

## Article

# TTCNN: A Breast Cancer Detection and Classification towards Computer-Aided Diagnosis Using Digital Mammography in Early Stages

Sarmad Maqsood <sup>1</sup>, Robertas Damaševičius <sup>1,\*</sup> and Rytis Maskeliūnas <sup>2</sup>

<sup>1</sup> Department of Software Engineering, Kaunas University of Technology, LT-51386 Kaunas, Lithuania; sarmad.maqsood@ktu.edu

<sup>2</sup> Department of Multimedia Engineering, Kaunas University of Technology, LT-51423 Kaunas, Lithuania; rytis.maskeliunas@ktu.lt

\* Correspondence: robertas.damasevicius@ktu.lt

**Abstract:** Breast cancer is a major research area in the medical image analysis field; it is a dangerous disease and a major cause of death among women. Early and accurate diagnosis of breast cancer based on digital mammograms can enhance disease detection accuracy. Medical imagery must be detected, segmented, and classified for computer-aided diagnosis (CAD) systems to help the radiologists for accurate diagnosis of breast lesions. Therefore, an accurate breast cancer detection and classification approach is proposed for screening of mammograms. In this paper, we present a deep learning system that can identify breast cancer in mammogram screening images using an “end-to-end” training strategy that efficiently uses mammography images for computer-aided breast cancer recognition in the early stages. First, the proposed approach implements the modified contrast enhancement method in order to refine the detail of edges from the source mammogram images. Next, the transferable texture convolutional neural network (TTCNN) is presented to enhance the performance of classification and the energy layer is integrated in this work to extract the texture features from the convolutional layer. The proposed approach consists of only three layers of convolution and one energy layer, rather than the pooling layer. In the third stage, we analyzed the performance of TTCNN based on deep features of convolutional neural network models (InceptionResNet-V2, Inception-V3, VGG-16, VGG-19, GoogLeNet, ResNet-18, ResNet-50, and ResNet-101). The deep features are extracted by determining the best layers which enhance the classification accuracy. In the fourth stage, by using the convolutional sparse image decomposition approach, all the extracted feature vectors are fused and, finally, the best features are selected by using the entropy controlled firefly method. The proposed approach employed on DDSM, INbreast, and MIAS datasets and attained the average accuracy of 97.49%. Our proposed transferable texture CNN-based method for classifying screening mammograms has outperformed prior methods. These findings demonstrate that automatic deep learning algorithms can be easily trained to achieve high accuracy in diverse mammography images, and can offer great potential to improve clinical tools to minimize false positive and false negative screening mammography results.

**Keywords:** biomedical image processing; computer aided diagnosis; breast cancer detection; digital mammograms; deep learning



**Citation:** Maqsood, S.; Damaševičius, R.; Maskeliūnas, R. TTCNN: A Breast Cancer Detection and Classification towards Computer-Aided Diagnosis Using Digital Mammography in Early Stages. *Appl. Sci.* **2022**, *12*, 3273. <https://doi.org/10.3390/app12073273>

Academic Editors: Won-Du Chang, Do-Won Kim, Youngjin Jung and Hyun Jun Park

Received: 14 February 2022

Accepted: 19 March 2022

Published: 23 March 2022

**Publisher's Note:** MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Copyright:** © 2022 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

Cancer is considered to be the gradual expansion and growth of abnormal cells in the human body. Breast cancer has been among the major cause of death among women worldwide [1,2]. Breast cancer has a higher death rate than any other disease, i.e., malaria or tuberculosis. The World Health Organization (WHO) cancer research institute (International agency for research on Cancer (IARC) and the American Cancer Society) reported that in 2018, there were 17.1 million cancer cases globally and it is expected to double by 2025 [3].

Although extensive research has been done by medical professionals and researchers, they cannot provide the best method of treating breast cancer to obtain the long-awaited treatment and guarantee the potential for reliable evidence for its prevention [4]. In addition, some important parts of cancerous tissue related to breast cancer are aggressive and pose the greatest threat to the lives of patients, because they are more likely to cause infections in other vital organs of the human body [5,6]. The enormous growth of the area of the mammary cells can cause tumors in women. Depending on the area, size, and location, these large tumor cells divide into cancer cells and non-cancer cells based on breast cancer reporting and data system (BI-RAD) scores [7,8]. Non-cancerous tumors are known as benign which is signified to the primary tumor area and the cancerous tumors are known as malignant and secondary tumor area. Benign tumors are treatable and their expansion can be restrained by taking the proper medications and will not endanger the lives of women [9,10]. Secondary tumors can spread to distant metastases or adjacent tissues. When cancer cells penetrate the lymphatic system or blood, malignant tumors can escalate to different parts of the body. The tumor stems from uncontrolled cell proliferation in the breast [11,12]. A malignant tumor can be treated only if the patient acquires appropriate treatment with surgery or radiation [13,14].

In the breast, cancer cells can increase to the lymph nodes and affect other body parts, i.e., lungs. Breast cancer usually begins with ductal dysfunction (invasive ductal carcinoma). Although, it can also originate in glandular tissues and other cells called lobules and breast tissue [15]. The researchers also found that changes in hormones, lifestyle, and the environment also increase the risk of breast cancer [16,17]. Low dose X-ray examination of the breast is utilized to envisage the internal structure of the breast. This process is medically called mammography. This is evaluated as the best suitable approach for the detection of breast cancer. Compared to previously used equipment, mammography shows the breast to a much lower radiation dose [18]. It is one of the most reliable screening tools and has been shown to be a significant approach for early breast cancer detection in recent times [19]. A mammogram of each breast was recorded in two different views, namely the medial-lateral oblique (MLO) view and the cranio-caudal (CC) view as shown in Figure 1. Estimated new cases recorded in 2021 were 281,550 and the percentage of all the new cancer cases was 14.8% in the U.S. The estimated deaths in 2021 was around 43,600 having a percentage of 7.2%. The 5-year relative survival rate of breast cancer from 2014 to 2019 was 92.16% in the U.S [20].

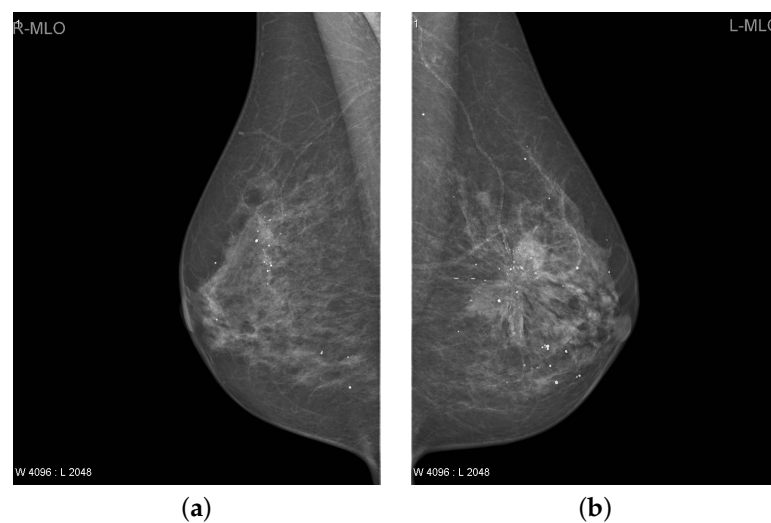
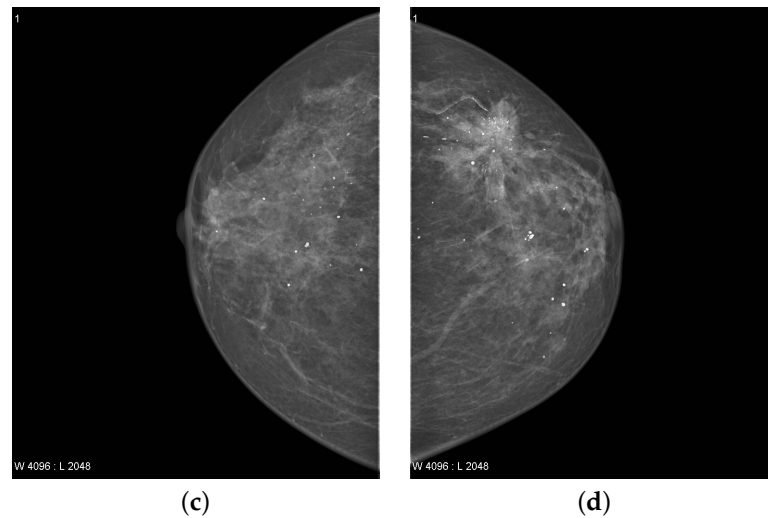


Figure 1. Cont.



**Figure 1.** The breast mammogram of a patient [21]. (a,b) Right and left medial-lateral oblique (MLO) view, (c,d) Right and left cranio-caudal (CC) view.

Many researchers focus on biomedical imaging to assist specialist radiologists. Early detection of tumors has a significant role in the diagnosis and analysis of breast cancer. Detection of breast cancer malignancies in early stages is important. Various biomedical imaging methods have been utilized to determine breast cancer, i.e., magnetic resonance imaging (MRI), ultrasound, and mammography [22]. Owing to the huge number of images, radiologists face many problems identifying suspicious areas or cancer. Therefore, to reduce the load, an efficient automated method is required to help radiologists significantly. A computer-aided diagnostic (CAD) model is used to assist radiologists detect malignant breast cancer tumors [23,24].

Recently, some research based on deep learning methods has addressed the concept of medical imaging on automated CAD systems [8,25–30]. In fact, deep learning approaches are the best option for medical image recognition and classification. A CAD system based on deep learning requires three successive phases: the pre-processing, parameters initialization, extraction of deep features and diagnosis [8,25]. Deep learning methods can be utilized to directly derive the large low–high levels deep hierarchical feature maps from the identical source breast image (mammogram) [8]. This shows that deep learning is the most efficient and dependable medical imaging method [8,25,31]. Recently, several CAD systems based on deep learning have been introduced to detect the breast lesions and are superior to traditional systems [32,33]. Automated detection of the breast lesions is the main task for the automatic evolution of CAD approaches to improve the diagnosis of breast cancer [8]. Accurate diagnosis of skeptical breast lesions plays an important part in obtaining a high true positive rate and enhancing the breast cancer final diagnosis [8,26]. Detecting breast lesions is a difficult task because these lesions are so diverse in shape, texture, size, and position. The use of image processing and deep learning methods to detect the breast lesions using traditional methods are proposed in [11,26]. Traditional technology relies on low-level manual functions and does not yet have the ability to perform identification tasks automatically [25]. Low contrast of mammogram images and detecting asymptomatic breast cancer through mammography is a difficult image classification task from the view of machine learning (ML) because the tumor itself only makes up a small part of the complete breast image. Recently, numerous new approaches based on deep learning recognition have been introduced to solve the problem of poor diagnosis performance of CAD approaches [8,31]. Classification of breast lesions is the last phase in the CAD model. The aim of this stage is to identify breast lesions that are considered benign or malignant [25,34]. The importance of the method of classification of breast lesions leans on the validity of the hypothetical characteristics which indicate the

leading characteristics [8]. Deep features must have an appropriate feature distribution in order to identify the malignant and benign breast lesions [8,9].

The existing breast cancer detection system is time consuming, expensive, and requires additional work to operate the equipment for radiologists. From an image processing perspective, accurate and automatic detection of breast cancer is not easy. Hence, it is deemed a requisite to diagnose breast cancer at an early stage and treated properly. However, breast cancer detection requires proper screening system and automation because of the following reasons [5]:

1. False diagnosis and prediction.
2. The tumor appears in an area with significantly lower contrast.
3. Lack of reliability of human experience in diagnosis.
4. Overload of the radiologists.
5. High complexity of memory.
6. Degree of human inaccuracy in diagnosis.
7. Current deep learning algorithms need large amounts of training data to overcome the problems of overfitting.
8. Existing breast cancer detection methods are computationally complex and require more treatment time to identify accurate tumors.

In addition, manual breast cancer diagnosis can take months and the localized stage of cancer infection can progress to a critical stage where survival cannot be achieved. Therefore, an accurate automated breast cancer detection method is required to diagnose cancer [35]. Today, deep learning approaches have been introduced for breast lesion detection and classification to enhance their accuracy, and have become an important element of CAD systems. The CAD techniques based on image processing are becoming more popular to accelerate the work of doctors for breast cancer diagnosis and reduce diagnosis time [8,36]. The proposed TTCNN system is effectual for early detection of breast cancer, however, the proposed system is effective for all stages of cancer. Detecting the small tumor in a whole breast is a more difficult task as compared to detecting the large area of tumor at a later stage. To this end, we propose a new approach of early breast cancer detection, which has the following contributions.

1. As a pre-processing step, a modified contrast-limited adaptive histogram equalization (CLAHE) algorithm is utilized to refine the edge details of the source image.
2. We propose a new transferable texture convolutional neural network (TTCNN) for the classification of breast cancer.
3. We use the energy layer (EL) in the proposed TTCNN framework. This allows us to preserve texture information, limit the size of the output vector, and refine the model's overall learning ability.
4. We perform the detailed functions of the most advanced CNN model for breast cancer classification i.e., InceptionResNet-V2, Inception-V3, VGG-16, VGG-19, GoogLeNet, ResNet-18, ResNet-50, and ResNet-101.
5. A convolutional sparse image decomposition approach is implemented for the feature fusion and entropy controlled firefly approach is employed to optimize the feature selection.
6. Aiming to check the stability of the proposed system, a comprehensive statistical analysis and comparison with the latest algorithms is performed.
7. The proposed approach minimizes the time required for radiologists to diagnose cancer while assuring reliable accuracy of detection. The proposed approach takes less processing time and provides accurate results.

The rest of the article is structured as follows: Section 2 briefly expresses the literature review involved in the detection and classification of breast cancer. The detailed methodology of the proposed system is presented in Section 3. Section 4 analyzes the proposed approach performance in comparison with other modern approaches. Finally, we conclude this work with the direction of future research in Section 5.

## 2. Related Work

Modern medical procedures actively use mammography images for breast cancer diagnosis [37]. This section systematically reviews the esteem achievement and performance on breast cancer detection and classification.

Over the past few years, deep learning methods in the field of image recognition, segmentation, detection, and computer vision have received much attention and interest [8,28,38]. In fact, deep learning has been utilized to address the inadequacies of traditional CAD approaches [8]. This is because conventional CAD approaches are based on manually developed functions and have limited diagnostic accuracy [35]. The similarity between benign and malignant lesions and their enormous changes in size, shape, color, and texture have challenged traditional methods [19,25]. CAD approaches have demonstrated the ability to use deep learning to discover complex hierarchical deep features automatically to improve diagnostic operation [8,32,35]. Many CAD approaches use conventional image processing and deep learning methods, but the use of deep learning detection and classification in breast imaging is inadequate for breast lesion diagnosis.

In recent years, AlexNet [39], ResNet [40], VGG16 [41], Inception [42], GoogleNet [43], DenseNet [44] deep learning models provide improve performance of classification as compared to the shallower ones. The overall classification accuracy of VGG16, ResNet50, and Inception-V3 reached 95%, 92.5%, and 95.5% respectively. The performance of classification of these deep learning methods is much better than that of the shallow model, but manually detection and memory complexity during training still remain challenges.

Several scholars have worked on the automatic detection and classification of breast cancer aiming to enhance the accuracy. For example, Al-antari et al. [4] evaluated the use of the YOLO detector to perceive breast injury from mammograms. Three deep learning classifiers, i.e., Inception ResNet-V2, ResNet-50, and feedforward CNN are used to assess classification of breast lesions using a digital database for screening mammography (DDSM) and the INbreast database. With these classifiers, the method acquired an accuracy rate of 95.32%. Al-antari et al. [8] used a CAD model (four-fold cross-validation) to estimate a full resolution convolutional network (FrCN) with X-ray mammograms using INbreast dataset. This approach achieved an accuracy of 95.96%, F1 score of 99.24%, and matthews correlation coefficient of 98.96%.

Chouhan et al. [19] proposed a diverse features (DFeBCD) method for breast cancer detection to classify mammograms as normal or abnormal. The IRMA mammography dataset uses two classifiers, a support vector machine (SVM) and an emotion learning-inspired integrated classifier (ELiEC). The ELiE classifier performance is superior to SVM and the accuracy rate reaches 80.30%. Muduli et al. [45] used the lifting wavelet transform (LWT) to obtain the region of interest features from the breast images. The size of the feature vectors diminish using a fusion of principal component analysis (PCA) and linear discriminant analysis (LDA) approaches. The extreme learning machine (ELM) and moth flame optimization approach is used for the classification using the DDSM and mammographic image analysis society (MIAS) databases. This approach attained an accuracy of 98.76% and 95.80% for DDSM and MIAS databases respectively. Junior et al. [46] presented a breast cancer method based on diversity analysis, geostatistics and alpha form using the SVM classifier on DDSM and MIAS datasets and achieved the detection accuracy of 96.30%. Ghosh et al. [47] presented an algorithm for segmentation utilizing intuitionistic fuzzy soft set and multigranulation rough set. The intuitionistic fuzzy soft set takes information from input images through various fuzzy membership operations. This approach deals with ambiguity between damaged and undamaged pixels when forming the membership function. This reduces the distant pixels that are not in the region of interest, the diseased tissue in the mammogram is separated by a rough approximation of the fuzzy concept of the multiregional granulation space. Zheng et al. [48] developed an efficient adaboost deep learning method for the recognition of breast cancer. This work focused on the combination of various deep learning approaches with feature extraction and selection using

classification and segmentation methods to evaluate the result and focus on finding the most suitable method and obtained an accuracy of 97.2%.

Harefa et al. [49] utilized the gray level co-occurrence matrix (GLCM) and SVM classifier to detect breast cancer abnormalities on the MIAS database and achieved an accuracy of 93.88% and outperformed k-Nearest Neighbour (kNN). Samala et al. [50] presented a use of the deep neural network for multilevel transfer learning for digital breast tomosynthesis. ImageNet knowledge first records mammogram information and then optimizes it during multi-step transmission of digital breast tomosynthesis information. Then the freezing of most of the convolution neural network (CNN) framework was compared to two transmission systems with the first convolution layer. Qi et al. [51] proposed a deep active learning approach for the classification of breast cancer with the aim of maximizing the precision of learning with very limited labels. This approach manually annotates the most valuable unlabeled samples and integrates them into the training set. Then update the system repeatedly as the training set grows. Two deep active learning framework selection strategies are proposed i.e., confidence boosting strategy and an entropy based strategy on the breast cancer histopathological (BreakHis) database. This approach reduced the cost of annotation by up to 66.67%, which is more accurate than the standard query strategy. Agarwal et al. [52] developed a patch-based CNN algorithm for the identification of lesions in breast in the full-field digital mammograms (FFDM) using a pre-trained CNN models (VGG16, ResNet50, and InceptionV3) for the extraction of features and attained the results of detection with a true positive rate (TPR) of 98%.

Irfan et al. [53] developed an algorithm based on dilated semantic segmentation network (DICNN) with morphological operation. A DenseNet201 is used for the feature extraction. The feature vectors acquired from DenseNet201 and 24 layer CNN employing parallel fusion were fused to categorize the nodules. These feature vectors are merged with the SVM classification and attained an accuracy of 98.9%. Rajnikanth et al. [54] developed a breast cancer detection method using breast thermal images. The two image features GLCM and local binary pattern (LBP) with heterogeneous weights are used to classify the breast thermal images into healthy and DCIS class by using the SVM and decision tree (DT) classifier. This algorithm had an accuracy of 92%. Kadry et al. [55] presented a joint thresholding (slime mould method and Shannon's entropy) and watershed segmentation method to improve and obtain the breast tumor from 2D MRI slices. The extracted breast tumor and ground truth image is executed and the essential image performance values (IPV) are calculated. Some of the existing works with dataset information and results are illustrated in Table 1.

**Table 1.** Detailed summary of recent related works on breast cancer detection and classification.

| References           | Approach and Methods Used  | Modality    | Dataset        | Results                      |
|----------------------|--|-------------|----------------|------------------------------|
| Al-antari et al. [4] | YOLO detector, Inception ResNet-V2, ResNet-50, and feedforward CNN | Mammography | DDSM, INbreast | Accuracy = 95.32%            |
| Al-antari et al. [8] | Full Resolution Convolutional Network (FrCN)                       | Mammography | INbreast       | Accuracy = 95.96%            |
| Chouhan et al. [19]  | Diverse Features (DFeBCD), SVM, and ELiEC                          | Mammography | IRMA           | Accuracy = 80.30%            |
| Muduli et al. [45]   | Lifting Wavelet Transform (LWT)                                    | Mammography | DDSM, MIAS     | Accuracy = 98.76% and 95.80% |
| Junior et al. [46]   | Diversity Analysis, Geostatistics, and Alpha Form                  | Mammography | DDSM, MIAS     | Accuracy = 96.30%            |
| Ghosh et al. [47]    | Intuitionistic fuzzy soft and Multigranulation rough set           | Mammography | MIAS, MedPix   | Accuracy = 87.22%            |
| Zheng et al. [48]    | Efficient Adaboost   | Mammography | ACRIN          | Accuracy = 97.2%             |

Table 1. Cont.

| References              | Approach and Methods Used     | Modality       | Dataset        | Results           |
|-------------------------|-------------------------------|----------------|----------------|-------------------|
| Harefa et al. [49]      | GLCM and SVM                  | Mammography    | MIAS           | Accuracy = 93.88% |
| Samala et al. [50]      | Multi-stage transfer learning | Mammography    | DDSM           | AUC = 0.92        |
| Qi et al. [51]          | Deep active learning          | Histopathology | BreakHis       | -                 |
| Agarwal et al. [52]     | Patch-based CNN               | Mammography    | DDSM, INbreast | TPR = 0.93        |
| Rajinikanth et al. [54] | GLCM and LBP                  | Thermography   | DITI-India     | Accuracy = 92%    |

From literature research, previous breast cancer detection and classification approaches have led to better information extraction. However, there are still numerous problems that require serious attention, i.e., (i) the tumor appears in a location with significantly lower contrast, (ii) memory complexity is high, (iii) existing approaches are computationally complex and require more treatment time to identify the accurate tumor, (iv) current deep learning approaches require a large amount of training data to overcome the problems of overfitting and high computational cost, and (v) practical implementation.

To resolve the above mentioned problems, we have proposed a new breast cancer detection and classification approach. This will be discussed in more detail in the next section.

### 3. The Proposed Transferable Texture Convolutional Neural Network (TTCNN) Method

In this paper, a deep learning based Transferable Texture Convolutional Neural Network (TTCNN) method is proposed for breast cancer diagnosis and classification using digital X-ray mammograms. Firstly, a pre-processing approach is executed for contrast enhancement on the source image to improve the contrast. After that, the proposed TTCNN is applied on the pre-processed images for breast cancer malignant and benign classification. The proposed TTCNN classification then perform on the latest deep features from the deep convolutional neural network (DCNN) models, i.e., InceptionResNet-V2, Inception-V3, GoogLeNet, VGG-16, VGG-19, ResNet-18, ResNet-50, and ResNet-101 for the feature extraction and executes transfer learning to redeem the selected databases. The proposed method contains six phases that include the materials and methods, contrast enhancement, proposed TTCNN architecture, transfer learning based deep feature extraction, feature fusion and the feature selection. These steps are elaborated in the following subsections. A schematic diagram of the proposed approach is presented in Figure 2.

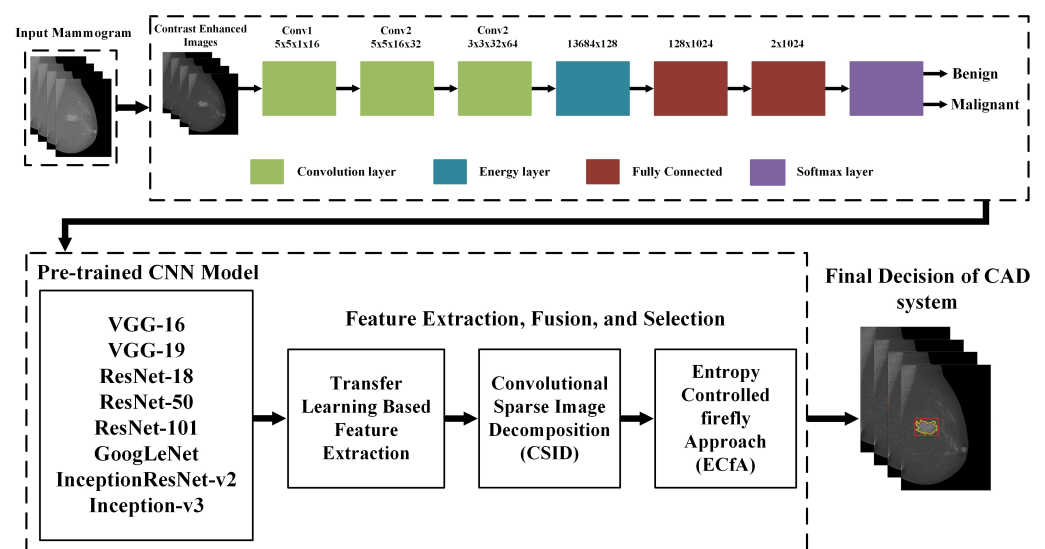


Figure 2. Framework of the proposed approach for the breast cancer detection and classification.

### 3.1. Materials and Methods

#### 3.1.1. Dataset

This work used three digital breast X-ray databases to assess the proposed CAD approach. The datasets are DDSM [56], INbreast [57] and MIAS [58].

##### DDSM

The DDSM database was compiled by a team of experts from the University of South Florida [56]. The DDSM dataset consists of digitized images in joint photographic experts group (jpeg) format compressed scanned mammography film. In this work, we have used CBIS-DDSM (Curated Breast Imaging Subset of DDSM) which is a modernized version of DDSM decompressed in DICOM format. The DDSM database includes 2620 breast cases divided into 43 different volumes. In each case, each breast uses two different slide views (MLO and CC) to collect four mammograms [7]. The average size of a DDSM database is  $3000 \times 4800$  pixels [4]. This contains the cancer type (benign or malignant) and the location of the wound [56]. Mammograms in the DDSM dataset are categorized in three sets i.e., normal, benign, and malignant by experts based on the BI-RAD score [7]. The CBIS-DDSM database consists of pixel-by-pixel annotations for regions of interest (ROI), i.e., tumor, calcification and disease pathology (benign or malignant). Additionally, each ROI is marked as calcification or mass. Most mammograms consist of only one ROI. The motive of this work is to anticipate the benign and malignant state of each image.

##### INbreast

The INbreast database consists of digital mammograms and was obtained from the Portuguese University Hospital (Centro Hospitalar de S. Joao [CHSJ], Breast Centre, Porto) with the consent of the National Data Protection Commission of Portugal and Hospital's Ethics Committee [57]. The average size of an INbreast database is  $3328 \times 4084$  pixels. A total of 115 cases or patients were taken from INbreast dataset using both breast views (MLO and CC). In the 90 affected cases of both breasts (i.e., 360 mammograms), each person had 4 mammograms, but in the other 25 mastectomy patients (i.e., 50 mammograms) in only two mammograms were performed [57]. Therefore, a total of 410 mammograms were generated with MLO and CC images from 115 patients. This includes normal, benign and malignant cases [57]. All mammograms with breast lesions (107 cases in a total) were used for evaluation from both the MLO and the CC perspectives. Multiple mammograms showed various lesions and the BI-RAD score was used to classify a total of 112 breast lesions. Therefore, by using the BI-RAD 36 mammograms were collected to indicate benign cases and 76 mammograms were collected to explain malignant cases using BI-RAD. This dataset consists of pixel-level batch annotations and histological information on cancer types. The dataset also contains some mammograms with multiple qualities.

##### MIAS

The MIAS dataset consists of 326 mammograms images and has three categories of tissue types (fatty, fatty-glandular, and dense-glandular). Among the 326 images, normal are 207, abnormal are 119, 68 were benign, and 51 were malignant. The size of all the images used in this dataset is  $1024 \times 1024$  pixels and are physically formatted in portable gray map (pgm) format.

After the cessation of this subsection, the proposed approach commences the second phase, which is explained in the next subsection.

### 3.2. Contrast Enhancement

Improving the contrast is an important preliminary step in the diagnostic process [59]. Due to the lack of illumination, the contrast of the source mammogram image is low. The histogram equalization method appears to be a more efficacious way of enhancing the low contrast images. The modified CLAHE [59] is implemented to adjust the contrast and maintain the standard brightness of the input image. This contrast enhancement method



affects small parts of the image (mosaics). The histogram in the output part harmonizes the identified histogram because the contrast of each mosaic is emphasized, rather than the absolute mammogram image. After leveling, use linear interpolation to connect adjacent tiles and remove artificial boundaries. CLAHE uses a custom clipping threshold to limit the enhancement of histogram clipping. The clipping level diminishes the noise level and sets the contrast level to enhance the histogram. We use 0 to 0.01 for this activity.

Firstly, we divide the input image into correlated areas that do not overlap. The total number of image tiles is  $R \times S$ . The histogram of the non-overlapping correlation area is calculated as the level of gray present in the image matrix. Equation (1) calculated the contrast limit histogram of the relevant area that does not overlap according to the clipping limit as follows.

$$In_{avg} = \frac{S_p \times S_q}{S_{gray}}, \tag{1}$$

where  $In_{avg}$  is the average number of pixels,  $S_{gray}$  is the number of related gray levels that do not overlap,  $S_p$  and  $S_q$  is the number of pixels in dimensions that do not overlap  $p$  and  $q$ . The clipping limit is given in Equation (2) as:

$$In_{LC} = S_{clipping} \times I_{av}, \tag{2}$$

where  $In_{LC}$  is the clipping limit,  $S_{clipping}$  is the normalized clipping limit in the range of [0, 1]. If the number of pixels is greater than  $In_{LC}$ , the pixels will be clipped. The remaining average pixels are allocated in each gray level as follows:

$$In_{av,gray} = \frac{S_{pcw}}{S_{gray}}, \tag{3}$$

where  $S_{pcw}$  denotes the integer number of pixels after truncation. Move the remaining pixels till all the pixels are connected. The pixel redistribution is given in Equation (4) as:

$$In_{pace} = \frac{S_{gray}}{S_{ip}}, \tag{4}$$

where  $S_{ip}$  is the number of remaining clipped pixels. Also, using the Rayleigh transformation in each part, the intensity values are rectified in Equation (5) as:

$$In_q = In_{mi} + \sqrt{2\zeta^2 \ln\left(\frac{1}{1 - C_{in}}\right)} I_{pace}, \tag{5}$$

where  $C_{in}$  represents the cumulative probability which develops the transfer function,  $In_{mi}$  is the lower limit of the pixel value, and  $\zeta$  is the scale parameter. The value of each intensity, the output probability density is given by Equation (6) as:

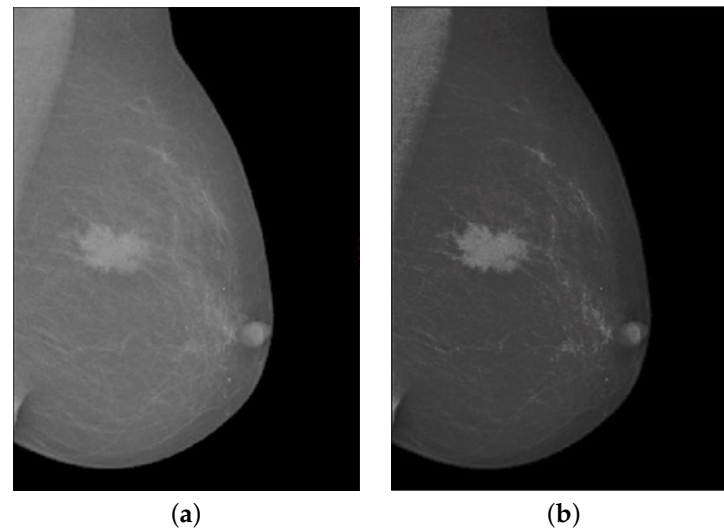
$$In_q = \frac{(In_q - In_{mi})}{\zeta^2} \cdot \exp\left(-\frac{(In_q - In_{mi})^2}{2\zeta^2}\right) \text{ for } In_q \geq In_{mi}. \tag{6}$$

Increasing the value of  $\zeta$  will significantly improve the contrast of the image, but will increase the saturation value and the noise level. Linear contrast stretching can be used to rearrange the output of the resulting transfer function to reduce the effects of sudden changes. The expansion of the linear contrast can be demonstrated in Equation (7) as:

$$In_x = \frac{I_q - w_{mi}}{w_{ma} - w_{mi}}, \tag{7}$$

where  $I_q$  is the transfer function obtained,  $w_{mi}$  and  $w_{ma}$  are the minimal and maximum transfer functions value respectively.  $In_x$  is used to extract the source image to acquire an image with enhanced contrast. Contrast enhancement refined the edges of the input image.

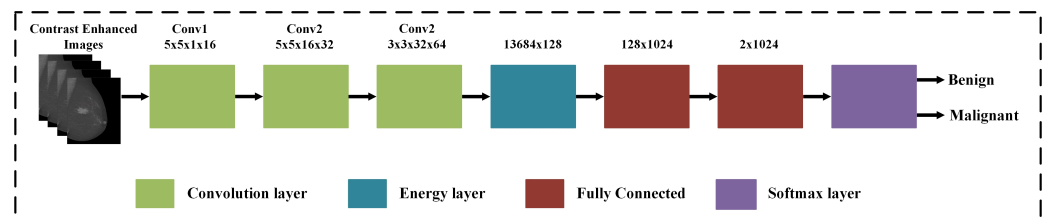
Figure 3 shows the contrast enhancement of the input image. In Figure 3a, the source mammogram image having a tumor is not visible as the tumor in the breast image has no sharp edges. Whereas in Figure 3b image shows that after using the modified contrast enhancement approach, the image gradient has been significantly improved and the disease part is prominently highlighted in the breast mammogram image. Upon closing of this phase, the proposed approach moves to the third phase, which is detailed in the next subsection.



**Figure 3.** The mammogram image for contrast enhancement. (a) source image, (b) contrast enhanced image using modified CLAHE.

### 3.3. Network Architecture

The TTCNN architecture is presented for breast cancer classification. The proposed deep CNN considers the three important features of the image: Firstly, the size of some description patterns is much smaller than the source image, but if their size is equal to the convolution filter mask size, the convolution filter can find the pattern. Secondly, certain shapes or patterns can be used in different parts of the mammogram image. These models can also be defined by convolving the entire source mammogram image. Thirdly, downsampled pixels are very important for the max-pooling layer and do not change the shape of the source mammogram image. The proposed TTCNN framework for the breast cancer classification is illustrated in Figure 4.



**Figure 4.** Framework of the proposed TTCNN for breast cancer classification.

The proposed TTCNN contains two convolution layers, followed by the pooling layers and the third convolution layer directs the EL. Then, a softmax layer is used with the fully connected (FC) layer. EL summarizes the feature maps of the last convolutional layer by averaging the rectified activation output. This returns a value for each feature map, equivalent to the energy response to a filter bank. In addition to reducing the number of layers, this architecture offers excellent performance in learning texture functions and requires less computation time and memory. EL enables this trade-off between performance and computation time. This layer is implemented to conserve the data flow of the original

layer. The flattened output of EL is redirected to the concatenation layer immediately after the last pooling layer. This connection creates a new flattened vector that contains information about the shape and texture of the image and spreads it to the fully connected layer. The complete description of the proposed network with both the input and output dimensions is given in Table 2. The output size of the convolution layer is mathematically computed in Equation (8) as:

$$Output = \frac{I_a - I_b + 2S}{\varrho + 1}, \tag{8}$$

where  $I_a$  and  $I_b$  represents the input and filter size respectively,  $S$  denotes the padding, and  $\varrho$  is the stride value.

**Table 2.** Proposed texture CNN architecture layers.

| Layers | Types                 | Input Size               | Kernel Size to Form Each Feature Map | Stride | Padding   | Output Size              |
|--------|-----------------------|--------------------------|--------------------------------------|--------|-----------|--------------------------|
| 1      | Convolutional Layer 1 | $64 \times 64 \times 1$  | $5 \times 5$                         | [1 1]  | [1 1 1 1] | $62 \times 62 \times 16$ |
| 2      | Max Pooling Layer 1   | $62 \times 62 \times 16$ | $2 \times 2$                         | [2 2]  | [1 1 1 1] | $32 \times 32 \times 16$ |
| 3      | Convolutional Layer 2 | $32 \times 32 \times 16$ | $5 \times 5$                         | [1 1]  | [1 1 1 1] | $30 \times 30 \times 32$ |
| 4      | ReLU                  |                          |                                      |        |           |                          |
| 5      | Max Pooling Layer 2   | $30 \times 30 \times 32$ | $2 \times 2$                         | [2 2]  | [1 1 1 1] | $16 \times 16 \times 32$ |
| 6      | Convolutional Layer 3 | $16 \times 16 \times 32$ | $3 \times 3$                         | [1 1]  | [1 1 1 1] | $16 \times 16 \times 64$ |
| 7      | ReLU                  |                          |                                      |        |           |                          |
| 8      | EL                    | $16 \times 16 \times 64$ | -                                    | -      | -         | $128 \times 1$           |
| 9      | Dropout               | $128 \times 1$           | -                                    | -      | -         | $128 \times 1$           |
| 10     | FC1                   | $128 \times 1$           | -                                    | -      | -         | $1024 \times 1$          |
| 11     | Dropout               | $1024 \times 1$          | -                                    | -      | -         | $1024 \times 1$          |
| 12     | FC2                   | $1024 \times 1$          | -                                    | -      | -         | $2 \times 1$             |
| 13     | Softmax Layer         |                          | -                                    | -      | -         |                          |
| 14     | Classification Layer  |                          | -                                    | -      | -         |                          |

Afterwards, the three convolution layers are utilized with kernel size of  $5 \times 5$  for the first two layers and the output of these channels are 16 and 32, respectively. The third convolution layer is examined as an intermediate layer for extracting texture properties with kernel size of  $3 \times 3$  having the output channels 64. Only 31,744 parameters can be learned from the convolution layer and these parameters are calculated using the following formulas in Equations (9) and (10) as:

$$\zeta_v = \zeta_V \times (I_k \times \varrho + 1), \tag{9}$$

$$\zeta_v = \zeta_V + X_k \times \varrho \times \zeta_V, \tag{10}$$

where  $\zeta_v$  denotes the CNN layer learnable parameters,  $I_k$  represents the kernel size, and  $\zeta_V$  denotes the channel number.

Each convolution layer computes the output of the neuron connected to the input. The calculation is the dot product between its weight and the smallest input field attached to it. The first convolution layer constructs a 16-kernel  $32 \times 32 \times 16$  size output. Equation (11) gives the output of the neurons in the first convolution layer as:

$$S_\vartheta = \sum_{\vartheta} C_\vartheta \times T_\vartheta + P_\vartheta, \tag{11}$$

where  $S_\theta$  represents the output feature maps,  $C_\theta$  represents the input feature maps, and  $T$  denotes the weighted map. Afterwards, an energy descriptor is utilized for the output of the final convolution layer. Energy layers are combined after the third convolution layer, taking into account the requirements of the energy descriptor. Its functionality is similar to a dense, messy texture descriptor. The connection is given in Equation (12) as:

$$EL(\xi, \theta) = \rho \left[ \sum_{i=1}^j T_i^\omega \theta_i + P \right], \tag{12}$$

where  $EL(\xi, \theta)$  represents the EL output layer,  $j$  represents the input connections, and  $T$  represents the EL weighted vector. Compared to the last conventional convolution layer interconnection, the interconnection between the EL and FC layers is much smaller and leads to a reduction in learnable parameters. In addition, EL retains energy information from the preceding layer and learns during forward and backward propagation. Furthermore, to reduce the vector size of the next FC layer, EL also improves the overall learning ability of the network and diminish the complication of the proposed system. Use Equation (13) to calculate the learnable parameters of EL as:

$$\xi_{EL} = \eta^m \times \eta^{m-1}, \tag{13}$$

where  $\xi_{EL}$  is the EL learnable parameters,  $\eta^m$  is the current FC layer neuron, and  $\eta^{m-1}$  is the previous FC layer neuron.

To speed up the training process the batch normalization and activation function is used between the convolution layer and rectified linear unit (ReLU) layer. Batch normalization is used to remove the internal covariate shift [50]. This can be done by normalizing the mean and standard deviation. The bulk normalization calculation uses the following Equations (14) and (15) to calculate mean and variance.

$$\tau_Q = \frac{1}{n} \sum_i^n l_i, \tag{14}$$

$$v_Q = \frac{1}{n} \times \sum_i^n (l_i - \tau_Q)^2, \tag{15}$$

where  $\tau_Q$  and  $v_Q$  represents the mean and variance respectively,  $n$  is the mini batch size of  $l_i$  element of features. In our work the value of  $n$  is 64. The batch normalization is calculated in Equation (16) as:

$$\lambda_i = \frac{\theta_i - \tau}{\sqrt{v^2 + \phi}} \alpha + A, \tag{16}$$

where  $\alpha$  and  $A$  are initial learnable parameter values of each output layer. ReLU is used as an activation function which is computed in Equation (17) and the output of the ReLU layer is calculated in Equation (18) as:

$$\lambda_{i,j,k} = \max(0, \theta_{i,j,k}), \tag{17}$$

$$\lambda_{ReLU} = ReLU(Bnorm(Conv(w, x))), \tag{18}$$

where  $\lambda_{i,j,k}$  denotes the output element features and  $\theta_{i,j,k}$  denotes the feature of the input element. Afterwards, the pooling layer reduces the size of feature maps, weights, and computations which shows overfitting of the control network. The max pooling layer is mathematically computed in Equation (19) as:

$$M_{pool} = \max \left( 0, \sum_Q \theta^{k-1} T_\theta \right), \tag{19}$$

where  $M_{pool}$  represents the output feature maps,  $\vartheta$  indicates the input feature maps,  $Q$  denotes the pooling size, and  $T$  represents the max pooling layer for kernel vector. In this work, two max pooling layers are used with each layer kernel size  $2 \times 2$ .

The dropout layer is used to remove a subset of random parameters repeatedly through the weighted update process to avoid overfitting training data. Drop editing is used to remove a subset of random parameters repeatedly during the weighted update process to avoid overfitting training data. FC layers have the most parameters across the network and are therefore subject to over-compatibility of training data. Hence, the dropout layer is determined after the FC layer. The softmax layer is utilized as a classifier which uses the loss function. The range of probability values for softmax is  $[0, 1]$ . The mathematical expression of loss function is given in Equation (20) as:

$$\kappa_l = \delta_j + \log \sum_i \exp(\delta_i), \quad (20)$$

where  $\kappa_l$  denotes the total loss and  $\delta_i$  having the class  $\delta$  which is  $i$ -th vector element. The purpose of the classifier is to reduce the probability difference between the real label and the estimated label which is calculated using the softmax function in Equation (21) as:

$$\lambda_i = \frac{\exp^{\delta_j}}{\sum_i \exp(\delta_i)}. \quad (21)$$

Upon completion of this stage, TTCNN proceeds to the fourth phase, which is explained in the following subsection.

### 3.4. Deep Feature Extraction Using Transfer Learning

Transfer learning is a deep learning (and machine learning) method that transfers expertise from one model to another. Transfer learning allows us to solve all or part of a specific task by using an already pre-trained model to complete another task. Machines use what they have learned from previous tasks to improve predictions for new tasks in transfer learning. Transfer learning has many benefits, but the most significant is that it reduces training time, enhances the performance of neural networks, and does not make large amounts of data available. Below is a concise description of the state-of-the-art CNN deep feature extraction model which is selected in our work.

#### 3.4.1. VGGNet

VGGNet is a CNN architecture presented by Andrew Zisserman and Karen Simonyan from the Oxford University in 2014 [41]. It is formulated by intensifying the depth of the accessible CNN architecture to 16 or 19 called VGG-16 and VGG-19, respectively as illustrated in Figure 5.

The VGG-16 architecture contains 138 million parameters and the VGG-19 contains 144 million parameters. The VGG-16 model has 13 convolutional layers, five max pooling layers, and three fully connected layers. VGG has a small filter ( $3 \times 3$ ) instead of a larger filter. It has the same effectual receptive field as having only one  $7 \times 7$  convolutional layer. The VGG-19 model has 16 convolutional layers, five max pooling layers, and three fully connected layers. In the two variants of VGGNet, there are two fully connected layers, each with 4096 channels, and then another fully connected layer with 1000 channels to anticipate 1000 labels.

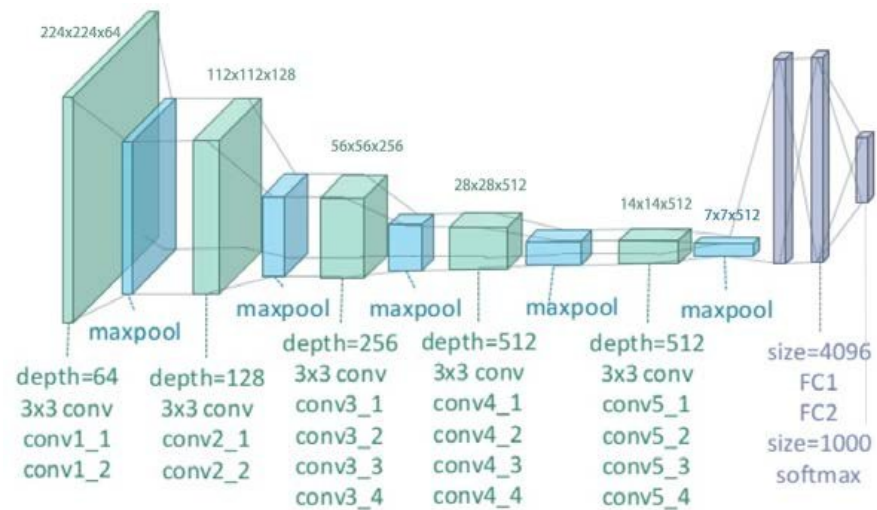


Figure 5. Framework of VGG-16 and VGG-19 architectures.

### 3.4.2. ResNet

ResNet was founded in 2015 by Kaiming He and the idea behind that is each layer of the framework is determined from the residuary function by referring to its own input layer. This network won the ILSVRC 2015 competition with a lower error rate of 3.57%. In this degree, the ResNet model is certainly optimized and the precision can be greatly improved [40]. For this task, we have used three pre-trained ResNet-18, ResNet-50 and ResNet-101 networks. Table 3 illustrated the basic architecture of all these networks.

The input size of the network is  $224 \times 224$ . In the three architectures described, the first convolutional layer and the final three layers, i.e., pooling, FC and the softmax are affixed. If the internal convolution layers number is increased, the depth of the deep network changes.

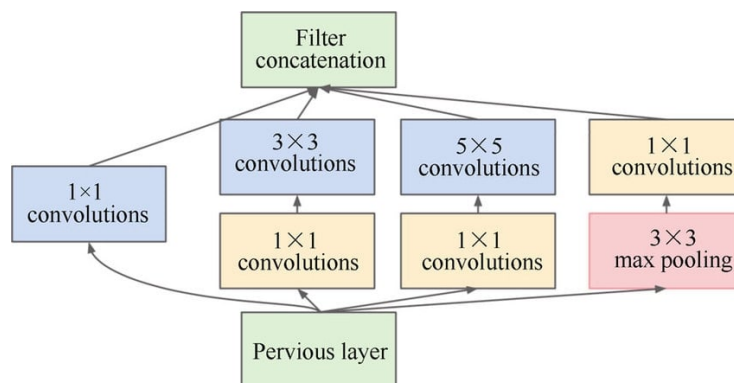
Table 3. The ResNet-18, ResNet-50 and ResNet-101 architectures.

| Layer Name | Size of Output          | ResNet-18   | ResNet-50   | ResNet-101   |
|------------|-------------------------|---|---|--|
| Conv1      | $112 \times 112$        | $7 \times 7, 64$ Stride 2<br>$3 \times 3$ maxpool, Stride 2                 | $7 \times 7, 64$ Stride 2<br>$3 \times 3$ maxpool, Stride 2                                     | $7 \times 7, 64$ Stride 2<br>$3 \times 3$ maxpool, Stride 2                                      |
| Conv2_x    | $56 \times 56$          | $\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 2$   | $\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$    | $\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$     |
| Conv3_x    | $28 \times 28$          | $\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 2$ | $\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$  | $\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$   |
| Conv4_x    | $14 \times 14$          | $\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 2$ | $\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$ | $\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$ |
| Conv5_x    | $7 \times 7$            | $\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 2$ | $\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$ | $\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$  |
| Pooling    | $1 \times 1 \times 512$ |   | Average pooling   |  |
| FC         | $512 \times 10$         |   | Fully connected   |  |
| Softmax    | 1000                    |   | Softmax   |  |

### 3.4.3. GoogLeNet

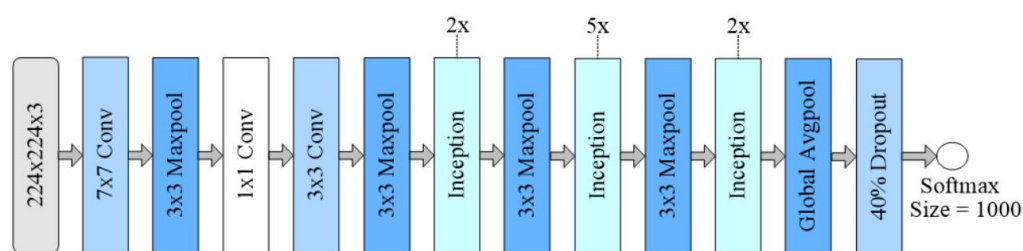
GoogLeNet [43] is a CNN based on the Inception architecture. GoogLeNet (Inception V1) won the ILSVRC 2014 challenge. It reached an error rate of 6.67% which is very

near to the human-level performance that challenge organizers must currently assess. The architecture of the Inception block is displayed in Figure 6. It uses the Inception module, which allows the network to choose between several convolution filter sizes for each block. The Inception network occasionally stacks these modules together and uses the maximum max-pooling layer with stride 2 to bisect the resolution of the grid.



**Figure 6.** Architecture of inception block.

The architecture of GoogLeNet consists of 22 deep CNN layers, but the parameters number is diminished to 4 million from 60 million (AlexNet). It consists of four simultaneous boughs. Convolution layers with kernel sizes of  $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$  are used in the first three boughs. The intricacy of the frame can be diminished by convolving two intermediate branches in the input channel with a window size of  $1 \times 1$ . Appropriate padding is employed for all four boughs so that the height and width of the inputs and outputs are the same. The final inception block is created after linking the output of each bough. It contains almost 6.8 million parameters. GoogLeNet architecture has nine inception blocks containing 6 convolutional layers, 3 convolutional layers of  $1 \times 1$  (for dimensionality reduction),  $3 \times 3$ , and  $7 \times 7$ , 4 layers of max-pooling, two layers of normalization layers, average pooling, and FC layer. The ReLU activation function is utilized by all the convolution layers and drop regularization is applied to the FC layer. The softmax function is used in the output layer. The block of GoogLeNet is shown in Figure 7.

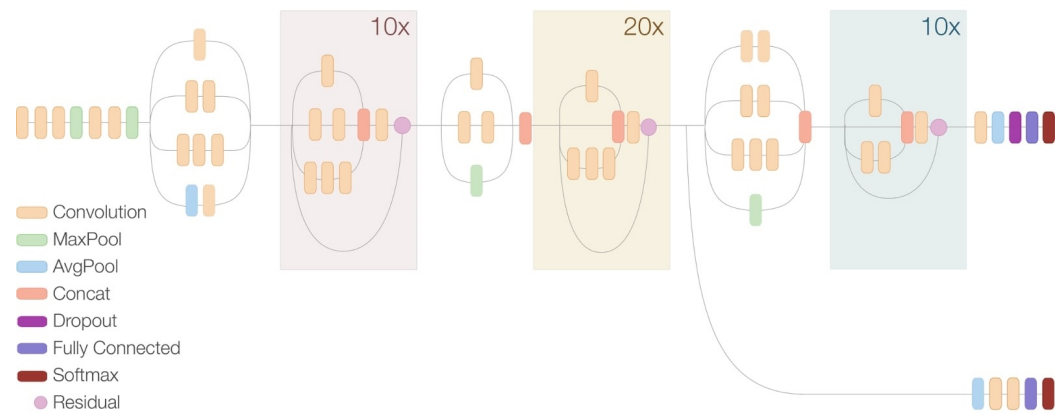


**Figure 7.** Architecture of GoogLeNet.

#### 3.4.4. Inception-ResNet-v2

This network is a form of Inception-v3 [43] which combines a certain design from the ResNet [40]. Inception-ResNet-v2 only uses batch normalization at the traditional layers. To increase the depth of the networks and the number of inception blocks the residual modules are used. The Inception-ResNet-v2 architecture contains one stem block (six convolutional blocks and one max pooling layer) and three different sets of inception blocks. The first block has five inception modules, each with seven convolution blocks. The first block has 10 inception modules, each with five convolution blocks, and the third and last block has five inception modules, each with four blocks of convolution, also contain two depletion blocks with different convolutional layers, average pooling, and FC layers.

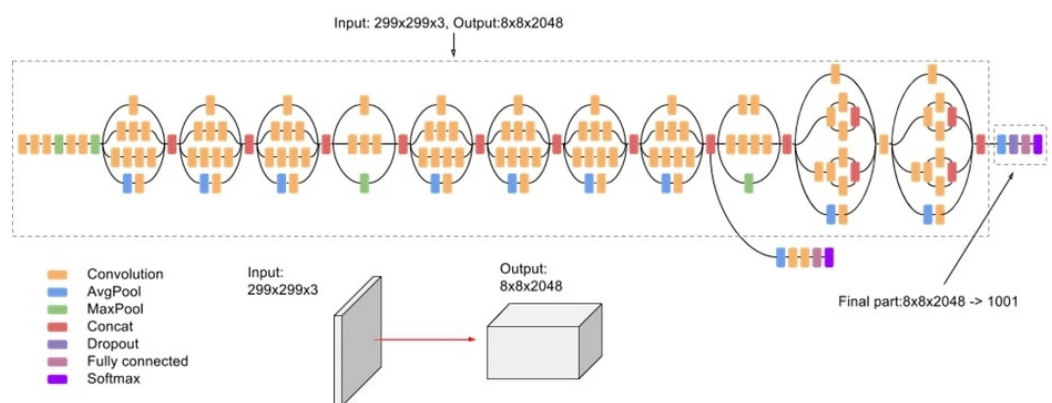
At the output layer softmax activation function is utilized. Figure 8 illustrates the model of Inception-ResNet-v2.



**Figure 8.** Architecture of Inception-ResNet-v2.

### 3.4.5. Inception-v3

This network is also the CNN architecture of the Inception series, including label smoothing,  $7 \times 7$  factoring convolution, and the use of auxiliary classifiers to further propagate label information over the network (batch use), and have been made some improvements, standardization for the sidehead layer. Inception-v3 used 1 million training images from thousands classes of ImageNet datasets for training. Inception v3 showed more than 78.1% accuracy rate on the ImageNet dataset. Over the years, this architecture is the climax of many ideas established by various researchers. The architecture comprises symmetric and asymmetric building blocks, such as convolutional layer, average pooling, max pooling layers, concatenation layer, dropout layer, and FC layer. Batchnorm is widely used throughout the architecture and is suitable for activation input. The loss is calculated using Softmax. This model truncates the number of parameters that can be learned and also truncates the complication of the network. The general framework of the Inception-v3 model is illustrated in Figure 9.



**Figure 9.** Framework of Inception-v3.

Upon completion of this stage, TTCNN moves to the fifth stage, which is elaborated in the following subsection.

### 3.5. Feature Fusion

This is the most highlighted field in the pattern recognition [60]. Feature fusion combines the features from different layers or branches. Feature fusion performed from



different operations i.e., concatenation and summation. In this regard, the convolutional sparse image decomposition (CSID) fusion approach [28] is employed to interconnect the selected feature vectors in the matrix to obtain feature vectors. The feature fusion is evaluated in Equation (22) as:

$$\chi_{o,p} = \min_{\zeta_{o,p}} \frac{1}{2} \left\| O_t - \sum_{p=1}^{Q_t} h_{t,p} * \zeta_{t,p} - \sum_{m=1}^{M_o} l_{o,p} * \zeta_{o,p} \right\|_2^2 + \diamond_o \sum_{p=1}^{Q_t} \|\zeta_{t,o}\|_1. \tag{22}$$

This operation continues till all pairs are compared.  $\chi_{o,p}$  is the final fused vector. Upon termination of this stage, TTCNN approach move to the sixth and last phase, which is discussed in the next subsection.

### 3.6. Feature Selection

In the last few years, feature selection methods have successfully enhanced the system efficiency and exhibit a prominent improvement in medical imaging [61]. Feature selection is employed to enhance the classification accuracy, annihilate the redundancy among the features, and use only robust features for perfect classification. This allows us to truncate prediction numbers and complete the testing process rapidly. In this regard, an entropy controlled firefly approach (ECfA) for optimal feature selection is used. First, the firefly approach selects the function, then proposes an entropy based activation function, and passes the function to the final selection stage. Yang et al. [62] proposed a firefly approach which is a modern and extensive metaheuristic optimization method caused by the glowing behavior of fireflies. Compared to the particle swarm optimization (PSO) and genetic approach (GA) [63], the firefly approach uses the flicker behavior of fireflies to optimize multimodal problems for robust performance.

The glow of a firefly with original brightness  $\gamma$  is expressed in Equation (23) as:

$$\gamma(u) = \gamma_{\diamond} e^{-\rho u}, \tag{23}$$

where  $\gamma_{\diamond}$  represents the original brightness,  $u$  represents the distance between two fireflies, and  $\rho$  represents the optical coefficient that causes the brightness and the number of individuals. As we all know, brightness is directly proportional to attractiveness. Therefore, the attractive force  $\beta$  is expressed in Equation (24) as:

$$\beta(u) = \beta_{\diamond} e^{-\rho u}, \tag{24}$$

when  $u = 0$ , the attraction is  $\beta_{\diamond}$ . The firefly attraction  $s$  and  $t$  is shown in Equation (25) as:

$$H_s^{i+1} = \beta_{\diamond} e^{-\rho u_{st}^2} (H_t^i - H_s^i) + \eta (Rand - \frac{1}{2}), \tag{25}$$

where  $\eta$  denotes the randomness of the parameters,  $i$  is the number of iterations, and  $Rand$  initiates random numbers from 0 to 1. The distance between  $sth$  and  $tth$  fireflies is represented by  $\zeta_{st}$  and is calculated in Equation (26) as:

$$\zeta_{st} = \|H_t - H_s\| = \sqrt{\sum_{j=1}^J (H_{sj} - H_{tj})^2}. \tag{26}$$

The minimum distance characteristic is evaluated. The next iteration is executed based on the error rate. We have chosen a total number of iterations  $n = 100$  in our work. After each iteration, the feature vectors  $\phi_N^{j1}$  and  $\phi_N^{j2}$  are obtained for the optimal vectors of  $N \times 1746$  and  $N \times 1822$  dimensions, respectively. The activation function based on entropy

is employed for all the features. At this phase, we use an entropy based activation function to further enhance the function. The activation function is determined as:

$$B(\varphi) = - \sum_{i=1}^j H(\varphi^i) \log_2 H(\varphi^i), \tag{27}$$

$$\widehat{\varphi}_N^{j1} = \{\forall \varphi_N^{j1} \geq B(\varphi)\}, \tag{28}$$

$$\widehat{\varphi}_N^{j2} = \{\forall \varphi_N^{j2} \geq B(\varphi)\}, \tag{29}$$

where  $j \in (j1, j2)$ ,  $\widehat{\varphi}_N^{j1}$  and  $\widehat{\varphi}_N^{j2}$  is an optimal feature for  $\varphi_N^{j1}$  and  $\varphi_N^{j2}$  respectively. In our work, the optimal length of the feature vector after applying the activation function is  $N \times 1346$  and  $N \times 1322$  respectively.

The ECfA based feature selection approach is also detailed in Algorithm 1.

---

**Algorithm 1** Feature Optimization based on Entropy Controlled Firefly

---

**Input:** Results from CSID fusion approach.

**Output:** Optimal features selection.

**Step 1.** Fitness function  $p(x)$ ,  $x = (x_1, x_2, x_3, \dots, x_m)$

**Step 2.** Generate initial firefly populations  $H_s$ , where  $s = 1, 2, 3, \dots, n$

**Step 3.** Calculate brightness  $\gamma \circ e^{-\rho u}$

**Step 4.** Absorption coefficient  $\rho$

**while** ( $r \leq$  Maximum Generation) **do**

**for**  $s = 1:t$  **do**

**for**  $t = 1:s$  **do**

**if** ( $\gamma_s > \gamma_t$ ) **then**

        ◦ Attraction depends on distance  $\zeta$ ,  $e^{\rho u}$

        ◦ Move the firefly  $s$  in the direction of  $t$  using  $H_s^{i+1} =$

$\beta \circ e^{-\rho u_{st}^2} (H_t^i - H_s^i) + \eta (Rand - \frac{1}{2})$

        ◦ Estimate new solutions and upgrade brightness

**end if**

**end for**  $s$

**end for**  $t$

**Step 5.** Find the best fireflies

**Step 6.** The activation function based on entropy is employed  $B(\varphi)$

**Step 7.** Best optimal feature vector are selected  $\widehat{\varphi}_N^j$

**end while**

**Step 8.** Processing and visualization of results

**end**

---

#### 4. Performance Evaluation

This section examines the experiment and validation of the proposed TTCNN approach for mammogram imaging. Descriptions of benchmark datasets, evaluation metrics and comparison with other state-of-the-art approaches have also been discussed.

##### 4.1. Simulation Setup

The experiments are implemented in MATLAB R2021a (MathWorks Inc., Natick, MA, USA) on a laptop for simulation results. The hardware platform comprises the Intel(R) Core(TM) i7 9750H CPU 2.59 GHz processor (Intel Corporation, San Francisco, CA, USA) with 12 GB RAM and GPU of NVIDIA GeForce GTX 1650 (NVIDIA Corporation, Santa Clara, CA, USA) in Microsoft Windows 11 (Microsoft, WA, USA).

### 4.2. Image Acquisition

In this study, the digitized mammogram images from DDSM [56], INbreast [57] and MIAS [58] datasets are used to evaluate the performance of the proposed method. In fact, all the databases are utilized to show the reliability and effectiveness of the proposed method in diagnosing breast cancer. The DDSM database consists of 1500 mammogram images having 519 normal images and 981 abnormal images. The INbreast database contains 336 mammogram images having 69 normal images and 269 abnormal images, whereas the MIAS database consists of 326 total mammogram images having 207 normal images and 119 abnormal images. In this work, a total of 1369 breast mammogram images were utilized for abnormal, in which the DDSM dataset have 479 benign and 502 malignant out of 981 mammogram images, the INbreast dataset contain 269 mammogram images having 220 benign and 49 malignant, and the MIAS dataset contain 119 mammogram images having 68 benign and 51 malignant [56]. To evaluate the supremacy and effectiveness of the proposed method. Experts collect ground truth labels and diagnose these cases (via labels) using a variety of tests including experimental screening and mammography. Tables 4 and 5 illustrated the mammography images distribution from the used databases. During the assessment, the images in each database are divided into two sets, a training set and a testing set.

**Table 4.** Database specifications (normal and abnormal).

| Database | Total Number of Images |          | Training |          | Testing |          |
|----------|------------------------|----------|----------|----------|---------|----------|
|          | Normal                 | Abnormal | Normal   | Abnormal | Normal  | Abnormal |
| DDSM     | 519                    | 981      | 311      | 589      | 208     | 392      |
| INbreast | 67                     | 269      | 40       | 162      | 27      | 107      |
| MIAS     | 207                    | 119      | 125      | 71       | 82      | 48       |

**Table 5.** Database specifications (benign and malignant).

| Database | Total Number of Images |           | Training |           | Testing |           |
|----------|------------------------|-----------|----------|-----------|---------|-----------|
|          | Benign                 | Malignant | Benign   | Malignant | Benign  | Malignant |
| DDSM     | 479                    | 502       | 287      | 302       | 192     | 200       |
| INbreast | 220                    | 49        | 132      | 29        | 88      | 20        |
| MIAS     | 68                     | 51        | 40       | 31        | 28      | 20        |

### 4.3. Performance Evaluation Criteria

The cross-validation approach is developed to enhance the efficiency, validity of performance and to verify the outcome of each database. To analyze the classification efficiency of our proposed approach several metrics are used i.e., accuracy (*Acc*), sensitivity (*Sen*), specificity (*Spe*), error rate (*Er*), matthews correlation coefficient (*Mcc*), Jaccard (*Jac*), positive predicted value (*PPV*), F1 score (*F1*), and area under receiver operating characteristic (ROC) curve also called the area under curve (AUC). These parameters are utilized as measurable elements to compare the proposed method performance with state-of-the-art algorithms. These measured values are defined as follows:

$$Accuracy(Acc) = \frac{(TP + TN)}{(TP + FP + TN + FN)} \times 100\%, \tag{30}$$

$$Sensitivity(Sen) = \frac{TP}{(TP + FN)} \times 100\%, \tag{31}$$

$$Specificity(Spe) = \frac{TN}{(TN + FP)} \times 100\%, \tag{32}$$

$$ErrorRate(Er) = \frac{FP + FN}{(TN + FP + FN + TN)} \times 100\%, \tag{33}$$

$$Mcc = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \times 100\%, \quad (34)$$

$$Jaccard(Jac) = \frac{TP}{(TP + FP + FN)} \times 100\%, \quad (35)$$

$$PPV = \frac{TP}{(TP + FP)} \times 100\%, \quad (36)$$

$$F1score(F1) = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (37)$$

where  $TP$  represents true positive values which correctly recognized the disease cases,  $TN$  stand for true negative values which correctly identified the healthy cases,  $FP$  stand for false positive values which incorrectly identified the disease cases, and  $FN$  stand for false negative values which incorrectly identified the healthy cases.

#### 4.4. Breast Cancer Classification Using Extracted Deep Features with TTCNN

The proposed TTCNN classification approach extracts deep CNN features. The results are acquired by the feature extraction from the optimal layers of every deep CNN model. We selected the best features for all deep CNN frameworks by determining the best feature extraction layer that can deliver the finest performance on TTCNN. To do this, we extract features from different layers, assess the performance of each deep CNN and identify the best layer. Table 6 illustrates the best performance results for deep feature extraction from the optimal layers.

**Table 6.** Best performance from optimal layers.

| Model              | Layer | Optimal Layer        |
|--------------------|-------|----------------------|
| InceptionResNet-V2 | 813   | 'activation_203'     |
| Inception-V3       | 313   | 'predictions'        |
| VGG-16             | 33    | 'fc6'                |
| VGG-19             | 39    | 'fc6'                |
| GoogLeNet          | 140   | 'pool5-7x7_s1'       |
| ResNet-18          | 64    | 'res5b_branch2b'     |
| ResNet-50          | 169   | 'activation_48_relu' |
| ResNet-101         | 342   | 'res5c'              |

The TTCNN classifier from optimal layers features are also analyzed using the evaluation metrics. Table 7 illustrates the classification result with features from the optimal layers.

**Table 7.** Best performance from optimal layers. Best values are emphasized in bold.

| Model              | Sensitivity (%) | PPV (%)      | Accuracy (%) | Error rate (%) |
|--------------------|-----------------|--------------|--------------|----------------|
| InceptionResNet-V2 | 82.69           | 85.53        | 81.74        | 18.26          |
| Inception-V3       | 85.43           | 87.69        | 84.03        | 15.97          |
| VGG-16             | 76.98           | 77.09        | 81.41        | 18.59          |
| VGG-19             | 77.83           | 81.51        | 84.38        | 15.62          |
| GoogLeNet          | 82.37           | <b>90.04</b> | <b>88.54</b> | <b>11.46</b>   |
| ResNet-18          | 87.25           | 83.02        | 87.68        | 12.32          |
| ResNet-50          | 83.38           | 87.31        | 82.87        | 18.58          |
| ResNet-101         | <b>89.51</b>    | 87.64        | 88.12        | 11.88          |

From the above Table 7, it can be perceived that the GoogLeNet provides the best performance than the other deep CNN models with an accuracy rate of 88.54% and an error rate of 11.46%. Although ResNet-101 is also very close to GoogLeNet with an accuracy rate of 88.12% and an error rate of 11.88%. ResNet-18 also provides very decent performance with an accuracy rate of 87.68% and an error rate of 12.32%. Among all the models, VGG-16 comes last with an accuracy rate of 81.41% and an error rate of 18.59%.

#### 4.5. Classification Results

The proposed TTCNN method classifies mammogram images of breast tumors as benign or malignant. The experiments are performed on DDSM, INbreast, and MIAS databases. The classification accuracy of the proposed TTCNN approach using three different databases is illustrated in Table 8. The total training samples of 287 benign and 302 malignant are selected and the total testing samples of 192 benign and 200 malignant are selected for the DDSM dataset and obtained an accuracy of 99.08%. Similarly, the total training samples selected for INbreast are 132 benign and 29 malignant while testing samples are 88 benign and 20 malignant and achieved an accuracy of 96.82%. The total training samples of 40 benign and 31 malignant are selected and the total testing samples of 28 benign and 20 malignant are selected for the MIAS dataset and achieved an accuracy of 96.57%.

**Table 8.** Summary of breast cancer detection.

| Database | Total Number of Images |           | Correctly Detected |           | Accuracy (%) |
|----------|------------------------|-----------|--------------------|-----------|--------------|
|          | Benign                 | Malignant | Benign             | Malignant |              |
| DDSM     | 479                    | 502       | 474                | 498       | 99.08        |
| INbreast | 220                    | 49        | 215                | 47        | 96.82        |
| MIAS     | 68                     | 51        | 66                 | 49        | 96.57        |

The quantitative comparison of the proposed TTCNN method is also compared with other existing state-of-the-art algorithms for each database. The proposed approach appears to outperform other state-of-the-art approaches with high values of accuracy, specificity, sensitivity, and F1 score as illustrated in Tables 9–11. The best value is highlighted in bold text.

**Table 9.** Performance comparison with respect to DDSM database. Best values are emphasized in bold. N.A.—the data is not provided.

| Authors                    | Method   | Accuracy (%) | Specificity (%) | Sensitivity (%) |
|----------------------------|----------|--------------|-----------------|-----------------|
| Sadad et al. [22]          | Ensemble | 93.1         | 88              | 95              |
| Muduli et al. [45]         | MFO-ELM  | 98.8         | 97.22           | 96.2            |
| Mohanty et al. [64]        | ANN      | 98.18        | N.A.            | N.A.            |
| Xie et al. [65]            | ELM      | 95.73        | 97.16           | 94.88           |
| Mohanty et al. [66]        | WC-SSA   | 98.63        | 97.37           | 98.77           |
| Zhang et al. [67]          | AdaBoost | 90.91        | 97.38           | 82.96           |
| <b>Our Proposed Method</b> | TTCNN    | <b>99.08</b> | <b>98.96</b>    | <b>99.19</b>    |

**Table 10.** Performance comparison with respect to INbreast database. Best values are emphasized in bold. N.A.—data not provided.

| Authors                    | Method   | Accuracy (%) | Specificity (%) | Sensitivity (%) |
|----------------------------|----------|--------------|-----------------|-----------------|
| Al-antari et al. [4]       | CNN      | 95.32        | N.A.            | N.A.            |
| Dhungel et al. [7]         | AlexNet  | 91.0         | N.A.            | N.A.            |
| Al-antari et al. [8]       | AlexNet  | 95.64        | N.A.            | N.A.            |
| Carneiro et al. [34]       | AlexNet  | 86           | N.A.            | N.A.            |
| Zhang et al. [67]          | AdaBoost | 87.93        | 97.73           | 57.20           |
| <b>Our Proposed Method</b> | TTCNN    | <b>96.82</b> | <b>97.68</b>    | <b>95.99</b>    |

**Table 11.** Performance comparison with respect to MIAS database. Best values are emphasized in bold. N.A.—data not provided.

| Authors                    | Method | Accuracy (%) | Specificity (%) | Sensitivity (%) |
|----------------------------|--------|--------------|-----------------|-----------------|
| Mohanty et al. [64]        | SVM    | 95.5         | 96.65           | 95.81           |
| Xie et al. [65]            | ELM    | 96.02        | 94.32           | 95.29           |
| Pezeshki et al. [68]       | ANN    | 61           | 77              | 95.03           |
| Tatikonda et al. [69]      | MLP    | 93.37        | 92.43           | 94.18           |
| Pashoutan et al. [70]      | ANN    | 94           | 96.67           | 90              |
| <b>Our Proposed Method</b> | TTCNN  | <b>96.57</b> | <b>97.03</b>    | <b>96.11</b>    |

Table 9 illustrates the quantitative comparison of our proposed approach with existing state-of-the-art algorithms for the DDSM database. The proposed method results yielded superior performance for the DDSM dataset and achieved an accuracy of 99.08%, specificity of 98.96%, and sensitivity of 99.19%. Mohanty et al. [66] also exhibit better performance than the remaining methods as it has the accuracy of 98.63%. However, our approach has better accuracy, specificity, and sensitivity when compared with other algorithms as indicated by the bold text. Table 10 illustrates the comparison for the INbreast database where the proposed approach showed enriched performance and attained an accuracy of 96.82%, specificity of 97.68%, and sensitivity of 95.99%. and outperformed other methods quantitatively. The accuracy result obtained by Al-antari et al. [8] is slightly better than the other approaches which is 95.64%. But still the proposed approach exhibits best performance for the INbreast database when compared with other approaches. Table 11 displays the comparison for the MIAS database where the proposed approach still reveals superior performance and surpasses all the existing state-of-the-art approaches and achieved an accuracy of 96.57%, specificity of 97.03%, and sensitivity of 96.11%. The proposed approach improves breast cancer detection and classification performance in comparison. The proposed approach can be utilized for real-time assessment and assist radiologists for automated analysis of mammogram images.

There may be performance differences when performing the same method on different datasets for specific reasons, i.e., background noise of a source image, illumination, occlusion, model overfitting, unrepresentative data samples, or the probability of the method. Inadequate evaluation of the model can lead to poor performing methods being pushed into production or suppressed under the assumption that the model is overfitted.

The performance of the proposed approach is also determined utilizing the ROC curves and Confusion Matrix. The confusion matrix of DDSM, INbreast, and MIAS datasets is illustrated in Figure 10. AUC is an important quantitative metric in the ROC curve. The ROC curves were drawn for false positive (1-specificity) and true positive (sensitivity) rates controlling the thresholds of the obtained probability maps. The AUC computations

values are estimated for the DDSM, INbreast, and MIAS datasets. The ROC curve graph is illustrated in Figure 11.

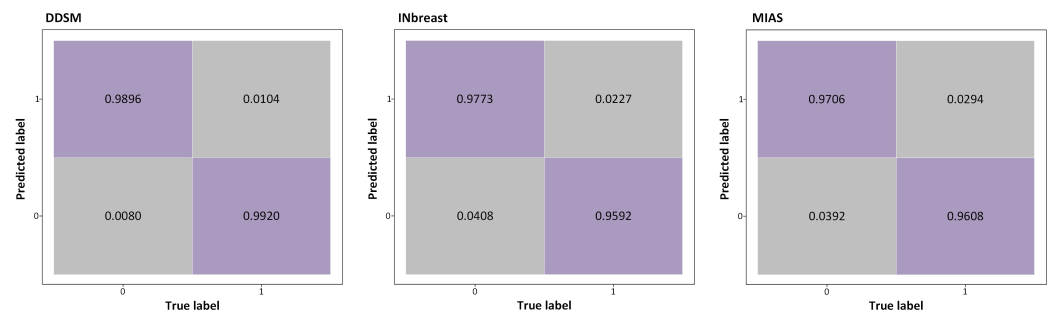


Figure 10. Confusion matrices for classification on DDSM, INbreast, and MIAS databases.

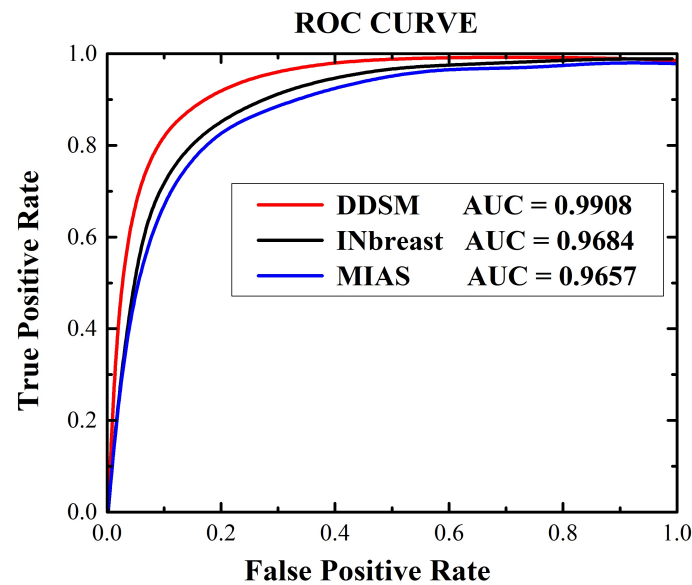


Figure 11. Receiver operating characteristic (ROC) plot for DDSM, INbreast, and MIAS datasets.

The results of breast cancer detection grading (with 95% confidence intervals) are presented in Table 12 that exhibits the data from aforesaid datasets (DDSM, INbreast, and MIAS). The proposed approach gives the PPV, F1, Jac, Mcc, Er, and AUC of 98.96%, 99.06%, 98.19%, 98.16%, 0.92%, and 99.08% on DDSM, 98.96%, 99.06%, 98.19%, 98.16%, 3.17%, and 99.08% on INbreast, and 97.06%, 96.58%, 93.39%, 93.14%, 3.43%, and 96.57% on MIAS datasets, respectively.

Table 12. Results of breast cancer detection and classification with 95% confidence interval (CI).

| Database | Accuracy (%)           | Sensitivity (%)        | Specificity (%)        | PPV (%)                | F1 score (%)           | Jaccard (%)            | Mcc (%)                | AUC (%)                |
|----------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|------------------------|
| DDSM     | 99.08<br>(99.05–99.11) | 99.19<br>(99.16–99.21) | 98.96<br>(98.94–98.99) | 98.96<br>(98.94–98.98) | 99.06<br>(99.03–99.09) | 98.19<br>(98.17–98.21) | 98.16<br>(98.12–98.19) | 99.08<br>(99.07–99.09) |
| INbreast | 96.82<br>(96.79–96.86) | 95.99<br>(95.96–96.12) | 97.68<br>(97.66–97.70) | 97.73<br>(97.70–97.76) | 96.85<br>(96.83–96.87) | 93.89<br>(93.86–93.92) | 93.66<br>(93.64–93.68) | 96.84<br>(96.80–96.88) |
| MIAS     | 96.57<br>(96.55–96.59) | 96.11<br>(96.10–96.13) | 97.03<br>(97.01–97.05) | 97.06<br>(97.03–97.09) | 96.58<br>(96.53–96.63) | 93.39<br>(93.35–93.43) | 93.14<br>(93.10–93.17) | 96.57<br>(96.54–96.61) |

Over the past decade, non-invasive applications such as breast cancer detection and classification have become very prevalent among medical professionals, making the di-

agnostic process simpler and more accurate [71,72]. TTCNN aims to improve clinical diagnosis by improving breast cancer detection and classification. Based on the accuracy produced by a certain algorithm, we obtained the opinion of two medical experts (one radiologists and one physician). These experts praised the improved results of TTCNN compared to other state-of-the-art approaches.

#### 4.6. Computational Efficiency of Deep Learning Classification

Table 13 shows the time execution (in seconds) of the training and testing for each dataset image. In the training set, the number of mammogram images directly alter the time it takes to finish the learning process. Using the TTCNN approach the DDSM dataset takes longer time to train than the INbreast and MIAS datasets. It takes approximately 3.2, 1.4 and 1.1 h to complete the 120 epoch training process on the DDSM, INbreast and MIAS datasets, respectively. For the testing process, the proposed approach took only 1.2 (s) for DDSM, 1.7 (s) for INbreast, and 0.4 (s) for MIAS to classify a single mammogram image. The runtime minimization will be further enhanced in future work, as our main goal is to enhance the detection and classification accuracy.

In general, the proposed approach improves the performance in comparison. The proposed approach can be employed for real-time assessment and assist the radiologists in automated evaluation of mammogram images.

**Table 13.** Computational time for breast cancer classification.

| Database | Time of Training/Epoch (s) | Time of Testing/Image (s) |
|----------|----------------------------|---------------------------|
| DDSM     | 102                        | 1.2                       |
| INbreast | 31                         | 1.7                       |
| MIAS     | 27                         | 0.4                       |

## 5. Conclusions

In the last few years, diagnostic computer systems based on image processing have been extensively used. This can help radiologists, and minimize the time of diagnosis. Several breast cancer detection and classification algorithms have been presented to enhance medical image analysis. These algorithms have various deficiency, i.e., false diagnosis and prediction; the tumor appears in a area with significantly lower contrast, high complexity of memory, computationally complex approach and require more treatment time to identify the accurate tumor, and current deep learning algorithm need extensive training data to overcome the problems of overfitting and high computational cost.

This article intended to resolve the aforementioned issues by the proposed TTCNN-based classification approach for breast cancer detection and classification. Firstly, the source mammogram images are pre-processed by employing the modification to the legacy CLAHE approach. Then, by using TTCNN architecture, the breast cancer regions that are malignant and benign are classified from the mammogram images and the EL examines the texture features and extracts the general information of shape, limit the size of the output vector, and refines the model's overall learning ability. Deep features are extracted from eight state-of-the-art DCNN models and optimal layers for feature extraction are determined by monitoring fluctuations in classification performance using a succession of experiments to select the best deep features. Afterwards, the features are fused using the CSID fusion algorithm, and for the robust feature selection an ECfA approach is used.

The proposed approach was applied to DDSM, INbreast, and MIAS datasets, and obtained an accuracy of 99.08%, 96.82%, and 96.57%, respectively. In addition, the proposed approach is visually gratifying, provides better results, and is more capable in the detection and classification of breast mammograms and outperforms other systems. Breast images are accurately detected and classified in less computation time and give pleasant results.



In the future, this work will be explored for other application areas of biomedical imaging such as brain tumor detection.

**Author Contributions:** Conceptualization: S.M.; methodology: S.M. and R.D.; software: S.M.; validation: S.M.; formal analysis: S.M. and R.D.; investigation: S.M., R.D. and R.M.; data curation: S.M. and R.D.; writing—original draft preparation: S.M.; writing—review & editing, R.D. and R.M.; supervision: R.D.; funding acquisition: R.D. 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 and program codes are available from the corresponding author upon reasonable request.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

- Xian, M.; Zhang, Y.; Cheng, H.D.; Xu, F.; Zhang, B. Automatic breast ultrasound image segmentation: A survey. *Pattern Recognit.* **2018**, *79*, 340–355. [[CrossRef](#)]
- Xu, C.; Gu, Z.; Liu, J.; Lin, X.; Wang, C.; Li, J.; Zhuang, Z. Adenosquamous carcinoma of the breast: A population-based study. *Breast Cancer* **2021**, *28*, 848–858. [[CrossRef](#)]
- Gardezi, S.J.S.; Elazab, A.; Lei, B.; Wang, T. Breast cancer detection and diagnosis using mammographic data: Systematic review. *J. Med. Internet Res.* **2019**, *21*, e14464. [[CrossRef](#)] [[PubMed](#)]
- Al-Antari, M.A.; Han, S.M.; Kim, T.S. Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms. *Comput. Methods Programs Biomed.* **2020**, *196*, 105584. [[CrossRef](#)] [[PubMed](#)]
- Punitha, S.; Al-Turjman, F.; Stephan, T. An automated breast cancer diagnosis using feature selection and parameter optimization in ANN. *Comput. Electr. Eng.* **2021**, *90*, 106958.
- Mao, Y.J.; Lim, H.J.; Ni, M.; Yan, W.H.; Wong, D.W.C.; Cheung, J.C.W. Breast Tumour Classification Using Ultrasound Elastography with Machine Learning: A Systematic Scoping Review. *Cancers* **2022**, *14*, 367. [[CrossRef](#)]
- Dhungel, N.; Carneiro, G.; Bradley, A.P. A deep learning approach for the analysis of masses in mammograms with minimal user intervention. *Med. Image Anal.* **2017**, *37*, 114–128. [[CrossRef](#)] [[PubMed](#)]
- Al-Antari, M.A.; Al-Masni, M.A.; Choi, M.T.; Han, S.M.; Kim, T.S. A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification. *Int. J. Med. Inform.* **2018**, *117*, 44–54. [[CrossRef](#)]
- Byra, M. Breast mass classification with transfer learning based on scaling of deep representations. *Biomed. Signal Process. Control* **2021**, *69*, 102828. [[CrossRef](#)]
- Assari, Z.; Mahloojifar, A.; Ahmadinejad, N. Discrimination of benign and malignant solid breast masses using deep residual learning-based bimodal computer-aided diagnosis system. *Biomed. Signal Process. Control* **2022**, *73*, 103453. [[CrossRef](#)]
- 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](#)] [[PubMed](#)]
- Jansson, D.; Dieriks, V.B.; Rustenhoven, J.; Smyth, L.C.; Scotter, E.; Aalderink, M.; Dragunow, M. Cardiac glycosides target barrier inflammation of the vasculature, meninges and choroid plexus. *Commun. Biol.* **2021**, *4*, 260. [[CrossRef](#)] [[PubMed](#)]
- Song, B.I. A machine learning-based radiomics model for the prediction of axillary lymph-node metastasis in breast cancer. *Breast Cancer* **2021**, *28*, 664–671. [[CrossRef](#)]
- Mukand, N.H.; Ko, N.Y.; Nabulsi, N.A.; Hubbard, C.C.; Chiu, B.C.H.; Hoskins, K.F.; Calip, G.S. The association between physical health-related quality of life, physical functioning, and risk of contralateral breast cancer among older women. *Breast Cancer* **2021**, *29*, 287–295. [[CrossRef](#)] [[PubMed](#)]
- Abdelrahman, L.; Ghamdi, M.A.; Collado-Mesa, F.; Abdel-Mottaleb, M. Convolutional neural networks for breast cancer detection in mammography: A survey. *Comput. Biol. Med.* **2021**, *131*, 104248. [[CrossRef](#)] [[PubMed](#)]
- Blakely, T.; Shaw, C.; Atkinson, J.; Cunningham, R.; Sarfati, D. Social inequalities or inequities in cancer incidence? Repeated census-cancer cohort studies, New Zealand 1981–1986 to 2001–2004. *Cancer Causes Control* **2011**, *22*, 1307–1318. [[CrossRef](#)] [[PubMed](#)]
- Lotter, W.; Diab, A.R.; Haslam, B.; Kim, J.G.; Grisot, G.; Wu, E.; Sorensen, A.G. Robust breast cancer detection in mammography and digital breast tomosynthesis using an annotation-efficient deep learning approach. *Nat. Med.* **2021**, *27*, 244–249. [[CrossRef](#)]
- Dembrower, K.; Wahlin, E.; Liu, Y.; Salim, M.; Smith, K.; Lindholm, P.; Strand, F. Effect of artificial intelligence-based triaging of breast cancer screening mammograms on cancer detection and radiologist workload: A retrospective simulation study. *Lancet Digit. Health* **2020**, *2*, 468–474. [[CrossRef](#)]

19. Chouhan, N.; Khan, A.; Shah, J.Z.; Hussain, M.; Khan, M.W. Deep convolutional neural network and emotional learning based breast cancer detection using digital mammography. *Comput. Biol. Med.* **2021**, *132*, 104318. [CrossRef]
20. Cancer Stat Facts: Female Breast Cancer. Available online: <https://seer.cancer.gov/statfacts/html/breast.html> (accessed on 10 January 2022).
21. Pacifici, S.; Murphy, A. Mediolateral Oblique View. Available online: <https://radiopaedia.org> (accessed on 29 December 2021).
22. Sadad, T.; Munir, A.; Saba, T.; Hussain, A. Fuzzy c-means and region growing based classification of tumor from mammograms using hybrid texture feature. *J. Comput. Sci.* **2018**, *29*, 34–45. [CrossRef]
23. Zeebaree, D.Q.; Abdulazeez, A.; Zebari, D.A.; Haron, H.; Hamed, H.N.A. Multi-level fusion in ultrasound for cancer detection based on uniform LBP features. *Comput. Mater. Contin.* **2021**, *66*, 3363–3382. [CrossRef]
24. Maqsood, S.; Damaševičius, R.; Shah, F.M.; Maskeliūnas, R. Detection of Macula and Recognition of Aged-Related Macular Degeneration in Retinal Fundus Images. *Comput. Inform.* **2021**, *40*, 957–987. [CrossRef]
25. Chougrad, H.; Zouaki, H.; Alheyane, O. Deep convolutional neural networks for breast cancer screening. *Comput. Methods Programs Biomed.* **2018**, *157*, 19–30. [CrossRef]
26. Kooi, T.; Litjens, G.; Ginneken, B.V.; Gubern-Mérida, A.; Sánchez, C.I.; Mann, R.; Karssemeijer, N. Large scale deep learning for computer aided detection of mammographic lesions. *Med. Image Anal.* **2017**, *35*, 303–312. [CrossRef]
27. Maqsood, S.; Javed, U. Multi-modal medical image fusion based on two-scale image decomposition and sparse representation. *Biomed. Signal Process. Control* **2020**, *57*, 101810. [CrossRef]
28. Muzammil, S.R.; Maqsood, S.; Haider, S.; Damaševičius, R. CSID: A Novel Multimodal Image Fusion Algorithm for Enhanced Clinical Diagnosis. *Diagnostics* **2020**, *10*, 904. [CrossRef]
29. Maqsood, S.; Javed, U.; Riaz, M.M.; Muzammil, M.; Muhammad, F.; Kim, S. Multiscale image matting based multi-focus image fusion technique. *Electronics* **2020**, *9*, 472. [CrossRef]
30. Jabeen, K.; Khan, M.A.; Alhaisoni, M.; Tariq, U.; Zhang, Y.D.; Hamza, A.; Mickus, A.; Damaševičius, R. Breast Cancer Classification from Ultrasound Images Using Probability-Based Optimal Deep Learning Feature Fusion. *Sensors* **2022**, *22*, 807. [CrossRef] [PubMed]
31. Al-Masni, M.A.; Al-Antari, M.A.; Choi, M.T.; Han, S.M.; Kim, T.S. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks. *Comput. Methods Programs Biomed.* **2018**, *162*, 221–231. [CrossRef] [PubMed]
32. Yassin, N.I.; Omran, S.; ElHouby, E.M.; Allam, H. Machine learning techniques for breast cancer computer aided diagnosis using different image modalities: A systematic review. *Comput. Methods Programs Biomed.* **2017**, *156*, 25–45. [CrossRef]
33. Wang, D.; Khosla, A.; Gargeya, R.; Irshad, H.; Beck, A.H. Deep Learning for Identifying Metastatic Breast Cancer. *arXiv* **2016**, arXiv:1606.05718.
34. 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] [PubMed]
35. Sechopoulos, I.; Teuwen, J.; Mann, R. Artificial intelligence for breast cancer detection in mammography and digital breast tomosynthesis: State of the art. *Semin. Cancer Biol.* **2021**, *72*, 214–225. [CrossRef]
36. Al-Masni, M.A.; Al-Antari, M.A.; Park, J.M.; Gi, G.; Kim, T.Y.; Rivera, P.; Kim, T.S. Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system. *Comput. Methods Programs Biomed.* **2018**, *157*, 85–94. [CrossRef] [PubMed]
37. Rezaei, Z. A review on image-based approaches for breast cancer detection, segmentation, and classification. *Expert Syst. Appl.* **2021**, *182*, 115204. [CrossRef]
38. Maqsood, S.; Damaševičius, R.; Maskeliūnas, R. Hemorrhage detection based on 3D CNN deep learning framework and feature fusion for evaluating retinal abnormality in diabetic patients. *Sensors* **2021**, *21*, 3865. [CrossRef]
39. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 1097–1105. [CrossRef]
40. 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 (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
41. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
42. Szegedy, C.; Ioffe, S.; Vanhoucke, V.; Alemi, A.A. Inception-v4, inception-resnet and the impact of residual connections on learning. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence, San Francisco, CA, USA, 4–9 February 2017.
43. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, MA, USA, 7–12 June 2015; pp. 1–9.
44. Huang, G.; Liu, Z.; Maaten, L.V.D.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
45. 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]
46. Junior, G.B.; da Rocha, S.V.; de Almeida, J.D.; de Paiva, A.C.; Silva, A.C.; Gattass, M. Breast cancer detection in mammography using spatial diversity, geostatistics, and concave geometry. *Multimed. Tools Appl.* **2019**, *78*, 13005–13031. [CrossRef]
47. Ghosh, S.K.; Mitra, A.; Ghosh, A. A novel intuitionistic fuzzy soft set entrenched mammogram segmentation under multigranulation approximation for breast cancer detection in early stages. *Expert Syst. Appl.* **2021**, *169*, 114329. [CrossRef]

48. Zheng, J.; Lin, D.; Gao, Z.; Wang, S.; He, M.; Fan, J. Deep learning assisted efficient AdaBoost algorithm for breast cancer detection and early diagnosis. *IEEE Access* **2020**, *8*, 96946–96954. [[CrossRef](#)]
49. Harefa, J.; Alexander, A.; Pratiwi, M. Comparison classifier: Support vector machine (SVM) and K-nearest neighbor (K-NN) in digital mammogram images. *J. Inform. Dan Sist. Inf.* **2017**, *2*, 35–40.
50. Samala, R.K.; Chan, H.P.; Hadjiiski, L.; Helvie, M.A.; Richter, C.D.; Cha, K.H. Breast cancer diagnosis in digital breast tomosynthesis: Effects of training sample size on multi-stage transfer learning using deep neural nets. *IEEE Trans. Med. Imaging* **2019**, *38*, 686–696. [[CrossRef](#)] [[PubMed](#)]
51. Qi, Q.; Li, Y.; Wang, J.; Zheng, H.; Huang, Y.; Ding, X.; Rohde, G.K. Label-efficient breast cancer histopathological image classification. *IEEE J. Biomed. Health Inform.* **2018**, *23*, 2108–2116. [[CrossRef](#)]
52. Agarwal, R.; Diaz, O.; Lladó, X.; Yap, M.H.; Martí, R. Automatic mass detection in mammograms using deep convolutional neural networks. *J. Med. Imaging* **2019**, *6*, 031409. [[CrossRef](#)]
53. Irfan, R.; Almazroi, A.A.; Rauf, H.T.; Damaševičius, R.; Nasr, E.A.; Abdelgawad, A.E. Dilated Semantic Segmentation for Breast Ultrasonic Lesion Detection Using Parallel Feature Fusion. *Diagnostics* **2021**, *11*, 1212. [[CrossRef](#)]
54. Rajinikanth, V.; Kadry, S.; Taniar, D.; Damaševičius, R.; Rauf, H.T. Breast-Cancer Detection using Thermal Images with Marine-Predators-Algorithm Selected Features. In Proceedings of the 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 25–27 March 2021; pp. 1–6.
55. Kadry, S.; Damaševičius, R.; Taniar, D.; Rajinikanth, V.; Lawal, I.A. Extraction of Tumour in Breast MRI using Joint Thresholding and Segmentation—A Study. In Proceedings of the 2021 Seventh International conference on Bio Signals, Images, and Instrumentation (ICBSII), Chennai, India, 25–27 March 2021; pp. 1–5.
56. Heath, M.; Bowyer, K.; Kopans, D.; Moore, R.; Kegelmeyer, W.P. The digital database for screening mammography. In Proceedings of the 5th International Workshop on Digital Mammography, Toronto, ON, Canada, 11–14 June 2000; pp. 212–218.
57. 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](#)]
58. Suckling, J.; Parker, J.; Dance, D.; Astley, S.; Hutt, I.; Boggis, C.; Ricketts, I.; Stamatakis, E.; Cerneaz, N.; Kok, S. The mammographic images analysis society digital mammogram database. *Excerpta Medica Int. Congr. Ser.* **1994**, *1069*, 375–378.
59. Fan, R.; Li, X.; Lee, S.; Li, T.; Zhang, H.L. Smart Image Enhancement Using CLAHE Based on an F-Shift Transformation during Decompression. *Electronics* **2020**, *9*, 1374. [[CrossRef](#)]
60. Zebari, D.A.; Ibrahim, D.A.; Zeebaree, D.Q.; Mohammed, M.A.; Haron, H.; Zebari, N.A.; Damaševičius, R.; Maskeliunas, R. Breast Cancer Detection Using Mammogram Images with Improved Multi-Fractal Dimension Approach and Feature Fusion. *Appl. Sci.* **2021**, *11*, 12122. [[CrossRef](#)]
61. Parsi, A.; Byrne, D.; Glavin, M.; Jones, E. Heart rate variability feature selection method for automated prediction of sudden cardiac death. *Biomed. Signal Process. Control* **2021**, *65*, 102310. [[CrossRef](#)]
62. Yang, X.S. Firefly algorithms for multimodal optimization. In *International Symposium on Stochastic Algorithms*; Springer: Berlin/Heidelberg, Germany, 2009; pp. 169–178.
63. Zang, H.; Zhang, S.; Hapeshi, K. A review of nature-inspired algorithms. *J. Bionic Eng.* **2010**, *7*, 232–237. [[CrossRef](#)]
64. Mohanty, F.; Rup, S.; Dash, B.; Majhi, B.; Swamy, M.N.S. Mammogram classification using contourlet features with forest optimization-based feature selection approach. *Multimed. Tools Appl.* **2019**, *78*, 12805–12834. [[CrossRef](#)]
65. Xie, W.; Li, Y.; Ma, Y. Breast mass classification in digital mammography based on extreme learning machine. *Neurocomputing* **2016**, *173*, 930–941. [[CrossRef](#)]
66. Mohanty, F.; Rup, S.; Dash, B.; Majhi, B.; Swamy, M.N.S. An improved scheme for digital mammogram classification using weighted chaotic salp swarm algorithm-based kernel extreme learning machine. *Appl. Soft Comput.* **2020**, *91*, 106266. [[CrossRef](#)]
67. Zhang, H.; Wu, R.; Yuan, T.; Jiang, Z.; Huang, S.; Wu, J.; Ji, D. DE-Ada: A novel model for breast mass classification using cross-modal pathological semantic mining and organic integration of multi-feature fusions. *Inf. Sci.* **2020**, *539*, 461–486. [[CrossRef](#)]
68. Pezeshki, H.; Rastgarpour, M.; Sharifi, A.; Yazdani, S. Extraction of spiculated parts of mammogram tumors to improve accuracy of classification. *Multimed. Tools Appl.* **2019**, *78*, 19979–20003. [[CrossRef](#)]
69. Tatikonda, K.C.; Bhuma, C.M.; Samayamantula, S.K. The analysis of digital mammograms using HOG and GLCM features. In Proceedings of the 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Bengaluru, India, 10–12 July 2018; pp. 1–7.
70. Pashoutan, S.; Shokouhi, S.B.; Pashoutan, M. Automatic breast tumor classification using a level set method and feature extraction in mammography. In Proceedings of the 2017 24th National and 2nd International Iranian Conference on Biomedical Engineering (ICBME), Tehran, Iran, 30 November–1 December 2017; pp. 1–6.
71. Aleem, J.; Qureshi, P.A.A.A.; Babar, N.; Sultan, A.; Rehman, A.U. Metastatic Choriocarcinoma of the Breast: A Rare Entity. *Cureus* **2022**, *14*, e22417. [[CrossRef](#)]
72. Shaban, M.; Raza, S.E.A.; Hassan, M.; Jamshed, A.; Mushtaq, S.; Loya, A.; Rajpoot, N. A digital score of tumour-associated stroma infiltrating lymphocytes predicts survival in head and neck squamous cell carcinoma. *J. Pathol.* **2022**, *256*, 174–185. [[CrossRef](#)]