identical blocks for spatial factorization and grid reduction (SFGR1, SFGR2). Figure 2 displays the Spatial Factorization and Grid Reduction (SFGR1) block. The SFGR1 block receives the characteristics of the CGNB layer output (CGN O). The vanishing gradient problem is solved by factoring the involved features into smaller sizes thanks to the benefit of auxiliary units that create a broader network. Stacking many filters of different sizes on the same level in parallel units instead of deep layers might reduce the network's dimensions and speed up computations. Spatial factorization is the outcome of splitting a convolution into smaller convolutions, which uses asymmetric convolutions to achieve optimal grid size reduction. The output of the normalised features A layered method is used to process CGN O from the Convolved Group Normalised (CGNB) block, utilizing four stack operations SOi=1–4. SO1 contains a convolution layer with a 1 × 1 filter size that allows for the decrease of dimensionality using non-linear features construction that facilitates squeezing an output to the appropriate level depth (the fifth equation).

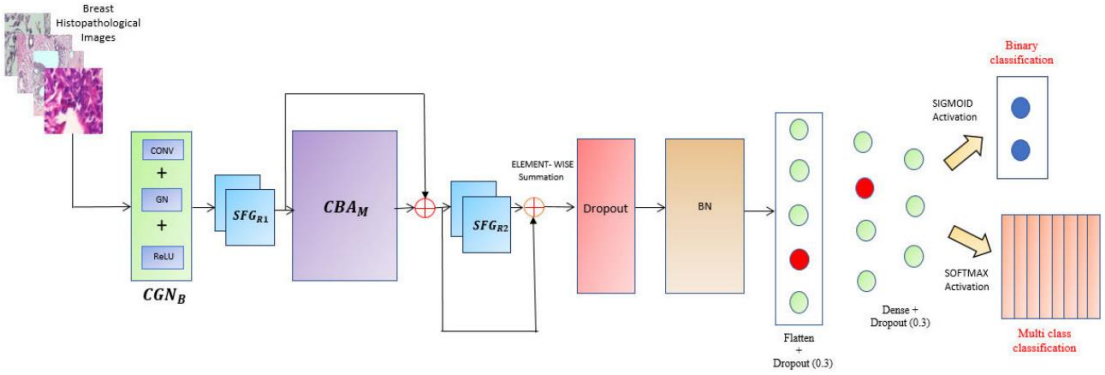


FIGURE 1. Overall workflow of the proposed system.

The results acquired from the Involved Group Regularized Block (CGN O) undergo convolutional layer (L) with a 1 x 1 filter, which minimised the SO1 characteristics by giving depth dimension more of an emphasis than spatial characteristics.

$$SO_1 = (\mathrm{C}_L\ (CGN_O))^{1\times1}$$

Equation 6 illustrates how the output of the Convolved Group Normalised Block (CGN O) is also subjected to max pooling, convolutional layer with 1 × 1 filter, and decreased SO2 features. Using max pooling (PM) and a convolution layer (CL) with a 1 × 1 filter size reduces the spatial dimensions of the feature maps by enhancing the complex patterns that occur from the non-linear interactions between the features.

$$SO_2 = (P_M(\mathrm{C}_L\ (CGN_O)))^{1\times1}$$

Using a 1 × 1 convolved layer (CL), SO3 maps the features. The resulting feature is then processed using dual filters of sizes 3 × 1 and 1 × 3, leading to spatial factorization into asymmetric convolutions as 3 × 3 is split into 1 × 3 and 3 × 1 with less complexity (equation7).

$$SO_3 = (\mathrm{C}_L(\mathrm{C}_L\ (CGN_O))^{1\times1})^{1\times3} + (\mathrm{C}_L(\mathrm{C}_L\ (CGN_O))^{1\times1})^{3\times1})$$

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

SO4 uses an auxiliary convolution layer of 1 × 1 to process the CGN O. This is followed by a 3 × 3 filter, which is further processed by a 1 × 3 factored spatial feature map (equation 8)

$$SO_4 = (\mathrm{C}_L(\mathrm{C}_L(\mathrm{C}_L\,(CGN_O))^{1\times1})^{3\times3})^{1\times3} + (\mathrm{C}_L(\mathrm{C}_L(\mathrm{C}_L\,(CGN_O))^{1\times1})^{3\times3})^{3\times1}$$

and produces a spatial structural pattern that is precise towards a certain orientation. SO4 uses an auxiliary convolution layer of 1 × 1 to process the CGN O. This is followed by a 3 × 3 filter, which is further processed by a 1 × 3 factored spatial feature map (equation 8) to provide a spatial structural pattern that is precise towards particular The concatenation of the depth features from SO1, complex patterns of SO2, SO3, and SO4 yields the expressive diversified feature FSFGR1. (equation 9) which balances the orientation of the shallow layers' weak contributions.

$$F_{SFG_{R1}} = SO_1 + SO_2 + SO_3 + SO_4$$

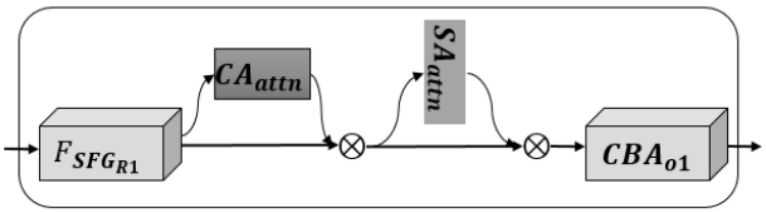


FIGURE 3. Convolution block attention (CBA_M).

D. MODULE FOR CONVOLUTION BLOCK ATTENTION (CBAM):

The Convolution Block Attention (CBAM) is depicted in Figure 3 and comprises the Channel Attention (CAattn) and Spatial Attention (SAattn) modules. This allows the network to focus more on discriminative features.1) CHANNEL ATTENTION MODULE (CAattn)

Figure 4 illustrates how the Channel Block Attention (CBattn) Block uses the feature map FSFGR1 from the Spatial Factorisation and Grid Reduction (SFGR) Block to take advantage of the inter-channel relationship of features.

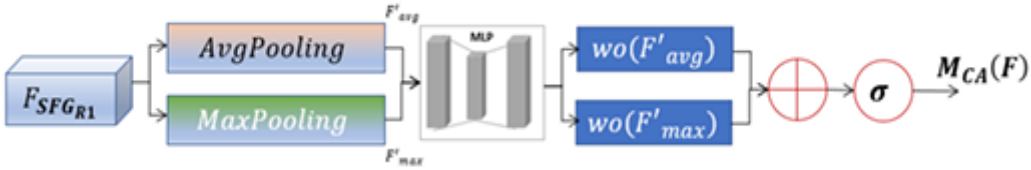


FIGURE 4. Channel attention (CA_{attn}) module.

To highlight the importance of the data included in the input feature, channel attention uses each channel in a feature map as a feature detector. The input feature map's spatial dimension is compressed in order to calculate the channel attention effectively.The average pooling and max-pooling procedures are used to first aggregate the spatial data of a feature map, producing average F′ avg and max pooled features F′max.

After feeding the F′ average and F′max to an MLP network, a one-dimensional channel attention map is produced, which is represented as

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

$M_{CA} \in \mathbb{R}^{C\times1\times1}$ where the number of channels is denoted by c.

The descriptors are each given a multilayer perceptron with a single hidden layer, and the resultant feature vectors are then combined using element wise summation.The concealed activation size is set to $\mathbb{R}^{C/r\times1\times1}$, where r is the reduction ratio, in order to lower parameter overhead.

Equation (10), which computes the channel attention regardless of the input presented, shows that σ is a logistic function ranging between 0 and 1.

$$W_0 \in \mathbb{R}^{C/r \times C} \text{ and } W_1 \in \mathbb{R}^{C\times/r}$$

$$M_{CA}(F) = \sigma(W_1(W_0(F'_{avg}) + W_1(W_0(F'_{max})))$$

2) **The Spatial Attention (SBattn) Block:**

Receives the channel information of a feature map MCA (F) and aggregates the results of max pooling and average pooling operations along the channel axis to produce average pooled features and max pooled features,

$$F'_{max} \epsilon \mathbb{R}^{1\times H\times W} \qquad F'_{avg} \epsilon \mathbb{R}^{1\times H\times W}.$$

They are then combined and convolved by a 7x7 convolution layer to create a two-dimensional spatial attention map. Equation (11) represents the spatial map, where f 7x7 denotes the convolution operation filter size 7 × 7.

$$M_{SA}(F) = \sigma(f^{7\times7}([F'_{avg}; F'_{max}]))$$

$$CBA_{o1} = M_{CA}(F) + M_{SA}(F)$$

The output of the (CBAM) Module, the produced CBAo1(equation12) features, is further processed using the filtered feature FSFGR1 to prevent the loss of any significant characters. Equation 13 shows that when the feature map FSFGR1 and the output from the element-wise multiplication of CBAo1 and FSFGR1 are combined, they produce (CBAO), which has better interpretability and robustness.

$$CBA_0 = F_{SFG_{R1}} + F_{SFG_{R1}} \otimes CBA_{o1}$$

The histopathology data's benign and malignant qualities can be accurately identified by generating an increased specific pattern of features using the channel and spatial features. The integration of channel and spatial information is made easier by a channel block attention module (CBAM), which produces feature representations that are more relevant and informative than those that are redundant or less informative. By dynamically analysing the features of the incoming data, this attention module concentrates on particular channels to recover more discriminative information, allowing the model to learn and efficiently exploit the intricate relationships within the data. The SFGR2 block receives the channel attention map CBAo1 and the feature map FSFGR1. The attention map CBA0 and the total output of SFGR2 are subjected to dropout with a probability of 0.3, which stops neurons from co-adapting. By applying batch normalisation, the training-induced change in the distributions of network activations was mitigated.

E. OUTPUT LAYERS

The resulting features are transformed into a single, long continuous vector by feeding the batch normalised attention map into the flatten layer. By choosing nodes to dropout with a 30% probability, dropout was included in the flatten layer, which enhanced model generalisation and decreased overfitting. The output is flattened and then processed to a dense layer using the ReLU activation function and a 30% dropout. The activation function used for multiclass classification is the softmax activation function, as shown in zl equation 14, where σ is the softmax, _z is the input vector, and e is the standard exponentiation function that refers to the vector of each input element. For binary classification, the activation function used is the sigmoid in the outer dense layer.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

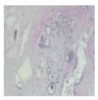
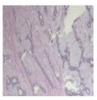
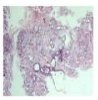
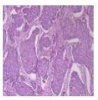
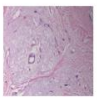
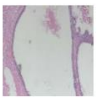
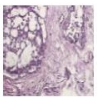
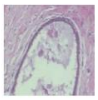
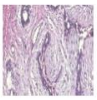
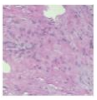
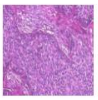
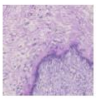
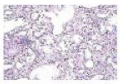
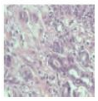
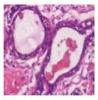
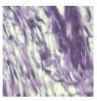
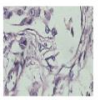
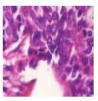
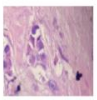
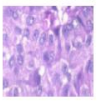
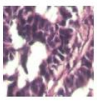
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

$$\sigma(\vec{z})_i = \frac{e^{z_l}}{\sum_{j=1}^{n} e^{z_m}}$$

TABLE 2. Images of eight different classes with four different resolutions from BreaKHis dataset.

	Adenosis	Fibroadenoma	Phyllodes Tumor	Tubular adenoma	Ductual carcinoma	Lobular carcinoma	Mucinous carcinoma	Papillary carcinoma
40x								
100x								
200x								
400x								

F. LEVENBERG MARQUARDT OPTIMIZATION AND HINGE LOSS FUNCTION

Levenberg Marquardt learning is used to improve the suggested model, and the error is calculated by contrasting the labelled data with the weight update that is improved by the combined Gauss Newton method and Error Backpropagation [44]. The Gauss Newton approach, which uses divergent approximation, improves the slow convergence of the Error Backpropagation process, which makes it inefficient. Despite having a significantly more complicated error surface, the Levenberg Marquardt method converges consistently and performs well in terms of speed and stability. Equations 15, 16, and 17 provide the weight updates for the Levenberg Marquardt, Gauss Newton, and steepest descent algorithms.

$$w_{i+1} = w_i - \lambda g_i \qquad (15)$$

$$w_{i+1} = w_i - (J_i^T J_i)^{-1} J_i e_i \qquad (16)$$

$$w_{i+1} = w_i - (J_i^T J_i + \lambda I)^{-1} J_i e_i \qquad (17)$$

where the words "I" stand for identity matrix, "J" for Jacobin matrix, "wi," and "wi+1" for weight, combination coefficient, and identity matrix, respectively. The ith iteration's error vector is denoted by e, while the error function's first order derivative is represented by g. If, after weight updating and the resultant error, the parameter produces less mistakes than the preceding one, the combination coefficient λ is decreased. By adaptively modifying step size and damping, the Levenberg-Marquardt optimization technique speeds up model training and promotes effective network convergence.

Because of its linear property (equation 18), the function employed in the model increased the model's training speed and produced superior convergence. Here, hl stands for hinge loss, av for actual value, and pv for predicted value.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

Equation 19 defines the hinge loss for multiclass classification, where wpv x and wav x are model parameters.

$$hl = \max(0, 1 - a_v.p_v) \qquad (18)$$

Equation 19 defines the hinge loss for multiclass classification, where wpv x and wav x are model parameters.

$$hl = \max(0, 1 + (w_{pv}x - w_{av}x)) \qquad (19)$$

IV. EXPERIMENTAL RESULTS AND DISCUSSIONS

This section describes the dataset description, augmentation techniques, ablation study, and performance evaluation of breast cancer utilising several performance metrics for the categorization of histological pictures.

A. DATASET DESCRIPTION

The suggested work makes use of a publicly accessible dataset called BreaKHis, which includes microscopic pictures of breast tissues taken from 82 people. 2,480 benign samples and 5429 malignant samples in four distinct magnification factors (40X, 100X, 200X, and 400X) are included in the collection.

Benign tumours are not cancerous, but malignant tumours are. Table 2 displays four different resolutions of representative images from the BreaKHis dataset, categorised into eight distinct classes. The following lists the quantity of photos used for each of the eight class classifications:

Adenosis images 444, Fibroadenoma photos 1014, Phyllodes tumour images 453, and tubular images 569 adenoma, 3457 Ductal carcinoma photos, 626 Lobular carcinoma images, 792 Mucinous carcinoma images, and 580 Papillary carcinoma images. Of the dataset, 80% is utilised for training and 20% is for testing.

B. EXPERIMENTAL SETUP

The suggested job was carried out using Google Colab notebooks, a cloud computing platform. A graphics processing unit (GPU) and tensor processing unit (TPU) are provided by Google Colab for building deep neural learning models. The training and efficient operation of the deep learning model were facilitated by this cloud-based solution. Python programming was used to create the LMHistNet model, and a number of necessary libraries were used, including Keras, matplotlib, OS, Tensor Flow, and sklearn. Table 3 displays the setup of the system that was used to test the suggested work.

TABLE 3. Computer configuration of the proposed LMHistNet.

Item	Configuration
Processor	Intel (R) Xeon (R) Gold 6230 CPU @2.10GHz
Graphics Card	NVIDIA Quadro RTX 5000 16 GB
Ram Size	64 GB
Hard-Disk Size	2 TB

C. DATA PREPROCESSING AND AUGMENTATIONS

The network structure's input picture is scaled to 224 X 224. The convergence of the LMHistNet model and the production of standard features will be greatly impacted by the size variation of the images. With data augmentation, you can use different operations, such as translation, rotation, shearing, etc., to artificially increase the size of the dataset.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

However, when compared to natural photographs, none of the data augmentation strategies would function as intended in the case of our images.

Therefore, care must be made when choosing the data augmentation methods. Images in our experiment are rotated with a rotation range of 15, and as pixels move outside the image, fill_mode to nearest fills the vacant space with the pixel closest to it.

D. HYPERPARAMETER TUNING

The gradient's learning rate, dropout rate, batch size, epochs, and optimization units were some of the LMHistNet models' hyperparameters that were adjusted using grid search. Experimentation revealed that a batch size of 32, and 0.001 learning rate were applied in the suggested job.LMHistNet. For each deep neural network, the Levenberg-Marquardt optimizer is utilised to fine-tune the network weights.

The intermediate layers employed the ReLU activation function.

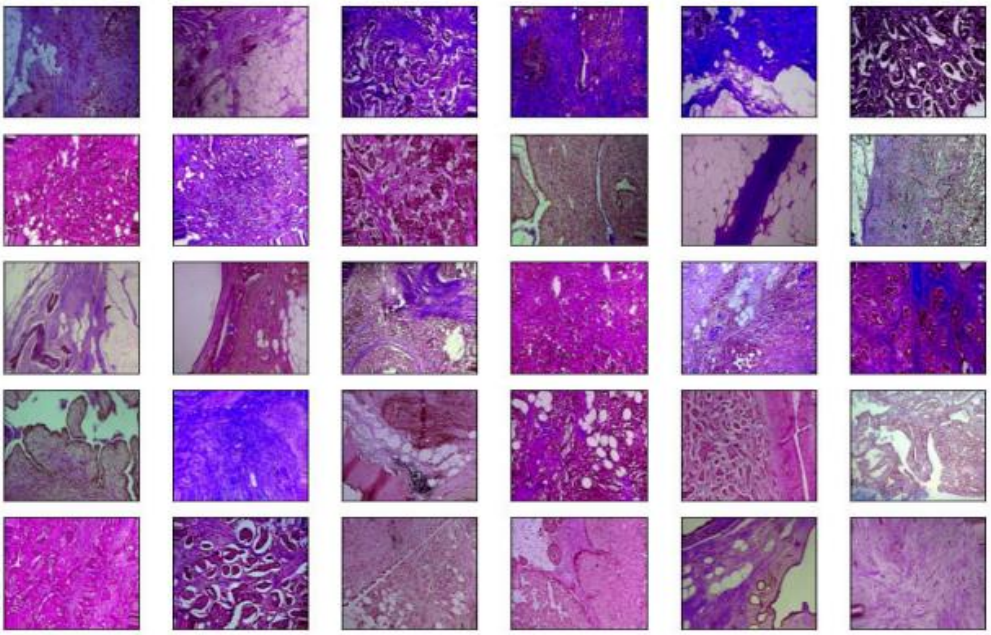


FIGURE 5. Sample data augmented images.

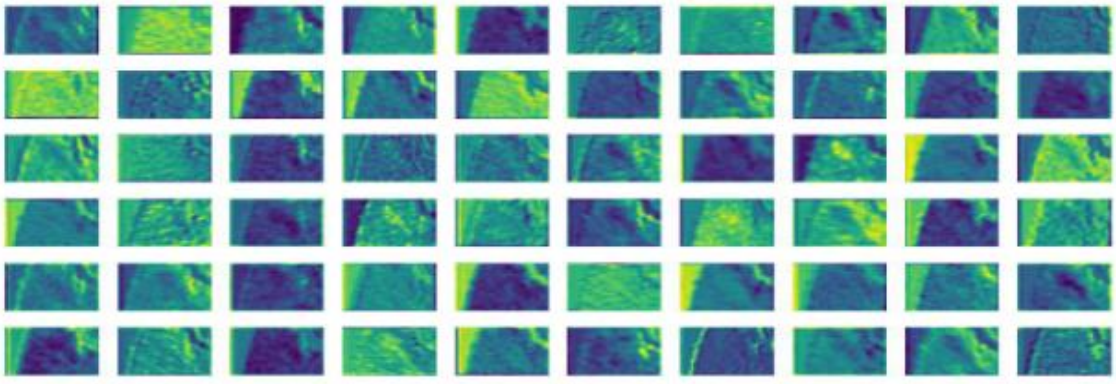


FIGURE 6. Visual interpretation of the intermediate features.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

TABLE 4. Classification accuracy of LMHistNet.

Proposed Model	Accuracy	
	Binary	Multiclass
CGN_B	95%	85.55%
$CGN_B + SFG_{R1}$	95.53%	86.33%
$CGN_B + SFG_{R1} + CBA_M + SFG_{R1}$	97.65%	87.65%
Proposed LMHistNet	99%	88%

E. VISUAL INTERPRETATION OF LMHistNet MODEL FEATURES

The features that the LMHistNet inner layers were able to recover during the training phase are shown in Figure 6. The model's training process for the extracted features is demonstrated visually by the intermediate feature layers. Our understanding of how the LMHistNet model internally understands the breast cancer histopathological pictures has significantly improved as a result of this representation.

F. EVALUATION METRICS OF THE PROPOSED ARCHITECTURE

A number of metrics, including accuracy, precision, recall, and f-score, are used to assess the suggested LMHistNet model.

In our experiment, the labels "positive" and "negative" are utilised for categorization purposes; the former denotes a malignant or cancerous picture. Below is the equation for each. Equation (20) is used to calculate accuracy, and equation (21) is used to get the f1-score. Equation (22) can be used to find precision for binary classification, and equation (23), for multiclass classification. The number of accurately diagnosed cases of breast cancer is represented by TP, the number of correctly classified cases of malignancy is represented by TN, and the number of malignant cases that are classified as benign is represented by FP. False negative, or FN, is the term used to describe the proportion of benign pictures that are mistakenly identified as malignant.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

$$F1 - score = \frac{2 \times precision \times recall}{precision + recall} \quad (21)$$

$$Precision\ for\ binary\ classification = \frac{TP}{TP + FP} \quad (22)$$

$$Precision\ for\ multiclass = \frac{\sum_i TP_i}{\sum_i TP_i + FP_i} \quad (23)$$

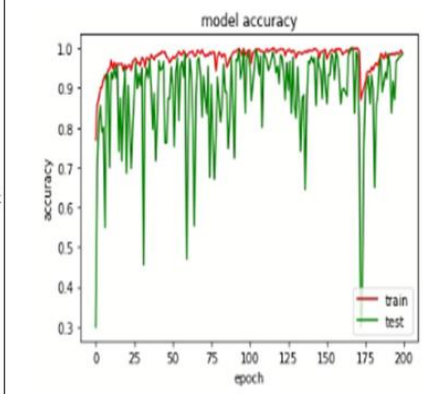
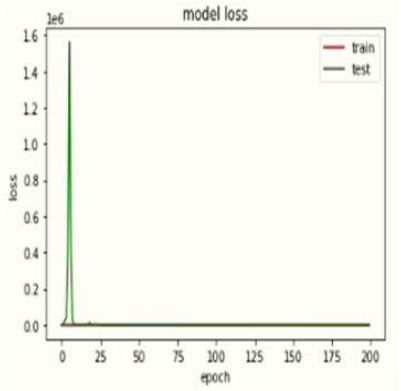
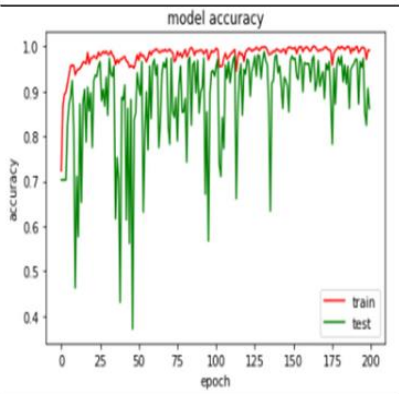
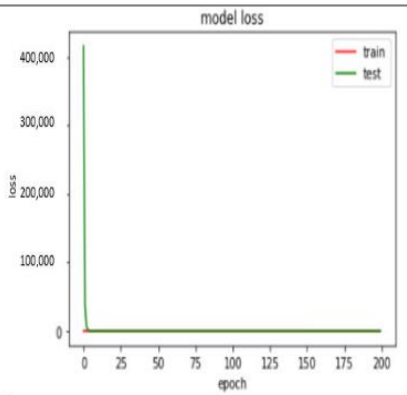
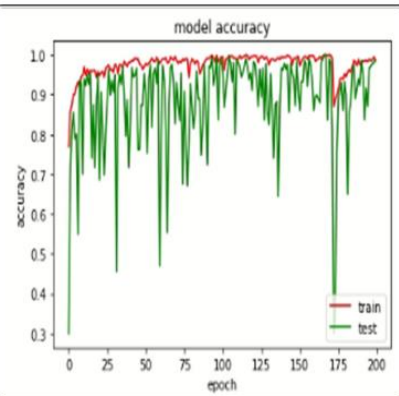
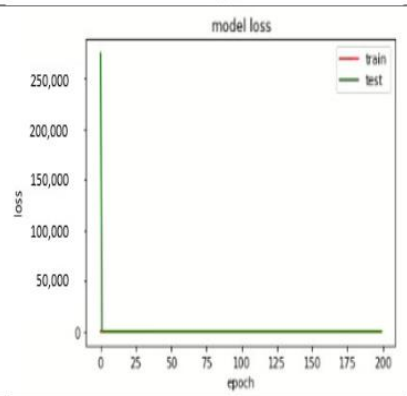
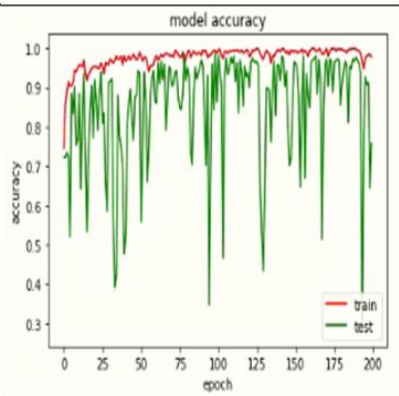
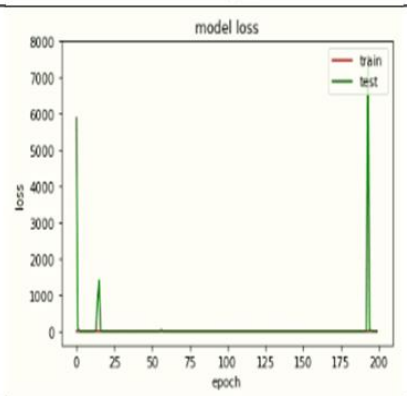
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

TABLE 5. Model accuracy and loss obtained for the binary classification of magnification-dependent and independent classes.

Image Resolution	Accuracy	Loss
Magnification Independent		
40X		
100X		
200X		

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

TABLE 5. *(Continued.)* **Model accuracy and loss obtained for the binary classification of magnification-dependent and independent classes.**

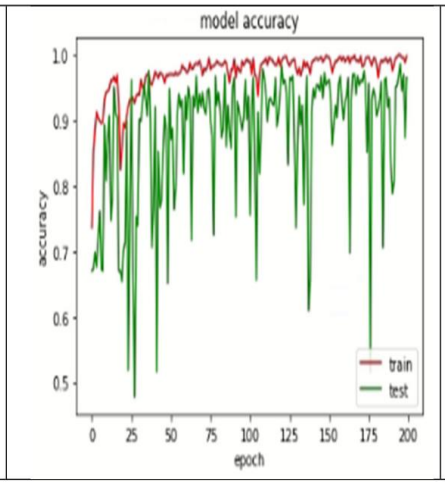
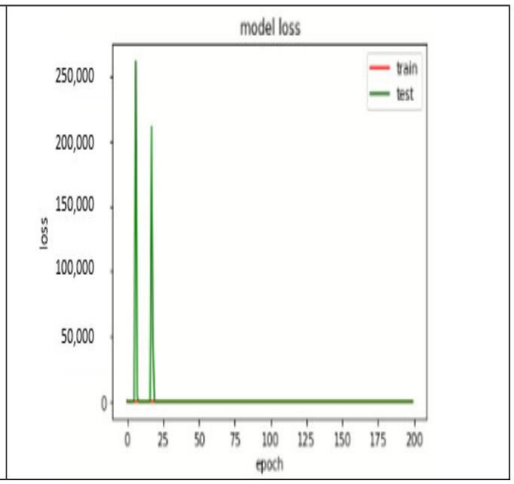
400X		

TABLE 6. **Image count of four different resolution.**

	Adenosis	Fibro Adenoma	Phyllodes Tumor	Tubular Adenoma	Ductual Carcinoma	Lobular Carcinoma	Mucinous Carcinoma	Papillary Carcinoma
40X	114	253	109	149	870	156	205	162
100X	113	260	121	150	903	170	222	142
200X	111	264	108	140	896	163	196	135
400X	106	237	115	130	788	137	169	141

G. ABLATION STUDIES

The impact of various LMHistNet model modules on the categorization of photos related to breast histopathology were investigated using ablation tests. It has been discovered that adding more blocks causes the model's performance to steadily increase.

Numerous criteria, including Accuracy, Precision, Recall, F1-score, and Confusion Matrix, were used to assess the LMHistNet model. The CGNB baseline block of the proposed LMHistNet model is made up of ReLU, group normalisation, and convolution. For binary classification, it achieved 95% classification accuracy, and for multiclass classification, 85.55% classification accuracy. The SFGR1 block is used in conjunction with CGNB, which produced binary and multiclass classification accuracy of 95.53% and 86.33%, respectively.

The CBAM module, which increased the model's discriminative power by suppressing the less informative channels, significantly increased the model's accuracy. A more intricate network architecture is produced by adding SFGR2 to the network. Levenberg-Marquardt further refines the model, yielding 99% and 88% accuracy for binary and multiclass classification of the histopathological breast cancer dataset, respectively. Table 4 presents the findings.

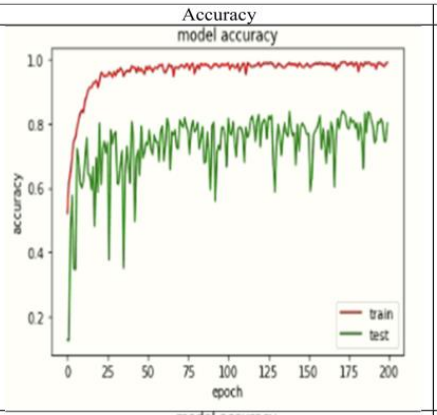
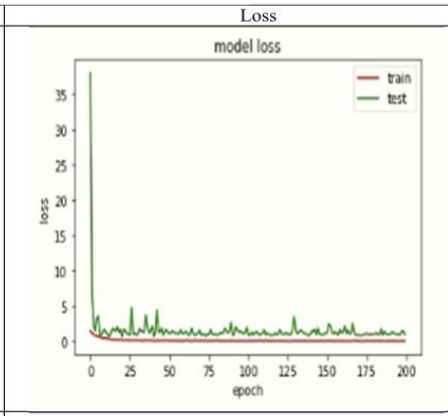
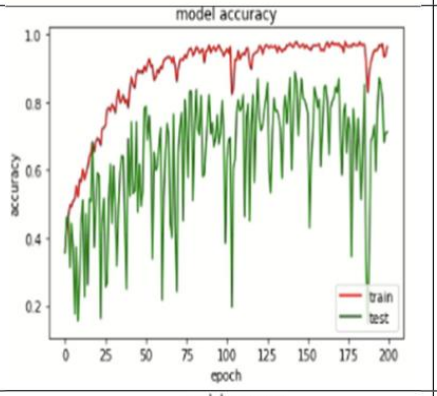
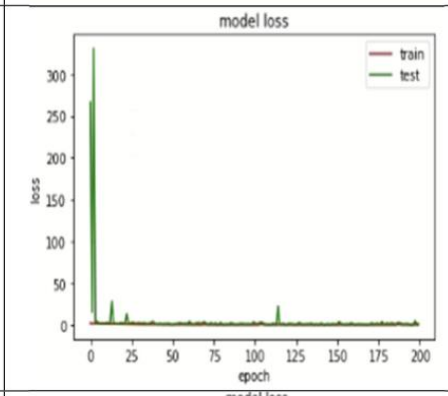
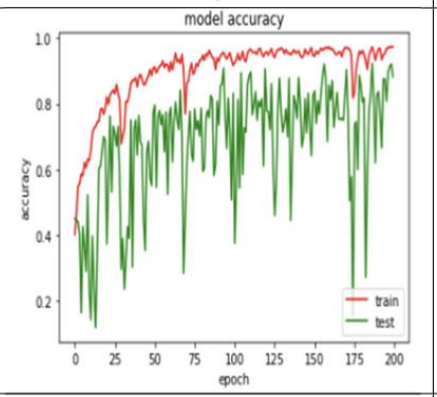
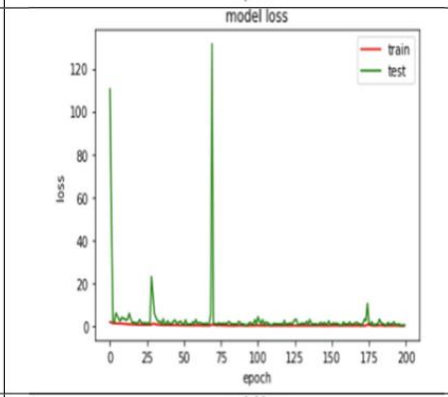
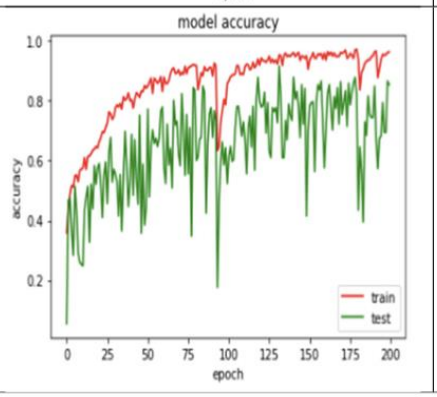
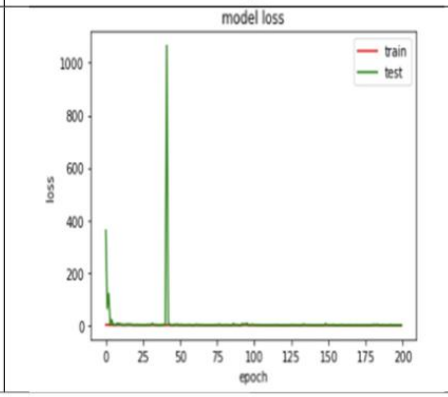
IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

TABLE 7. **Model accuracy and loss for multiclass classification of magnification-dependent and magnification-independent classes.**

Image Resolution	Accuracy	Loss
Magnification-Independent		
40X		
100X		
200X		

DOI: 10.17148/IJARCCE.2024.13812

H. MAGNIFICATION DEPENDENT AND INDEPENDENT BINARY CLASSIFICATION

The images are divided into two classes: benign and malignant, regardless of their resolution. Using the suggested LMHistNet model, a total of 2480 benign and 5428 malignant pictures were examined. Similarly, training and testing was done on magnification-dependent images with 625 benign and 1393 malignant of 40X resolution, 644 benign and 1437 malignant of 100X resolution, 623 benign and 1390 malignant of 200X resolution, and 588 benign and 1253 malignant of 400X resolution. Four distinct magnifications of the histopathological images—40x, 100x, 200x, and 400x—have been employed.varying features in images at varying magnifications lead to increased learning and improved classification accuracy.

The use of random initialisations, varied dropout rates, and batch normalisation led to different convergence pathways, which caused oscillations in the training process, even when the unbalanced data is augmented and regularised to improve generalisation. In order to improve accuracy at all magnifications, the learning rate and step size were adjusted to minimise variations. Table 5 displays the model's accuracy and loss for the binary classification of both classes that depend on magnification and those that don't.

I. MAGNIFICATION DEPENDENT AND INDEPENDENT MULTICLASS CLASSIFICATION

The suggested LMHistNet model was utilised to analyse the classifications into eight classes, both independent and dependent on magnification. The following are the eight distinct classes and the total number of photos that were trained and tested:

Adenosis images (444), fibroadenoma images (1014), phyllodes tumour images (453), tubular adenoma images (569), mucinous carcinoma images (792), lobular carcinoma images (626), ductal carcinoma images (3457), and papillary carcinoma images (580). Table 6 lists the quantity of photos from eight distinct classes at four different resolutions. Table 7 displays the model accuracy and loss obtained for both independent and magnification dependent multiclass categorisation.

J. PERFORMANCE ANALYSIS

For histopathological pictures with a resolution of 100X, LMHistNet achieved a 99% accuracy for the magnification-dependent binary classification. This is greater than the 86%, 85%, and 96% accuracies obtained for images with resolutions of 400X, 200X, and 40X, correspondingly. The pathological 100X resolution photos displayed the highest level of accuracy.99 for both F1 score and recall. Model LMHistNet obtained an F1 score of 88, recall of 86, and accuracy of 88.85 for the binary categorization of the magnification-dependent pictures with a 40X resolution, but the accuracy, recall and F1 scores for the histopathological were 82, 82, and 81 pictures at a 200X resolution. Regarding the binary categorization of histopathological images that is dependent on magnification .

The histopathological analysis yielded an accuracy of 88%, precision, recall, and F1 score of 81, 71, and 73 correspondingly.Using LMHistNet for magnification-dependent multiclass classification, pictures with a resolution of 100X achieved the best classification accuracy. Depending on magnification multiple class categorization, 200X resolution photos possessed the second-highest accuracy (85%), after pictures.At 400X and 40X resolutions, exhibiting accuracy of 71% and 70%, in that order. For the reliant on magnification multiple class categorization of 40X-resolution pictures.The LMHistNet model produced recall of 81% and accuracy of 71, an F1 score of 73, and corresponding values of 85, 86, and 85 for the photos' F1 score, recall, and precision with a 200X resolution. The model attained the recall and precision.

K. COMPARISON OF THE PROPOSED LMHistNet WITH STATE-OF-THE-ART ARCHITECTURES

Binary histopathological analysis was carried out by Chowdhury et al. [21],picture classification with a 99.58% accuracy rate using the BreaKHis dataset. Using Gour et al. [23], a binary 92.5% classification accuracy and the same dataset were utilised. Alzubaidi and colleagues [24] achieved a categorization of four classes.97.4% accuracy rate.

Anuranjeeta and colleagues [20] and Mahmoud [27] also employed the BreaKHis dataset for histopathology image analysis, with 86.8% accuracy achieved. Additionally, 85.7%. [28] Alom et al. and [26] Bardou et al. achieved multiclass and binary classifications, with accuracy of 99.05 for the multiclass classification and 98.33% for the binary 88.23% and 98.59% for the multiclass classification, respectively,categorization in turn. In Gecer et al.'s [29] work, five categorizations of histological pictures and acquired a 55% accuracy rate.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 8, August 2024

DOI: 10.17148/IJARCCE.2024.13812

TABLE 9. Comparison of LMHistNet with state of the art architectures.

Reference	Accuracy	
	Binary classification	Multiclass classification
[21]	99.58	-
[23]	92.5	-
[26]	98.33	88.23%(eight class)
[24]		97.4%
[27]	83.33	-
[29]	-	55%(five class)
[30]	83.3	77.8(four class)
[31]	83.25	-
[32]	-	93.2%(eight class)
[33]	-	88.89%(four class)
[34]	99.70%	97.81%
[35]	99.16	-
[37]	-	99.1(three class)
[38]	86.21%	-
[39]	-	99.17%(three class)
[40]	-	99.32%(three class)
[42]	0.9642	-
[43]	-	97.6%(EIGHT class)
LMHistNet	99%	88%

V. CONCLUSION AND FUTURE WORK

The main focus is on the effective and efficient classification of breast histopathology images, which will aid in the diagnosis procedure. The suggested LMHistNet model was applied to both multiclass and binary classification. The integration of the convolution block attention module with the LMHistNet model helps the model focus more on discriminative characteristics. Through efficient weight optimization, the Levenberg–Marquardt optimization technique enhanced the characteristics. Between these stacked clusters, batch normalisation was used to enhance the feature obtained, enabling better categorization.

The suggested LMHistNet produced encouraging results for binary classification; however, the results for multiclass classification were relatively low due to a number of factors, including the classification into eight classes, which is a large number of classes, and the low number of image samples in each individual class. Using the BreaKHis dataset, classifications that are both magnification-dependent and independent have been carried out. In multiclass classification, the LMHistNet achieved up to 85% accuracy, whereas in binary classification, it achieved up to 99% accuracy.

DOI: 10.17148/IJARCCE.2024.13812

Future research can concentrate on refining pre-processing methods to improve the quality of the images. One major barrier to computer-aided diagnostics is the availability of large datasets. One can look into some shot learning techniques to get around the constraints of small datasets. It is possible to apply deep learning models for multiclass categorization that resemble TNM (tumour node, metastasis) and BRAIDS (breast imaging reporting and data system).

REFERENCES

[1] F.Bray,J.Ferlay,I.Soerjomataram,R.L.Siegel,L.A.Torre,andA.Jemal, ‘‘Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,’’ *CA Cancer J. Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

[2] R.L.Siegel,K.D.Miller,N.S.Wagle,andA.Jemal,‘‘CancerStatistics,’’ *CA Cancer J. Clinicians*, vol. 67, no. 27, pp. 7–30, 2017.

[3] K. L. Britt, J. Cuzick, and K.-A. Phillips, ‘‘Key steps for effective breast cancer prevention,’’ *Nature Rev. Cancer*, vol. 20, no. 8, pp. 417–436, Aug. 2020

[4] Y.-S.Sun,Z.Zhao,Z.-N.Yang,F.Xu,H.-J.Lu,Z.-Y.Zhu,W.Shi,J.Jiang, P.-P. Yao, and H.-P. Zhu, ‘‘Risk factors and preventions of breast cancer,’’ *Int. J. Biol. Sci.*, vol. 13, no. 11, pp. 1387–1397, 2017

[5] Centers for Disease Control and Prevention. *Content Source: Division of Cancer Prevention and Control*. Accessed: Apr. 20, 2022. [Online]. Available: https://www.cdc.gov/cancer/dcpc/about/index.htm

[6] N. Takkar, S. Kochhar, P. Garg, A. Pandey, U. Dalal, and U. Handa, ‘‘Screening methods(clinical breast examination and mammography) to detect breast cancer in women aged 40–49 years,’’ *J. Mid-Life Health*, vol. 8, no. 1, p. 2, 2017.

[7] A. B. Nassif, M. A. Talib, Q. Nasir, Y. Afadar, and O. Elgendy, ‘‘Breast cancer detection using artificial intelligence techniques: A systematic literature review,’’ *Artif. Intell. Med.*, vol. 127, May 2022, Art. no. 102276.

[8] S. Indolia, A. K. Goswami, S. P. Mishra, and P. Asopa, ‘‘Conceptual under- standing of convolutional neural networks—A deep learning approach,’’ *Proc. Comput. Sci.*, vol. 132, pp. 679–688, Jan. 2018.

[9] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, ‘‘CBAM: Convolutional block attention module,’’ in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 3–19

[10] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, ‘‘A dataset for breast cancer histopathological image classification,’’ *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016.

[11] S. Bacha and O. Taouali, ‘‘A novel machine learning approach for breast cancer diagnosis,’’ *Measurement*, vol. 187, Jan. 2022, Art. no. 110233.

[12] B. Dai, R.-C. Chen, S.-Z. Zhu, and W.-W. Zhang, ‘‘Using random forest algorithm for breast cancer diagnosis,’’ in *Proc. Int. Symp. Comput., Consum. Control (IS3C)*, Dec. 2018, pp. 449–452.

[13] S. Kabiraj, M. Raihan, N. Alvi, M. Afrin, L. Akter, S. A. Sohagi, and E. Podder, ‘‘Breast cancer risk prediction using XGBoost and random forest algorithm,’’ in *Proc. 11th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, Jul. 2020, pp. 1–4.

[14] D. A. Omondiagbe, S. Veeramani, and A. S. Sidhu, ‘‘Machine learning classification techniques for breast cancer diagnosis,’’ *IOP Conf. Ser., Mater. Sci. Eng.*, vol. 495, Jun. 2019, Art. no. 012033.

[15] R. Vijayarajeswari, P. Parthasarathy, S. Vivekanandan, and A. A. Basha, ‘‘Classification of mammogram for early detection of breast cancer using SVM classifier and Hough transform,’’ *Measurement*, vol. 146, pp. 800–805, Nov. 2019.

[16] J. Wu and C. Hicks, ‘‘Breast cancer type classification using machine learning,’’ *J. Personalized Med.*, vol. 11, no. 2, p. 61, Jan. 2021.

[17] E.Michael,H.Ma,H.Li,andS.Qi,‘‘Anoptimizedframeworkforbreast cancer classification using machine learning,’’ *BioMed Res. Int.*, vol. 2022, pp. 1–18, Feb. 2022.

[18] S. Ara, A. Das, and A. Dey, ‘‘Malignant and benign breast cancer classification using machine learning algorithms,’’ in *Proc. Int. Conf. Artif. Intell. (ICAI)*, Apr. 2021, pp. 97–101.

[19] M.A.Naji,S.E.Filali,K.Aarika,E.H.Benlahmar,R.A.Abdelouhahid, and O.Debauche, ‘‘Machine learning algorithms for breast cancer prediction and diagnosis,’’ *Proc. Comput. Sci.*, vol. 191, pp. 487–492, Jan. 2021

[20] A. Anuranjeeta, K. K. Shukla, A. Tiwari, and S. Sharma, ‘‘Classification of histopathological images of breast cancerous and non cancerous cells based on morphological features,’’ *Biomed. Pharmacol. J.*, vol. 10, no. 1, pp. 353–366, Mar. 2017.

[21] D. Chowdhury, A. Das, A. Dey, S. Sarkar, A. D. Dwivedi, R. R. Mukkamala, and L. Murmu, ‘‘ABCanDroid: A cloud integrated Android app for noninvasive early breast cancer detection using transfer learning,’’ *Sensors*, vol. 22, no. 3, p. 832, Jan. 2022.

[22] Y.-D. Zhang, S. C. Satapathy, D. S. Guttery, J. M. Górriz, and S.-H. Wang, ‘‘Improved breast cancer classification through combining graph convolutional network and convolutional neural network,’’ *Inf. Process. Manage.*, vol. 58, no. 2, Mar. 2021, Art. no. 102439.

DOI: 10.17148/IJARCCE.2024.13812

[23] M. Gour, S. Jain, and T. Sunil Kumar, ''Residual learning based CNN for breast cancer histopathological image classification,'' *Int. J. Imag. Syst. Technol.*, vol. 30, no. 3, pp. 621–635, Sep. 2020.

[24] L. Alzubaidi, O. Al-Shamma, M. A. Fadhel, L. Farhan, J. Zhang, and Y. Duan, ''Optimising the performance of breast cancer classification by employing the same domain transfer learning from hybrid deep convolutional neural network model,'' *Electronics*, vol. 9, no. 3, p. 445, Mar. 2020.

[25] F. F. Ting, Y. J. Tan, and K. S. Sim, ''Convolutional neural network improvement for breast cancer classification,'' *Expert Syst. Appl.*, vol. 120, pp. 103–115, Apr. 2019.

[26] D.Bardou,K.Zhang,andS.M.Ahmad,''Classification Of Breast Cancer based on histology images using convolutional neural networks,'' *IEEE Access*, vol. 6, pp. 24680–24693, 2018.

[27] M. Mahmoud, ''Breast cancer classification in histopathological images using convolutional neural network,'' *Int. J. Adv. Comput. Sci. Appl.*, vol. 9, no. 3, pp. 64–68, Apr. 2018.

[28] M. Z. Alom, C. Yakopcic, M. S. Nasrin, T. M. Taha, and V. K. Asari, ''Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network,'' *J. Digit. Imag.*, vol. 32, no. 4, pp. 605–617, Aug. 2019.

[29] B. Gecer, S. Aksoy, E. Mercan, L. G. Shapiro, D. L. Weaver, and J.G.Elmore, ''Detection and classification of cancer in whole slide breast histopathology images using deep convolutional networks,'' *Pattern Recognit.*, vol. 84, pp. 345–356, Dec. 2018.

[30] T.Araìujo,G.Aresta,E.Castro,J.Rouco,P.Aguiar,C.Eloy,A.Polìonia, and A. Campilho, ''Classification of breast cancer histology images using convolutional neural networks,'' *PloS One*, vol. 12, no. 6, 2017, Art. no. e0177544.

[31] N.Bayramoglu,J.Kannala,andJ.Heikkilä,''Deeplearningformagnifi- cation independent breast cancer histopathology image classification,'' in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Dec. 2016, pp. 2440–2445.

[32] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, ''Breast cancer multi- classification from histopathological images with structured deep learning model,'' *Sci. Rep.*, vol. 7, no. 1, p. 4172, Jun. 2017.

[33] Y. Li, J. Wu, and Q. Wu, ''Classification of breast cancer histology images using multi-size and discriminative patches based on deep learning,'' *IEEE Access*, vol. 7, pp. 21400–21408, 2019.

[34] H. Aljuaid, N. Alturki, N. Alsubaie, L. Cavallaro, and A. Liotta, ''Computer-aided diagnosis for breast cancer classification using deep neu- ral networks and transfer learning,'' *Comput. Methods Programs Biomed.*, vol. 223, Aug. 2022, Art. no. 106951, doi: 10.1016/j.cmpb.2022.106951.

[35] S. Majumdar, P. Pramanik, and R. Sarkar, ''Gamma function based ensemble of CNN models for breast cancer detection in histopathology images,'' *Expert Syst. Appl.*, vol. 213, Mar. 2023, Art. no. 119022.

[36] M. B. Tayel, M. A. A. Mokhtar, and A. F. Kishk, ''Breast cancer diagnosis using histopathology and convolution neural network CNN method,'' in *Proc. Int. Conf. Innov. Comput. Commun.* Singapore: Springer, 2023, pp. 585–600.

[37] K.Jabeen,M.A.Khan,M.Alhaisoni,U.Tariq,Y.-D.Zhang,A.Hamza, A.Mickus, and R.Damaševičius, ''Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion,'' *Sensors*, vol. 22, no. 3, p. 807, Jan. 2022, doi: 10.3390/s22030807.

[38] X.Wang,I.Ahmad,D.Javeed,S.A.Zaidi,F.M.Alotaibi,M.E.Ghoneim, Y. I. Daradkeh, J. Asghar, and E. T. Eldin, ''Intelligent hybrid deep learning model for breast cancer detection,'' *Electronics*, vol. 11, no. 17, p. 2767, Aug. 2022, doi: 10.3390/electronics11172767.

[39] A. Sahu, P. K. Das, and S. Meher, ''High accuracy hybrid CNN classifiers for breast cancer detection using mammogram and ultrasound datasets,'' *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104292.

[40] A. Altameem, C. Mahanty, R. C. Poonia, A. K. J. Saudagar, and R. Kumar, ''Breast cancer detection in mammography images using deep convolutional neural networks and fuzzy ensemble modeling techniques,'' *Diagnostics*, vol. 12, no. 8, p. 1812, Jul. 2022.