

Article

CSA-Net: Channel and Spatial Attention-Based Network for Mammogram and Ultrasound Image Classification

Osama Bin Naeem ^{1,2,*}  and Yasir Saleem ² ¹ Department of Electrical Engineering, University of Engineering and Technology, Lahore-Narowal Campus, Narowal 51600, Pakistan² Department of Computer Engineering, University of Engineering and Technology, Lahore 39161, Pakistan; yasir@uet.edu.pk

* Correspondence: osamabinnaeem@uet.edu.pk or osamabinnaeem@gmail.com

Abstract: Breast cancer persists as a critical global health concern, emphasizing the advancement of reliable diagnostic strategies to improve patient survival rates. To address this challenge, a computer-aided diagnostic methodology for breast cancer classification is proposed. An architecture that incorporates a pre-trained EfficientNet-B0 model along with channel and spatial attention mechanisms is employed. The efficiency of leveraging attention mechanisms for breast cancer classification is investigated here. The proposed model demonstrates commendable performance in classification tasks, particularly showing significant improvements upon integrating attention mechanisms. Furthermore, this model demonstrates versatility across various imaging modalities, as demonstrated by its robust performance in classifying breast lesions, not only in mammograms but also in ultrasound images during cross-modality evaluation. It has achieved accuracy of 99.9% for binary classification using the mammogram dataset and 92.3% accuracy on the cross-modality multi-class dataset. The experimental results emphasize the superiority of our proposed method over the current state-of-the-art approaches for breast cancer classification.

Keywords: attention mechanism; breast cancer; binary classification; classification; deep learning; mammogram; ultrasound; mass; multi-classification



Citation: Naeem, O.B.; Saleem, Y. CSA-Net: Channel and Spatial Attention-Based Network for Mammogram and Ultrasound Image Classification. *J. Imaging* **2024**, *10*, 256. <https://doi.org/10.3390/jimaging10100256>

Received: 12 August 2024

Revised: 12 October 2024

Accepted: 14 October 2024

Published: 16 October 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

There are 37.2 trillion cells in the human body [1], which undergo a continuous cycle of division, proliferation, and eventual death. This process undergoes strict regulation and continuous operation throughout life. On average, the human body replaces its cells every seven to ten years, although the lifespan of specific organs may vary [2]. When ageing, if cells fail to undergo programmed cell death and instead continue to divide and proliferate uncontrollably, they can form tumors. However, there is an exception in the case of leukemia, which does not produce tumors. This condition poses a serious health risk and could lead to life-threatening consequences [3].

Breast cancer ranks among the most common cancers worldwide in women, although its incidence varies from region to region. According to the World Health Organization's (WHO) 2022 report and research conducted by the International Agency for Research on Cancer, cancer stands as the leading cause of abnormality and mortality globally, affecting both genders; see Figure 1A. In Figure 1B, the distribution of female cancer cases globally in 2022 is depicted in terms of the percentile distribution. Women represent the majority of the reported instances, accounting for 23.8% (2,296,840 cases) of the total. Additionally, Figure 1C documents these cancer cases in Pakistan, comprising 31.3% (30,682 cases) of all cancer cases [4,5].

Computer-aided detection (CAD) systems in mammography have historically faced performance challenges due to their reliance on radiometric features, leading to unnecessary treatments and costs. Deep learning (DL) promises to address these limitations, although it

requires substantial training data. Medical imaging data, adhering to big data principles, presents both opportunities and hurdles. DL has the potential to improve performance and reduce costs through enhanced initial screenings. Ultimately, DL could improve early breast cancer diagnosis and healthcare efficiency [6].

Population	Total	Percent.
Lung	2 480 675	12.4%
Breast	2 296 840	11.5%
Prostate	1 467 854	7.3%
Colon	1 142 286	5.7%
Others	12 588 844	63.0%

Total : 19 976 499

(A) Breast Cancer Cases-Both Sexes (Worldwide)-2022

Population	Total	Percent.
Breast	2 296 840	23.8%
Lung	908 630	9.4%
Cervix uteri	662 301	6.9%
Thyroid	614 729	6.4%
Others	5 182 389	53.6%

Total : 9 664 889

(B) Breast Cancer Cases-Females (Worldwide)-2022

Population	Total	Percent.
Breast	30 682	31.3%
Lip, oral cavity	5 170	5.3%
Ovary	4 987	5.1%
Cervix uteri	4 762	4.9%
Others	52 579	53.6%

Total : 98 180

(C) Breast Cancer Cases-Females (Pakistan)-2022

Population	Total	Percent.
Breast	15 552	26.1%
Oesophagus	4 207	7.0%
Ovary	3 492	5.8%
Lip, oral cavity	3 365	5.6%
Others	33 081	55.4%

Total : 59 697

(D) Breast Cancer Deaths-Females (Pakistan)-2022

Population	Total	Percent.
Breast	666 103	15.4%
Lung	584 228	13.5%
Cervix uteri	348 874	8.1%
Colon	254 383	5.9%
Others	2 459 960	57.0%

Total : 4 313 548

(E) Breast Cancer Deaths-Females (Worldwide)-2022

Comparison of Breast Cancer Cases from 2018,2020 and 2022 GLOBOCON Survey						
Description	GLOBOCON-2018		GLOBOCON-2020		GLOBOCON-2022	
	cases	%age	cases	%age	cases	%age
Breast cancer cases among both sexes (Worldwide)	2,088,849	11.6	2,261,419	11.7	2,296,840	11.5
Breast cancer cases among Females (Worldwide)	2,088,849	24.2	2,261,419	24.5	2,296,840	23.8
Breast cancer cases among Females (Pakistan)	34,066	19.6	25,928	28.7	30,682	31.3
Mortality rate from Breast cancer (Worldwide)	626,679	15	684,996	15.5	666,103	15.4
Mortality rate from Breast cancer (Pakistan)	17,158	29.9	13,724	24.3	15,552	26.1

Figure 1. Cancer cases according to GLOBOCON in 2018, 2020, and 2022—WHO [4,5,7,8].

Despite the prevalence of breast cancer, traditional detection methods fall short in terms of performance compared to emerging technologies like deep learning. The application of deep learning in medical imaging, particularly for breast cancer detection, is hindered by the lack of large-scale training data. Although deep learning shows promise in pattern recognition through neural network layers, its effectiveness in this domain is constrained by data availability. Therefore, there is a significant opportunity to enhance screening accuracy through further research and development of automated breast cancer detection systems [9]. Our study aims to contribute to these efforts by utilizing deep learning models for the classification of breast abnormalities, thereby supporting ongoing endeavours to improve cancer screening efficiency and accuracy.

The proposed research architecture assimilates the pre-trained EfficientNet-B0 model, renowned for its ability to extract intricate features from diverse datasets. By integrating channel and spatial attention mechanisms, the model enhances its discriminative capabilities, focusing on salient features that are crucial for accurate breast cancer classification.

Channel and spatial attention are used as a convolutional block attention module (CBAM), allowing the model to concentrate on relevant spatial regions and effectively address class imbalances. Overall, this comprehensive approach aims to elevate the performance in the complex task of breast cancer classification by strategically incorporating attention mechanisms and focal loss.

This paper is organized into sections as follows. The literature work is cited in Section 2. Then, Section 3 covers the methodology used in obtaining the desired results. Section 4 covers the results, where Section 4.1 covers the datasets used, the system configuration used for the experimentation, and the evaluation measures, and Section 5 covers the analysis conducted in response to the experimentation conducted. Finally, Section 6 displays the conclusions in response to the results generated.

2. Related Work

Deep neural networks are widely explored and applied in breast cancer diagnosis due to their robust feature extraction capabilities. Numerous methods utilizing convolutional neural networks (CNNs) for feature extraction have been introduced to enhance the efficiency of breast cancer diagnosis [10–12].

Two main strategies have emerged to improve the efficiency of breast cancer detection with CNNs, capitalizing on their capability to learn complex features and label data in detail, particularly in mammogram images. One approach integrates information from both the complete mammogram image and labeled image patches. A deep fusion model is introduced to tackle anomaly detection in mammogram images by incorporating patch information alongside the broader image context. This method entails two key steps: initially, region of interest (ROI) images are trained, followed by investigating the channel's interrelationships within the block [13].

In [14], the authors proposed a method to improve the classification accuracy, where the CNN was initially pre-trained using patch data, followed by fine-tuning on complete mammography images. While this approach proved to yield more detailed lesion characteristics, it required additional annotations for patch labels. In [15], the authors introduced another approach that integrated both cranio-caudal (CC) and medio-lateral oblique (MLO) views, along with their corresponding segment annotations of the same breast, as input to neural networks. This method aimed to enhance the classification of mammogram images. Additionally, they proposed an automated approach to classifying MLO views and unregistered CC views. This method improved the effectiveness of breast cancer diagnosis by integrating images from both views with their corresponding segmentation maps of breast lesions [16].

Ref. [17] compiled a substantial dataset of mammography images to train a deep CNN designed for breast cancer diagnosis. Each examination encompassed four distinct image views, and the deep CNN utilized these comprehensive examinations as input to predict the potential class. This approach is advantageous as it integrates a larger amount of input data, thereby enhancing the performance. Both strategies contributed to the improved effectiveness of breast cancer diagnosis by harnessing additional data, such as multi-view and patch label information. However, it is noteworthy that higher-grade datasets might necessitate more supplementary information. Additionally, the process of compiling and annotating a high-quality medical dataset requires significant effort, time, and specialized expertise.

2.1. CNN-Based Models

In recent studies, several researchers have proposed that enhancing the feature extraction process could improve breast cancer diagnosis. By focusing on the extraction of discriminating features, these techniques aim to enhance the output while reducing the need for manual labor. For example, researchers proposed a group-max pooling structure based on regions to efficiently extract discriminating features from the data [18]. The authors in [19] implemented a waterfall technique using generative and discriminative approaches to obtain discriminating features. In this method, abnormal tissue was treated as an outlier, while normal tissue was accurately modeled. Enabling the feature extraction component to collect discriminative features from images, several researchers redesigned the loss function from conventional softmax to an angular space and integrated loss restraints. For instance, ref. [20] pioneered the SphereFace technique, which utilizes a hyperparameter m to regulate

the distribution of various classes using a novel loss function. The hyperparameter m defines the dimensions of the distribution space. Building upon this concept, refs. [21,22] developed ArcFace and CosFace, expanding on these principles.

Although initially developed for face recognition tasks, the direct application of these methods to mammogram image classification has led to a significant decrease in effectiveness. Yan Wang et al. extended the utilization of the angular softmax loss function from face recognition to mammography image classification in their research methods. This approach bore similarities to previous studies, such as the utilization of dropout-based Bayesian uncertainty measures by Leibig et al. to enhance the detection of diabetic retinopathy from fundus images [23,24]. To improve the classification performance, the authors integrated calibration and uncertainty within the framework of Bayesian deep learning models [25].

Analyzing mammograms often leads to unnecessary biopsies, resulting in wasted time and resources. To address this issue, the authors in [6] utilized neural networks, along with techniques like noise reduction, thus benefiting both the patient and the healthcare system.

In [26], the researchers endeavored to categorize breast mammography images into normal, benign, and malignant classes. Morphological techniques were applied to adjust the image characteristics, while preprocessing methods like median and mean filters were implemented to enhance the image quality and mitigate noise.

The methodology outlined in [27] employed adaptive median filtering, edge detection, Sobel filtering, and histogram equalization to enhance the image quality and remove unwanted elements like pectoral muscles and impulsive noise. Transfer learning was then applied to extract relevant features from mammograms. Succeeding this, reduced dimensionality and principal component analysis (PCA) were utilized to extract features, balancing computational efficiency with information preservation.

In another study, the author [9] employed median and Wiener filters as filtering techniques and opted for contrast adaptive histogram equalization (CLAHE) due to its effectiveness in enhancing breast images. The classification was performed using a CNN model, ResNet50, with transfer learning applied for feature extraction. Data augmentation techniques were utilized to expand the dataset.

2.2. Attention-Based CNN Models

Inspired by the human cognitive process of selectively filtering out irrelevant information to focus on important aspects, researchers have developed attention mechanisms in deep learning algorithms. These attention mechanisms mimic the human ability to prioritize and focus on essential components while processing large volumes of data. By dynamically allocating attention to specific parts of the input data, these algorithms effectively filter out noise and irrelevant details, allowing them to extract and emphasize the most relevant features. This approach enhances the efficiency and effectiveness of sorting through homogeneous datasets by directing the computational resources towards the most significant aspects of the data, similar to the way in which human attention operates in daily information exploration [28].

The CNN has proven its effectiveness in applications. Moreover, attention mechanisms, explored in image classification research, have been found to further improve the performance of CNNs [29].

In a recent study [30], the authors suggested integrating the ResNet network with a convolutional block attention module (CBAM) to tackle classification problems. The proposed model consisted of a classification module, attention mechanisms, and a backbone comprising ResNet structures. CBAM integration involved incorporating channel and spatial attention maps into ResNet blocks, enhancing the feature representation. The model also addressed the spatial characteristics of CT scan images, focusing particularly on the consolidations and ground glass opacities commonly observed in COVID-19 cases. This approach enabled the model to effectively capture both prominent and subtle features present in grayscale levels.

In [31], skin lesions were classified using the DenseNet-121 model, specifically chosen for its dense connections, facilitating the skip connection by the transmission of lesion features through the layers. DenseNet-121 employed various numbers of densely linked convolutional layers to extract multi-scale lesion features. To enhance the feature map information, class-wise attention was incrementally applied after specific dense blocks, chosen based on high-level semantic information. This attention mechanism enabled the learning of class-specific properties from regions before the final classification.

There has been a notable transition from manual feature extraction to the automated identification, selection, and classification of breast abnormalities. Deep learning algorithms have emerged as the preferred choice over traditional feature-based classifiers due to their numerous advantages, leading to widespread adoption in the biomedical field for breast cancer diagnosis [32,33]. In deep learning, the process of feature extraction and selection is seamlessly incorporated within the network architecture, eliminating the necessity for manual intervention. Consequently, deep learning approaches demonstrate enhanced capabilities in handling complex data and are regarded for their reliability. However, the intricate nature of deep neural networks often results in less interpretable outcomes, posing challenges in assessing the diagnostic significance of imaging data and the individual contribution of each feature to the classification performance.

The authors in [34] introduced a novel multi-scale convolutional neural network (CNN) to classify breast masses as benign or malignant. The architecture leverages two distinct CNN models: Res2Net for the extraction of global features from larger images and DenseNet for the capture of fine-grained details from smaller patches. Both networks were equipped with a convolutional block attention module (CBAM), which enhanced the model's ability to focus on the most relevant features while disregarding less useful information. After feature extraction, the global and local features were merged in a fully connected layer to perform the final classification. This multi-scale approach enabled the model to integrate detailed and broader contextual information, significantly improving the accuracy and robustness of the classification task.

In [35], the INbreast dataset was used to classify breast masses as benign or malignant. They developed a custom preprocessing pipeline, the BDP method, which enhanced the image quality by applying noise reduction, contrast improvement through CLAHE, and data augmentation techniques. The classification was performed using an enhanced ResNet50 model, called ECA-Net50, which integrated an efficient channel attention (ECA) module to improve the feature extraction. To handle class imbalances and challenging samples, the focal loss was used during training, whereas transfer learning fine-tuned the pre-trained model on the mammogram dataset to optimize the performance further.

In [36], the author used thermal imaging from the DMR-IR dataset. The study involved preprocessing the dataset of grayscale thermal images by cropping, normalization, and resizing, followed by data augmentation to enhance the dataset's diversity. A convolutional neural network (CNN) was utilized for feature extraction. Furthermore, a bidirectional long short-term memory (BLSTM) layer was introduced to leverage temporal information. At the same time, three types of attention mechanisms—self-attention, soft attention, and hard attention—were integrated to emphasize critical areas within the thermal images. The model's efficiency was assessed, with the results indicating that the models incorporating attention mechanisms performed significantly better, indicating the importance of attention in enhancing breast cancer detection.

2.3. Breast Cancer Classification Challenges

Comparing the different techniques mentioned in the literature is challenging due to inconsistencies in reporting the outcomes. The accuracy, precision, sensitivity, specificity, F1-score, and area under the curve (AUC) are common evaluation metrics. It is essential for a paper to include all relevant performance evaluation measures to comprehensively describe an algorithm. Solely presenting the accuracy and AUC may be misleading, as high accuracy can be achieved by trading off low sensitivity and high specificity. However,

in clinical settings, low sensitivity is undesirable as it correlates with a high rate of false negative detections.

The early diagnosis of breast cancer significantly reduces unnecessary stress by allowing for timely treatment interventions when the tumor is still small and localized. Research indicates that diagnosing breast cancer at an early stage improves the prognosis and reduces the overall medical costs by minimizing the need for extensive treatments and procedures later on [17,37–42]. Despite this, many patients are referred for biopsy even when the likelihood of a malignancy is low, leading to unnecessary surgical procedures that are not only painful but also carry negative psychological and financial burdens. Studies have shown that only a small percentage of these referred cases are malignant, thus highlighting the need for more accurate and efficient diagnostic methods to minimize unnecessary interventions and to improve patients' care.

To improve the effectiveness of mammogram interpretation, it is imperative to enhance the diagnostic efficiency [38,41–52] while simultaneously reducing the occurrence of false alarms [38,40,47,49,50]. This approach aims to optimize the readability of mammograms by ensuring that legitimate abnormalities are correctly identified, while minimizing unnecessary concerns raised by false positive findings.

3. Methodology

This section presents an overview of the proposed methodology.

3.1. Proposed Architecture Overview

The CSA-NET architecture represents a sophisticated framework designed for precise breast cancer diagnosis by embedding advanced deep learning techniques and specialized attention mechanisms tailored to the analysis of mammograms and ultrasound images. The workflow begins with preprocessing steps applied to mammograms, which include standardization, resizing, and augmentation techniques. These steps are crucial in improving the dataset quality and class heterogeneity and essential for robust model training. These are important aspects, especially when facing the class imbalance challenge, which is often encountered in medical datasets and is addressed through strategies like upsampling.

Vital to CSA-NET's architecture is the utilization of transfer learning with the EfficientNet-B0 model, which is selected based on a performance evaluation of different CNN models trained on the mammogram dataset, as evaluated in Table 1. Leveraging pre-trained models such as EfficientNet-B0 [53], initially pre-trained on ImageNet, provides a significant advantage by harnessing learned features from a large-scale dataset. By integrating attention mechanisms into the EfficientNet-B0 backbone, CSA-NET enhances its capacity to capture salient features from mammogram images effectively. Please refer to Figure 2.

The attention mechanisms in CSA-NET consist of two key components: channel attention and spatial attention. The channel attention mechanism optimizes the feature channel importance, allowing the model to prioritize the most discriminative aspects of the input mammogram images. This recalibration enhances the model's ability to discern crucial features for accurate classification.

Following channel attention, the spatial attention mechanism further refines the feature representations by highlighting pertinent spatial regions within the mammogram images. This refinement process enhances the model's sensitivity to diagnostically relevant areas while minimizing noise from less informative regions. Such hierarchical attention mechanisms empower CSA-NET to adaptively focus on both high-level abstract features and fine-grained details that are critical for mammogram classification.

The post-attention mechanism processing features undergo aggregation via global average pooling (GAP), enabling the integration of spatial information across the entire image. Regularization techniques like dropout are applied to prevent overfitting, hence bolstering the model's generalization capabilities.

The optimization of CSA-NET employs the NADAM optimizer with the focal loss function to mitigate class imbalances, steering the model learning towards minority classes. Training strategies such as early stopping and adaptive learning rate adjustments ensure efficient convergence and prevent overfitting.

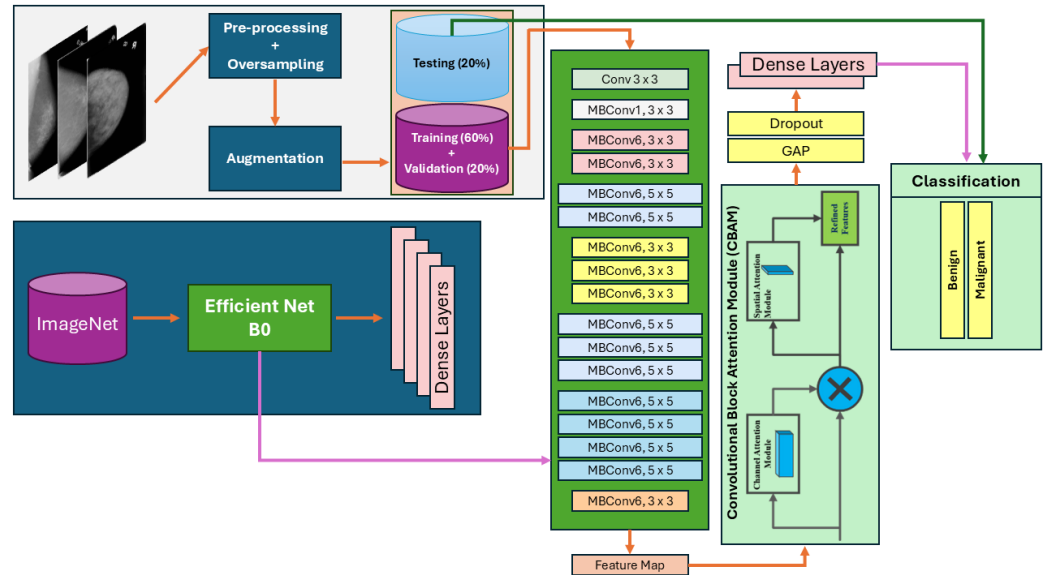


Figure 2. Proposed CSA-Net architecture.

The evaluation metrics encompass a comprehensive range including the accuracy, precision, recall, sensitivity, ROC curve, and F1-score. These metrics collectively reflect CSA-NET's effectiveness in classifying both mammogram and ultrasound images, underscoring its potential to advance breast cancer diagnosis through early detection and precise characterization (Algorithm 1).

Algorithm 1 Algorithm of Proposed Model

Input: E : Number of Epochs; α : Learning Rate; b : Batch Size;
 Breast cancer classification: X_{train} : Training dataset; X_{test} : Test dataset;
Output: A_{test} : Classification evaluation scores on test dataset

Data Preprocessing:

$X_{train} \leftarrow \text{Pre-process}(X_{train})$,
 $X_{test} \leftarrow \text{Pre-process}(X_{test})$

Training of Breast Cancer Classification:

$\delta \leftarrow DL - \text{Model}(X_{train}, E)$,
while δ has not converged **do**
 for local epoche $\epsilon \leftarrow 1$ to E **do**
 for $b = (x, y) \in$ random batch from X_{train} **do**
 Apply Channel and Spatial Attention
 Update the breast classification model weights and biases
 $\delta_i \leftarrow \delta_i - \alpha(\Delta(L(\delta_i; b)))$
 end for
 end for
end while

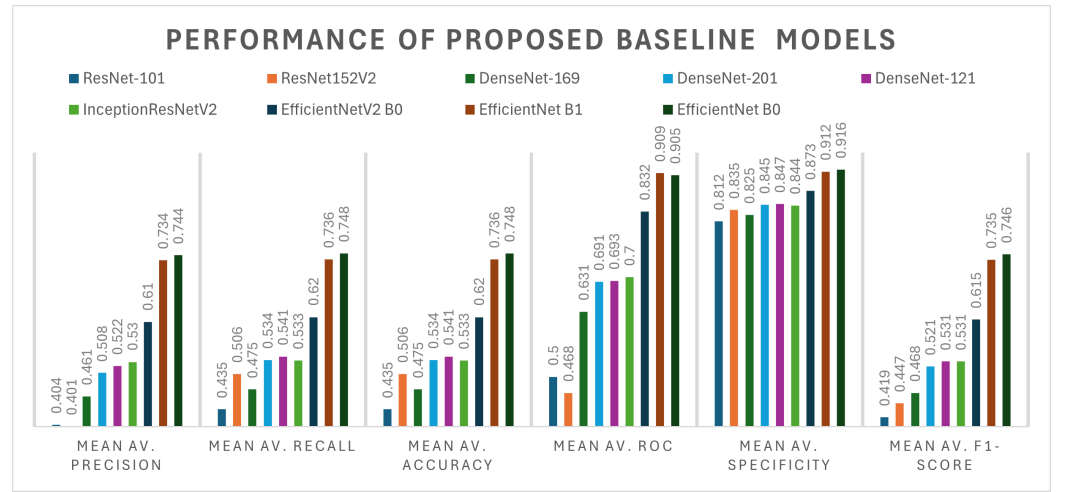
Evaluation Scores of Breast Cancer Classification Model

$A_{test} \leftarrow \text{calculateScore}(\delta, X_{test})$

return: A_{test}

Table 1. Baseline model selection.

Network	Mean Av. Precision	Mean Av. Recall	Mean Av. Accuracy	Mean Av. ROC	Mean Av. Specificity	Mean Av. F1-Score
DenseNet-121 [54]	0.522	0.541	0.541	0.693	0.847	0.531
DenseNet-169 [54]	0.461	0.475	0.475	0.631	0.825	0.468
DenseNet-201 [54]	0.508	0.534	0.534	0.691	0.845	0.521
ResNet-101 [55]	0.404	0.435	0.435	0.500	0.812	0.419
ResNet152V2 [55]	0.401	0.506	0.506	0.468	0.835	0.447
InceptionResNetV2 [56]	0.530	0.533	0.533	0.700	0.844	0.531
EfficientNet-B0 [53]	0.744	0.748	0.748	0.905	0.916	0.746
EfficientNet-B1 [53]	0.734	0.736	0.736	0.909	0.912	0.735
EfficientNetV2-B0 [53]	0.610	0.620	0.620	0.832	0.873	0.615



3.1.1. Channel and Spatial Attention Mechanisms

The CBAM block leverages both channel and spatial attention mechanisms; see Figure 3. Channel attention Equation (1) is executed by applying shared dense layers to the global average-pooled and global max-pooled representations of the input feature. The resulting attention weights are combined and applied to the input feature. Spatial attention—see Equation (2)—begins by concatenating the average- and max-pooled representations, followed by a convolutional layer that generates spatial attention weights. The input feature is then multiplied by these weights, allowing the model to focus on crucial spatial regions [57].

$$\begin{aligned}
 M_{ch}(F) &= \sigma(MLP_{ch}(Av.Pool(F)) + MLP_{ch}(Max.Pool(F))) \\
 &= \sigma(W_{1_{ch}}(W_{0_{ch}}(F_{av}^c)) + W_{1_{ch}}(W_{0_{ch}}(F_{max}^c)))
 \end{aligned} \quad (1)$$

where σ refers to the sigmoid function, $W_{0_{ch}} \in R^{C/r \times C}$ and $W_{1_{ch}} \in R^{C \times C/r}$. Here, $W_{0_{ch}}$ and $W_{1_{ch}}$ are the weights of the multi-layer perceptron (MLP) and are shared for both inputs. Additionally, the ReLU activation function is applied following $W_{0_{ch}}$.

$$\begin{aligned}
 M_{sp}(F) &= \sigma(f^{3 \times 3}([Av.Pool(F); Max.Pool(F)])) \\
 M_{sp}(F) &= \sigma(f^{3 \times 3}([F_{av}^s; F_{max}^s]))
 \end{aligned} \quad (2)$$

where σ is the sigmoid function, and a filter size of 3×3 is employed for the convolution operation $f^{3 \times 3}$.

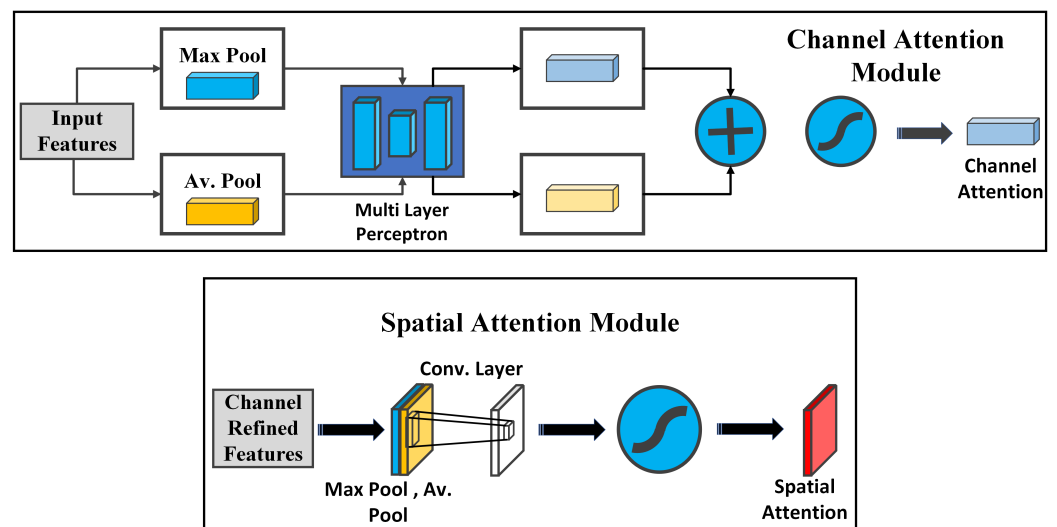


Figure 3. Channel and spatial attention mechanisms.

This comprehensive architecture is precisely designed to reinforce the model's capacity to capture relevant features through the strategic incorporation of spatial and channel attention techniques. The overall objective is to elevate the performance in tasks such as breast cancer classification. See Figure 2.

3.1.2. Loss Function

Focal loss serves as a powerful tool to address the class imbalance by assigning distinct weights to different classes, and its behavior is controlled by the parameters alpha and gamma. The focal loss function [31] is defined in Equation (3):

$$Loss = -\frac{1}{B} \sum_{i=1}^B (1 - pl(s_i))^{\gamma} \log(pl(s_i)) \quad (3)$$

where i indicates the index of the i th sample within a batch of size B , and the final dense layer outputs the probability p for the sample s_i . The ground truth is denoted by l , representing the predicted value. Here, γ is a test sample's weighting factor, setting its value to 2.

4. Results

In this section, the results that are generated by the proposed CSA-Net model are first discussed and then evaluated against the state-of-the-art models.

4.1. Materials

This section provides an overview of the datasets utilized, the experimental setup employed, the data augmentation techniques applied, and the evaluation measures used in our study. We will investigate each of these aspects to offer a comprehensive understanding of the methodology applied.

4.1.1. Dataset

This section covers the details of the datasets used for this study. The datasets include INbreast [58], MIAS [59], CBIS-DDSM [60], and the combination of CBIS-DDSM, MIAS, and INbreast (CIM) [61]. The reason for using more than one dataset is to verify the effectiveness of the proposed methodology and compare its results on multiple datasets.

A description of each dataset is given in Table 2.

Table 2. Details of binary class datasets.

Dataset	Classes	Class Names	Total Images	Benign	Malignant
INbreast	2	BM/MM	7632	2520	5112
MIAS	2	BM/MM	3816	2376	1440
CIM ¹	2	BM/MM	24,576	10,866	13,710
CBIS-DDSM	2	BM/MM ²	13,128	5970	7158

¹ CIM: CBIS-DDSM + INbreast + MIAS datasets. ² BM: Benign Mass, MM: Malignant Mass.

The datasets CBIS-DDSM, MIAS, and INbreast are composed of Dicom images. For the INbreast dataset, out of 410 case images, 106 mass cases were selected, preprocessed, and augmented. The images were preprocessed using CLAHE and augmented to a total of 7632 images of binary classes [61]. Each image belonged to the binary class of benign or malignant breast cancer. The numbers of images in the benign and malignant classes after augmentation were 2520 and 5112, respectively. The same procedure was applied for the selection of images in the CBIS-DDSM and MIAS datasets.

The Breast Ultrasound dataset [62] comprises three classes: normal, benign, and malignant. A total of 781 images are included, with 134 images for normal, 437 for benign, and 210 for malignant. The dataset is used for the cross-modality multi-class classification of the proposed method.

Summarizing this dataset overview, we provide an initial image distribution across classes in the multi-class datasets before any augmentation techniques are applied.

4.1.2. Experimental Setup

This research explores the application of deep learning alongside enhanced data preprocessing techniques for breast cancer classification using benchmark datasets. The issue of class imbalance is addressed by using the upsampling approach. The study balanced minority classes by replicating images, ensuring equitable representation in the training set. The dataset is then divided into 60% for training, 20% for validation, and 20% for testing.

During the training phase, the focal loss [63] and the Nadam optimizer are utilized for 50 epochs. To counteract overfitting, early stopping is incorporated. Furthermore, the learning rate is dynamically adjusted based on the validation performance, beginning at 0.001 and decreasing by a factor of 0.25 after every five epochs. The backend framework utilized in this study is Keras 3.2.1 and TensorFlow 2.17.0, while an NVIDIA K80 GPU is employed for the training of different model configurations.

To optimize the training process, the focal loss function is adopted as in [31]. The focal loss proves advantageous in addressing class imbalances and classifying samples, making it particularly well suited for medical image classification tasks like breast cancer diagnosis.

In terms of optimization, the Nadam optimizer is selected, blending the merits of the Adam and Nesterov accelerated gradient techniques. Renowned for its efficiency in training deep neural networks, the Nadam optimizer facilitates enhanced generalization and accelerated convergence.

To augment the diversity of the training dataset and fortify the model's resilience and generalization capabilities, various data augmentation strategies are introduced. These include random rotations, flips, zooms, and shifts, contributing to dataset enrichment and mitigating the risk of overfitting.

4.1.3. Data Augmentation

The data augmentation process incorporates a suite of transformations aimed at enriching the training dataset. The specific augmentation parameters include a rotation range of 180 degrees, allowing for comprehensive adjustments in the image orientation. To accommodate slight variations in object positioning, a width shift range of 0.1 and a height shift range of 0.1 are introduced for modest translations in the vertical and horizontal

positions, thus enhancing the model's adaptability. The inclusion of a zoom range of 0.1 enables the diverse scaling of the images, leading to dataset variability. Horizontal and vertical flips are both set to True, introducing additional image variations by mirroring along the corresponding axes. The 'nearest' fill mode ensures that any voids resulting from these transformations are filled with the closest pixel values available. Collectively, these augmentation techniques—rotation, shifting, zooming, and flipping—contribute to an augmented dataset with increased diversity, reinforcing the model's capacity for generalization and its long-term reliability.

Addressing the issue of limited data, a data augmentation strategy is employed. This strategy includes random rotations ranging from 0 to 180 degrees, horizontal and vertical flips, and adjustments in width, height, zoom, and brightness. Subsequently, image resolution is standardized to 224×224 pixels.

4.1.4. Evaluation Measures

Whenever a model undergoes training, its efficiency requires some means of measurement. Based on this, the model is evaluated as efficient or not. To evaluate the model's efficiency, the trained model is tested for its precision, recall, accuracy, specificity, and area under the curve (AUC). These efficiency measurement mechanisms are further dependent on the prediction outcomes of the confusion matrices for breast cancer images: true positive (TP_B), true negative (TN_B), false positive (FP_B), and false negative (FN_B). The evaluation measures mentioned above provide results based on certain criteria, as explained below.

Accuracy: The accuracy is the number of true predicted classification instances out of the total occurred instances [64]. It is calculated as seen in Equation (4):

$$Accuracy = \frac{TP_B + TN_B}{TP_B + TN_B + FP_B + FN_B} \quad (4)$$

Precision: Precision is the ability of the model to avoid false negatives. It is the ratio of true positive (TP_B) or correctly positively classified predictions to all positive predictions [64]. See Equation (5). Its calculation plays a vital role, especially in medical diagnostics, due to identifying false positive (FP_B) cases. It is worth mentioning that the effect of minimizing the FP is valuable in many aspects of the process of diagnosis, from costs to diagnostic procedural complexity.

$$Precision = \frac{TP_B}{TP_B + FP_B} \quad (5)$$

Sensitivity or Recall: Sensitivity is a measure of the true positive predictions that are positives; see Equation (6). The success of a model in identifying a positive case is reflected by the sensitivity. This measure is quite helpful in minimizing false negative cases [64].

$$Sensitivity = \frac{TP_B}{TP_B + FN_B} \quad (6)$$

Specificity: Specificity is defined as the fraction of true negatives accurately predicted by the model. It quantifies the model's ability to correctly identify instances that do not belong to the positive class, offering a measure of its accuracy in recognizing true negatives within the dataset [64]. See Equation (7):

$$Specificity = \frac{TN_B}{TN_B + FP_B} \quad (7)$$

F1-score: The F1-score, also known as the F-measure, serves as a pivotal metric in evaluating the effectiveness of classification models, particularly in scenarios involving binary classification tasks. The F1-score, which combines the precision and recall, offers

a balanced assessment of the model’s performance, ensuring a fair evaluation of both its accuracy and reliability.

$$F1-Score = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (8)$$

Area Under the Curve (AUC): The receiver operating characteristic (ROC) curve’s assessment outcome, commonly denoted by the area under the curve (AUC), provides valuable insights into the model’s performance. The AUC serves as a key performance metric for classification models, providing insights into their effectiveness across different threshold values. This metric is instrumental in gauging the model’s discriminatory power and overall performance in classifying instances within the dataset [64].

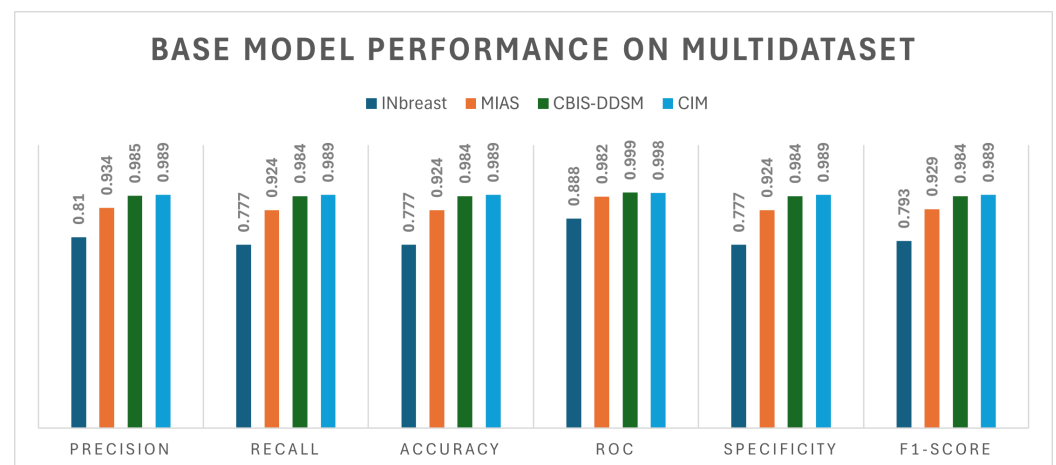
4.2. Ablation Study

The ablation study integrates three diverse datasets—CBIS-DDSM, INbreast, and MIAS—for distinct tasks within the realm of medical image classification. EfficientNet-B0 [53] is the base model used to check the performance of the binary and multi-class classification on mammogram images. The model has undergone a three-fold training procedure, and the performance evaluation of the base model is described in Table 3.

Table 3. Performance evaluation of base model.

Dataset	Classes	Class Names	Precision	Recall	Accuracy	ROC	Specificity
INbreast	2	BM/MM ²	0.810	0.777	0.777	0.888	0.777
MIAS	2	BM/MM	0.934	0.924	0.924	0.982	0.924
CBIS-DDSM	2	BM/MM	0.985	0.984	0.984	0.999	0.984
CIM ¹	2	BM/MM	0.989	0.989	0.989	0.998	0.989

¹ CIM: CBIS-DDSM + INbreast + MIAS datasets. ² BM: Benign Mass, MM: Malignant Mass.



4.2.1. Binary Classification

An ablation study is conducted to enhance the generalization of the proposed model, thus resulting in the reinforcement of the robustness and diversity of the training dataset. This study involves two classifications, utilizing the datasets CBIS-DDSM, INbreast, and MIAS for two classes, benign and malignant. Due to the uneven distribution of the classes in the experiments, we employ data upsampling techniques. Additionally, data augmentation methods are applied, including random rotation and vertical and horizontal flipping. For optimization during backpropagation, the Nadam method is considered. In our investigation of the base models, we found that a pre-trained EfficientNet-B0 [53] network achieved the highest accuracy score among all examined models; refer to Table 1.

Consequently, we use EfficientNet-B0 as the base model to evaluate the performance of binary and multi-class classification on mammogram images.

4.2.2. Cross-Modality Multi-Class Classification Results Using Ultrasound Images

This study incorporates ultrasound images from the Breast Ultrasound dataset [62] for a cross-modality analysis. This analysis necessitates evaluating the model's generalization capabilities across different imaging modalities, specifically ultrasound and mammography. The model is trained on mammography images from the CBIS-DDSM, INbreast, MIAS, and CIM datasets and subjected to evaluation using ultrasound images to assess its robustness and performance in a cross-modality context. The outcomes obtained from this evaluation offer valuable insights into the model's capacity for cross-modality generalization, providing crucial information about its adaptability and effectiveness when applied to ultrasound images, in addition to its original training modality of mammography.

This comprehensive investigation strategically leverages the strengths inherent in each imaging modality and dataset to address specific challenges in medical image classification. Through the incorporation of diverse datasets, the model gains exposure to intricate patterns associated with various imaging modalities and clinical scenarios. This holistic approach culminates in the development of a more adaptable and reliable classification model tailored to the assessment of breast health. The amalgamation of information from different sources enhances the model's capacity to discern refined features and variations, thus contributing to a robust and versatile solution for medical image analysis in the context of breast health assessment.

5. Discussion

5.1. Comparison of Results before and after Using Attention Mechanism

Analyzing the effect on the outcomes before and after the integration of the attention mechanisms yields significant insights into the influence of the mechanisms on the model's overall performance. The attention mechanism is intricately crafted to augment the model's focus on particular regions or features within the input data, with the potential to enhance its efficiency in capturing pertinent information. This comparative evaluation allows for a detailed understanding of how the attention mechanisms contribute to the model's ability to discern and prioritize crucial aspects within the dataset, ultimately influencing its overall performance.

Table 4 presents a comparative analysis of the performance metrics across different datasets with and without the implementation of the channel and spatial attention (CSA) mechanism. Specifically, the datasets examined include Inbreast, MIAS, CBIS-DDSM, and CIM (comprising a combination of the MIAS, CBIS-DDSM, and INbreast datasets). Evaluation metrics such as the precision, recall, accuracy, ROC, specificity, and F1-score are considered for each dataset.

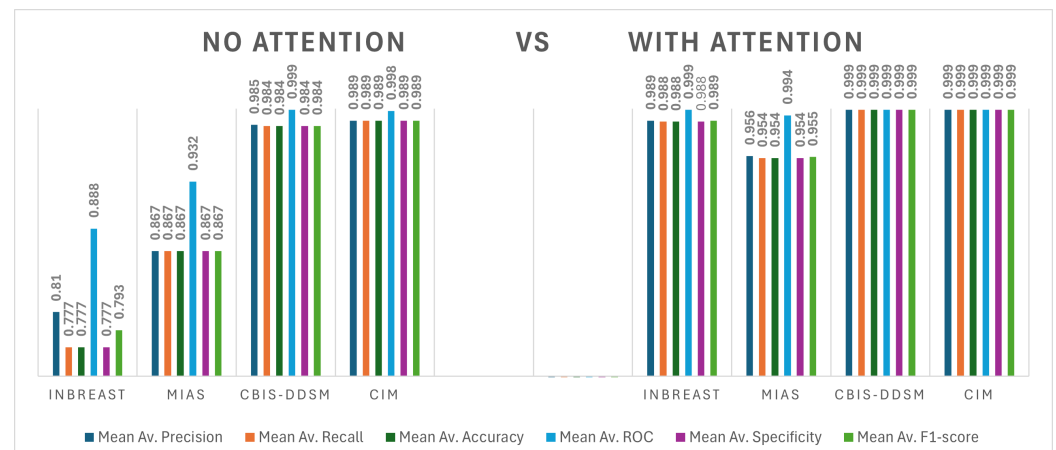
The results' analysis indicates a consistent enhancement in performance across all datasets when CSA is applied. Notably, on the Inbreast dataset, CSA leads to improvements in accuracy, sensitivity, and precision. Similarly, significant enhancements are observed in the MIAS and CBIS-DDSM datasets, particularly in parameters like the ROC score, sensitivity, and precision.

Moreover, when evaluating the combined CIM dataset, the benefits of CSA become more pronounced, with notable improvements in the ROC score, accuracy, sensitivity, and precision. These findings underscore the effectiveness of the CSA mechanism in augmenting the model's capability to extract crucial spatial and channel-wise features, consequently boosting the classification performance across diverse datasets. Overall, the inclusion of channel and spatial attention contributes to enhanced classification accuracy and model robustness across multiple datasets.

Table 4. Performance evaluation of proposed model with and without attention mechanism on binary datasets.

Dataset	Mean Av. Precision	Mean Av. Recall	Mean Av. Accuracy	Mean Av. ROC	Mean Av. Specificity	Mean Av. F1-Score
Without Attention						
Inbreast	0.810	0.777	0.777	0.888	0.777	0.793
MIAS	0.867	0.867	0.867	0.932	0.867	0.867
CBIS-DDSM	0.985	0.984	0.984	0.999	0.984	0.984
CIM	0.989	0.989	0.989	0.998	0.989	0.989
With Attention ¹						
Inbreast	0.989	0.988	0.988	0.999	0.988	0.989
MIAS	0.956	0.954	0.954	0.994	0.954	0.955
CBIS-DDSM	0.999	0.999	0.999	0.999	0.999	0.999
CIM	0.999	0.999	0.999	0.999	0.999	0.999

¹ CSA: Channel and Spatial Attention.



An important outcome after implementing CSA is the notable improvement in the accuracy, sensitivity, and F1-score across all datasets. By selectively attending to meaningful channels and spatial locations, the model becomes more proficient in detecting significant patterns and refinements in the data, resulting in more accurate classifications. Specifically, in datasets like MIAS and CBIS-DDSM, where the precision and F1-score values approach perfection after applying CSA, this enhancement is particularly striking.

Furthermore, utilizing CSA enhances the model's overall performance and discriminatory ability by improving the accuracy and area under the ROC curve (AUC). The achievement of near-perfect accuracy when employing CSA on the CBIS-DDSM dataset suggests that the attention mechanisms effectively capture the complex patterns and variations present in the data.

Moreover, the incorporation of CSA improves the model's generalization and resilience across diverse datasets. By focusing on both channel-wise and spatial-wise features, the model becomes more robust to the variations and complexities inherent in the different datasets, resulting in consistent performance improvements across datasets like INbreast, MIAS, and CBIS-DDSM.

In this study, we conduct a thorough comparison of the proposed CSA-Net approach with several existing breast cancer diagnosis methods across three datasets: CBIS-DDSM, INbreast, and MIAS (Table 5 and Figures 4 and 5). To ensure a fair and accurate comparison, identical experimental setups on all datasets are utilized. Additionally, comprehensive performance measures are provided, including the F1-score, ROC, accuracy, precision, recall, and specificity.

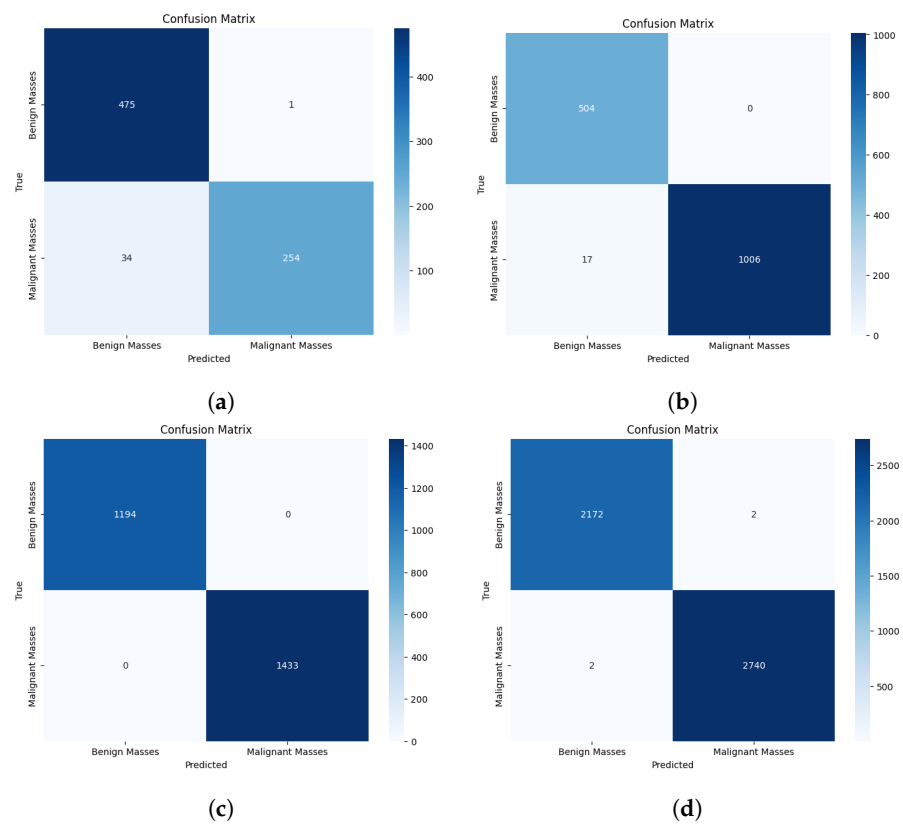


Figure 4. Confusion matrix of proposed model on different datasets. (a) Confusion matrix of MIAS dataset. (b) Confusion matrix of INbreast dataset. (c) Confusion matrix of CBIS-DDSM dataset. (d) Confusion matrix of CIM dataset.

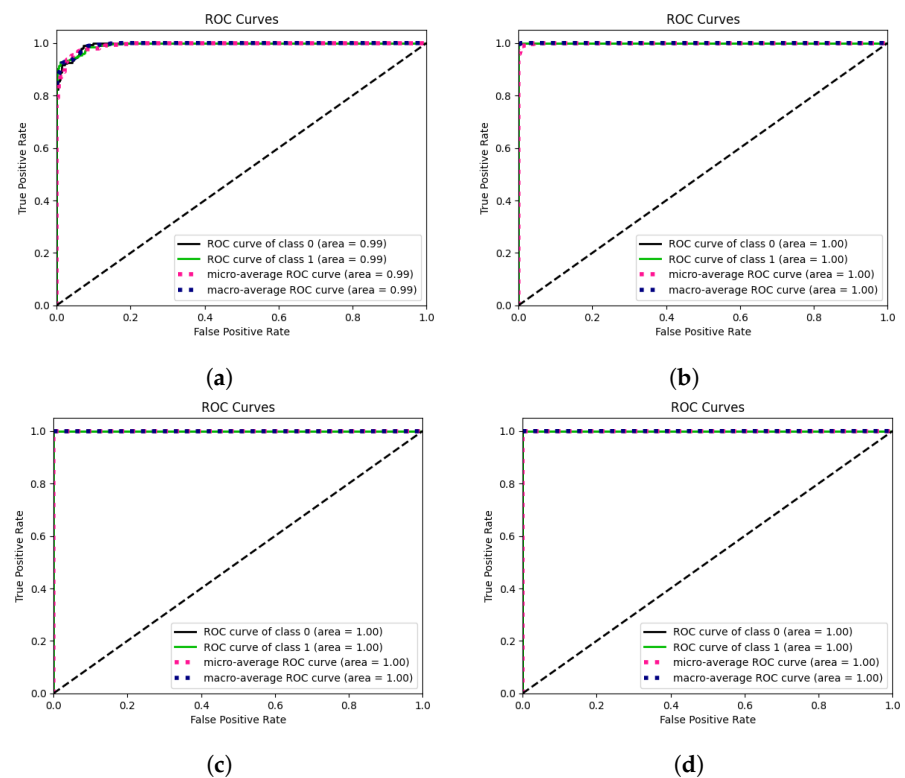
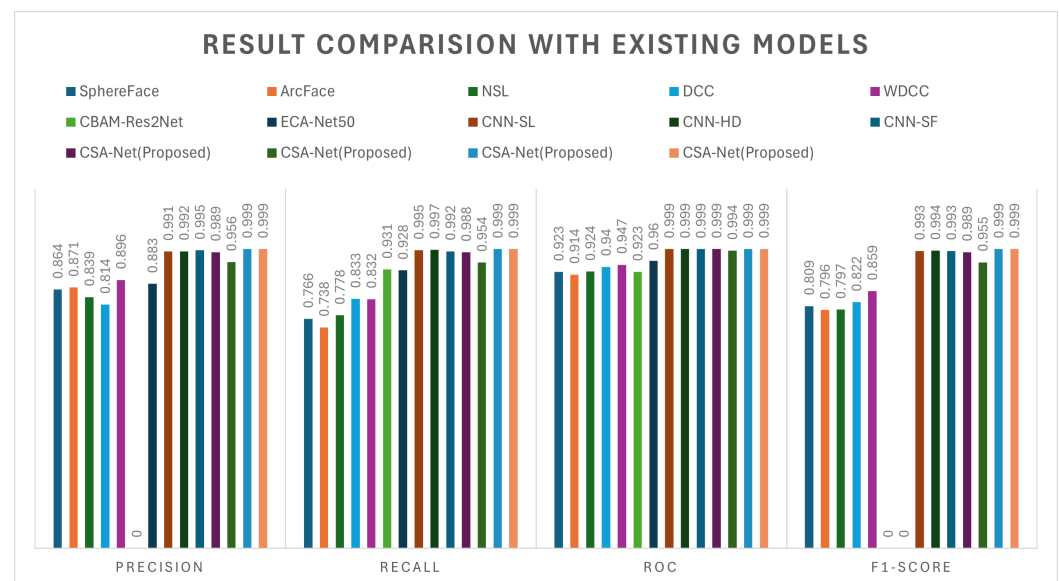


Figure 5. ROC of proposed model on different datasets. (a) ROC on MIAS dataset. (b) ROC on INbreast dataset. (c) ROC on CBIS-DDSM dataset. (d) ROC on CIM dataset.

Table 5. Result comparison of binary classification with existing models.

Citation	Dataset	Precision	Recall	Accuracy	ROC	Specificity	F1-Score
SphereFace [20]	INbreast	0.864	0.766	0.912	0.923	-	0.809
ArcFace [21]	INbreast	0.871	0.738	0.912	0.914	-	0.796
NSL [22]	INbreast	0.839	0.778	0.905	0.924	-	0.797
DCC [23]	INbreast	0.814	0.833	0.912	0.940	-	0.822
WDCC [23]	INbreast	0.896	0.832	0.934	0.947	-	0.859
CBAM-Res2Net [34]	INbreast/DDSM	-	0.931	0.933	0.923	-	-
ECA-Net50 [35]	INbreast	0.883	0.928	0.929	0.960	-	-
CNN-SL [36]	DMR-IR	0.991	0.995	0.993	0.999	0.995	0.993
CNN-HD [36]	DMR-IR	0.992	0.997	0.994	0.999	0.997	0.994
CNN-SF [36]	DMR-IR	0.995	0.992	0.993	0.999	0.992	0.993
CSA-Net	INbreast	0.989	0.988	0.988	0.999	0.988	0.989
CSA-Net	MIAS	0.956	0.954	0.954	0.994	0.954	0.955
CSA-Net	CBIS-DDSM	0.999	0.999	0.999	0.999	0.999	0.999
CSA-Net	CIM	0.999	0.999	0.999	0.999	0.999	0.999



Through multiple experiments conducted on three publicly accessible datasets, our proposed approach consistently shows improvements in its classification performance. Notably, the integration of the spatial and channel-wise attention mechanisms results in state-of-the-art outcomes, further validating the effectiveness of the proposed approach.

5.2. Cross-Modality Multi-Class Classification Results Using Breast Ultrasound Dataset

The results provided in Table 6 showcase the model's performance on the Breast Ultrasound dataset, encompassing three distinct classes: normal, benign, and malignant. The model demonstrates commendable results, achieving 92.9% precision and 92.3% recall, indicating its capacity to identify accurately positive instances while minimizing false positives. With overall accuracy of 92.3%, the model showcases reliable classification across the dataset. The high ROC score of 96.0% underscores the model's proficiency in distinguishing between classes. Additionally, the specificity score of 96.1% reflects the model's effectiveness in correctly identifying true negative cases. These results, combined

with the F1-score of 92.1%, affirm the model's robust performance in accurately classifying breast lesions in ultrasound images across diverse categories.

Table 6. Cross-modality results on Breast Ultrasound dataset.

Dataset	Classes	Precision	Recall	Accuracy	ROC	Specificity	F1-Score
Breast Ultrasound	3	0.929	0.923	0.923	0.960	0.961	0.921

CSA-Net achieves high classification accuracy, as evident in Figure 6, indicating that it effectively pinpoints lesions and assigns them to the appropriate categories. However, there are instances where the expected samples do not yield accurate predictions. In cases where only one lesion is present among the incorrectly predicted samples, our approach typically prioritizes the lesion area for classification.

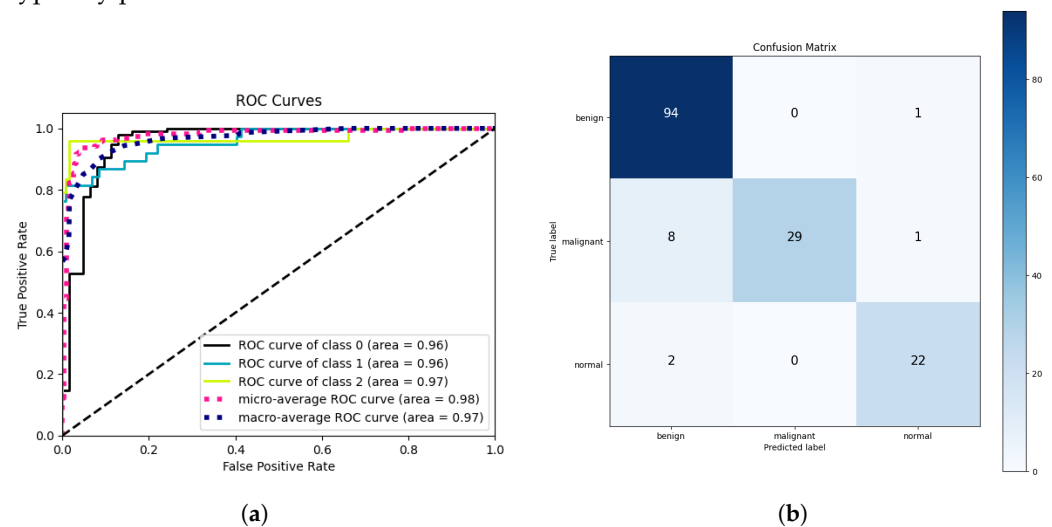


Figure 6. ROC graph and confusion matrix for multi-class cross-modality. (a) ROC on Breast Ultrasound dataset. (b) Confusion matrix of Breast Ultrasound dataset.

6. Conclusions

This study illustrates the efficiency of employing a deep learning model with channel and spatial attention mechanisms for breast cancer classification. By achieving promising outcomes on both mammogram and ultrasound datasets, the model exhibits potential for practical implementation in clinical settings. In binary classification tasks, the model shows notable improvements, especially upon the integration of channel and spatial attention mechanisms, reaching 99.9% accuracy on a combined dataset of CBIS-DDSM, INbreast, MIAS, and CIM. This enhancement emphasizes the model's capacity to discern critical features within the data, thereby enhancing its classification performance. The model's versatility across different imaging modalities is evident in its successful evaluation with breast cancer ultrasound images. This capability highlights the model's adaptability and reliability in varying clinical scenarios.

Through a cross-modality multi-class analysis involving breast cancer ultrasound images, the model demonstrates its proficiency in generalizing across diverse imaging modalities, showing accuracy of 92.3% on three classes. Its high precision, recall, accuracy, ROC score, and specificity collectively affirm its robustness in accurately classifying breast lesions, further solidifying its potential in real-world clinical applications.

Author Contributions: O.B.N.: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Data Curation, Writing—Original Draft. Y.S.: Conceptualization, Methodology, Resources, Writing—Review and Editing, Supervision, Project Administration. 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 datasets include INbreast [58], MIAS [59], CBIS-DDSM [60], and the combination of CBIS-DDSM, MIAS, and INbreast (CIM) [61].

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

References

1. Bianconi, E.; Piovesan, A.; Facchin, F.; Beraudi, A.; Casadei, R.; Frabetti, F.; Vitale, L.; Pelleri, M.C.; Tassani, S.; Piva, F.; et al. An estimation of the number of cells in the human body. *Ann. Hum. Biol.* **2013**, *40*, 463–471. [CrossRef]
2. Sciencefocus. What Cells in the Human Body Live the Longest? Available online: <http://www.sciencefocus.com/the-human-body/what-cells-in-the-human-body-live-the-longest/> (accessed on 4 April 2020).
3. Cancercenter. What Is Cancer? Available online: <https://www.cancercenter.com/what-is-cancer> (accessed on 26 December 2019).
4. Ferlay, J.; Ervik, M.; Lam, F.; Laversanne, M.; Colombet, M.; Mery, L. *Global Cancer Observatory: Cancer Today*, Version 1.1; International Agency for Research on Cancer: Lyon, France, 2024. Available online: <https://gco.iarc.who.int/today> (accessed on 23 March 2024).
5. Naeem, O.B.; Saleem, Y.; Khan, M.; Khan, A.R.; Saba, T.; Bahaj, S.A.; Ayesha, N. Breast Mammograms Diagnosis Using Deep Learning: State of Art Tutorial Review. In *Archives of Computational Methods in Engineering*; Springer: Berlin, Germany, 2024; pp. 1–19.
6. Ahmed, L.; Iqbal, M.M.; Aldabbas, H.; Khalid, S.; Saleem, Y.; Saeed, S. Images data practices for semantic segmentation of breast cancer using deep neural network. *J. Ambient. Intell. Humaniz. Comput.* **2023**, *14*, 15227–15243. [CrossRef]
7. International Agency for Research on Cancer. *Global Cancer Observatory*; International Agency for Research on Cancer: Lyon, France, 2018.
8. Ferlay, J.; Ervik, M.; Lam, F.; Colombet, M.; Mery, L.; Piñeros, M.; Znaor, A.; Soerjomataram, I.; Bray, F. *Global Cancer Observatory: Cancer Today*; International Agency for Research on Cancer: Lyon, France, 2020.
9. Heenaye-Mamode Khan, M.; Boodoo-Jahangeer, N.; Dullull, W.; Nathire, S.; Gao, X.; Sinha, G.; Nagwanshi, K.K. Multi-class classification of breast cancer abnormalities using Deep Convolutional Neural Network (CNN). *PLoS ONE* **2021**, *16*, e0256500. [CrossRef]
10. Hinton, B.; Ma, L.; Mahmoudzadeh, A.P.; Malkov, S.; Fan, B.; Greenwood, H.; Joe, B.; Lee, V.; Kerlikowske, K.; Shepherd, J. Deep learning networks find unique mammographic differences in previous negative mammograms between interval and screen-detected cancers: A case-case study. *Cancer Imaging* **2019**, *19*, 41. [CrossRef]
11. 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]
12. Al-Antari, M.A.; Al-Masni, M.A.; Kim, T.S. Deep learning computer-aided diagnosis for breast lesion in digital mammogram. In *Deep Learning in Medical Image Analysis: Challenges and Applications*; Springer: Berlin, Germany, 2020; pp. 59–72.
13. Yu, X.; Pang, W.; Xu, Q.; Liang, M. Mammographic image classification with deep fusion learning. *Sci. Rep.* **2020**, *10*, 14361. [CrossRef]
14. 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]
15. Carneiro, G.; Nascimento, J.; Bradley, A.P. Unregistered multiview mammogram analysis with pre-trained deep learning models. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; Springer: Berlin/Heidelberg, Germany, 2015; pp. 652–660.
16. Carneiro, G.; Nascimento, J.; Bradley, A.P. Automated analysis of unregistered multi-view mammograms with deep learning. *IEEE Trans. Med. Imaging* **2017**, *36*, 2355–2365. [CrossRef]
17. Wu, N.; Phang, J.; Park, J.; Shen, Y.; Huang, Z.; Zorin, M.; Jastrzębski, S.; Févry, T.; Katsnelson, J.; Kim, E.; et al. Deep neural networks improve radiologists' performance in breast cancer screening. *IEEE Trans. Med. Imaging* **2019**, *39*, 1184–1194. [CrossRef]
18. Shu, X.; Zhang, L.; Wang, Z.; Lv, Q.; Yi, Z. Deep neural networks with region-based pooling structures for mammographic image classification. *IEEE Trans. Med. Imaging* **2020**, *39*, 2246–2255. [CrossRef]
19. Zhang, F.; Luo, L.; Sun, X.; Zhou, Z.; Li, X.; Yu, Y.; Wang, Y. Cascaded generative and discriminative learning for microcalcification detection in breast mammograms. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 12578–12586.
20. Liu, W.; Wen, Y.; Yu, Z.; Li, M.; Raj, B.; Song, L. Sphereface: Deep hypersphere embedding for face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 212–220.
21. Deng, J.; Guo, J.; Xue, N.; Zafeiriou, S. Arcface: Additive angular margin loss for deep face recognition. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 4690–4699.

22. Wang, H.; Wang, Y.; Zhou, Z.; Ji, X.; Gong, D.; Zhou, J.; Li, Z.; Liu, W. Cosface: Large margin cosine loss for deep face recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 5265–5274.
23. Wang, Y.; Wang, Z.; Feng, Y.; Zhang, L. WDCCNet: Weighted Double-Classifer Constraint Neural Network for Mammographic Image Classification. *IEEE Trans. Med. Imaging* **2022**, *41*, 559–570.
24. Leibig, C.; Allken, V.; Ayhan, M.S.; Berens, P.; Wahl, S. Leveraging uncertainty information from deep neural networks for disease detection. *Sci. Rep.* **2017**, *7*, 17816. [[CrossRef](#)] [[PubMed](#)]
25. Carneiro, G.; Pu, L.Z.C.T.; Singh, R.; Burt, A. Deep learning uncertainty and confidence calibration for the five-class polyp classification from colonoscopy. *Med. Image Anal.* **2020**, *62*, 101653. [[CrossRef](#)] [[PubMed](#)]
26. 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.
27. S. R. Sannasi Chakravarthy, N.B.; Rajaguru, H. Multi-Deep CNN based Experimentations for Early Diagnosis of Breast Cancer. *IETE J. Res.* **2023**, *69*, 7326–7341.
28. He, K.; Gan, C.; Li, Z.; Rekik, I.; Yin, Z.; Ji, W.; Gao, Y.; Wang, Q.; Zhang, J.; Shen, D. Transformers in medical image analysis. *Intell. Med.* **2023**, *3*, 59–78.
29. Wang, J.; Lu, L.; Zhang, Z.; Slam, N. A novel deep convolution neural network model for ct image classification based on covid-19. In Proceedings of the 2022 3rd International Conference on Computer Vision, Image and Deep Learning & International Conference on Computer Engineering and Applications (CVIDL & ICCEA), Changchun, China, 20–22 May 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 15–20.
30. Deeb, A.; Debow, A.; Mansour, S.; Shkodyrev, V. COVID-19 diagnosis with Deep Learning: Adjacent-pooling CTScan-COVID-19 Classifier Based on ResNet and CBAM. *Biomed. Signal Process. Control* **2023**, *86*, 105285.
31. Naveed, A.; Naqvi, S.S.; Khan, T.M.; Razzak, I. PCA: Progressive class-wise attention for skin lesions diagnosis. *Eng. Appl. Artif. Intell.* **2024**, *127*, 107417. [[CrossRef](#)]
32. Loizidou, K.; Elia, R.; Pitris, C. Computer-aided breast cancer detection and classification in mammography: A comprehensive review. *Comput. Biol. Med.* **2023**, *153*, 106554.
33. Zou, L.; Yu, S.; Meng, T.; Zhang, Z.; Liang, X.; Xie, Y. A technical review of convolutional neural network-based mammographic breast cancer diagnosis. *Comput. Math. Methods Med.* **2019**, *2019*, 6509357. [[CrossRef](#)]
34. Niu, J.; Li, H.; Zhang, C.; Li, D. Multi-scale attention-based convolutional neural network for classification of breast masses in mammograms. *Med. Phys.* **2021**, *48*, 3878–3892. [[CrossRef](#)] [[PubMed](#)]
35. Lou, Q.; Li, Y.; Qian, Y.; Lu, F.; Ma, J. Mammogram classification based on a novel convolutional neural network with efficient channel attention. *Comput. Biol. Med.* **2022**, *150*, 106082. [[CrossRef](#)] [[PubMed](#)]
36. Alshehri, A.; AlSaeed, D. Breast cancer detection in thermography using convolutional neural networks (cnns) with deep attention mechanisms. *Appl. Sci.* **2022**, *12*, 12922. [[CrossRef](#)]
37. Rashed, E.; El Seoud, M.S.A. Deep learning approach for breast cancer diagnosis. In Proceedings of the 8th International Conference on Software and Information Engineering, Cairo, Egypt, 9–12 April 2019; pp. 243–247.
38. Montelius, L.; Rezkalla, G. Providing Mass Context to a Pretrained Deep Convolutional Neural Network for Breast Mass Classification, 2019, Thesis. Available online: <https://kth.diva-portal.org/smash/record.jsf?pid=diva2%3A1354169&dswid=-9683> (accessed on 1 July 2024).
39. Ragab, D.A.; Sharkas, M.; Marshall, S.; Ren, J. Breast cancer detection using deep convolutional neural networks and support vector machines. *PeerJ* **2019**, *7*, e6201. [[CrossRef](#)]
40. Kumar, A.; Mukherjee, S.; Luhach, A.K. Deep learning with perspective modeling for early detection of malignancy in mammograms. *J. Discret. Math. Sci. Cryptogr.* **2019**, *22*, 627–643. [[CrossRef](#)]
41. Gao, Y.; Geras, K.J.; Lewin, A.A.; Moy, L. New frontiers: An update on computer-aided diagnosis for breast imaging in the age of artificial intelligence. *AJR Am. J. Roentgenol.* **2019**, *212*, 300. [[CrossRef](#)]
42. Falconí, L.G.; Pérez, M.; Aguilar, W.G. Transfer learning in breast mammogram abnormalities classification with mobilenet and nasnet. In Proceedings of the 2019 International Conference on Systems, Signals and Image Processing (IWSSIP), Osijek, Croatia, 5–7 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 109–114.
43. Abdelhafiz, D.; Yang, C.; Ammar, R.; Nabavi, S. Deep convolutional neural networks for mammography: Advances, challenges and applications. *BMC Bioinform.* **2019**, *20*, 281. [[CrossRef](#)]
44. Tsochatzidis, L.; Costaridou, L.; Pratikakis, I. Deep learning for breast cancer diagnosis from mammograms—A comparative study. *J. Imaging* **2019**, *5*, 37. [[CrossRef](#)]
45. Rampun, A.; López-Linares, K.; Morrow, P.J.; Scotney, B.W.; Wang, H.; Ocaña, I.G.; Maclair, G.; Zwigglelaar, R.; Ballester, M.A.G.; Macía, I. Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network. *Med. Image Anal.* **2019**, *57*, 1–17. [[CrossRef](#)]
46. Xi, P.; Shu, C.; Goubran, R. A Unified Deep Learning Framework for Multi-Modal Multi-Dimensional Data. In Proceedings of the 2019 IEEE International Symposium on Medical Measurements and Applications (MeMeA), Istanbul, Turkey, 26–28 June 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 1–6.
47. Ting, F.F.; Tan, Y.J.; Sim, K.S. Convolutional neural network improvement for breast cancer classification. *Expert Syst. Appl.* **2019**, *120*, 103–115. [[CrossRef](#)]

48. Panda, S.; Jangid, M. Improving the model performance of deep convolutional neural network in MURA dataset. In Proceedings of the Smart Systems and IoT: Innovations in Computing: Proceeding of SSIC 2019 ; Springer: Singapore, 2020; pp. 531–541.
49. Li, C.; Wang, X.; Liu, W.; Latecki, L.J.; Wang, B.; Huang, J. Weakly supervised mitosis detection in breast histopathology images using concentric loss. *Med. Image Anal.* **2019**, *53*, 165–178. [CrossRef] [PubMed]
50. Chougrad, H.; Zouaki, H.; Alheyane, O. Multi-label transfer learning for the early diagnosis of breast cancer. *Neurocomputing* **2020**, *392*, 168–180. [CrossRef]
51. Castro, E.; Pereira, J.C.; Cardoso, J.S. Weight Rotation as a Regularization Strategy in Convolutional Neural Networks. In Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Berlin, Germany, 23–27 July 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 2106–2110.
52. Rampun, A.; Wang, H.; Scotney, B.; Morrow, P.; Zwiggelaar, R. Classification of mammographic microcalcification clusters with machine learning confidence levels. In Proceedings of the 14th International Workshop on Breast Imaging (IWBI 2018), Atlanta, GA, USA, 8–11 July 2018; SPIE: Kuala Lumpur, Malaysia, 2018; Volume 10718, pp. 345–352.
53. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Long Beach, CA, USA, 9–15 June 2019; PMLR; 2019; pp. 6105–6114.
54. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
55. 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; pp. 770–778.
56. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the AAAI conference on artificial intelligence, San Francisco, CA, USA, 4–9 February 2017; Volume 31.
57. Woo, S.; Park, J.; Lee, J.Y.; Kweon, I.S. Cbam: Convolutional block attention module. In Proceedings of the European Conference on Computer Vision (ECCV), Munich, Germany, 8–14 September 2018; pp. 3–19.
58. Moreira, I.C.; Amaral, I.; Domingues, I.; Cardoso, A.; Cardoso, M.J.; Cardoso, J.S. Inbreast: Toward a full-field digital mammographic database. *Acad. Radiol.* **2012**, *19*, 236–248. [CrossRef]
59. Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Ricketts, I.; Stamatakis, E.; Cerneaz, N.; Kok, S.; et al. *Mammographic Image Analysis Society (Mias) Database v1*. 21; Apollo—University of Cambridge Repository: Cambridge, UK, 2015.
60. Sawyer-Lee, R.; Gimenez, F.; Hoogi, A.; Rubin, D. Curated Breast Imaging Subset of Digital Database for Screening Mammography (CBIS-DDSM) [Data set]. The Cancer Imaging Archive, 2016. Available online: <https://www.cancerimagingarchive.net/collection/cbis-ddsm/> (accessed on 1 July 2024).
61. Huang, M.L.; Lin, T.Y. Dataset of breast mammography images with masses. *Data Brief* **2020**, *31*, 105928.
62. Al-Dhabyani, W.; Gomaa, M.; Khaled, H.; Fahmy, A. Dataset of breast ultrasound images. *Data Brief* **2020**, *28*, 104863. [CrossRef]
63. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.
64. Grandini, M.; Bagli, E.; Visani, G. Metrics for multi-class classification: An overview. *arXiv* **2020**, arXiv:2008.05756.

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.