



Article

Attention-Based Ensemble Network for Effective Breast Cancer Classification over Benchmarks

Su Myat Thwin¹, Sharaf J. Malebary² , Anas W. Abulfaraj³ and Hyun-Seok Park^{1,*}

¹ Department of Computer Science and Engineering, Ewha Womans University, Seoul 03760, Republic of Korea; sumyatthwin@ewhain.net

² Department of Information Technology, Faculty of Computing and Information Technology, King Abdulaziz University, P.O. Box 344, Rabigh 21911, Saudi Arabia; smalebary@kau.edu.sa

³ Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, P.O. Box 344, Rabigh 21911, Saudi Arabia; awabulfaraj@kau.edu.sa

* Correspondence: neo@ewha.ac.kr

Abstract: Globally, breast cancer (BC) is considered a major cause of death among women. Therefore, researchers have used various machine and deep learning-based methods for its early and accurate detection using X-ray, MRI, and mammography image modalities. However, the machine learning model requires domain experts to select an optimal feature, obtains a limited accuracy, and has a high false positive rate due to handcrafting features extraction. The deep learning model overcomes these limitations, but these models require large amounts of training data and computation resources, and further improvement in the model performance is needed. To do this, we employ a novel framework called the Ensemble-based Channel and Spatial Attention Network (ECS-A-Net) to automatically classify infected regions within BC images. The proposed framework consists of two phases: in the first phase, we apply different augmentation techniques to enhance the size of the input data, while the second phase includes an ensemble technique that parallelly leverages modified SE-ResNet50 and InceptionV3 as a backbone for feature extraction, followed by Channel Attention (CA) and Spatial Attention (SA) modules in a series manner for more dominant feature selection. To further validate the ECS-A-Net, we conducted extensive experiments between several competitive state-of-the-art (SOTA) techniques over two benchmarks, including DDSM and MIAS, where the proposed model achieved 96.50% accuracy for the DDSM and 95.33% accuracy for the MIAS datasets. Additionally, the experimental results demonstrated that our network achieved a better performance using various evaluation indicators, including accuracy, sensitivity, and specificity among other methods.

Keywords: attention module; breast cancer; deep learning; ensemble technique; feature extraction



Citation: Thwin, S.M.; Malebary, S.J.; Abulfaraj, A.W.; Park, H.-S. Attention-Based Ensemble Network for Effective Breast Cancer Classification over Benchmarks. *Technologies* **2024**, *12*, 16. <https://doi.org/10.3390/technologies12020016>

Academic Editor: Lipo Wang

Received: 13 December 2023

Revised: 8 January 2024

Accepted: 17 January 2024

Published: 23 January 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Breast cancer (BC) is considered the second most widespread deadly disease worldwide, where abnormal cells in the breast begin to spread uncontrollably, threatening the lives of women specifically aged 25 to 50. There are two common types of breast cancer: ductal and lobular cancer. Ductal cancer begins from the nipple, and lobular cancer starts in the gland that produces breast milk. The risk factor of breast cancer is a combination of factors which are genetic and environmental. Genetics is the factors that cannot be changed and controlled, such as family history, early menstruation, late menopause, and dense breast tissue. Secondly, the environmental and lifestyle risk factors that can be changed and controlled are being overweight or obese, lack of physical activity, and drinking alcohol. In 2019, the World Health Organization (WHO) reported 9.8 million deaths related to cancer [1].

According to the International Agency for Research on Cancer, lung cancer for the male gender and BC for the female gender are considered the most deadly diseases

worldwide [2]. BC is the most threatening type of cancer in women, and its early detection can increase the survival rate, which reduces the death ratio around the globe, as proved by medical reports. The main challenge in controlling breast cancer is to be able to provide comprehensive services in diagnosis and treatment. One of the most commonly used imaging modalities used to detect breast cancer in the early stages is mammography, in which normal (benign) and abnormal (malignant) have different categories. However, the detection of benign and malignant tumors in the early stages is very challenging and time-consuming. Various studies have been conducted to understand tumor detections in mammography images. Numerous researchers have used traditional texture analysis [3,4], extreme learning machine (ELM) [5–7], and random forest classifier [7,8] to separate the mammography image from the selected dataset, whether the image is benign or malignant. Many researchers have used the IRMA dataset to classify the mammogram using a histogram-oriented gradient (HOG) and local configuration pattern (LCP). Currently, the Deep Learning Convolution Neural Network (DL-CNN) shows remarkable performance in early BC detection, which provides an end-to-end mechanism for feature extraction [9]. Several DL-CNN-based techniques are developed to analyze medical imaging modalities that classify BC in mammographic images [10]; these techniques include the DL-CNN to classify breast tumors into normal and abnormal tumors. However, this task is quite challenging due to the wide-ranging light and pixel similarity among the affected and non-affected areas of the breast.

From the recent literature, we have observed that there is very limited literature available that can claim to classify BC accurately and efficiently in mammographic images. According to the literature, most of the work in this area has been conducted by traditional machine learning-based methods which require extra expertise in the feature extraction domain. The DL-based techniques overcome the problems associated with traditional ML approaches by providing automatic feature extraction; however, these techniques pose several limitations, including heavyweight, time complexity, and poor performance. In order to overcome these limitations, in this paper, we employ an ensemble technique based on attention modules for effective and efficient BC detection within mammographic data. Further, we conducted a comprehensive evaluation, comparing the proposed ECS-A-Net with SOTA ML and DL techniques using different evaluation metrics for assessment, such as accuracy, sensitivity, and specificity.

The major contributions of the study are available in the following bullet points:

- The publicly available datasets seem limited, which may affect the model performance for effective BC classification. Therefore, this study significantly contributes to employing data augmentation techniques, for instance rotation, flipping, scaling, etc. The data augmentation strategy plays a significant role in enhancing the diversity and quantity of the training and testing sets, leading to improved model performance in terms of accuracy, sensitivity, and specificity.
- This paper presents a novel approach named an Ensemble-based Channel and Spatial Attention Network (ECS-A-Net) for effective BC classification. An ECS-A-Net parallel leverages modified SE-ResNet50 and InceptionV3 for meaning pattern selections, resulting in enhancing model performance for effective BC classification.
- We incorporated Channel Attention (CA) and Spatial Attention (SA) modules in a serial manner to acquire relevant features from the data. In this study, the CA module is responsible for highlighting the specific channel feature map based on relevance to the task, while the SA module is applicable for emphasizing the spatial location within the feature maps.
- To evaluate the generalization capability of the proposed ECS-A-Net, we performed extensive experiments between numerous SOTA techniques using publicly available benchmarks based on various evaluation indicators, including accuracy, sensitivity, and specificity, whereas the comparative analysis indicates that the proposed network achieved high performance for BC classification among SOTA techniques.

The remainder of this paper is organized as follows: Section 2 presents major trends in literature. Section 3 provides a detailed explanation about the novel framework used in the study; in addition, the experimental setup, datasets discussion, and detailed comparison are provided in Section 4. Finally, the conclusions and future direction of the study are discussed in Section 5.

2. Literature Review

The DL-based networks play a significant role in various challenging domains, especially in medical image analysis [11,12]. Therefore, different techniques have been developed to detect BC in its early stage; however, these methods require modification to enhance performance in terms of accuracy, sensitivity, and specificity for effective BC classification and detection. For instance, Aboutalib et al. [13] used the DL-CNN to classify breast cancer using mammography images into malignant, negative, and benign categories. For the experiments, they combined two datasets and achieved 76% to 91% accuracy. Platania et al. [14] used a new framework based on CNN for automated breast cancer detection and classification. They used a Digital Database for Screening Mammography (DDSM) consisting of three classes and achieved 90% detection and 93.5% classification accuracy. The next approach presented by Jannesari et al. [15] used fine-tuned pre-trained deep neural networks. They used the ResNetV1-50 pre-trained model to predict four types of cancers, including breast, bladder, lung, and lymphoma. They used a benchmark dataset, consisting of 7909 images of 82 patients by name of the BreakHis database. Khuriwal et al. [16] used a DL model for breast cancer diagnosis. Their method includes three basic steps, such as (1) pre-processing (data collection and filtration), (2) splitting the data into train and test sets and drawing a graph for visualization. Finally, a model is developed, and data is fed into it. The next approach presented by Li et al. [17] used an improved DenseNet model. In their approach, they used a mammographic images dataset for the training of four different models, i.e., AlexNet, VGGNet, DenseNet, and GoogleNet. In the experiments, the DensNet model achieved better accuracy compared to the others.

Khan et al. [18] used a DL framework for the detection and classification of breast cancer using transfer learning. They extracted the features using different pre-trained CNN models, namely GoogLeNet, VGGNet, and ResNet; they used a fully connected layer for the classification of malignant and benign cells using average pooling classification and achieved better results compared to others. The next approach presented by Tsochatzidis et al. [19] used the DL-CNN for breast cancer classification from mammograms with mass lesions. They used pre-trained models and achieved mismatch accuracy using two different datasets, DDSM and the Curated Breast Imaging Subset of DDSM (CBIS-DDSM). The next approach presented by Zhou et al. [20] used the DL model to anticipate lymph node metastasis in patients with clinically lymph node-negative breast cancer based on images of primary breast cancer. They used three pre-trained DL architectures, i.e., Inception V3, Inception-ResNetV2, and ResNet-101, and achieved an accuracy of 0.89%. Next, Shen et al. [21] used a CNN model for the distributed screening of mammogram images, which attained excellent results in comparison with previous state-of-the-art techniques. They used a single model and reached per-image accuracy of 88%, and the average of the four models improved the AUC to 91%. The authors used two datasets in this research, namely full-field digital mammography (FFDM) and CBIS-DDSM.

The next approach was presented by Ismail et al. [22], where they employed two different DL models for breast cancer classification into normal and abnormal tumors. They used VGG16 and ResNet50 models, which achieved 94% and 91.4% accuracy, respectively. They used the IRMA database for their experiments. Charan et al. [23] used a DL neural network to classify normal and abnormal breast cancer. They used the Mammograms-MIAS dataset in this research to train the model; firstly, they classified the normal and abnormal types of cancer and then they divided the abnormal class into a further six classes and found other types of abnormalities in the breast. The next method presented by Sha et al. [24] used a CNN model for the optimal diagnosis of breast cancer by mammographic images.

They used CNN to segment the region of interest (RoI), and several features were extracted to decrease time complexity and increase precision. The extracted feature was fed to SVM classifiers for classification. They performed their experiments on two benchmark datasets, MIAS and DDSM, and reached an accuracy of 92%. Kaur et al. [25] used a pre-processing technique and built-in feature extraction using K-mean clustering for Speed-Up Robust Features (SURFs) selection. They used the automated DL method using K-mean clustering with MSVM, which is better than using a decision tree model. Ten-fold cross-validation was used, and the acquired results for the SVM, K-nearest neighbor (KNN), linear discriminant analysis (LDA), and decision tree were 96.9%, 93.8%, 89.7%, and 88.7%, individually. They used the Mini MIAS dataset in this research.

Arefan et al. [26] used a DL model which anticipates short-term breast cancer risk using normal screening mammogram images. They used the GoogleNet model and achieved 73% accuracy. This paper showed that the GoogleNet-LDA model outperformed the end-to-end GoogLeNet model. These two DL models achieved a better result than the mammographic breast density. In this approach [27], the authors provided comparative results of two deep learning models, such as VGG16 and ResNet50, by using mammographic images. In this experiment, the VGG16 model reveals good results in terms of accuracy compared to ResNet50. The accuracy of VGG16 and ResNet50 reached 94% and 91.7%, respectively. Similarly, Dhahbi et al. [28] proposed an innovative framework for feature extraction, focusing on curvelet transform and moment theory for effective BC classification. Rocha et al. [29] introduced a novel texture analysis technique based on Local Binary Patterns (LBPs) to generate diverse representations of region of interest (ROI) images. These representations, concatenating methods like histograms and co-occurrence matrices, enable comprehensive texture analysis. Additionally, ecological indexes (Shannon, McIntosh, Simpson, Gleason, Menhinick) are applied to these representations as unique texture descriptors. Eun et al. [4] presented a texture analysis followed by random forest for BC detection using MRI modality. A CNN with multiscale feature extraction capability was proposed by Ranjbarzadeh et al. [30]; they used MIAS and DDSM datasets and achieved an average accuracy of 89.90% and 91.70%, respectively. Furthermore, Saran et al. [31] employed the ResNet50 model with a data augmentation approach utilizing the used MIAS and DDSM datasets and obtained average accuracies of 94.0% and 96.0%, respectively. Besides that, Tsochatzida et al. [32] presented a novel framework for highlighting the infected region in mammographic images. Their study is focused on exploring the integration of content-based image retrieval (CBIR) into computer-aided diagnosis (CADx) to enhance radiologists' decision making when characterizing mammographic masses.

As follow-up research, Bokade et al. [7] fused the deep learning-based features extractor with a random forest classifier for an accurate BC classification. Aslan et al. [33] proposed a hybrid model, integrated a CNN and Bidirectional Long Short Term Memories (BiLSTMs), for breast cancer diagnosis without the region of interest requirement, and they obtained a higher performance. A self-attention-based hybrid model with a random forest classifier for effective BC classification is proposed in [8]. A lightweight hybrid model integrated with depth-wise separable, multi-head self-attention for BC classification was proposed by Zhou et al. [34]. Similarly, Iqbal et al. [35] presented a lightweight transformer model with feature fusion capability for BC segmentation and classification. Sapate et al. [36] introduced a hybrid model based on connected component labeling and adaptive fuzzy region growing techniques for effective BC detection. To validate the proposed network, they utilized two publicly available datasets, where the results indicated that the proposed model achieved a high performance in terms of sensitivity, specificity, and FPsI. Similarly, another study also focused on BC detection; after detecting the infected region, the major aim of their proposed framework was to classify masses into two different classes using multi-resolution analysis, and these classes include benign and malignant [37]. However, their performance requires further enhancement in terms of accuracy for both classes.

Al-Antari et al. [38] developed an automatic system based on Deep Belief Network (DBN) for BC detection in mammogram data. They validated the performance of their proposed network using a digital mammograph database. In addition, their proposed network is focused on two ROI extraction techniques, including multiple mass ROI and whole mass ROI. Moreover, Muramatsu et al. [39] proposed a CNN-based method for BC classification, where the authors applied a data augmentation technique to enlarge the size of the data, resulting in optimal performance in terms of sensitivity, specificity, and accuracy. Furthermore, an ensemble classifier model, which utilizes eigenvalue, bagging, KNN, SVM, and Naive Bayes classifiers, was used to determine whether a mammogram contains malignant, benign, or normal tumors, proposed in [40,41]. However, the limited performance and handcrafted feature extraction mechanisms restricted these systems from real-world implementations. Several CNN-based methods for benign and malignant tumor classification are proposed in [42–44]. However, the network structure and limited performance in terms of high false positive rates need to be improved.

Jahangeer et al. [45] used a series of various networks and VGG-16 for effective BC segmentation. Their proposed study is mainly focused on image preprocessing, where they applied different preprocessing strategies, including salt and pepper filtering techniques, resulting in optimal performance. Similarly, Saber et al. [46] proposed a CNN network using transfer learning techniques for accurate BC detection and classification. Another study presented by Byra et al. [47] developed a technique based on deep meta-learning for effective BC mass classification using ultrasound images. Muralikrishnan et al. [48] proposed a study where they provided a comparison between two different networks within an effective BC classification domain; although, this study requires further enhancement in performance in terms of accuracy, precision, recall, and F1-score.

As evident from the existing literature, the recent approaches reveal several limitations, for instance, DL-based approaches for effective BC classification pose substantial challenges; these approaches have high time complexity and a high number of parameters, making them unsuitable for real-time implementation. Additionally, attention-based techniques, with their potential to enhance model performance for BC detection, have been inadequately explored, particularly in terms of achieving accurate BC detection. Furthermore, the existing ML- and DL-based approaches need further improvement in terms of accuracy, precision, recall, and F1-score for effective performance. In addition, publicly available datasets exhibit several limitations, with one of the limitations being the scarcity of data. This challenge necessitates the implementation of a creative data augmentation strategy to address the data scarcity problem. Therefore, in this study, we propose an efficient and effective network using data augmentation that can accurately classify and localize BC in mammographic images. The detailed explanation of the proposed network is discussed in the subsequent section.

3. The Proposed Methodology

This section presents a detailed explanation of a novel framework that leverages modified SE-ResNet50 and InceptionV3 networks for feature extraction and attention modules, including CA and SA and dominant feature selection, as presented in Figure 1. In this study, the proposed framework comprises two main phases, including data preprocessing and model training/testing, aimed at effective BC classification.

Initially, a preprocessing strategy is employed, involving data augmentation techniques such as rotation, horizontal flip, contrast stretching, and scaling, aiming to enhance the size of the publicly available datasets for achieving high performance. In addition, the input data are then processed to an ensemble technique, utilizing modified SE-ResNet50 and InceptionV3 networks for meaningful feature selection. Subsequently, the outputs from both networks are then fused and passed to the CA and SA modules for dominant feature selection. Finally, we incorporated two fully connected and a classification layer with Softmax as an activation function for final classification. Detailed explanations about the ensemble technique and attention modules are provided in the subsequent sections.

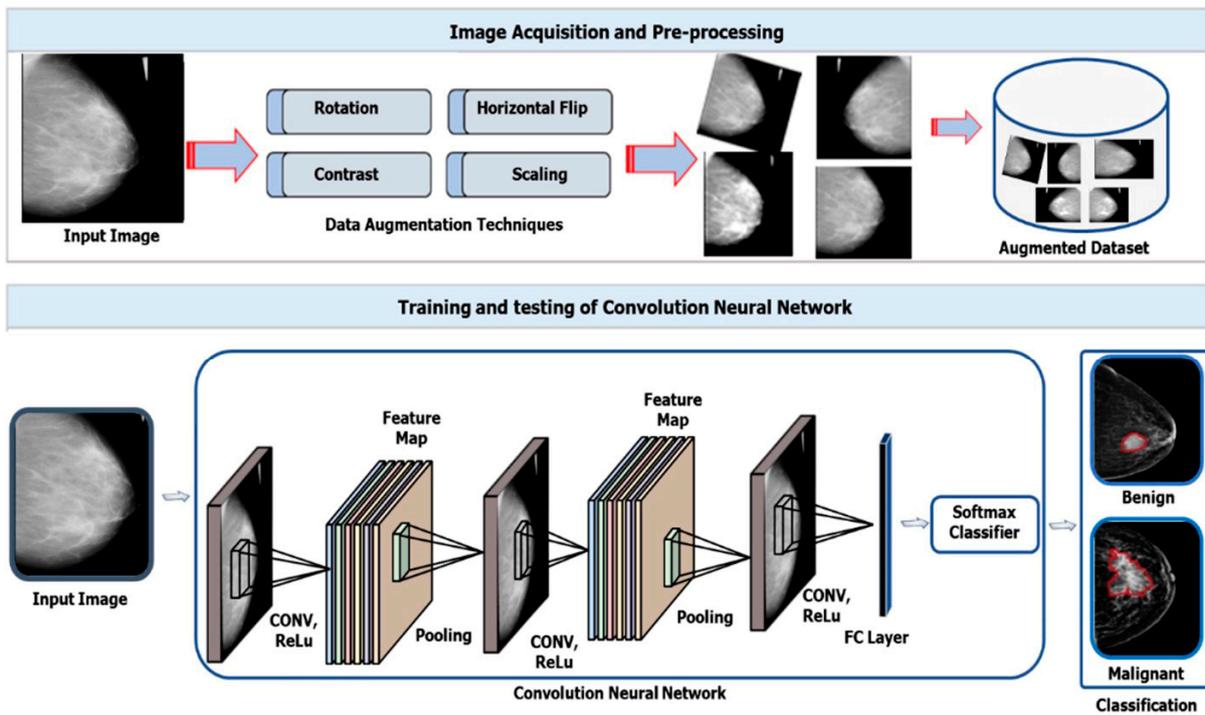


Figure 1. A high-level framework of the ECS-A-Net for accurate BC classification.

3.1. The SE-ResNet50 Network

The Squeeze-and-Excitation Residual Network 50 (SE-ResNet50) is a pre-trained deep neural network (DNN) with 50 layers and is considered a tiny network among other versions of residual networks. The SE-ResNet50 has an extra identity mapping capability compared to other DP learning algorithms. This network is based on delta, which is essential to access the final prediction from layer one to the next layers [49–51]; in addition, it also plays a vital role in resolving the vanishing gradient problem during training of a network. The internal design of SE-ResNet50 enables the network to acquire complex patterns and also enhances the network performance for image recognition tasks. Similarly, the utilization of identity mapping in ResNet architecture allows the network to avoid irrelevant CNN weight layers, which helps in overfitting problems to the training set.

The SE block is a novel approach in ResNet50, allowing networks to improve the quality representation of the features produced by the network [52]. The SE block is incorporated in the internal architecture of the current ResNet50 network by placing it after the nonlinearity following each convolution. In this modification, SE block transformation serves as the non-identity branch of a residual module. Similarly, this block also enhances channel interdependencies with minimal computational cost overhead, leading to optimal performance. The internal mechanism of ResNet50 and the incorporation of SE-block within ResNet50 are shown in Figure 2.

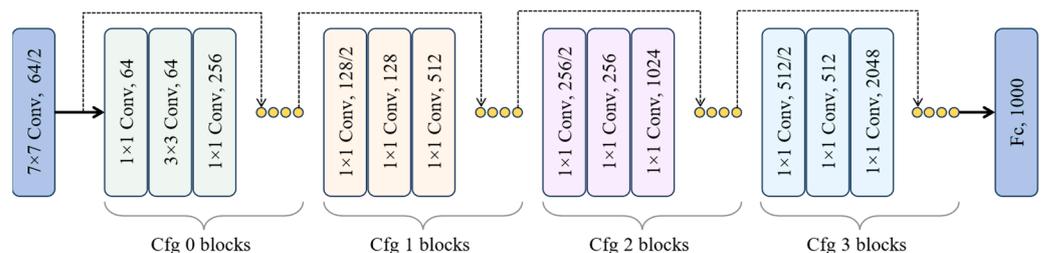


Figure 2. The internal architecture of ResNet50.

3.2. InceptionV3 Architecture for BC Classification

The Google research team aims to reduce the time complexity of CNNs without compromising on its producible results. Therefore, Christian et al. developed a new module in CNN models named the Inception module, which includes 1×1 , 3×3 , and 5×5 convolutional filters. These layers are used in parallel in the network and the sub-sampling (pooling) layers in each block to decrease the time complexity. The Google research team has also developed several more complex models, but in this, InceptionV3 attains good results. InceptionV3 consists of nine inception module blocks that include different convolution filters, as mentioned above. In InceptionV3, there is a total of four max-pool layers, one average pool, one fully connected, and one SoftMax layer, as depicted in Figure 3.

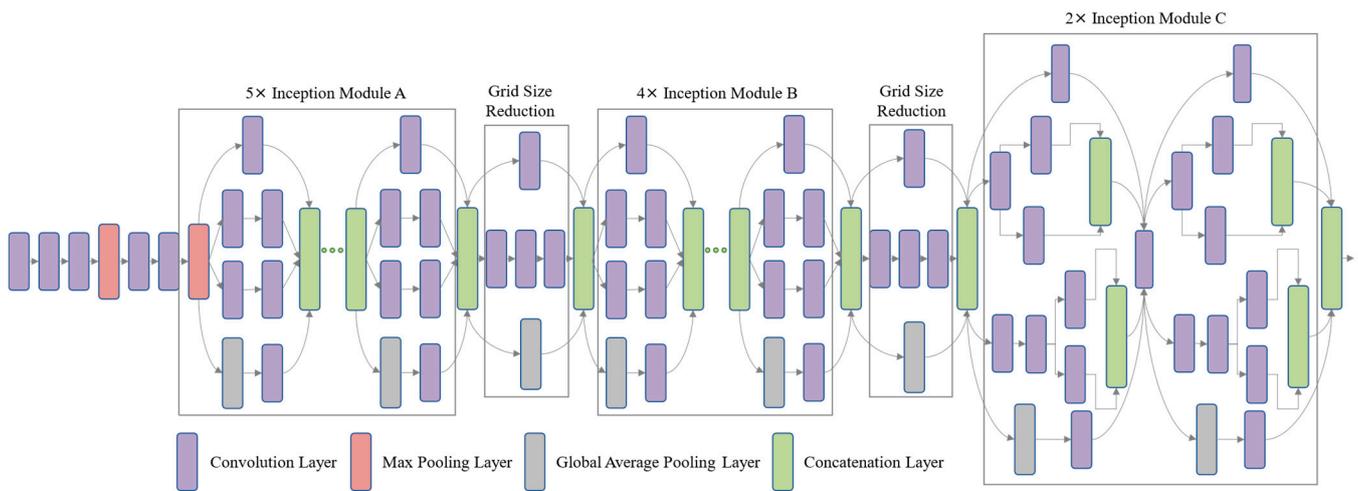


Figure 3. A detailed internal diagram of the pre-trained Inception module [53].

3.3. Channel Attention Module

Various studies have highlighted the importance of attention modules in human perception [54–57]. In recent decades, attention modules have played a significant role in the computer vision domain and have become essential components across diverse tasks. Significantly, CA modules have been integrated into different computer vision applications to enhance the generalization capability of the proposed network [58].

In the recent literature, the deep learning approaches in the visual domain have begun to significantly involve basic convolution operations. Therefore, these approaches lead to the loss of global patterns, and they are only capable of extracting local patterns from input data. To cope with these issues, in this study, we incorporated a channel module to highlight the relevant features before passing it to the SA module, as presented in Figure 4a.

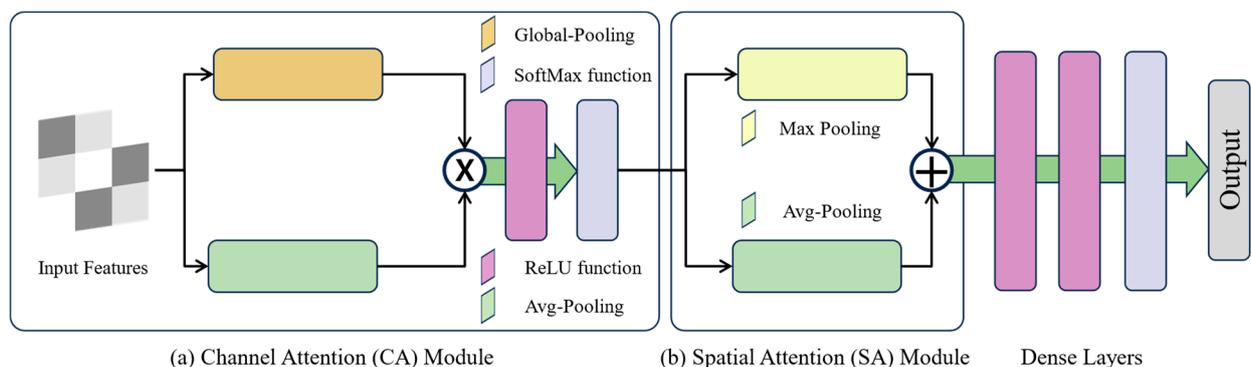


Figure 4. Visual representation of CA, SA, and fully connected layers.

In the attention module, Global Average Pooling (GAP) enables the network to extract channel-wise statistical information. For every channel within an input feature map, the statistical characteristic of the channel is acquired through the following equation:

$$z_c = GAP(x_c) = 1/(H \times W) \sum_{i=1}^H \sum_{j=1}^W x_c(i, j) \quad (1)$$

In the above equation, z_c indicates the statistical information of a channel; similarly, x_c represents the input features map. In addition, H and W demonstrate the height and width of the input features map, respectively, while i and j indicate the spatial coordinates of the input feature map. It is observed that the Channel Attention module utilizing GAP is equivalent to the squeeze-and-excitation (SE) block [52].

The obtained information, represented as z , can be utilized to assess interconnections among channels using a combination of two fully connected layers with ReLU activation and the Softmax nonlinear activation function.

$$s = \sigma(W_2 \delta(W_1 z)) \quad (2)$$

In the given equation, σ and δ indicate Sigmoid and ReLU nonlinear activation functions, respectively. Similarly, $W_1 \in R^{C \times C/r}$ and $W_2 \in R^{C/r \times C}$ are considered the weights of dense layers. We employ a reduction ratio r to decrease the node count in W_1 and enhance the feature map count by C in W_2 . Specifically, in this case, a channel reduction factor of 16 is used. The scaling factors, determined through the above equation, are then multiplied with the input feature map for each channel.

$$\hat{x}_c = s_c \times x_c \quad (3)$$

Here, s_c represents the c -th channel of the scaling factors, and \hat{x}_c denotes the rescaled feature map's c -th channel. In our study, the channel module plays a significant role in controlling global information within the ensemble technique.

3.4. Spatial Attention Module

In the internal architecture of the proposed network, we incorporated the Spatial Attention (SA) module to further evaluate the intermediate pattern of the ensemble technique. Generally, spatial maps are created by analyzing the inter-spatial relationship within the features. Unlike CA, the SA module significantly emphasizes the dominant elements within spatial areas. In this study, the internal architecture of the SA module is focused on a parallel integration of max pooling and average pooling layers, as presented in Figure 4b. The outputs of both layers are then significantly concatenated to produce an effective feature descriptor. The utilization of pooling operations has given priority to relevant or critical regions within features, resulting in the enhancement of the model generalization capability. The parallel incorporation of max pooling and average operations is an effective way to emphasize the relevant information within the features. The mathematical calculation of max and average pooling is given in Equations (1) and (2).

$$\Omega M_a \sum = avg - P(f) \quad (4)$$

$$\Omega M_a \lambda = max - P(f) \quad (5)$$

In the above equations, \sum indicates avg pooling and another notation, while λ indicates max. In addition, the output feature maps of both pooling operations are fused by addition operations and passed to the convolution layer, followed by a ReLU activation function, which is responsible for making a two-dimensional SA feature map. The mathematical process is given below:

$$M_a P(f) = \left(\times \left(f^{1 \times 1} \left(\times \left(f^{3 \times 3} \Omega M_a \sum + \Omega M_a \lambda \right) \right) \right) \right) \quad (6)$$

In this equation, f demonstrates the size of the filter within the convolution layers of the SA module. The MSA map represented as $M_aP(f)_{GAP}$ is obtained by performing the Global Average Pooling (GAP) operation on the feature maps of $M_aP(f)$. Afterward, the result of the GAP operation is combined with the output of the function f through concatenation, as presented in Equations (7) and (8).

$$M_aP(f)_{GAP} = GAP(f) \quad (7)$$

$$f_s p a f = \times [M_aP(f)_{GAP}, f] \quad (8)$$

After the concatenation step, the generated feature maps referred to as $f_s p a f$ are subjected to batch normalization. Subsequently, these normalized feature maps are integrated with Ω to produce $f_s p a f$.

$$f_s p a f = \times [BatchNormalization, \Omega] \quad (9)$$

Finally, the output of SA is transferred to three fully connected layers, with 64, 32, and 2 numbers of neurons, respectively. In this study, we utilized a SoftMax activation function to categorize the input images based on their corresponding classes, as shown in Figure 4.

3.5. The Proposed Model

The proposed model is focused on a parallel connection of DL-based pre-trained networks as a backbone for meaningful feature extraction from the input data, followed by CA and SA modules in a serial manner to further emphasize the dominant features for effective BC classification. In this study, the proposed ECS-A-Net is based on modified SE-ResNet50 and InceptionV3, where SE-ResNet50 is employed to capture intricate patterns within the input data of BC images. This involved a careful modification process that included the exclusion of certain layers deemed less relevant to our task. This selective modification enhances the network's ability to discern subtle features crucial for accurate BC classification. Leveraging the power of residual connections and channel-wise feature recalibration, our modified SE-ResNet50 ensures the effective learning of hierarchical features, thereby significantly improving the model's discrimination capabilities. InceptionV3, on the other hand, is another advanced CNN architecture characterized by its inception modules. These modules employ multiple convolutions with different kernel sizes in parallel, enabling the network to capture features at various scales. InceptionV3 exhibits at capturing both fine-grained details and global patterns within the images. The utilization of InceptionV3 into our framework enables the proposed network to understand the diverse texture, shapes, and structure of cancer within input data, and it plays a significant role in further boosting the model performance for classification tasks, especially in the medical domain. Within InceptionV3, specific layers were excluded to streamline the architecture for BC classification. This adjustment optimizes the model's ability to capture diverse textures, shapes, and structures indicative of cancer within input data. InceptionV3's advanced architecture, characterized by inception modules employing multiple convolutions in parallel, is harnessed more effectively after these modifications.

To fuse the distinctive features captured by SE-ResNet50 and InceptionV3, their outputs are concatenated. Subsequently, this concatenated feature map undergoes a two-step refinement process. First, the CA module is applied to dynamically recalibrate channel-wise patterns, emphasizing crucial features and suppressing less relevant ones. Additionally, we then incorporated the SA module to selectively highlight significant spatial regions, refining the feature map further. By incorporating these modules, the proposed ECS-A-Net not only benefits from the diverse capabilities of SE-ResNet50 and InceptionV3 but also fine-tunes the feature representation to focus on the most relevant information for accurate breast cancer classification. Lastly, the proposed ECS-A-Net incorporates fully connected layers, enabling the network to learn complex relationships among the extracted features. These layers perform the final classification, mapping the learned features to specific BC classes.

The combination of SE-ResNet50 and InceptionV3, along with the attention mechanisms and fully connected layers, allows the network to learn both local and global features within the input data.

The step-by-step procedure of implementation is highlighted in Algorithm 1, which is significantly proposed for the ECS-A-Net. Initially, we applied various augmentation techniques on publicly available datasets to enhance the size of the data, resulting in optimal performance. We then categorized the augmented data into training, validation, and testing sets and passed them into an ensemble network, which leverages InceptionV3 and SE-ResNet50 parallelly as a backbone for feature extraction. The output of both the networks is then fed to attention-based modules; these modules include CA and SA in a serial manner for further dominant pattern selection. Lastly, we incorporated a fully connected layer for the final effective BC classification. The proposed model was determined using different evaluation metrics, such as precision, recall, F1-score, and accuracy. In conclusion, this comprehensive analysis facilitates precise and reliable BC classification, making our model more robust and well suited for medical diagnosis and decision making.

Algorithm 1: Training and Testing Procedure of ECS-A-Net

```

1: Start
2: Load the dataset  $\mathcal{D}$  from repository, split  $\zeta$ , Flipping  $\mathcal{F}$ , Scaling  $\mathcal{S}$ , Rotation  $\mathcal{R}$ 
3: Channel Attention  $c$ , Spatial Attention  $s$ 
4:  $\lceil$  = data preprocessing [Flipping  $\mathcal{F}$  [ $\mathcal{D}$ ], Scaling  $\mathcal{S}$  [ $\mathcal{D}$ ], Rotation  $\mathcal{R}$  [ $\mathcal{D}$ ]]
5: Load pre-trained weights = InceptionV3 and SE-ResNet50
6: ECS-A-Net = InceptionV3 + SE-ResNet50 +  $c$  +  $s$ 
7: Initialization:  $\mathbb{C}$  = (ECS-A-Net,  $\zeta$ [0.7, 2.0, 0.1]), parameter = {epochs = 30, (optimizer = SGD,
learning rate = 0.001, momentum = 0.9, validation = per epoch, batch size = 32)}
8: [training data:  $\mathcal{J}_\lceil$ , validation data:  $\mathcal{V}_\lceil$ , testing data:  $\mathcal{T}_\lceil$ ]
9:  $(\mathcal{J}_\lceil, \mathcal{V}_\lceil, \mathcal{T}_\lceil) = (\zeta[0], \zeta[1], \zeta[2])$ 
10:  $\mathcal{J}_\lceil = \text{random}(\lceil, \mathcal{J}_\lceil)$ 
11:  $\mathcal{V}_\lceil = \text{random}(\lceil, \mathcal{V}_\lceil)$ 
12:  $\mathcal{T}_\lceil = \text{random}(\lceil, \mathcal{T}_\lceil)$ 
13: Return  $\mathcal{J}_\lceil, \mathcal{V}_\lceil, \mathcal{T}_\lceil$ 
14:  $[\lceil, \uparrow, \mathbb{E}] = \mathbb{C}(\mathcal{V}_\lceil, \mathcal{T}_\lceil, \text{parameter})$ 
15:  $[\text{precision}, \text{recall}, \text{F1 - score}, \text{accuracy}] = \lceil(\mathcal{T}_\lceil)$ 
16: Output  $[\text{precision}, \text{recall}, \text{F1 - score}, \text{accuracy}]$ 
17: End

```

4. Results

This section provides a detailed explanation of the experimental setup, performance parameters, dataset, and a comparison between SOTA techniques to validate the generalization capability of the proposed network. In this work, we performed all the experiments on NVidia GPU GTX 2060 GPU with 16 GB, 2.8 GHz processor, and 512 GB memory space; similarly, we utilized a DL framework Keras with backend TensorFlow 2.8.0 and python version 3.7.0, as detailed in Table 1.

Table 1. Software specification for the proposed framework.

Name	Configuration
Operating System	Window 10
Integrated Development Environment	Jupyter Notebook, Python 3.7.2
Libraries	Matplotlib, Keras, TensorFlow, NumPy
Imaging Libraries	Scikit-Image, Scikit-Learn, OpenCV

To fairly determine the performance of the proposed ECS-A-Net, in this work, we performed extensive experiments between different SOTA techniques using three var-

ious evaluation metrics, including accuracy, sensitivity, and specificity, collecting from a confusion matrix, which is based on four parameters such as true positive (TP), true negative (TN), false positive (FP), and false negative (FN). Additionally, the mathematical formulation of accuracy, precision, recall/sensitivity, and specificity can be seen in Equations (10)–(12), and detailed explanation is given in [59–62].

$$Accuracy = \left(\frac{\text{Number of accurate classification}}{\text{Total number of classification}} \right) \quad (10)$$

$$Precision = \left(\frac{TP}{TP + FP} \right) \quad (11)$$

$$Recall/Sensitivity = \left(\frac{TP}{TP + FN} \right) \quad (12)$$

4.1. Datasets Description

In this work, we used two publicly available benchmarks, including DDSM and MIAS. The DDSM was taken from the University of South Florida, maintaining the accessibility of the dataset on the web. The dataset contains 55,890 total images and includes three different classes, for instance benign, malignant, and normal, as shown in Figure 5. For the experimental results, we utilized a small version of the DDSM [40] for classification. Similarly, the MIAS [40] dataset contains three different classes, including normal, benign, and malignant, where it contains 209, 62, and 51 numbers of images for each class, respectively. This dataset also represents details for ground-truth information in the mammogram images, for instance, abnormality center coordinates, background tissue, tumor type, and approximate radius for enclosing the abnormality circle. The visualization samples of the MIAS datasets are presented in Figure 6. In the training phase, we utilized both datasets with a ratio of 70% for training, 20% for validation, and 10% for testing.

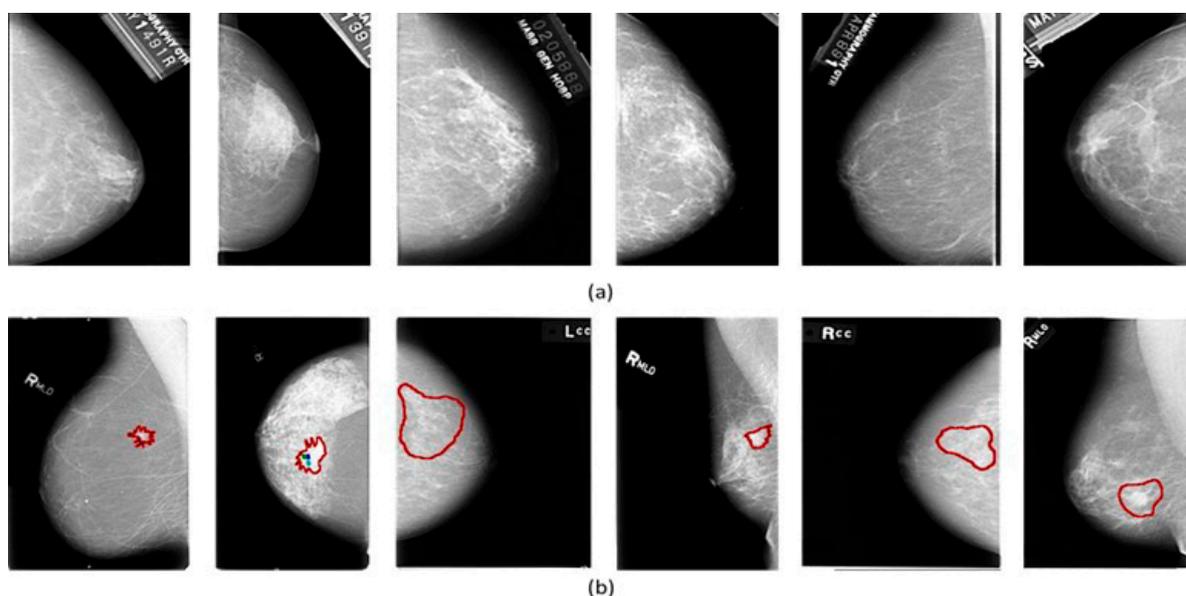


Figure 5. Sample images from DDSM dataset: (a) show benign images and (b) malignant.

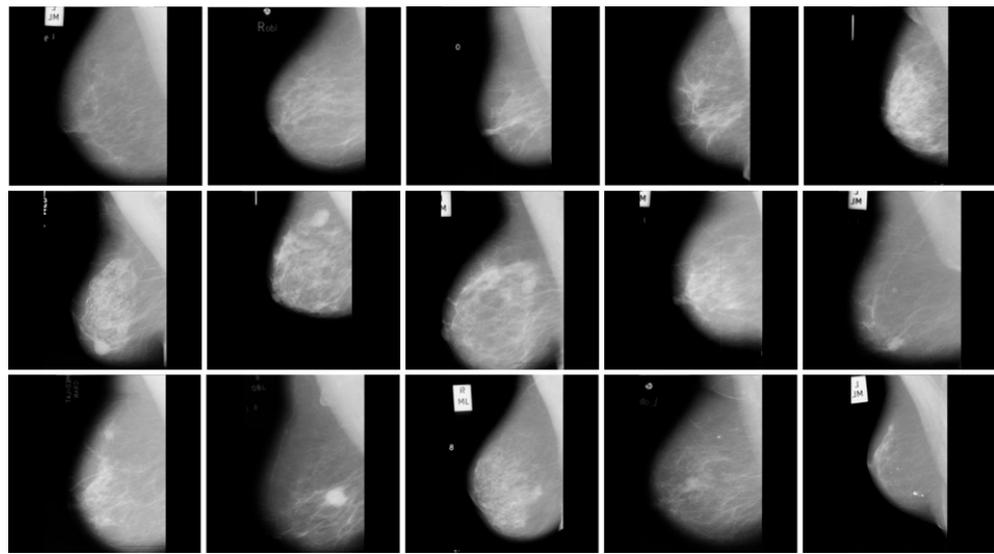


Figure 6. Sample images from MIAS dataset.

4.2. Results and Discussion

In this section, we provided the experimental results of the proposed model performance using two benchmarks, focusing on the accuracy, loss graphs, and confusion matrix. For a fair evaluation of both datasets, we trained the proposed network for 30 epochs with batch size and learning rate momentum set at 32, 0.001, and 0.9, respectively.

4.2.1. Comparative Analysis of DDSM Dataset

In Figure 7a,b, the horizontal axis represents the number of epochs, while the vertical axis reflects the corresponding accuracy and loss ratios. The blue color shows training accuracy and the orange color indicates validation accuracy; notably, our model achieved a training accuracy of 0.99 and a validation accuracy of 0.97. In Figure 7c, we provide the confusion matrix, which offers insights into the model’s accurate and erroneous predictions. The confusion matrix provides a class-wise classification report of the proposed model using a test set from the DDSM dataset, as depicted in Figure 7c. Here, we observe that the benign class achieves an accuracy level of 0.98, the malignant class 0.97, while the normal class obtains an accuracy of 0.95. In addition, the model performance for each class individually over DDSM is presented in Table 2.

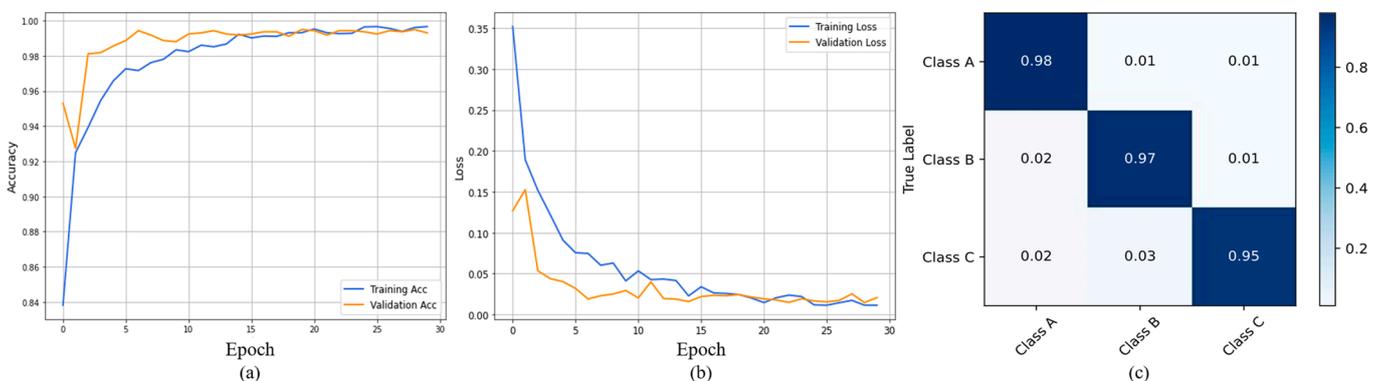


Figure 7. Represents training and validation accuracy: (a) training and validation loss (b) and confusion matrix (c) for DDSM dataset.

Table 2. Classification report of the proposed network using DDSM dataset.

Class	Precision	Recall	F1-Score
1	0.98	0.97	0.98
2	0.97	0.96	0.97
3	0.95	0.97	0.96

To determine the generalization capability of the proposed network, we provided a detailed comparison among various deep learning models, including DBM [38], DL and Ensemble Learning [63], CapsNet [64], CNN [31], CNN-Multi feature extractor [30], Ensemble-classifier [40], and the proposed model. As shown in Table 3, the CapsNet [64] and CNN-Multi feature extractor [30] attained the worst performance in terms of accuracy for BC classification, which are 77.78% and 91.70%, respectively. Similarly, DL and Ensemble Learning [63] and CNN [31] are considered the second best classifiers for BC classification; both of the models achieved 96.0% on average in terms of accuracy. Compared to the proposed model, named Ensemble, with an attention mechanism, it achieved 96.50% accuracy using DDSM; our model surpassed the second-best model by obtaining a 0.50 higher value for average accuracy.

Table 3. Comparison of the proposed network between different classifiers using DDSM dataset.

S: No	Ref	Classes	Methods	Accuracy
1	[38]	Normal, Benign, Malignant	DBM	92.86
2	[63]	Normal, Benign, Malignant	DL and Ensemble Learning	96.00
3	[64]	Normal, Benign, Malignant	CapsNet	77.78
4	[31]	Normal, Benign, Malignant	CNN	96.00
5	[30]	Normal, Benign, Malignant	CNN-Multi feature extractor	91.70
6	[40]	Normal, Benign, Malignant	Ensemble-classifier	93.26
7	ECS-A-Net	Normal, Benign, Malignant	Ensemble-CNN with Attention mechanism	96.50

In short, the proposed network demonstrates superior performance in terms of accuracy, when it is compared with different classifiers using DDSM, as tabulated in Table 3. Additionally, the experimental results validate the proposed network's robustness and generalization, affirming its suitability for highly effective breast cancer classification.

4.2.2. Comparative Analysis of MIAS Dataset

In this subsequent section, the proposed model is evaluated using the MIAS dataset. In Figure 8a, the model achieved a training accuracy of 0.96 and a validation accuracy of 0.95. Similarly, Figure 8b shows the loss graphs; our model reached the minimum training and validation loss up to 0.2 and 0.1, respectively. The confusion matrix provides a class-wise classification report of the proposed model using a test set from the MIAS dataset as shown in Figure 8c. Here, we observe that the benign class achieves an accuracy level of 0.98, the malignant class 0.95, and the normal class 0.93. Moreover, the ECS-A-Net also achieved a better performance for each class individually, as tabulated in Table 4.

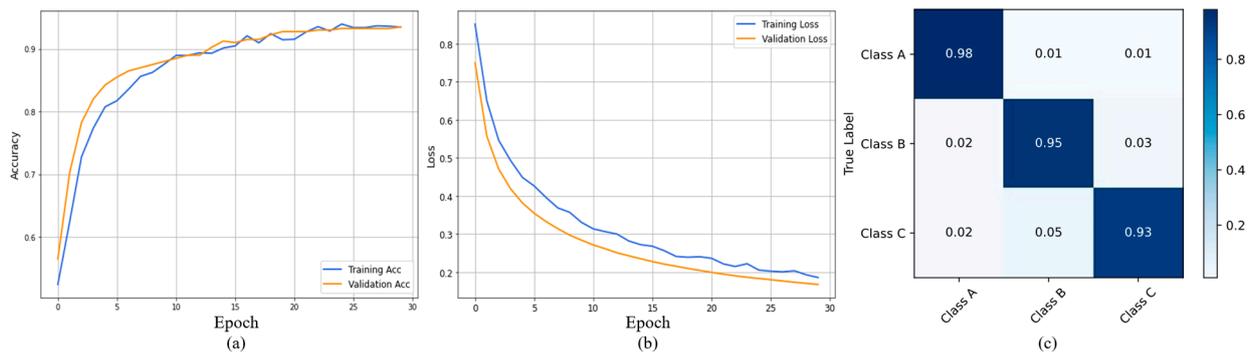


Figure 8. Represents training and validation accuracy (a), training and validation loss (b), and confusion matrix (c) for the MIAS dataset.

Table 4. Classification report of the proposed network over MIAS.

Class	Precision	Recall	F1-Score
1	0.98	0.96	0.97
2	0.95	0.94	0.95
3	0.93	0.96	0.94

In Table 5, a comprehensive comparison between our model and state-of-the-art methods is presented. It is evident that the CNN-Multi feature extractor [30] obtained the lowest performance, while the proposed model outperformed all others, obtaining the highest performance. Notably, the proposed model surpassed the second-ranked model [63] and the third-ranked model [40] by achieving higher accuracy rates of 0.33 and 4.33%, respectively. These results demonstrate that our model exhibits promising performance in comparison to the benchmark models, establishing it as a viable and accurate choice for the specific task at hand.

Table 5. Comparison of the proposed network between different classifiers using the MIAS dataset.

S: No	Ref	Classes	Classifier	Accuracy
1	[44]	Normal, Benign, Malignant	CNN	90.50
2	[63]	Normal, Benign, Malignant	DL and Ensemble Learning	95.00
3	[30]	Normal, Benign, Malignant	CNN-Multi feature extractor	89.90
4	[40]	Normal, Benign, Malignant	Ensemble-classifier	91.00
5	[31]	Normal, Benign, Malignant	CNN	94.00
6	ECS-A-Net	Normal, Benign, Malignant	Ensemble-CNN with Attention mechanism	95.33

4.3. Ablation Study

This section provides a detailed ablation analysis between the proposed ECS-A-Net and various pre-trained architectures with and without attention mechanisms in terms of evaluation accuracy over DDSM and MIAS. As given in Table 6, the Resnet50 with CA and SA mechanisms exhibited better performance when compared to the Resnet50, ResNet50 with CA, and ResNet50 with SA mechanism. Similarly, the InceptionV3 network with CA and SA modules outperformed among plain InceptionV3, InceptionV3-CA, and InceptionV3-SA, and it also surpassed ResNet50-CA and SA mechanisms and is considered the second-best model among others, where it attained 94.08% accuracy using DIAS and

93.13% accuracy using MIAS datasets, as tabulated in Table 6. Similarly, we followed the same strategies for the proposed network as well. The proposed ECS-A-Net is trained with and without an attention module and compared with the ResNet50 and InceptionV2 network over the DDSM dataset, where the experimental results determined that the proposed ensemble network with CA and SA modules achieved optimal results among others. As demonstrated in Table 6, the plain Ensemble achieved 94.45% and 93.70% accuracy using DDSM and MIAS datasets, respectively; similarly, Ensemble with CA further improved the performance, where it achieved 95.89% performance using DDSM and 94.30% performance using the MIAS dataset. In addition, Ensemble-SA exhibited 98.77% and 94.23% performance in terms of accuracy for the DDSM and MIAS datasets, respectively. The Ensemble-CA and SA modules outperformed compared to others and attained the highest accuracy for both DDSM and MIAS datasets, which is 96.50% and 95.33%. Hence, the extensive and thorough analysis of the results demonstrates the efficacy of the ECS-A-Net ensemble-based network for accurate BC classification, particularly in dealing with complex data. Through this comprehensive evaluation, it becomes evident that the developed Ensemble model with attention modules exhibits an ability to generalize, making it a robust solution for addressing the challenges posed by complex data in the realm of BC classification.

Table 6. Detailed comparative analysis of each module used in the proposed network using DDSM and MIAS datasets. × indicate the module is not included whereas the ✓ indicated the module is used.

S: No	Model	Channel Attention	Spatial Attention	Accuracy	
				DDSM	MIAS
1	Resnet50	×	×	92.65	91.00
2		✓	×	93.89	92.50
3		×	✓	93.67	92.33
4		✓	✓	94.08	93.13
5	InceptionV3	×	×	93.15	92.21
6		✓	×	94.42	93.00
7		×	✓	94.28	92.90
8		✓	✓	95.03	93.55
9	Ensemble	×	×	94.45	93.70
10		✓	×	95.89	94.30
11		×	✓	98.77	94.23
12		✓	✓	96.50	95.33

4.4. Model Parameter Investigation

In the realm of computer vision, the real-time implementation of deep learning (DL)-based networks is pivotal across diverse domains such as healthcare, surveillance, and driver distraction analysis. Given the critical role of real-time processing in these applications, a comparison between the proposed Enhanced Cancer Screening Attention Network (ECS-A-Net) and other competitive networks becomes essential, particularly focusing on the model's parameters and size.

Our comprehensive analysis, as given in Table 7, reveals findings about the efficiency and resource requirements of various competitive models. Notably, the approaches published in [45,46] are considered a heavyweight model, with a number of parameters of 145.05 and 138.4, respectively. Similarly, the model sizes of both approaches are large. In contrast, InceptionResNetv2 [48] exhibits a more streamlined architecture with 55.9 parameters and a model size of 215, which is the second-ranking model in our comparison. However, our model has 6.4 million lower parameters and 53 megabytes of reduced size. Similarly, ResNet152V2 has 60.4 million parameters and a model size of 232. Additionally, the proposed ECS-A-Net is considered to be the best, with 49.5 million parameters and a model

size of 162. This efficiency in parameter and model size positions the ECS-A-Net favorably against competitive networks, demonstrating its capacity to obtain robust performance.

Table 7. Comparison between the proposed ECS-A-Net and other networks in terms of parameters and model size.

References	Methods	Number of Parameters (Million)	Model Size (Mega Byte)
[45]	Hybrid-VGG-16	145.05	529
[46]	VGG-16	138.4	528
[48]	InceptionResNetv2	55.9	215
[65]	ResNet152V2	60.4	232
Proposed ECS-A-Net		49.5	162

The justification for these findings extends beyond mere numerical values. The proposed ECS-A-Net's streamlined architecture not only optimizes computational resources but also showcases its potential for real-time implementation on edge devices. This strategic balance between model complexity and efficiency underscores the generalizability and practical applicability of the ECS-A-Net in real-world scenarios, affirming its suitability for resource-constrained environments and demanding real-time decision-making tasks.

5. Conclusions and Future Work

BC remains a serious global health concern and a leading cause of mortality among women. The challenge lies in the elusive nature of symptoms during the early stages, resulting in a staggering 80% of diagnoses happening at an advanced stage. Recently, researchers have developed various methods for BC classification; however, their limited performance and the large number of learning parameters hinder these models from implementation. To cope with this, we introduced an innovative framework named the ECS-A-Net, designed for proficient BC classification and detection. The proposed framework contains two different phases, preprocessing and model training. In the first phase, the preprocessing strategy, called data augmentation, was applied to enhance the size of the input data. Meanwhile, in the second phase, we incorporated a modified SE-ResNet50 and InceptionV3 parallelly as a backbone for feature extractions, followed by a CA and SA module to further evaluate the task-related features. We performed a detailed experimental study to validate the performance of the proposed network and also compared it with SOTA techniques over benchmark datasets. The ECS-A-Net network revealed high performance for all datasets in terms of accuracy, precision, recall, and F1-measure. Thus, the proposed model is an effective solution for early BC classification with the capability to lower model complexity and size.

In the future, we will use different object detection algorithms such as FastRCNN, FasterRCNN, MaskRCNN, etc., to detect the exact location of the tumors. Furthermore, we also aim to use the knowledge distillation concept to train the larger model as a teacher and recall their knowledge in the student (lightweight) model for the generation of more efficient results.

Author Contributions: Conceptualization, S.M.T. and H.-S.P. Funding acquisition, H.-S.P. Investigation and methodology, S.M.T. and S.J.M. Writing of the original draft, S.M.T. Writing of the review and editing, S.M.T., S.J.M. and A.W.A. Validation, S.M.T., S.J.M., A.W.A. and H.-S.P. Formal analysis, S.M.T., S.J.M., A.W.A. and H.-S.P. Data curation, S.M.T., S.J.M. and A.W.A. Visualization, S.M.T. and H.-S.P. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data used for this research is publicly available.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Zerouaoui, H.; Idri, A.J.J.o.M.S. Reviewing Machine Learning and Image Processing Based Decision-Making Systems for Breast Cancer Imaging. *J. Med. Syst.* **2021**, *45*, 8. [[CrossRef](#)] [[PubMed](#)]
- ElOuassif, B.; Idri, A.; Hosni, M.; Abran, A. Classification techniques in breast cancer diagnosis: A systematic literature review. *Comput. Methods Biomech. Biomed. Eng. Imaging Vis.* **2021**, *9*, 50–77. [[CrossRef](#)]
- Chitalia, R.D.; Kontos, D. Role of texture analysis in breast MRI as a cancer biomarker: A review. *J. Magn. Reson. Imaging* **2019**, *49*, 927–938. [[CrossRef](#)] [[PubMed](#)]
- Eun, N.L.; Kang, D.; Son, E.J.; Youk, J.H.; Kim, J.-A.; Gweon, H.M. Texture analysis using machine learning-based 3-T magnetic resonance imaging for predicting recurrence in breast cancer patients treated with neoadjuvant chemotherapy. *Eur. Radiol.* **2021**, *31*, 6916–6928. [[CrossRef](#)] [[PubMed](#)]
- Aslan, M.F.; Celik, Y.; Sabanci, K.; Durdu, A. Breast Cancer Diagnosis by Different Machine Learning Methods Using Blood Analysis Data. *Int. J. Intell. Syst. Appl. Eng.* **2018**, *6*, 289–293. [[CrossRef](#)]
- Muduli, D.; Dash, R.; Majhi, B. Automated breast cancer detection in digital mammograms: A moth flame optimization based ELM approach. *Biomed. Signal Process. Control.* **2020**, *59*, 101912. [[CrossRef](#)]
- Bokade, A.; Shah, A. Breast Mass Classification with Deep Transfer Feature Extractor Model and Random Forest Classifier. In Proceedings of the 2021 International Conference on Recent Trends on Electronics, Information, Communication & Technology (RTEICT), Bangalore, India, 27–28 August 2021; pp. 634–641.
- Li, J.; Shi, J.; Chen, J.; Du, Z.; Huang, L. Self-attention random forest for breast cancer image classification. *Front. Oncol.* **2023**, *13*, 1043463. [[CrossRef](#)] [[PubMed](#)]
- Li, X.; Qin, G.; He, Q.; Sun, L.; Zeng, H.; He, Z.; Chen, W.; Zhen, X.; Zhou, L. Digital breast tomosynthesis versus digital mammography: Integration of image modalities enhances deep learning-based breast mass classification. *Eur. Radiol.* **2020**, *30*, 778–788. [[CrossRef](#)]
- Debelee, T.G.; Schwenker, F.; Ibenthal, A.; Yohannes, D. Survey of deep learning in breast cancer image analysis. *Evol. Syst.* **2019**, *11*, 143–163. [[CrossRef](#)]
- Yar, H.; Abbas, N.; Sadad, T.; Iqbal, S. Lung nodule detection and classification using 2D and 3D convolution neural networks (CNNs). In *Artificial Intelligence and Internet of Things*; CRC Press: Boca Raton, FL, USA, 2021; pp. 365–386.
- Munsif, M.; Ullah, M.; Ahmad, B.; Sajjad, M.; Cheikh, F.A. Monitoring neurological disorder patients via deep learning based facial expressions analysis. In *Artificial Intelligence Applications and Innovations. AIAI 2022 IFIP WG 12.5 International Workshops*; Springer: Cham, Switzerland, 2022.
- Aboutalib, S.S.; Mohamed, A.A.; Berg, W.A.; Zuley, M.L.; Sumkin, J.H.; Wu, S. Deep Learning to Distinguish Recalled but Benign Mammography Images in Breast Cancer Screening. *Clin. Cancer Res.* **2018**, *24*, 5902–5909. [[CrossRef](#)]
- Platania, R.; Shams, S.; Yang, S.; Zhang, J.; Lee, K.; Park, S.-J. Automated Breast Cancer Diagnosis Using Deep Learning and Region of Interest Detection (BC-DROID). In *Proceedings of the 8th ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics, Boston, MA, USA, 20–23 August 2017*; Association for Computing Machinery: New York, NY, USA, 2017; pp. 536–543. [[CrossRef](#)]
- Jannesari, M.; Habibzadeh, M.; Aboulkheyr, H.; Khosravi, P.; Elemento, O.; Totonchi, M.; Hajirasouliha, I. Breast Cancer Histopathological Image Classification: A Deep Learning Approach. In Proceedings of the 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Madrid, Spain, 3–6 December 2018; pp. 2405–2412.
- Khuriwal, N.; Mishra, N. Breast cancer detection from histopathological images using deep learning. In Proceedings of the 2018 3rd International Conference and Workshops on Recent Advances and Innovations in Engineering (ICRAIE), Jaipur, India, 22–25 November 2018.
- Li, H.; Zhuang, S.; Li, D.-A.; Zhao, J.; Ma, Y. Benign and malignant classification of mammogram images based on deep learning. *Biomed. Signal Process. Control.* **2019**, *51*, 347–354. [[CrossRef](#)]
- Khan, S.; Islam, N.; Jan, Z.; Din, I.U.; Rodrigues, J.J.P.C. A novel deep learning based framework for the detection and classification of breast cancer using transfer learning. *Pattern Recognit. Lett.* **2019**, *125*, 1–6. [[CrossRef](#)]
- Tsochatzidis, L.; Costaridou, L.; Pratikakis, I. Deep Learning for Breast Cancer Diagnosis from Mammograms—A Comparative Study. *J. Imaging* **2019**, *5*, 37. [[CrossRef](#)] [[PubMed](#)]
- Zhou, L.-Q.; Wu, X.-L.; Huang, S.-Y.; Wu, G.-G.; Ye, H.-R.; Wei, Q.; Bao, L.-Y.; Deng, Y.-B.; Li, X.-R.; Cui, X.-W.; et al. Lymph Node Metastasis Prediction from Primary Breast Cancer US Images Using Deep Learning. *Radiology* **2020**, *294*, 19–28. [[CrossRef](#)] [[PubMed](#)]
- Shen, L.; Margolies, L.R.; Rothstein, J.H.; Fluder, E.; McBride, R.; Sieh, W. Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Sci. Rep.* **2019**, *9*, 12495. [[CrossRef](#)]
- Ismail, N.S.; Sovuthy, C. Breast cancer detection based on deep learning technique. In Proceedings of the 2019 International UNIMAS STEM 12th Engineering Conference (EnCon), Kuching, Malaysia, 28–29 August 2019.

23. Charan, S.; Khan, M.J.; Khurshid, K. Breast Cancer Detection in Mammograms Using Convolutional Neural Network. In Proceedings of the 2018 International Conference on Computing, Mathematics and Engineering Technologies (iCoMET), Wuhan, China, 7–8 February 2018; pp. 1–5.
24. Sha, Z.; Hu, L.; Rouyendegh, B.D. Deep learning and optimization algorithms for automatic breast cancer detection. *Int. J. Imaging Syst. Technol.* **2020**, *30*, 495–506. [[CrossRef](#)]
25. Kaur, P.; Singh, G.; Kaur, P. Intellectual detection and validation of automated mammogram breast cancer images by multi-class SVM using deep learning classification. *Inform. Med. Unlocked* **2019**, *16*, 100151. [[CrossRef](#)]
26. Arefan, D.; Mohamed, A.A.; Berg, W.A.; Zuley, M.L.; Sumkin, J.H.; Wu, S. Deep learning modeling using normal mammograms for predicting breast cancer risk. *Med. Phys.* **2020**, *47*, 110–118. [[CrossRef](#)]
27. e Silva, D.; Cortes, O. On Convolutional Neural Networks and Transfer Learning for Classifying Breast Cancer on Histopathological Images Using GPU. In *Progress of the CBEB 2020: XXVII Brazilian Congress on Biomedical Engineering, Vitória, Brazil, 26–30 October 2020*; Springer: Cham, Switzerland, 2020.
28. Dhahbi, S.; Barhoumi, W.; Zagrouba, E. Breast cancer diagnosis in digitized mammograms using curvelet moments. *Comput. Biol. Med.* **2015**, *64*, 79–90. [[CrossRef](#)]
29. da Rocha, S.V.; Braz Junior, G.; Corrêa Silva, A.; de Paiva, A.C.; Gattass, M. Texture analysis of masses malignant in mammograms images using a combined approach of diversity index and local binary patterns distribution. *Expert Syst. Appl.* **2016**, *66*, 7–19. [[CrossRef](#)]
30. Ranjbarzadeh, R.; Sarshar, N.T.; Ghouschi, S.J.; Esfahani, M.S.; Parhizkar, M.; Pourasad, Y.; Anari, S.; Bendeche, M. MRFE-CNN: Multi-route feature extraction model for breast tumor segmentation in Mammograms using a convolutional neural network. *Ann. Oper. Res.* **2023**, *328*, 1021–1042. [[CrossRef](#)]
31. Saran, K.B.; Sreelekha, G.; Sunitha, V.C. Breast Density Classification to Aid Clinical Workflow in Breast Cancer Detection Using Deep Learning Network. In Proceedings of the TENCON 2023—2023 IEEE Region 10 Conference (TENCON), Chiang Mai, Thailand, 31 October–3 November 2023; pp. 444–449.
32. Tsochatzidis, L.; Zagoris, K.; Arikidis, N.; Karahaliou, A.; Costaridou, L.; Pratikakis, I. Computer-aided diagnosis of mammographic masses based on a supervised content-based image retrieval approach. *Pattern Recognit.* **2017**, *71*, 106–117. [[CrossRef](#)]
33. Aslan, M.F. A hybrid end-to-end learning approach for breast cancer diagnosis: Convolutional recurrent network. *Comput. Electr. Eng.* **2023**, *105*, 108562. [[CrossRef](#)]
34. Addo, D.; Zhou, S.; Sarpong, K.; Nartey, O.T.; Abdullah, M.A.; Ukwuoma, C.C.; Al-Antari, M.A. A hybrid lightweight breast cancer classification framework using the histopathological images. *Biocybern. Biomed. Eng.* **2024**, *44*, 31–54. [[CrossRef](#)]
35. Iqbal, A.; Sharif, M. Memory-efficient transformer network with feature fusion for breast tumor segmentation and classification task. *Eng. Appl. Artif. Intell.* **2024**, *127*, 107292. [[CrossRef](#)]
36. Sapate, S.G.; Mahajan, A.; Talbar, S.N.; Sable, N.; Desai, S.; Thakur, M. Radiomics based detection and characterization of suspicious lesions on full field digital mammograms. *Comput. Methods Programs Biomed.* **2018**, *163*, 1–20. [[CrossRef](#)] [[PubMed](#)]
37. Chakraborty, J.; Midya, A.; Rabidas, R. Computer-aided detection and diagnosis of mammographic masses using multi-resolution analysis of oriented tissue patterns. *Expert Syst. Appl.* **2018**, *99*, 168–179. [[CrossRef](#)]
38. Al-Antari, M.A.; Al-Masni, M.A.; Park, S.-U.; Park, J.; Metwally, M.K.; Kadah, Y.M.; Han, S.-M.; Kim, T.-S. An Automatic Computer-Aided Diagnosis System for Breast Cancer in Digital Mammograms via Deep Belief Network. *J. Med. Biol. Eng.* **2017**, *38*, 443–456. [[CrossRef](#)]
39. Muramatsu, C.; Nishio, M.; Goto, T.; Oiwa, M.; Morita, T.; Yakami, M.; Kubo, T.; Togashi, K.; Fujita, H. Improving breast mass classification by shared data with domain transformation using a generative adversarial network. *Comput. Biol. Med.* **2020**, *119*, 103698. [[CrossRef](#)]
40. Yan, F.; Huang, H.; Pedrycz, W.; Hirota, K. Automated breast cancer detection in mammography using ensemble classifier and feature weighting algorithms. *Expert Syst. Appl.* **2023**, *227*, 120282. [[CrossRef](#)]
41. Milosevic, M.; Jankovic, D.; Peulic, A. Comparative analysis of breast cancer detection in mammograms and thermograms. *Biomed. Eng. Biomed. Tech.* **2014**, *60*, 49–56. [[CrossRef](#)] [[PubMed](#)]
42. Lou, M.; Wang, R.; Qi, Y.; Zhao, W.; Xu, C.; Meng, J.; Deng, X.; Ma, Y. MGBN: Convolutional neural networks for automated benign and malignant breast masses classification. *Multimedia Tools Appl.* **2021**, *80*, 26731–26750. [[CrossRef](#)]
43. Li, B.; Ge, Y.; Zhao, Y.; Guan, E.; Yan, W. Benign and malignant mammographic image classification based on Convolutional Neural Networks. In Proceedings of the ICMLC 2018: 2018 10th International Conference on Machine Learning and Computing, Macau, China, 26–28 February 2018.
44. Ting, F.F.; Tan, Y.J.; Sim, K.S. Convolutional neural network improvement for breast cancer classification. *Expert Syst. Appl.* **2018**, *120*, 103–115. [[CrossRef](#)]
45. Jahangeer, G.S.B.; Rajkumar, T.D. Early detection of breast cancer using hybrid of series network and VGG-16. *Multimedia Tools Appl.* **2020**, *80*, 7853–7886. [[CrossRef](#)]
46. Saber, A.; Sakr, M.; Abo-Seida, O.M.; Keshk, A.; Chen, H. A Novel Deep-Learning Model for Automatic Detection and Classification of Breast Cancer Using the Transfer-Learning Technique. *IEEE Access* **2021**, *9*, 71194–71209. [[CrossRef](#)]
47. Byra, M.; Karwat, P.; Ryzhankow, I.; Komorowski, P.; Klimonda, Z.; Fura, L.; Pawlowska, A.; Zolek, N.; Litniewski, J. Deep meta-learning for the selection of accurate ultrasound based breast mass classifier. In Proceedings of the 2022 IEEE International Ultrasonics Symposium (IUS), Venice, Italy, 10–13 October 2022; pp. 1–4.

48. Muralikrishnan, M.; Anitha, R. Comparison of breast cancer multi-class classification accuracy based on inception and inceptionresnet architecture. In *Emerging Trends in Computing and Expert Technology*; Springer: Berlin/Heidelberg, Germany, 2020.
49. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, Las Vegas, NV, USA, 27–30 June 2016.
50. Khan, Z.A.; Hussain, T.; Ullah, F.U.M.; Gupta, S.K.; Lee, M.Y.; Baik, S.W. Randomly Initialized CNN with Densely Connected Stacked Autoencoder for Efficient Fire Detection. *Eng. Appl. Artif. Intell.* **2022**, *116*, 105403. [[CrossRef](#)]
51. Ali, H.; Farman, H.; Yar, H.; Khan, Z.; Habib, S.; Ammar, A. Deep learning-based election results prediction using Twitter activity. *Soft Comput.* **2021**, *26*, 7535–7543. [[CrossRef](#)]
52. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*; IEEE: Piscataway, PA, USA, 2018.
53. Kristiani, E.; Yang, C.-T.; Huang, C.-Y.; Chan, Y.-W.; Fathoni, H. *Image Classification Model Using Deep Learning on the Edge Device, in Innovative Computing*; Springer: Berlin, Germany, 2020; pp. 11–22.
54. Yar, H.; Hussain, T.; Agarwal, M.; Khan, Z.A.; Gupta, S.K.; Baik, S.W. Optimized Dual Fire Attention Network and Medium-Scale Fire Classification Benchmark. *IEEE Trans. Image Process.* **2022**, *31*, 6331–6343. [[CrossRef](#)]
55. Yar, H.; Ullah, W.; Khan, Z.A.; Baik, S.W. An Effective Attention-based CNN Model for Fire Detection in Adverse Weather Conditions. *ISPRS J. Photogramm. Remote. Sens.* **2023**, *206*, 335–346. [[CrossRef](#)]
56. Khan, Z.A.; Hussain, T.; Ullah, W.; Baik, S.W. *A Trapezoid Attention Mechanism for Power Generation and Consumption Forecasting. IEEE Transactions on Industrial Informatics*; IEEE: Piscataway, PA, USA, 2023; pp. 1–13. [[CrossRef](#)]
57. Khan, Z.A.; Hussain, T.; Baik, S.W. Dual stream network with attention mechanism for photovoltaic power forecasting. *Appl. Energy* **2023**, *338*, 120916. [[CrossRef](#)]
58. Khan, T.; Choi, G.; Lee, S. EFFNet-CA: An Efficient Driver Distraction Detection Based on Multiscale Features Extractions and Channel Attention Mechanism. *Sensors* **2023**, *23*, 3835. [[CrossRef](#)] [[PubMed](#)]
59. Khan, T.; Aslan, H.İ. Performance Evaluation of Enhanced ConvNeXtTiny-based Fire Detection System in Real-world Scenarios. In Proceedings of the Eleventh International Conference on Learning Representations, Kigali, Rwanda, 1–5 May 2023.
60. Yar, H.; Khan, Z.A.; Ullah, F.U.M.; Ullah, W.; Baik, S.W. A modified YOLOv5 architecture for efficient fire detection in smart cities. *Expert Syst. Appl.* **2023**, *231*, 120465. [[CrossRef](#)]
61. Hijji, M.; Yar, H.; Ullah, F.U.M.; Alwakeel, M.M.; Harrabi, R.; Aradah, F.; Cheikh, F.A.; Muhammad, K.; Sajjad, M. FADS: An Intelligent Fatigue and Age Detection System. *Mathematics* **2023**, *11*, 1174. [[CrossRef](#)]
62. Jan, H.; Yar, H.; Iqbal, J.; Farman, H.; Khan, Z.; Koubaa, A. Raspberry pi assisted safety system for elderly people: An application of smart home. In Proceedings of the 2020 First International Conference of Smart Systems and Emerging Technologies (SMARTTECH), Riyadh, Saudi Arabia, 3–5 November 2020.
63. Malebary, S.J.; Hashmi, A. Automated Breast Mass Classification System Using Deep Learning and Ensemble Learning in Digital Mammogram. *IEEE Access* **2021**, *9*, 55312–55328. [[CrossRef](#)]
64. Soulami, K.B.; Kaabouch, N.; Saidi, M.N. Breast cancer: Classification of suspicious regions in digital mammograms based on capsule network. *Biomed. Signal Process. Control.* **2022**, *76*, 103696. [[CrossRef](#)]
65. Shehzad, I.; Zafar, A. Breast Cancer CT-Scan Image Classification Using Transfer Learning. *SN Comput. Sci.* **2023**, *4*, 789. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.