Article

Detecting Breast Arterial Calcifications in Mammograms with Transfer Learning

Rimsha Khan and Giovanni Luca Masala *

School of Computing, University of Kent, Canterbury CT2 7NZ, UK
* Correspondence: g.masala@kent.ac.uk; Tel.: +44-01227816045

Abstract: Cardiovascular diseases, which include all heart and circulatory diseases, are among the major death-causing diseases in women. Cardiovascular diseases are not subject to screening programs, and early detection can reduce their mortal effect. Recent studies have shown a strong association between severe Breast Arterial Calcifications and cardiovascular diseases. The aim of this study is to use the screening programs for breast cancer to detect the high severity of BACs and therefore to obtain indirect information about coronary diseases. Previous attempts in the literature on the detection of BACs from digital mammograms still need improvements to be used as a standalone technique. In this study, a dataset of mammograms with BACs is divided into 4 grades of severity, and this study aims to improve their classification through a transfer learning approach to overcome the need for a large dataset of training. The performances achieved in this study by using pre-trained models to detect four Breast Arterial Calcifications severity grades reached an accuracy of 94% during testing. Therefore, it is possible to benefit from the advantage of Deep Learning models to define a rapid marker of BACs along Brest Cancer screening programs.

Keywords: breast arterial calcifications; cardiovascular diseases; coronary artery disease; deep learning; transfer learning

1. Introduction

Coronary Artery Disease (CAD) is among women's major reasons of death. The ratio of death due to this disease is more in women, and it is around 1 out of 7 in the U.S alone [1]. The obstruction of arteries that are the source of blood to the heart (called coronary arteries) and other body parts is the main reason for CAD [2]. These channels (arteries) are right on top of the heart muscle. This blockage is caused by the build-up of plaque, which is the deposition of cholesterol, calcium, fats, fibrin, and other substances along the arteries that narrow down the path. When these arteries are narrowed down heart does not get enough blood to function properly [3]. Presently, the disease is not detected until it is very late, so early detection of CAD among women is very important.

Breast Arterial Calcification (BAC) is a calcium deposition on the peripheral arterioles, which is known as Monckeberg medial calcific sclerosis [4]. Monckeberg's arteriosclerosis is a type of vessel hardening, where calcium deposits are found in the muscular middle layer of the walls of arteries [5]. BACs detection does not require any additional X-ray procedure besides the traditional mammography, which is taken by the majority of women older than 40–45 years, reducing health care costs and the body radiation rate in patients [6]. An estimate shows that approximately 40,000,000 mammography exams are executed per annum within the US [6]. BACs are a common finding that shows up as white areas in breast arteries on mammograms. They can appear in varying shapes and sizes, usually appearing as parallel or tabular tracks on mammograms [7]. Figure 1 shows BACs on a mammogram. It can be caused by several factors including normal aging, past trauma, inflammation, etc. [8].
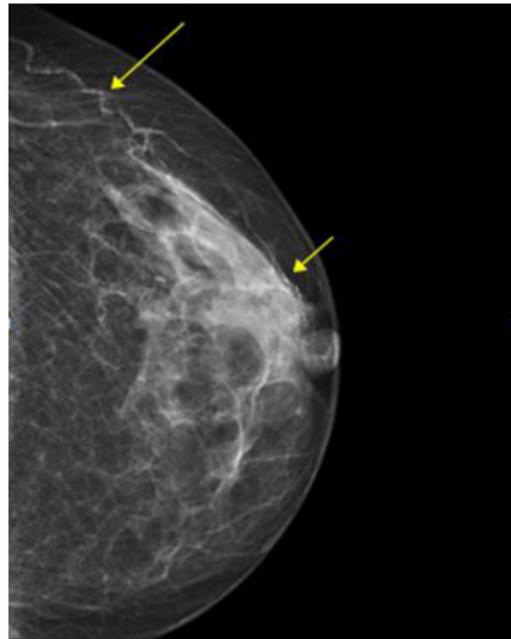
Figure 1. A digital mammogram. The yellow arrows are showing examples of Breast Arterial Calcifications (BACs).

By taking advantage of mammograms, which are used in breast cancer screening programs, we can detect BACs. At present, the final mammogram reports do not include BACs findings as they are considered irrelevant [9] and treated as false positives for the detection of breast cancer. However, rather than discarding this information from mammograms, it can be used to automatically detect BACs.

Mostafavi et al. (2015) [10] identified the relationship between mammographic breast arterial calcifications and coronary artery disease based on coronary computed tomography angiography results (CTA). The study analyzed 100 women who took both mammography and coronary CTA. The degree of BACs was extracted from the mammogram exams, being graded based on their length, extent, and severity. A large proportion of 83% of females diagnosed with moderate or advanced BACs on mammography was diagnosed by the coronary CTA with moderate to severe CAD.

Several studies succeeded in showing a positive association between CAD and the severity of BACs [10–12], and they proved that Breast Arterial Calcifications (BACs) identified on mammograms can be a beneficial risk indicator to show the existence of coronary artery disease [7,13–15]. With technological advancements and with getting more aware of the significance of Breast Arterial Calcifications, the automatic detection of BACs is becoming essential [16,17] Rather than having invasive and expensive approaches, it is possible to use routine mammography for this detection [6,16]. Hence, several studies are focusing on developing techniques that can be used for the automatic detection of BACs.

Deep learning is employed successfully for medical image analysis, and some contributions are exploring its potential in this domain [13,16–18]. Transfer learning is a branch of deep learning that has proved efficient in other medical image analyses [4,18,19] but has not yet been explored for the problem of BACs detection. Transfer learning techniques, as proposed in this study, can offer a real-time solution and strong performances, using a small dataset of training.

The research work in this paper attempts to detect the severity of BACs on mammograms as it is considered a valuable risk marker to identify CAD. The technical contribution of this work is related to the feasibility evaluation of pre-trained deep learning models in this emerging medical domain; the goal is the achievement of high performances, in comparison with the literature, with the constraint of a relatively small dataset of training and a more detailed description of the BACs severity.

The article is structured in this way: In Section 2 the most relevant articles on the topic are considered; in Section 3 the transfer learning methods and the dataset used in this paper are presented; in Section 4 the results of pre-trained models are shown with a comparison with the literature; and finally, in Section 5 the conclusions are summarized.

2. Literature Review

2.1. Previous Approaches

Some previous studies tried to show how BACs and CAD risk factors are related to one another to determine how well we can use BACs to determine CAD. Mostafavi et al.(2015) [10] attempted to evaluate Logistic regression for predicting CAD in patients. In the studied population, the prevalence of BACs was 12% and CAD was 29%. The results showed that 10 of 12 patients with BACs had CAD. Similar to this study, one other study [20] conducted performed a meta-analysis and systematic literature review showing a positive relation between BACs and cardiovascular events. Authors [11] also showed a confirmatory association between Coronary Artery Calcifications and BACs. Almost all the research depicts a relation between CAD and the presence of BACs, be it severe or mild.

Breast Arterial Calcifications appear in different shapes and sizes [7]; they can appear as elongated paths or short and inconsistent paths but can vary considerably. There have been many attempts to find an accurate method to quantify and detect BACs from mammograms; some methods can be used as an initial tool but due to these varying shapes and sizes of BACs, the detection becomes even more difficult. One study in this direction is presented [21]; this study started with image preparation by removing the black region around the breast, image resolution was also reduced, and then images were passed through a Gaussian filter and line Strength Algorithm to make images more clear and smooth. The images were then analyzed by region growing algorithm, and the white levels were used to decide the grade and presence of calcification. The highest accuracy (81%) was achieved for determining whether the calcification was mild or severe. Another study [22] in this direction used a two-step procedure. The first step was edge detection, and the second step was image segmentation. It used the line strength algorithm and region growing algorithm to make vessels more clear and highlight them. The results [22] showed that this method can accurately quantify BACs. Authors in [23] used a learning-based method for the segmentation of vessels. To capture variation in vessel patterns, it created a pool for several features that included local features along with Gabor, and Haar features that were taken from mammograms. This resulted in a very high-dimensional feature space. To handle a huge number of training data and high dimensional feature space, it applied forest along with boosting and tree for segmentation of vessels.

As the interest grew in this direction, numerous other algorithms were also presented in the literature for the detection of BACs. References. [24–26] proposed a process by seeing together both the vessel-ness clues and calcifications. The process had two steps; the first step generated several sampling paths by using an uncertainty system, and the second step grouped BAC paths to BACs using another assembling and connecting algorithm.

2.2. Previous Approaches Using Deep Learning

Article [18] was the first step to detecting BACs from mammograms using deep learning. A 12-layer CNN architecture was considered, pixel-wise, and a patch-based approach was used to detect BACs. The tissue region was determined as the biggest region that was connected having an intensity value more than the average intensity of the complete image. To detect BAC, thresholding and morphology operations are applied and the resultant BAC is identified as an overlapping region from the latter two approaches. The results showed an accuracy of 62.61%.

Several studies used U-Net with numerous alternations to detect different diseases. Xiaomeng Li [27] suggested an H-Dense U-Net that contracted the path of a U-Net with dense connections as a 2-dimensional Dense U-Net followed by a 3-dimensional Dense

U-Net for liver and tumor segmentation, Guan et al. [28] joined U-Net with Dense Net to eradicate the pieces of images. In the meantime, Chen et al. [29] proposed a method using dense connection blocks to extend the U-Net architecture for biomedical segmentation.

Inspired by results from many U-Net models in numerous segmentations and Dense Net in semantic segmentation, a study [17] used the U-Net model with dense connectivity for the automatic detection of BACs in mammograms. This method combines both short and long skip connections. The short skip connections were used to stop the model from learning unwanted features and increase the flow of information, whereas the long skip connections recover the information that was lost during encoding. It used a summation operation at the terminating point of the long skip connections. The results from this method were quite accurate, giving an accuracy of 91.47%.

Another deep learning approach [13] for the similar purpose of segmenting and quantifying Breast Arterial Calcifications was used. The study used a method for the segmentation of vessels called Simple Context U-Net (SCU-Net). The images were very large due to which they were divided into patches and then stitched back to obtain the final output. The study used five quantifiable metrics to determine the development of BACs. The results show around 95% correlation between the predicted mask of SCU-Net in comparison to the ground truth and measurement of breast calcifications on computed tomography.

Despite all these approaches, the accurate detection of BACs from mammograms is still an unresolved problem. There were issues in the deep learning model detection like the detection of small vessels. Severity detection in this research domain is currently an open area. Besides giving an excellent performance, the algorithms lagged in a few areas. The narrow and varied appearance of BACs makes it much more challenging. Some using the pixel-wise approach [18] limited the data used, some models proposed had very complex architecture [27], and some required large data sets. Moreover, some may produce wrong results when BACs are non-continuous [17]. There are no large, annotated datasets of BACs available, and the use of large images makes the processing significantly difficult. Furthermore, the training of deep learning models from scratch with such limitations makes the problem challenging to resolve.

The research work presented in this paper with respect to previously cited studies works on mammograms from screening programs without the need for additional input parameters to improve the performances. This study employs transfer learning where the models are already trained on a huge dataset, which reduces the time and resources required for model training, building a much lighter and more accurate model. Furthermore, it addresses the problem of the exact grade of calcifications, considered in [10,21], that can be used to define an automatic index associated with the severity of BACs.

3. Methodology

Transfer learning is an effective deep learning technique. Its main focus is to use the knowledge gained from one problem to solve other problems, it uses pre-trained models as a starting point for new tasks [30]. It is sometimes also used to overcome problems or constraints of traditional machine learning like the time required for training from scratch, or computational costs, or requiring a large amount of data [31].

Deep learning models have layered architectures; initial layers are used to learn generic features, whereas final layers are used to learn higher-order features or features related to a specific task. This layered architecture allows using pre-trained models without their final layers to extract features from other tasks [32]. Another approach is to freeze certain layers and retrain some other layers that help achieve better performance [32]. Most of the pre-trained models are trained on a large dataset (Image-Net dataset).

This research used pre-trained models and shows a comparison between different pre-trained models to define which model performs better in the classification of the grade of severity of BACs and can be used as a potential marker.

As for the implementation details, this study used Google Colaboratory (Google Research) as it provides free online GPU service and python with Keras, Tensor flow, and several python libraries (e.g., Augmentor). Pre-trained models were imported from Keras. Google Colab hosts Jupyter notebook services including GPU. The GPU provides the distribution of training processes that enables significant machine-learning operations.

3.1. Pre-Trained Models

In this research, the pre-trained models that were used on the pre-processed data are as follows:

3.1.1. VGG-19

VGG stands for Visual Geometry Group. The architecture of VGG is described in detail by Kaushik [33]. The architecture used in this model consisted of 19 layers. All the layers are sequentially arranged. Convolutional layers use a kernel of 3×3 with a stride size of 1 pixel, max pooling is applied over a 2×2 window with stride 2 for downsampling, and Rectified linear Unit (ReLu) is used to introduce nonlinearity. 4096 neurons exist in the fully connected layers and the final layer is a SoftMax function.

3.1.2. ResNet50

ResNet stands for Residual Networks. The architecture is explained in detail by Kaushik [34]. It is a CNN model consisting of 50 layers [35]. It makes use of skip connections. The reason for adding skip connections is to avoid a vanishing gradient. The architecture contains 48 convolutional layers, an average pool layer, and a max pool. After every convolutional layer, a batch normalization layer is attached.

3.1.3. DenseNet121

DenseNet121 stands for densely connected convolutional network. In this network, each layer is linked to every other upcoming layer. The architecture of DenseNet can be seen in [36]. DenseNet121 ensures that maximum information flows from input to output by a feed-forward network that joins the output of each layer to the next after applying composite operations. The architecture prevents redundant learning and provides feature reuse. After each dense block, transition layers are applied including convolutional operations, batch normalization, and activation function.

3.1.4. InceptionV3

The model comprises two portions. The first part extracts features using CNN and the second is the classification part using fully connected and softmax layers. The architecture for InceptionV3 is 48 layers deep with an input of 299×299. The main focus of Inception V3 was to use less computational power and prevent the number of parameters from growing too large when the network goes deep.

The architecture of Inception V3 shown by [37] uses factorized convolutional layers to reduce the parameters required for the network, smaller convolutions that result in faster training, and an auxiliary classifier to improve convergence and deal with vanishing gradient problems so it acts as a regularizer.

3.1.5. Mobile Net

Mobile Net [38] is an efficient and lightweight model. It is aimed at devices having limited computing power and memory by having a small network without compromising the speed of the model and is primarily built from depth-wise separable convolutions in place of convolutional layers. Each of these layers consists of depth-wise convolution and pointwise convolution. The model uses two new global parameters. The first one is the width parameter that helps in constructing smaller and computationally cheap models. The second is the resolution multiplier that helps to decrease the resolution of the input image.

3.2. Dataset

To date, there is no official benchmark per BACs classification. Large annotated datasets of mammograms used for breast cancer have only the region of interest related to tumors without details related to BACs. The images used in this study are originally provided by Peninsula Radiology Academy in Plymouth, UK, and have already been used in [21]. The data set consisted of a total of 104 mammograms from 26 female patients. For each patient, the dataset has a set of four mammograms from different projections, two for the right breast and two for the left breast. The dataset is unbiased and in equilibrium. The 26 patients were split into four groups based on the grade of calcification proposed in [10]. As shown in Figure 2, each grade represents a different level of severity. Grade 1 represents the lowest severity (no BACs) and Grade 4 represents the highest severity of calcification. Besides Grade 4, all other grades consist of data from seven patients, whereas grade 4 contains data from five patients. In this paper, we use the term 2-class classification for determining the severity of BACs (mild and severe) and 4-class classification is used when the classification refers to determining the actual grade of calcification (Grade 1, Grade 2, Grade 3, Grade 4).
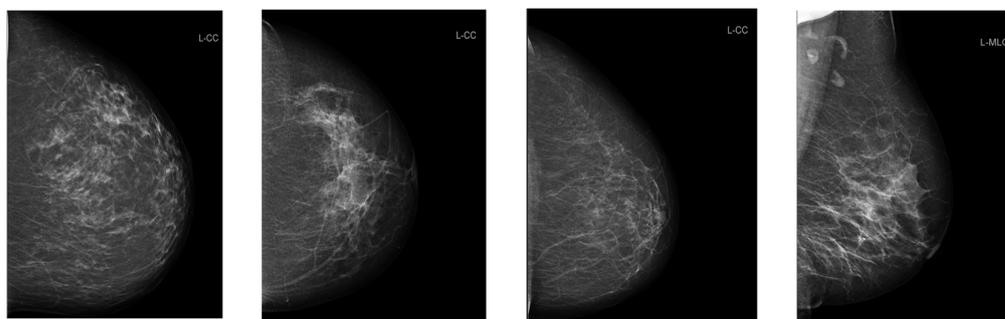

Figure 2. In the picture there are BAC examples from grade 1 until grade 4.

The dataset consisted of two sizes of images. Some images were of size 2082 × 2800 pixels, whereas the other images were of size 2800 × 3518 pixels. So, the images needed to be standardized before going any further. A set of images was separated from training and validation data to be used only for blind testing. For 4-class classification a total of 200 augmented images were used in the training set, however, the original images were divided into 58 and 46 for validation and test set, respectively. However, for 2-class classification the original images i.e., a total of 104 images were used for validation, whereas augmented images were divided as 200 and 50 for training and test sets, respectively.

3.3. Methods

While exploring the dataset, it was seen that each mammogram had an extra black area as a part of it, which does not contribute to the detection of calcifications, and it increases the size of the image unnecessarily. These black backgrounds were removed from all images, cropping out this area for each image.

The images in each grade folder consisted of a pair of right and left breasts. The study aimed to accurately detect the severity of Breast Arterial Calcifications and not be concerned about the left and right breast, so all images were flipped in one direction to improve similarity between images.

In addition, it was found that the dataset had images of two different dimensions. To deal with these issues, the data set was resized and the training was done with different image dimension datasets to see which dimensions produced the best results. The images of size 300 × 400 pixels were selected for final processing.

3.4. Data Augmentation Settings

The next step was to make sure that there was enough data to do the processing so that the model can be trained efficiently. Data augmentation increases the data set artificially by

several techniques like flipping, rotating, zooming, shearing, etc. [39]. For this purpose, a separate python library was used called Augmentor. This library took a folder for each grade of images and output a specified number of augmented images. Some of the results from augmented images are shown in Figure 3. The data augmentation technique is a very useful tool in avoiding over-fitting and it can be applied to both the training and the validation datasets. However, it is not indicated to use the data generator on the testing dataset as the model must be tested on real images. For images of size 300×400 pixels, 50 augmented images were generated for each grade only for the training set, whereas for testing and validation sets the original resized images were divided.
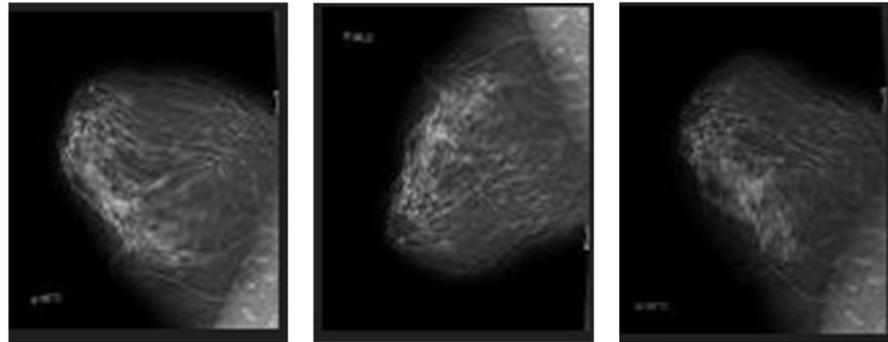
Figure 3. These images show augmented images extracted from a patient with a BAC of grade 1.

For this study, the following data augmentation operations were applied.
Rotation—As the name suggests it rotates the images to a certain degree specified.
Shearing—It changes the orientation of images.
Zooming—As the name suggests it either zooms in or out.
Shift range—This function shifts the image in either width or height.
The table below (Table 1) summarizes the augmentation parameters provided for the training data.

Table 1. Augmentation parameters set for the generating images.

Parameters	Values
Rescale	1/255
Height shift	0.2
Width shift	0.2
Zoom	0.2
Shear	0.2
Rotation	40
Horizontal flip	True

3.5. Transfer Learning Optimizations

Transfer learning allows building a model on top of what is already there by changing some of the layers of existing models and freezing a few layers [32]. For this study, pre-trained models were used for feature extraction and the output of the model was then used as input to the new classifier. So, the pre-trained model is integrated with a few extra layers so the weights of the pre-trained model are frozen.

When loading, model parameters are adjusted to load the model from the last convolution layer or pooling layer directly and prevent the loading of fully connected layers. The shape of the images is also specified before loading the model. The multi-dimensional output is flattened to linear output and then fed to dense layers to generate the output.

For this research purpose, the sequential model was used. The sequential model allows building the model layer by layer [40]. The pre-trained model is integrated into this sequential model on top of which a flattened layer and two dense layers are added.

3.6. Training of All Models

After defining the model, the next step is to set hyper-parameters that are required to get the best results and also match the processing capability. Batch size denotes the number of samples that are taken at a time in a single epoch [41]. For this research, a batch size of 32 was used.

The second important parameter that had a major impact on the results was the optimizer. It helps to reduce the overall loss and improve accuracy [42]. ADAGrad was used as it produced the best results; one possible reason for ADAgrad performing well was that it uses different learning rates depending on the parameters [43]. Figure 4 attached shows the results with Adagrad (Figure 4). Figure 5 shows the performance of the Adagrad optimizer using a confusion matrix during the validation of Inception V3 on determining the severity of BAC.
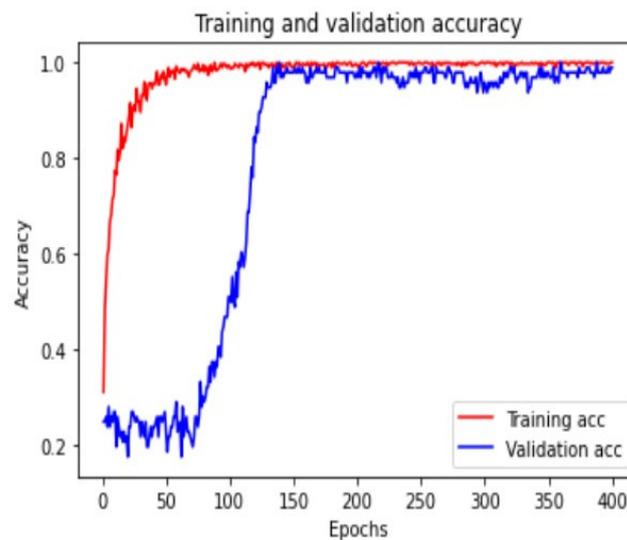
Figure 4. This figure shows training on Inception-V3 done using ADAgrad as an optimizer, which produced stable results.
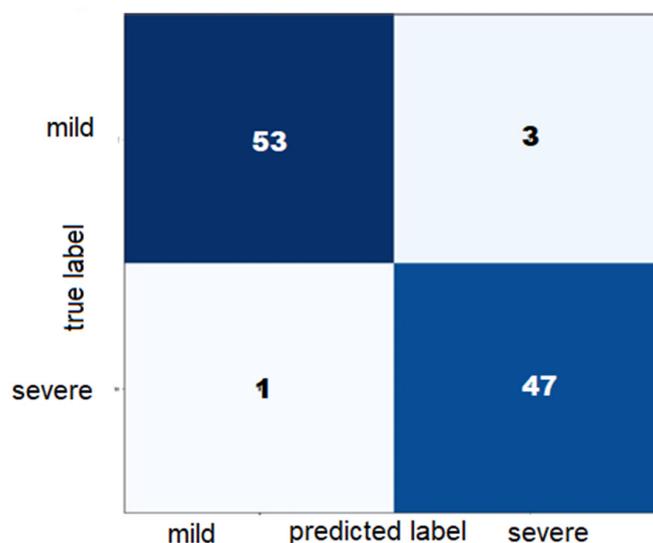
Figure 5. This figure shows the confusion matrix on Inception-V3 validation data in a 2-class classification.

Epochs are the number of rounds for which we pass over the entire dataset. For this study 400 to 600 epochs were used to make results stable and to avoid overfitting [44].

Another parameter adjusted is the loss function. Categorical cross-entropy was employed in this research. Categorical cross-entropy is used for multi-class classification. This cross-entropy assumes that among different classes, only one class is true and the remaining are false. Below is the table (Table 2) to summarize the hyper-parameters adjusted for the study.

Table 2. Hyper-parameters set for this model.

Hyper-Parameters	Values
Batch size	16
Epochs	400/600
Optimizer	Adagrad
Learning rate	Default (0.001)
Loss function	Categorical Cross-entropy

4. Results and Discussion

This section evaluates the performance of the pre-trained model in classifying BACs from mammograms in the considered dataset.

4.1. Results of Model

This section shows the result of the test data set. Tables 3 and 4 show the best results that were obtained after the adjustments of all hyper-parameters for the pre-trained models. The confusion matrices for the Tables 3 and 4 are placed below them as Figures 6 and 7 respectively. As it can be seen from both tables, the results for determining the severity are better than for determining the exact grade for calcification; this can be due to an increased dataset in the case of 2-class classification. However, looking at them individually, the performance is good in both cases of 2-class classification and 4-class classification.

The study not only evaluates the accuracy but other important indexes on the test set that are needed to determine the reliability of a model for medical use. For this study, the results are excellent for all the pre-trained models and some of them achieve over 94% in all indexes for Tables 3 and 4, which proves their reliability on this application.

The Comparison of accuracies for the training, validation, and blind testing is also shown in Figures 8 and 9. The majority of the images in the test set are correctly classified, which shows that the pre-trained models learned well and are capable of classifying new data.

Table 3. Comparison of pre-trained models for 2 Class Classification.

	VGG-19	ResNet50	Mobile Net	DenseNet-121	Inception V3
Training time (avg.) ms/epoch	9.733	8.673	7.533	8.631	7.640
Val. Accuracy (%)	100	93.75	95.83	95.83	97.92
Val. loss	1.88	11.92	22.41	13.07	2.42
Test Accuracy	96.15	97.11	97.11	98.07	97.08
Test Sensitivity (Recall)	100	100	100	100	98
Test Specificity	93	95	95	96	95
Test Precision	92	94	94	96	94
Figure No.	6	7	8	9	10

Table 4. Comparison of pre-trained models for 4 Class Classification.

Model	VGG-19	ResNet50	Mobile Net	DenseNet-121	Inception V3
Training time (avg.) ms/epoch	9.733	8.676	6.523	7.598	7.588
Val. Accuracy (%)	90.62	93.75	84.38	93.75	96.88
Val. loss	42.05	7.45	58.84	43.20	13.49
Test Accuracy	90	92	94	92	93
Test Sensitivity	89	93	94	93	94
Test Precision	90	93	94	93	94
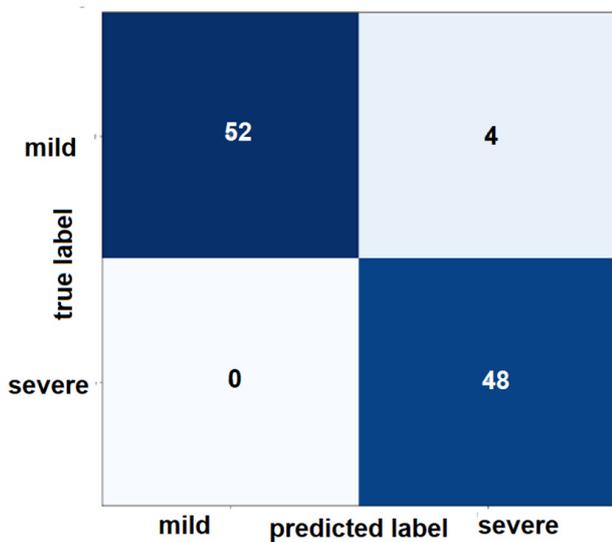
(**a**) VGG-19
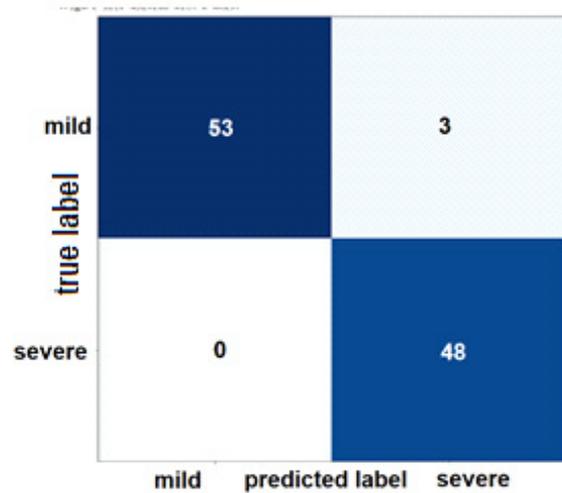
(**b**) ResNet50
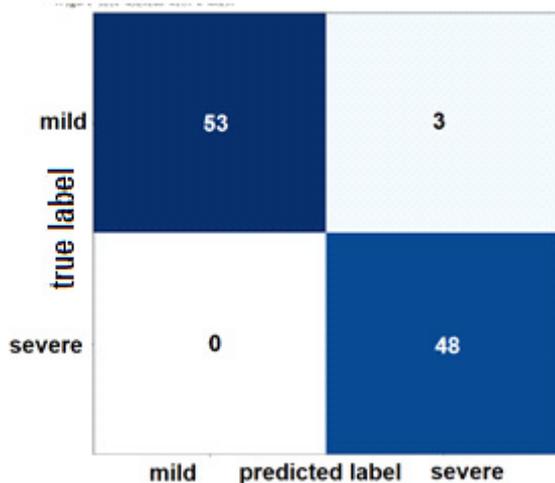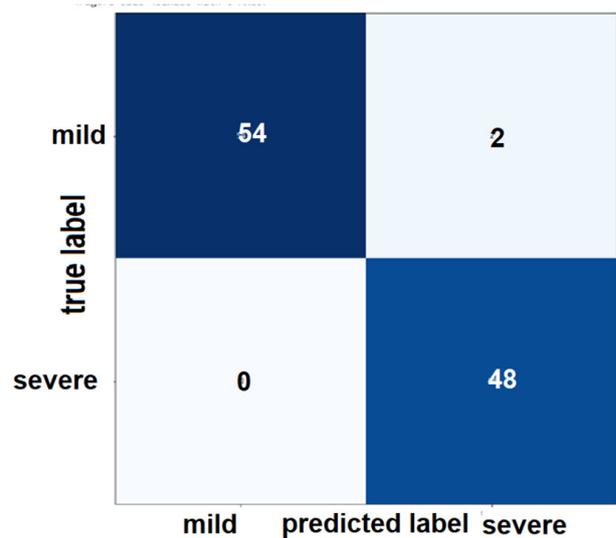
(**c**) MobileNet

(**d**) DenseNet-121

Figure 6. Confusion matrices for 2–class classification.
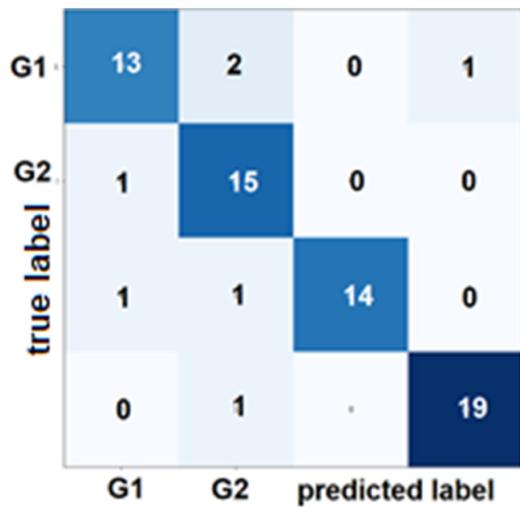
(**a**) VGG-19
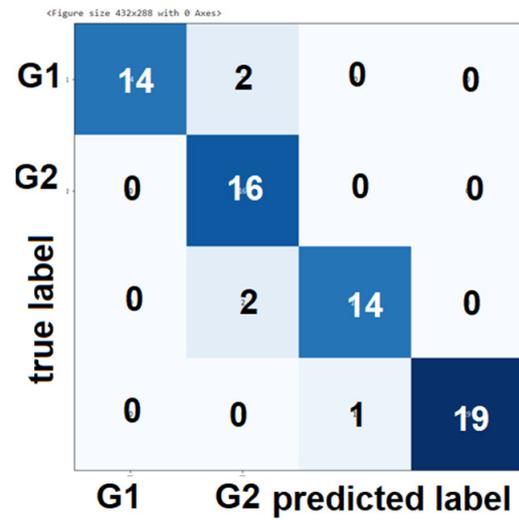
(**b**) ResNet50
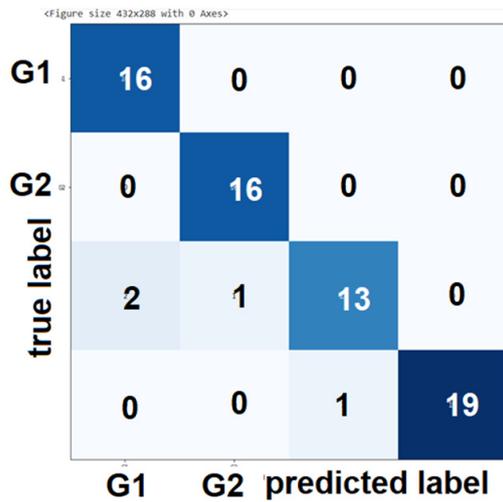
(**c**) MobileNet
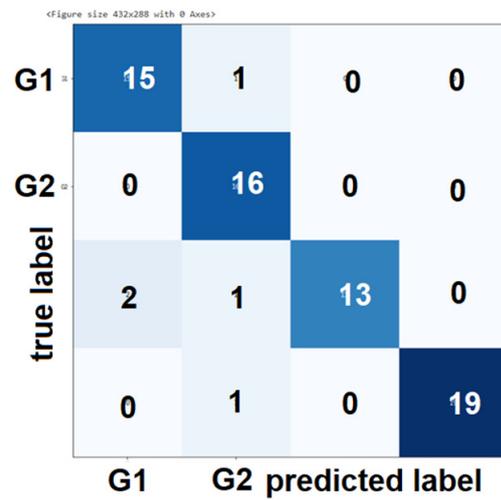
(**d**) DenseNet-121
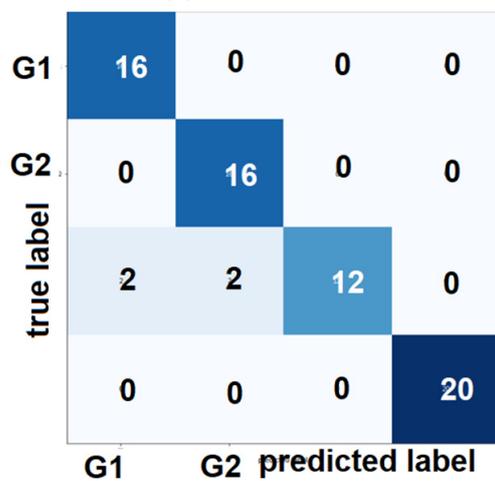
(**e**) InceptionV3

Figure 7. Confusion matrices for 4 class classification.
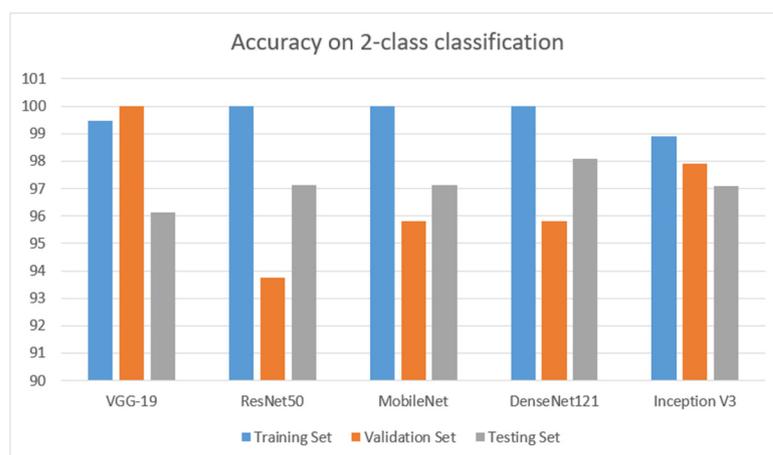
**Figure 8.** The plot shows accuracies compared for the 2-class classification.
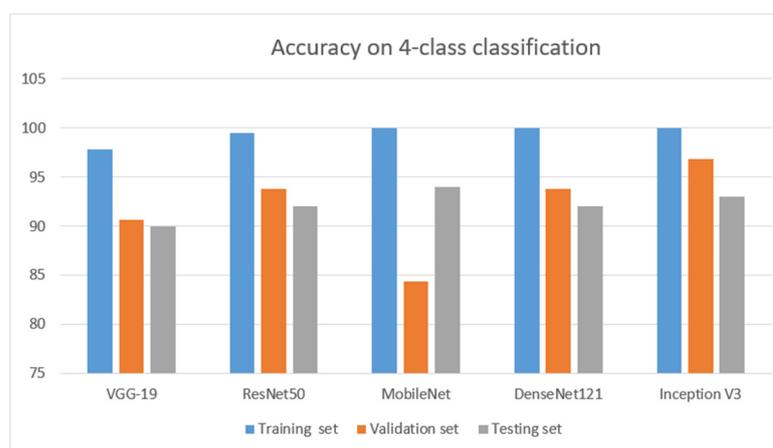
Figure 9. The plot shows accuracies compared for the 4-class classification.

4.2. Verification of Results on Large Population

To verify the model and the adjustments that were done to the algorithm and the images, the Stratified K fold method was used to divide data into three parts and in each loop used one part for validating and two parts for training. After going through the three loops, the resultant model was then tested with the test data set that was kept separate and did not become a part of training or validation. To estimate the accuracy of the pre-trained methods over a large population, starting from the sample of the 104 images in our dataset, a 95% confidence interval was evaluated using the Wilson method [45]. Such prediction of accuracy on the proposed models, excluding DenseNet121, can be considered between 87.13–100% in all cases, whether it uses 2-class classification or 4-class classification, which shows that the pre-trained model can be a good method of classifying the Breast Arterial Calcifications accurately.

Furthermore, the results produced by this study can be compared with the previous studies that are summarized below in Table 5.

Comparing our proposed model to the previous studies shown in Table 5, it can be seen that the results from our pre-trained models are more accurate than most of the previous studies. Along with accuracies on the severity of BACs, the pre-trained models also provide the exact grade of calcifications. Furthermore, this study employs transfer learning where the models are already trained on a huge dataset, which reduces the time and resources required for model training, building a model that is much lighter and more accurate.

Table 5. Summary of the quantitative recent approaches in the literature.

Reference	Objective	Classification Type	Method	Year	Accuracy %
[17]	Detect BACs in real time using mammograms	2 class	Simple context U-net (SCU-Net)	2021	95
[46]	Compute the quantity of BACs and also determine vessel calcification in the breast as a detector	4 class	Remove the extra black space from the mammograms and apply different algorithms.	2019	81.8
[13]	Determine the performance of deep learning for vascular extraction	2 class	A retrospective analysis was used to determine the performance of the 2 classes	2020	98
[18]	Detect the presence and severity of BACs from mammograms	2 class	CNN	2017	62.61
[21]	Detect the presence and severity of 2 class BACs from mammograms	2 class	A handcrafted method with medical imaging technique	2019	82

5. Conclusions

To summarize, in this research paper, the use of transfer learning for the detection of BACs on mammograms was examined, which is considered a useful risk marker for the detection of CAD. This work tested the results on the given dataset and has also shown a comparison between different pre-trained models to see which model performs well. In this case, most of the pre-trained models we used, i.e., DenseNet, InceptionV3, and Mobile Net, perform very well in terms of accuracy, specificity, and sensitivity. Mobile Net with around 97% of accuracy in 2-class classification and 94% of accuracy in 4-class classification shows that the performance of pre-trained models is, in general, better than previous studies in literature, and they optimize the time and resources required for model training. Furthermore, the performances are robust also considering a 95% confidence interval for the sample of the images considered, having predicted accuracy on large populations between 87.13–100%.

The results direct that this approach can be evolved further to be used as an automatic mechanism to identify the severity of BACs (along with the exact grade of BACs), integrated in a Computer Aided Detection System for Breast Cancer. Such a system can be used along breast cancer screening programs to provide an indirect screening of Coronary Artery Disease for the same target population.

References

1. Mozaffarian, D.; Benjamin, E.J.; Go, A.S.; Arnett, D.K.; Blaha, M.J.; Cushman, M.; Das, S.R.; De Ferranti, S.; Després, J.-P.; Fullerton, H.J. Heart disease and stroke statistics—2016 update: A report from the American Heart Association. *Circulation* **2016**, *133*, e38–e360. [CrossRef] [PubMed]
2. Libby, P.; Theroux, P. Pathophysiology of coronary artery disease. *Circulation* **2005**, *111*, 3481–3488. [CrossRef] [PubMed]
3. Schoepf, U.J.; Becker, C.R.; Ohnesorge, B.M.; Yucel, E.K. CT of coronary artery disease. *Radiology* **2004**, *232*, 18–37. [CrossRef] [PubMed]
4. Havaei, M.; Davy, A.; Warde-Farley, D.; Biard, A.; Courville, A.; Bengio, Y.; Pal, C.; Jodoin, P.-M.; Larochelle, H. Brain tumor segmentation with deep neural networks. *Med. Image Anal.* **2017**, *35*, 18–31. [CrossRef] [PubMed]
5. Castling, B.; Bhatia, S.; Ahsan, F. Mönckeberg's arteriosclerosis: Vascular calcification complicating microvascular surgery. *Int. J. Oral Maxillofac. Surg.* **2015**, *44*, 34–36. [CrossRef] [PubMed]
6. US Food and Drug Administration. *FDA Approves First 3-D Mammography Imaging System*; FDA: Silver Spring, MD, USA, 2011.
7. Ge, J.; Chan, H.-P.; Sahiner, B.; Zhou, C.; Helvie, M.A.; Wei, J.; Hadjiiski, L.M.; Zhang, Y.; Wu, Y.-T.; Shi, J. Automated detection of breast vascular calcification on full-field digital mammograms. In Proceedings of the Medical Imaging 2008: Computer-Aided Diagnosis, San Diego, CA, USA, 16–21 February 2008; pp. 379–385.
8. Marks, H. Breast Calcifications. *WebMD*. 27 August 2022. Available online: https://www.webmd.com/women/guide/breast-calcification-symptoms-causes-treatments (accessed on 30 December 2022).
9. Molloi, S.; Mehraien, T.; Iribarren, C.; Smith, C.; Ducote, J.L.; Feig, S.A. Reproducibility of breast arterial calcium mass quantification using digital mammography. *Acad. Radiol.* **2009**, *16*, 275–282. [CrossRef]
10. Mostafavi, L.; Marfori, W.; Arellano, C.; Tognolini, A.; Speier, W.; Adibi, A.; Ruehm, S.G. Prevalence of coronary artery disease evaluated by coronary CT angiography in women with mammographically detected breast arterial calcifications. *PLoS ONE* **2015**, *10*, e0122289. [CrossRef]
11. Matsumura, M.E.; Maksimik, C.; Martinez, M.W.; Weiss, M.; Newcomb, J.; Harris, K.; Rossi, M.A. Breast artery calcium noted on screening mammography is predictive of high risk coronary calcium in asymptomatic women: A case control study. *Vasa* **2013**, *42*, 429–433. [CrossRef]
12. Pecchi, A.; Rossi, R.; Coppi, F.; Ligabue, G.; Modena, M.G.; Romagnoli, R. Association of breast arterial calcifications detected by mammography and coronary artery calcifications quantified by multislice CT in a population of post-menopausal women. *La Radiol. Med.* **2003**, *106*, 305–312.
13. Guo, X.; O'Neill, W.C.; Vey, B.; Yang, T.C.; Kim, T.J.; Ghassemi, M.; Pan, I.; Gichoya, J.W.; Trivedi, H.; Banerjee, I. SCU-Net: A deep learning method for segmentation and quantification of breast arterial calcifications on mammograms. *Med. Phys.* **2021**, *48*, 5851–5861. [CrossRef]
14. Kataoka, M.; Warren, R.M.L.; Sala, E. How predictive is breast arterial calcification of cardiovascular disease and risk factors when found at screening mammography? *Am. J. Roentgenol.* **2006**, *187*, 73–80. [CrossRef] [PubMed]
15. Rotter, M.A.; Schnatz, P.F.; Currier, A.A., Jr.; O'Sullivan, D.M. Breast arterial calcifications (BACs) found on screening mammography and their association with cardiovascular disease. *Menopause* **2008**, *15*, 276–281. [CrossRef] [PubMed]
16. Iribarren, C.; Go, A.S.; Tolstykh, I.; Sidney, S.; Johnston, S.C.; Spring, D.B. Breast vascular calcification and risk of coronary heart disease, stroke, and heart failure. *J. Women's Health* **2004**, *13*, 381–389. [CrossRef] [PubMed]
17. AlGhamdi, M.; Abdel-Mottaleb, M.; Collado-Mesa, F. DU-Net: Convolutional network for the detection of arterial calcifications in mammograms. *IEEE Trans. Med. Imaging* **2020**, *39*, 3240–3249. [CrossRef] [PubMed]
18. Wang, J.; Ding, H.; Bidgoli, F.A.; Zhou, B.; Iribarren, C.; Molloi, S.; Baldi, P. Detecting cardiovascular disease from mammograms with deep learning. *IEEE Trans. Med. Imaging* **2017**, *36*, 1172–1181. [CrossRef] [PubMed]
19. Al Ghamdi, M.; Li, M.; Abdel-Mottaleb, M.; Abou Shousha, M. Semi-supervised transfer learning for convolutional neural networks for glaucoma detection. In Proceedings of the ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Brighton, UK, 12–17 May 2019; pp. 3812–3816.
20. Hendriks, E.J.; De Jong, P.A.; van der Graaf, Y.; Willem, P.T.M.; van der Schouw, Y.T.; Beulens, J.W. Breast arterial calcifications: A systematic review and meta-analysis of their determinants and their association with cardiovascular events. *Atherosclerosis* **2015**, *239*, 11–20. [CrossRef] [PubMed]
21. Mazidi, N.; Roobottom, C.; Masala, G. Automatic Quantification of Breast Arterial Calcification on Mammographic Images. In *Innovation in Medicine and Healthcare Systems, and Multimedia*; Springer: New York, NY, USA, 2019; pp. 283–292.
22. Nava, E.; Barba, I.; Sendra, F.; Gómez–Rebollo, C. Quantification of vascular calcifications on digitized mammograms. In *International Workshop on Digital Mammography*; Springer: New York, NY, USA, 2010; pp. 183–190.
23. Cheng, E.; McLaughlin, S.; Megalooikonomou, V.; Bakic, P.R.; Maidment, A.D.; Ling, H. Learning-based vessel segmentation in mammographic images. In Proceedings of the IEEE International Conference on Healthcare Informatics, Imaging and Systems Biology, HISB 2011, San Jose, CA, USA, 26–29 July 2011; pp. 315–322.
24. Cheng, J.-Z.; Chen, C.-M.; Cole, E.B.; Pisano, E.D.; Shen, D. Automated delineation of calcified vessels in mammography by tracking with uncertainty and graphical linking techniques. *IEEE Trans. Med. Imaging* **2012**, *31*, 2143–2155. [CrossRef] [PubMed]
25. Cheng, J.-Z.; Chen, C.-M.; Shen, D. Identification of breast vascular calcium deposition in digital mammography by linear structure analysis. In Proceedings of the 2012 9th IEEE International Symposium on Biomedical Imaging (ISBI), Barcelona, Spain, 2–5 May 2012; pp. 126–129.

26. Cheng, J.-Z.; Cole, E.B.; Pisano, E.D.; Shen, D. Detection of arterial calcification in mammograms by random walks. In Proceedings of the 27th International Conference, IPMI 2021, Virtual Event, 28–30 June 2021; pp. 713–724.

27. Li, X.; Chen, H.; Qi, X.; Dou, Q.; Fu, C.-W.; Heng, P.-A. H-DenseUNet: Hybrid densely connected UNet for liver and tumor segmentation from CT volumes. *IEEE Trans. Med. Imaging* **2018**, *37*, 2663–2674. [CrossRef]

28. Guan, S.; Khan, A.A.; Sikdar, S.; Chitnis, P.V. Fully dense UNet for 2-D sparse photoacoustic tomography artifact removal. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 568–576. [CrossRef]

29. Chen, L.; Bentley, P.; Mori, K.; Misawa, K.; Fujiwara, M.; Rueckert, D. DRINet for medical image segmentation. *IEEE Trans. Med. Imaging* **2018**, *37*, 2453–2462. [CrossRef]

30. Baheti, P. A Newbie-Friendly Guide to Transfer Learning. *v7*, 19 July 2022.

31. Seldon. Transfer Learning for Machine Learning. *Seldon*, 29 June 2019.

32. Sarkar, D.D. A Comprehensive Hands-on Guide to Transfer Learning with Real-World Applications in Deep Learning. *Towards Data Science*. 2018. Available online: https://towardsdatascience.com/a-comprehensive-hands-on-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a (accessed on 30 December 2022).

33. Kaushik, A. *Understanding the VGG19 Architecture*, OpenGenus IQ: Computing Expertise & Legacy. 2022.

34. Kaushik, A. *Understanding ResNet50 architecture*, OpenGenus IQ: Computing Expertise & Legacy. 2022.

35. Boesch, G. *Deep Residual Networks (ResNet, ResNet50)—Guide in 2022*; Datagen: New York, NY, USA, 2022.

36. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.

37. Feng, C.; Zhang, H.; Wang, S.; Li, Y.; Wang, H.; Yan, F. Structural damage detection using deep convolutional neural network and transfer learning. *KSCE J. Civ. Eng.* **2019**, *23*, 4493–4502. [CrossRef]

38. Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* **2017**, arXiv:1704.04861.

39. Shah, D. The Essential Guide to Data Augmentation in Deep Learning. Available online: https://www.v7labs.com/blog/data-augmentation-guide (accessed on 19 July 2022).

40. Fchollet: The Sequential Model. *Keras*, 4 December 2020.

41. Brownlee, J. Difference Between a Batch and an Epoch in a Neural Network. *Machine Learning Mastery*, 20 July 2018.

42. Doshi, S. Various Optimization Algorithms For Training Neural Network. *Towards Data Science*, 13 January 2019.

43. Mayanglambam, G. Deep Learning Optimizers. *Towards Data Science*, 18 November 2020.

44. Afaq, S.; Rao, S. Significance of epochs on training a neural network. *Int. J. Sci. Technol. Res.* **2020**, *9*, 485–488.

45. Newcombe, R.G. Two-sided confidence intervals for the single proportion: Comparison of seven methods. *Stat. Med.* **1998**, *17*, 857–872. [CrossRef]

46. Trimboli, R.M.; Codari, M.; Bert, A.; Carbonaro, L.A.; Maccagnoni, S.; Raciti, D.; Bernardi, D.; Clauser, P.; Losio, C.; Tagliafico, A.; et al. Breast arterial calcifications on mammography: Intra-and inter-observer reproducibility of a semi-automatic quantification tool. *Radiol. Med.* **2018**, *123*, 168–173. [CrossRef]