Available online at www.ijournalse.org

Emerging Science Journal

(ISSN: 2610-9182)

Vol. 8, No. 6, December, 2024

Breast Cancer Prediction Using Transfer Learning-Based Classification Model

Sheeba Armoogum [1], Kezhilen Motean [2], Deshinta Arrova Dewi [3*], Tri Basuki Kurniawan [4], Jureerat Kijsomporn [5]

[1] *Faculty of Information, Communication and Digital Technologies, University of Mauritius, Moka, Mauritius.*

[2] *Spoon Consulting, Moka, Mauritius.*

[3] *INTI International University, Nilai, Malaysia.*

[4] *Faculty of Science and Technology, Universitas Bina Darma Palembang, 30264, UBD, Palembang, Indonesia.*

[5] *Nursing Faculty, Shinawatra University, Bang Toei, Thailand.*

Abstract

Breast cancer is currently the most prevalent type of cancer in women, with a growing number of fatalities worldwide. Different imaging methods like mammography, computed tomography, Magnetic Resonance Imaging, ultrasound, and biopsies assist in detecting breast cancer. Recent developments in deep learning have revolutionized breast cancer pathology by facilitating accurate image categorization. This study introduces a novel approach to enhance detection and classification using the Convolutional Neural Network Deep Learning method and Transfer Learning to create a high-speed, accurate image classification model. The model is trained on pre-processed data subjected to thorough analysis and augmentation to ensure the quality of inputs. The experimental results from the Breast Ultrasound Image dataset indicate that our model, with a 0.1 test size ratio, outperforms its counterparts. It achieved an accuracy of 90.12%, with a loss of 0.2641, validation accuracy of 90.15%, and validation loss of 0.31, evidencing its superior classification capability. This research introduces an innovative approach to the automated diagnosis of breast cancer. By combining CNN, Transfer Learning, and data augmentation, we have developed a desktop application that expedites the classification process and significantly improves accuracy. This advancement represents a key development in machine learning applications for breast cancer prognostics and diagnostics.

Keywords:

Breast Cancer;
Machine Learning;
Deep Learning; Transfer Learning;
CNN; BUSI; Process Innovation;
Public Health.

Article History:

Received:	22	May	2024
Revised:	21	November	2024
Accepted:	27	November	2024
Published:	01	December	2024

1- Introduction

Breast cancer is a complicated condition identified by the uncontrollable multiplication of cells in the breast tissue and has become widespread among women. According to a 2023 report from the World Health Organization [1], about 2.3 million women were diagnosed with breast cancer, and this number is rising annually. It is projected that the incidence of breast cancer will escalate to approximately 3 million cases per year by 2040, leading to an estimated one million deaths annually. The prevalence and death rates related to breast cancer highlight the critical requirement for greater awareness, early identification, and enhanced treatment choices. Breast cancer remains a significant global health challenge, affecting millions of women and causing substantial mortality rates. Both men and women can develop breast

---

* **CONTACT**: deshinta.ad@newinti.edu.my

DOI: http://dx.doi.org/10.28991/ESJ-2024-08-06-014

cancer, but females are at a higher risk due to various factors such as hormonal differences and genetic susceptibility. Only 0.5-1% of cases are concerned with men, which makes at least 99% of the cases accountable for women. Numerous factors increase the likelihood of breast cancer among women, including age, family history, and having dense breasts. In Mauritius, as of 2022, 617 out of the 1814 diagnosed cancer cases were associated with breast cancer. Additionally, among the 827 recorded cancer-related deaths, 219 were attributed to this specific type of cancer [2].

The timely and accurate identification of breast cancer through routine screening and diagnostic examinations significantly improves the chances of survival for individuals with the disease. Detecting the condition before symptoms appear, particularly in its early stages when treatment is most effective, significantly increases the likelihood of successful therapy [3]. Various imaging techniques, such as mammography, computed tomography, magnetic resonance imaging, ultrasound, and biopsies, are essential diagnostic modalities that can detect breast cancer in its early stages [4]. Histopathological examination, involving the examination of a biopsy or the microscopic analysis of tissue samples, is widely considered the primary standard for diagnosing breast cancer. However, its precision may be affected by subjective analyses from pathologists and the quality of the images depicting tissue samples. Variations in human decision-making and differences in image analysis and texture can result in inconsistencies in diagnosing medical conditions. This highlights the importance of standardized image quality and the possible incorporation of objective, computer-assisted diagnostic systems to aid pathologists in reducing bias and enhancing uniformity in diagnosing breast cancer [3].

Applying advanced machine learning methods in analyzing diagnostic images has enhanced breast cancer detection and classification precision, ultimately improving patient outcomes. Machine learning (ML) and deep learning (DL) models have shown promise in achieving high levels of accuracy, often surpassing manual detection in specific tasks. However, their effectiveness can depend on factors including the quality and diversity of the data they've been trained on and the algorithms utilized [5]. Machine learning algorithms have demonstrated the potential to attain high levels of precision, sometimes exceeding manual detection in specific duties. Nonetheless, their effectiveness may differ depending on factors such as the quality and diversity of the training data and the particular methods used. It should be emphasized that aspects like dataset size and quality, extracted features, and selected classifiers also influence the performance of machine learning models. Incorporating machine learning into diagnostic procedures can enhance breast cancer identification's effectiveness, precision, and impartiality. Instead of entirely supplanting manual detection, it functions most effectively as a supplementary tool that enhances healthcare practitioners' knowledge. ML is vital in aiding medical professionals in detecting crucial details that might not be apparent to the human eye, assisting with early diagnosis and predicting breast cancer prognosis [4]. Deep learning techniques such as convolutional neural networks (CNNs) have demonstrated considerable promise in the analysis of medical images in recent years. These techniques are mainly applied to recognizing intricate patterns in visual data that may be difficult for the human eye to perceive. CNNs can be trained using extensive datasets of mammographic or ultrasound images to accurately differentiate between benign and malignant lesions in breast cancer detection [6]. Transfer Learning, where a model developed for one task is repurposed on a second related task, has allowed for the efficient application of deep learning in medical contexts. This implies that a previously trained CNN can be adjusted with a smaller collection of specialized medical images, resulting in faster deployment and reduced computational resources [7]. By employing these advanced techniques, researchers and clinicians can significantly improve the speed and reliability of breast cancer diagnosis. Consequently, this can lead to earlier detection and treatment, which is crucial in improving survival rates for breast cancer patients. Integrating deep learning into diagnostic workflows holds promise for enhancing the capabilities of medical imaging technologies and supporting healthcare professionals in their decision-making processes [6, 7].

By utilizing deep learning techniques and analyzing breast cancer images, this study aims to improve the accuracy and efficiency of a breast cancer diagnosis. This study seeks to create a deep-learning framework that effectively classifies breast cancer into benign, malignant, or normal categories using distinct criteria and characteristics. CNN and Transfer Learning will be applied to the BUSI dataset, and different test size ratios will be used for assessment. Ultrasound imaging will be utilized in developing the model, as ultrasonography is becoming increasingly crucial in lessening the worldwide impact of breast cancer due to its wide availability, cost-efficiency, and safety [8]. With resources in high demand and mammography impractical and inaccessible in low- and middle-income nations, this innovative method can serve as a valuable diagnostic tool for identifying conditions early [9].

The rest of this document is organized as follows: Section 2 includes an extensive literature review summarizing the studies on related works in classifying breast cancer images. Section 3 outlines the research methods employed in this study, detailing the approach and techniques used. The study results and conclusion will be presented in sections 4 and 5.

2- Literature Review

A multitude of research studies have been conducted in the field of breast cancer diagnosis, including various methods for prediction and early detection. These studies have contributed to advancements in understanding the disease and improving patient outcomes. With advancements in medical research, numerous works have emerged to improve the detection and classification of breast cancer. Pavithra et al. [10] developed a Computer-Aided Diagnosis (CAD) system to enhance breast cancer detection accuracy using methods such as noise removal, segmentation (active contour-based segmentation), feature extraction, and classification on the BUSI dataset. The extracted features were utilized with classifiers, including the K-Nearest Neighbours (KNN) algorithm, Decision tree algorithm, and Random Forest classifier. It was found that Random Forest outperformed other classifiers with an 88% accuracy rate. Similarly, Hijab et al. [11] employed ultrasound images but used different training strategies for their CNN model: baseline training from scratch, Transfer Learning using a pre-trained VGG-16 CNN, and fine-tuned learning to address overfitting resulted in a fine-tuned model achieving superior performance with 97% accuracy and an AUC value of 0.98. These research studies demonstrate the effectiveness of Computer-Aided Diagnosis systems in improving the accuracy of breast cancer detection.

Uysal et al. [12] investigated utilizing three distinct CNN structures—ResNet50, ResNeXt50, and VGG16—to categorize benign, malignant, and normal cases. The BUSI dataset underwent pre-processing to enhance the quality of data input for accurate analysis. Augmentation techniques were also applied to expand the dataset while maintaining its integrity. The augmented dataset was then divided into a 70:30 ratio for training and validation purposes to ensure robust model performance across different subsets of data. Findings revealed that while ResNeXt50 achieved the highest accuracy rate, ResNet50 exhibited enhanced stability and more reliable outcomes due to its unique architecture suited for medical imaging analysis. In another research work led by Arooj et al. [13], a customized CNN Alex Net model with Transfer Learning on ultrasound and histopathology images was built to automate the detection process. The datasets (A, B, C, and A2) consisted of ultrasound images in A, histopathology images in B and C, and a subset of A with two classes in A2. After the pre-processing and the data-splitting (80-20 split for training and testing) phase, the model attained an accuracy of 99.4% on A, 99.66% on B, 99.11% on C, and 100% on A2. They deduced that their proposed system outperformed the existing A, B, C, and A2 models.

Noel et al. [14] stated that ultrasound breast cancer datasets have limited availability, leading to challenges in efficient detection due to blurry and indistinct regions of interest in the images. They thoroughly researched various image decomposition methods to enhance the information inputted into a pre-trained model, explicitly addressing this issue. The model's performance metrics are accuracy 93%, sensitivity 95%, specificity 88%, F1-Score 93%, and AUC 97%. Another recent study that specifically examined ultrasound breast images was conducted by Zeimarani et al. [15]. They developed a customized CNN with regularization techniques and employed Transfer Learning using a dataset comprising 641 cases. The model assessment was carried out using a 5-fold cross-validation approach. Before augmentation and regularization, the accuracy and AUC were reported at 85.98% and 0.94, respectively; post-augmentation and regularization improved significantly to 92.05% for accuracy and 0.97 for AUC. Lastly, the pre-trained model achieved an accuracy of 87.07% and an Area Under the Curve of 96%.

Bahmani et al. [16] proposed another innovative and effective way to construct a reliable machine-learning model, enhancing its capabilities and robustness for real-world applications. After pre-processing the image using Contourlet Transformation, the feature extraction method, namely the dependent model, was applied. Before classification with Machine Learning, feature reduction using PCA was employed. According to the performance analysis, it was deduced that the Decision Tree Classifier showed the best performance. Raza et al. [17] introduced a novel deep-learning model, DeepBreastCancerNet, to detect, classify, and predict outcomes for breast cancer patients with high accuracy using advanced machine-learning techniques. The BUSI dataset was pre-processed and augmented before being fed to the model, comprising 24 layers with specialized activation functions and normalization operations. With a classification accuracy of 99.35%, the model was superior during the comparative analysis. Another publicly available dataset was used to validate the model, and an accuracy of 99.63% was attained.

Dan et al. [18] developed a novel framework to improve breast cancer detection. The study presented a deep learning model for classifying breast masses in ultrasound images, which included integrating a unique data augmentation approach using Generative Adversarial Networks (GAN) and Transfer Learning (TL) for feature extraction. Their research utilized the BUSI dataset for training and testing after pre-processing, specifically histogram equalization. The proposed framework achieved an accuracy of 99.6%, outperforming existing methods and demonstrating the efficacy of GAN and TL in enhancing classification accuracy.

Ashurov et al. [19] introduced a novel breast cancer histopathological image classification approach. It leverages modified pre-trained CNN models and attention mechanisms to enhance model interpretability and robustness, emphasizing localized features and enabling accurate discrimination of complex cases. Their research involved Transfer Learning with deep CNN models—Xception, VGG16, ResNet50, MobileNet, and DenseNet121—augmented with the convolutional block attention module (CBAM). The pre-trained models are fine-tuned, and the two CBAM models are incorporated at the end of the pre-trained models. The test accuracy rates for the attention mechanism (AM) using the Xception model on the “BreakHis” breast cancer dataset are encouraging at 99.2% and 99.5%. The test accuracy for DenseNet121 with AMs is 99.6%.

Reshan et al. [5] implemented an automated system for predicting breast cancer using various features and ensemble machine-learning methods to improve the accuracy of clinical detection. The study aimed to identify and classify two types of breast cancer by utilizing the Wisconsin Diagnostic Breast Cancer benchmark feature set, focusing on minimizing the number of features required to attain optimal diagnostic accuracy. The approach employs sophisticated ensemble machine learning methods for identifying and categorizing breast cancer. EML methods comprise voting, bagging, stacking, boosting, and recursive feature elimination. The suggested approach resulted in the stacking model achieving an exceptional average accuracy of 99.89%. Additionally, their model achieved high scores for sensitivity, specificity, F1-score, precision, and area under the curve (AUC/ROC), all at 1.00% or 99.9%.

Huang et al. [6] introduced a novel deep learning model, BM-Net, that integrates MobileNet-V3 with a bilinear structure to improve the analysis of Whole Slide Images (WSI) for breast cancer. The WSIs are divided into segments, and a streamlined BM-Net is employed to capture characteristics of breast cancer from these segments, ensuring a balance between precision and computational speed. Data augmentation methods such as random flipping, random rotation, random translation, random center-cropping, and color variation enhance dataset variability and size. This improves the accuracy and specificity of the network. The focal loss approach aims to tackle the imbalance between classes in the training data, primarily due to the more significant number of invasive carcinoma patches. The BM-Net model showed excellent efficacy in identifying breast cancer within whole slide images. The model attained an accuracy of 0.88 in classifying patches to distinguish normal, benign, in situ carcinoma, and invasive carcinoma tissue within WSIs.

Machine learning and deep learning methods have displayed significant promise in detecting breast cancer, especially in analyzing and categorizing ultrasound images. Numerous scientific research studies have been conducted on the classification and recognition of cancer tumors utilizing different models, although they have certain constraints. There is a scarcity of publicly accessible benchmark datasets for breast cancer, leading to limited studies in this area. Moreover, current literature exhibits several discernible gaps in breast cancer detection and classification using Transfer Learning. AI and machine learning models must be more effectively incorporated into clinical processes, with inadequate documentation of healthcare professionals' real-world applicability and adoption of such models. Advancing AI in medicine is crucial, and addressing these gaps can significantly enhance the efficiency and effectiveness of patient care.

3- Research Methodology

This section will explore the methodology for integrating Machine Learning using the Convolutional Neural Network model into early breast cancer diagnosis. Figure 1 presents a visual representation of the research, illustrating the sequential stages of the study in detail. Subsequent subsections will elaborate on each process step, providing a comprehensive understanding.

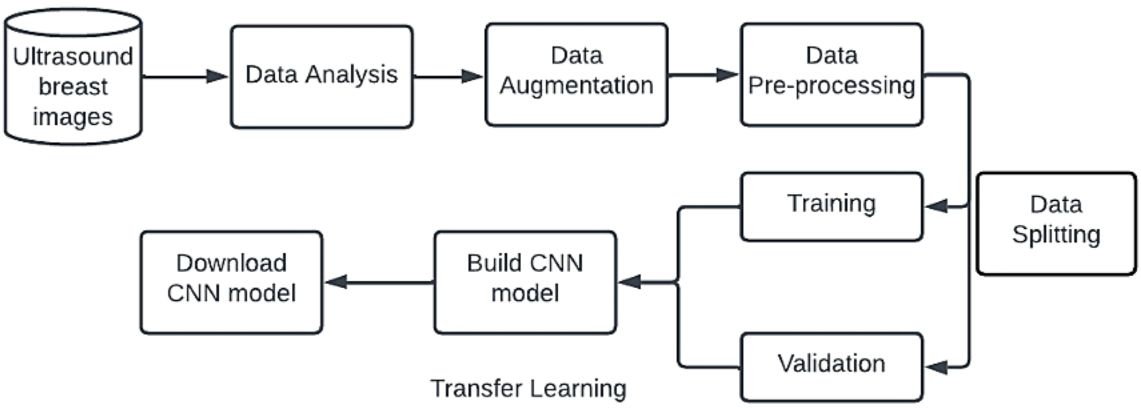


Figure 1. Image-based Classification Workflow

3-1- Dataset Description

The research used the publicly available Breast Ultrasound Image (BUSI) dataset [20], comprising a set of breast ultrasound images. The dataset contains ultrasound images taken from female individuals around the age range of 35 to 75 years. Ultrasound images of different types are typically utilized in medical imaging to evaluate breast tissue. The collection comprises 780 images in PNG format, with an average dimension of 500x500 pixels, offering a substantial dataset for analysis and training models. The dataset classifies images into three primary groups: normal, non-cancerous, and cancerous, which is essential for creating accurate classification models to identify healthy breast tissue or detect benign or malignant tumors. The dataset contains unprocessed ultrasound images as well as accompanying ground truth images. These supplementary images are annotated to indicate the existence and positioning of abnormalities or other characteristics pertinent to diagnosing breast cancer. The BUSI dataset frequently requires pre-processing procedures before utilization in training deep learning models. These procedures encompass resizing, normalization, and augmentation to ready the data for improved model performance.

3-2- Dataset Analysis and Augmentation

The analysis and enhancement of the BUSI dataset encompass a series of procedures aimed at readying the data for practical application in machine learning algorithms used for classifying breast cancer. The dataset is first filtered to include only the essential ultrasound and ground-truth images required for model training and testing. As depicted in Figure 2, the dataset analysis indicated an imbalance in classes, with the 'Normal' and 'Malignant' categories being less prevalent than the 'Benign' categories. This imbalance can harm the model's performance, leading to potential bias towards the more commonly occurring class. Data augmentation techniques were applied to address the class imbalance issue and enhance the dataset. These techniques artificially expand the dataset by creating modified versions of the existing images, including rotation, height shift, width shift, shear range, zoom range, horizontal flip, and fill mode [21].

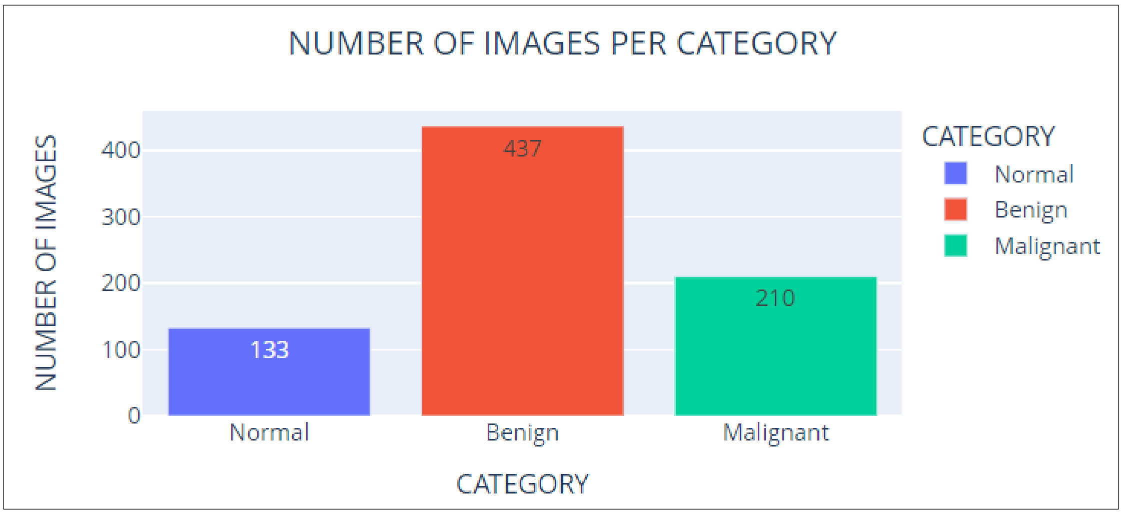


Figure 2. Class Distribution of Images

3-3- Data Pre-Processing

This section comprises image enhancement techniques that manipulate and transform the images before further processing and classification.

3-3-1- Image Resizing

Resizing pictures is crucial for standardizing their measurements before feeding them into the model. This standardization process also helps enhance the model's accuracy and optimize training time. Saponara et al. [22] further highlighted that image resizing is necessary due to the fixed-size input requirement of Deep learning models. Bilinear interpolation is applied for real-time image resizing and is suitable for both enlarging and reducing the size of an image. Bilinear interpolation entails assessing the surrounding 2x2 array of given pixel values near the computed location of the unidentified pixel. It then executes a linear estimation initially in one orientation and subsequently in the opposite direction to achieve the ultimate pixel value. The resizing process is intended to balance computational efficiency and retain sufficient detail for the model to classify information accurately. This is crucial for improving training time and boosting the model's effectiveness.

3-3-2- Image Denoising

During data preprocessing, denoising is a fundamental step for ensuring that images used for training machine learning models are of good quality and free from noise that could affect performance. Denoising eliminates undesired noise from images while retaining essential features, such as edges and textures, which are crucial for further analysis. The integration of denoising during the initial processing stage is especially critical for medical images, as the distortions resulting from noise may result in incorrect diagnoses and subsequent inappropriate interventions. Eventually, the Gaussian Filter removed the noise from images while preserving the larger-scale features.

3-3-3- Image Contrast Enhancement

Image contrast significantly impacts the quality of medical images, particularly in detecting calcifications in breast tissue. Improving the contrast of a breast medical image is crucial for facilitating accurate diagnosis and interpretation. Histogram Equalization was employed to enhance the image's contrast by redistributing its predominant intensity values. This method entails using the image's histogram, which visually represents the distribution of pixel intensities.

3-4- Classification

After processing the images, they were separated into two parts based on the test-size ratio. Different test size ratios (0.1 - 0.5) were used to split the dataset. The CNN model was constructed using the VGG-16 architecture with integrated Transfer Learning. After initializing the base model, a customized top layer was added, excluding the original top layer responsible for classifying images into ImageNet classes. The new fully connected layer in the model includes a dropout layer, a dense layer, and activation functions such as ReLU and softmax. To preserve pre-trained weights in the base model, only the weights of this new layer needed to be updated while keeping extraction layers frozen. Callback methods such as early stopping and model checkpoint were utilized to prevent overfitting and save the best validation performance. Finally, multiple training sessions were conducted with different test-size ratios, and the one that produced optimal results was selected. Figure 3 illustrates these steps in implementing the training of our model.

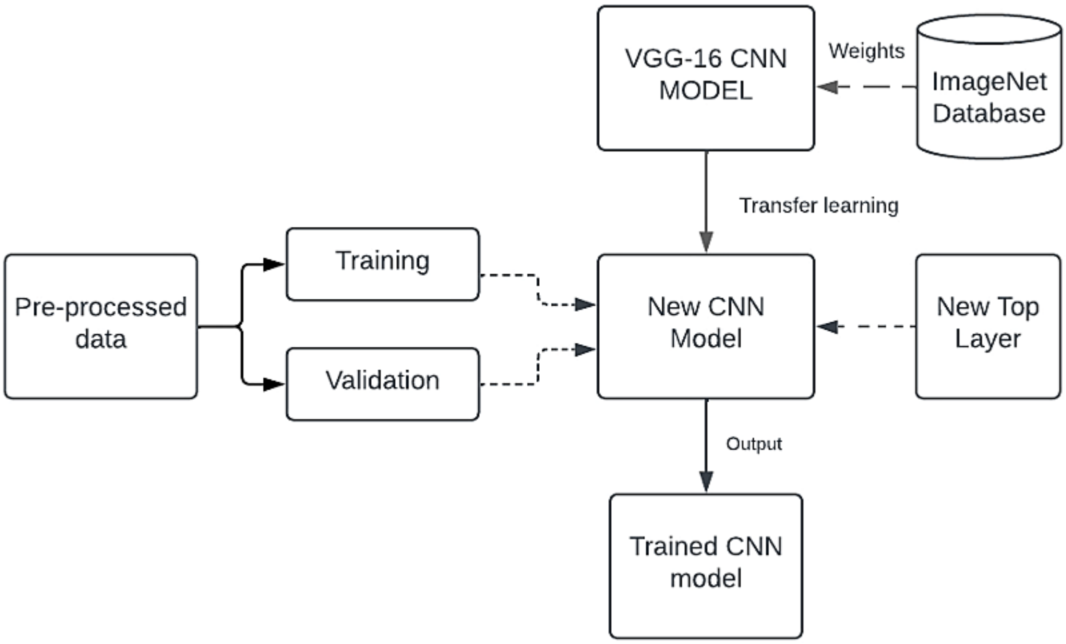


Figure 3. Transfer Learning Implementation Steps

Transfer Learning is employed to enhance the performance of CNN models in the medical domain [23]. The researchers suggested employing Transfer Learning to improve the efficacy of CNN models in the medical field, mainly because there is a lack of large datasets for training from the ground up. In the medical domain, acquiring a well-labeled dataset poses a considerable challenge for clinicians. Transfer Learning enables CNN models to acquire essential foundational features and patterns across various modalities or domains by utilizing pre-trained models on extensive datasets like ImageNet. This method allows the models to attain greater precision and generate forecasts based on medical data, even when working with constrained or small sets of labeled data. Data augmentation was employed in the study to address dataset imbalance and mitigate potential bias. Augmenting medical datasets (small) can present both benefits and challenges. Generating diverse variations from the same image through multiple augmentations does not necessarily enhance diversity. Consequently, the model might need to generalize novel data more effectively because it is being trained using a similar original image with various modifications. However, appropriate data augmentation techniques can help fully utilize the potential of CNN models and capture the full scope of image modality features [24].

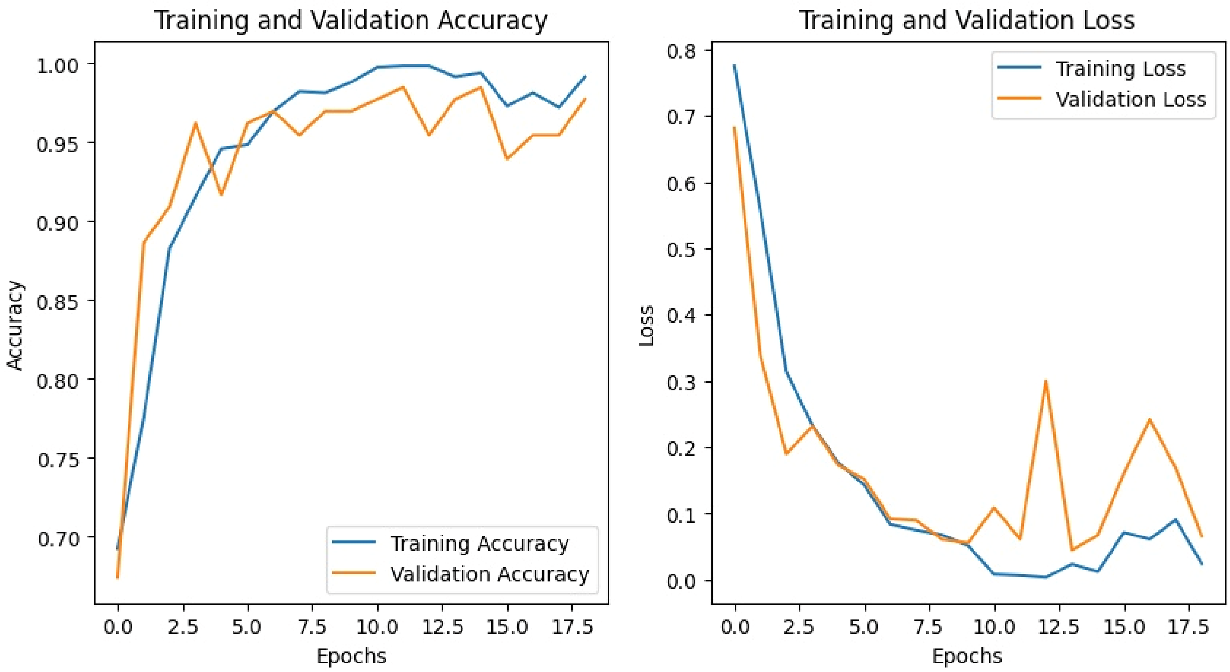
4- Results and Discussion

The results of this study demonstrate the effectiveness of using Transfer Learning-based classification models for breast cancer prediction. The models showed promising results in predicting breast cancer.

A proprietary CNN was developed without utilizing the VGG-16 architecture and implementing Transfer Learning. The outcomes are as follows: loss: 0.0119—Accuracy: 0.9992—Val_Loss: 0.3643—Val_Accuracy: 0.9318. The significant difference in precision between the training and validation sets could suggest a potential issue with overfitting. Overfitting can be particularly challenging when working with small datasets, as models may quickly memorize the limited information instead of capturing broader patterns. Transfer learning is an effective strategy to address this issue.

The training accuracy gradually improved with each epoch as the model became more familiar with the training data, as shown in Figure 4. The validation accuracy initially increased, indicating that the model has good generalization ability for the validation set. However, the validation accuracy decreased beyond a certain point, suggesting that the model must be more balanced with the training data and adapt more effectively to new information.

The test was conducted on the VGG-16 CNN using Transfer Learning to address the abovementioned difficulties. Several training sessions of different CNN models with varying test-size ratios were conducted, resulting in various outcomes, as detailed in Table 1. The selection of the train-test split ratio proved to be pivotal and had an impact on the overall performance of the system. The model's assessment relied on its accuracy and validation accuracy, where high values indicated superior quality. Other metrics, such as loss and validation loss, were also evaluated



.

Figure 4. CNN Model without Transfer Learning Results

Table 1. CNN with transfer learning results

Test-Size Ratio	Accuracy	Loss	Val_Acc	Val_Loss
0.1	0.9012	0.2641	0.9015	0.3115
0.2	0.8869	0.2876	0.8598	0.3382
0.3	0.8664	0.3366	0.8203	0.4757
0.4	0.8885	0.2695	0.9127	0.3133
0.5	0.9225	0.2216	0.8055	0.4941

Figure 5 presents a visual illustration to help readers better understand and interpret the findings. The optimal model was chosen between Model 1 and Model 5, both of which were trained using varying proportions of the dataset (90% and 50%, respectively).

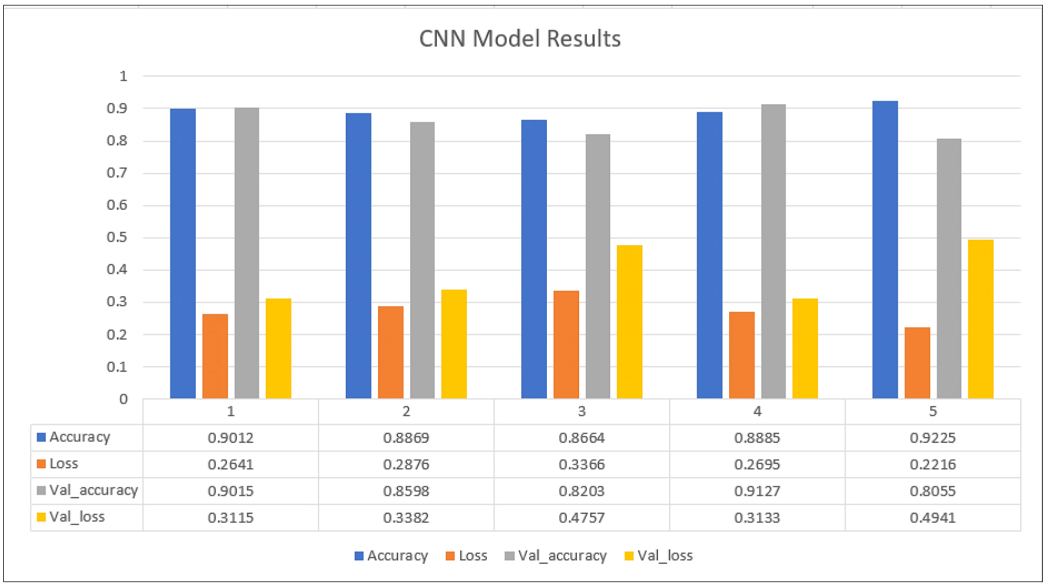


Figure 5. CNN Model with Transfer Learning Results

Figure 5 clearly shows that the best model options were narrowed down to either Model 1 with a test-size ratio of 0.1 or Model 5 with a test-size ratio of 0.5. According to the evaluation metrics, Model 5 attained higher overall accuracy than Model 1. However, it is essential to highlight that Model 1 outperformed Model 5 by approximately 10% in terms of validation accuracy. Additionally, the results for validation loss also indicate the superior performance of Model 1.

Considering these factors, model 1 was ultimately selected and acquired because of its superior ability to generalize and capture patterns in unfamiliar data. Learning curves were employed to further evaluate the model's performance, as depicted in Figure 6. This was done to confirm the convergence and stability of the model during its training process. A confusion matrix was generated (refer to Figure 6) to extract True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN).

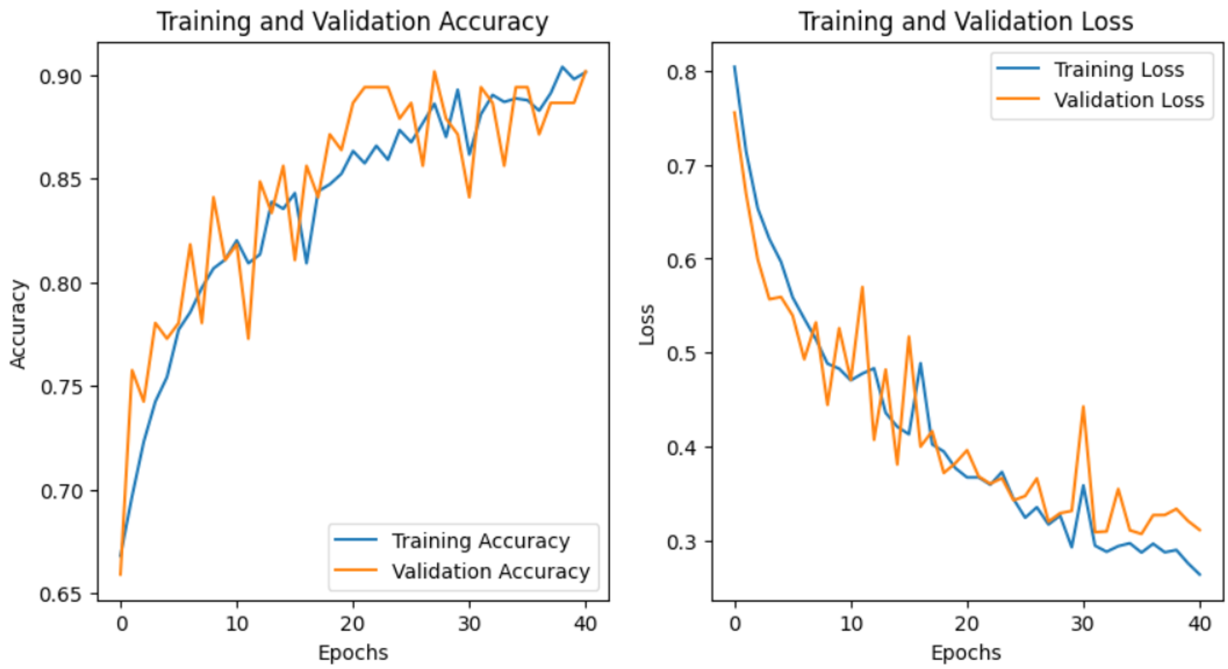


Figure 6. **Learning Curves for the CNN Model with Transfer Learning**

An increasing and consistently improving pattern in training and validation accuracy indicates that the model continuously enhances its ability to make precise forecasts on the data it was trained on and unseen validation data. As both training and validation loss continue to decrease, the model becomes more proficient at minimizing prediction errors, leading to an overall enhancement in performance. These trends collectively point towards a highly optimized and well-generalized machine learning model with robust predictive capabilities.

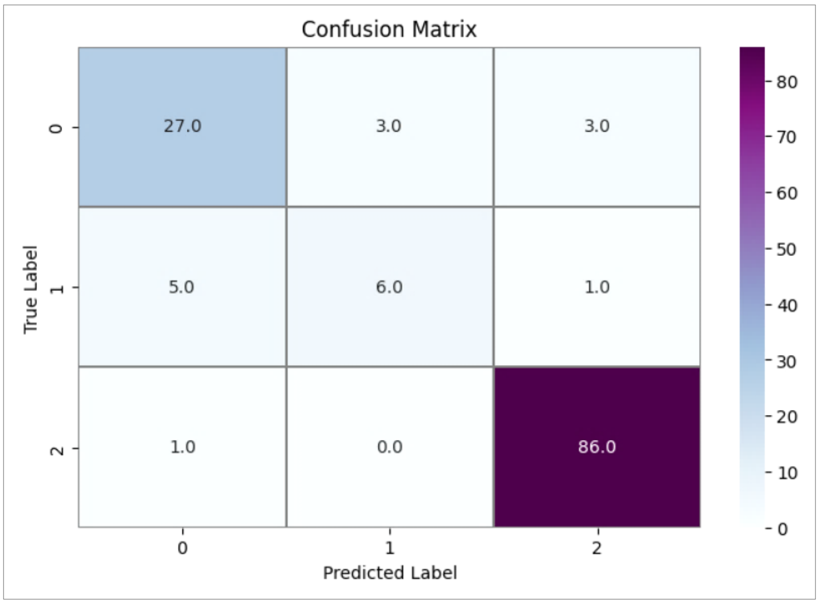


Figure 7. **Confusion Matrix**

From the confusion matrix in Figure 7, a classification report (Figure 8) can be generated outlining the model's precision, recall, and F1 score. Also, the accuracy for the classes 0 (Benign), 1 (Malignant), and 2 (Normal) is 0.9091, 0.9318, and 0.9621, respectively.

	precision	recall	f1-score	support
benign	0.82	0.82	0.82	33
Malignant	0.67	0.50	0.57	12
Normal	0.96	0.99	0.97	87
accuracy			0.90	132
macro avg	0.81	0.77	0.79	132
weighted avg	0.89	0.90	0.90	132

Figure 8. Classification Report

After examining the classification report, it can be inferred from the accuracy that a high level of precision in categorizing instances into their respective classes has been attained. The consistent precision, recall, and F1 results [25] for the ‘Benign’ class demonstrate that the model effectively identifies instances of this class. However, for the ‘Malignant’ class, although the precision indicates correct identification in a significant proportion of cases, there is potential for improvement. Furthermore, the recall and F1 scores imply that there may be difficulty in identifying all actual malignant instances. The ‘Normal’ class obtained high scores, signifying the accurate identification of regular cases. While achieving good results overall, further enhancements are needed to improve its discriminatory capabilities [26].

Once the model was deemed appropriate, it was downloaded for local use in the desktop application. When a user submits a raw breast ultrasound image for classification, it undergoes similar pre-processing techniques and is input into the model. As a result, the image is categorized as ‘Benign’, ‘Normal’, or ‘Malignant’. A simplified representation of the process can be seen in Figure 9.

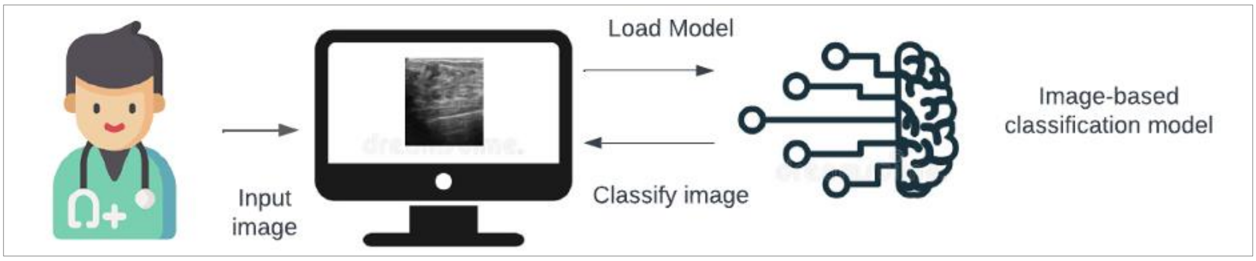


Figure 9. Classification Process

The desktop application was built using Python, and Figure 10 shows the application's user interface.

A comprehensive evaluation process was conducted after the model was integrated into the app. Before training, 15 images (5 from each class) were carefully chosen for assessment. These visual representations were utilized to gauge the model's effectiveness in predicting diagnoses. In addition to the model's predictions, expert diagnoses from a radiologist with over 25 years of experience were also obtained. The findings presented in Table 2 were compared to evaluate the research work's performance.

Table 2. Accuracy of Model concerning Radiologist

	No. of Correct Diagnosis	No. of incorrect diagnosis	Accuracy
System	13	2	86.7%
Radiologist	13	2	86.7%

From the comparative analysis, it can be deduced that the CNN model with Transfer Learning yielded promising results. Its performance highlights the model's potential as a reliable tool for assisting healthcare professionals in diagnosing breast cancer.

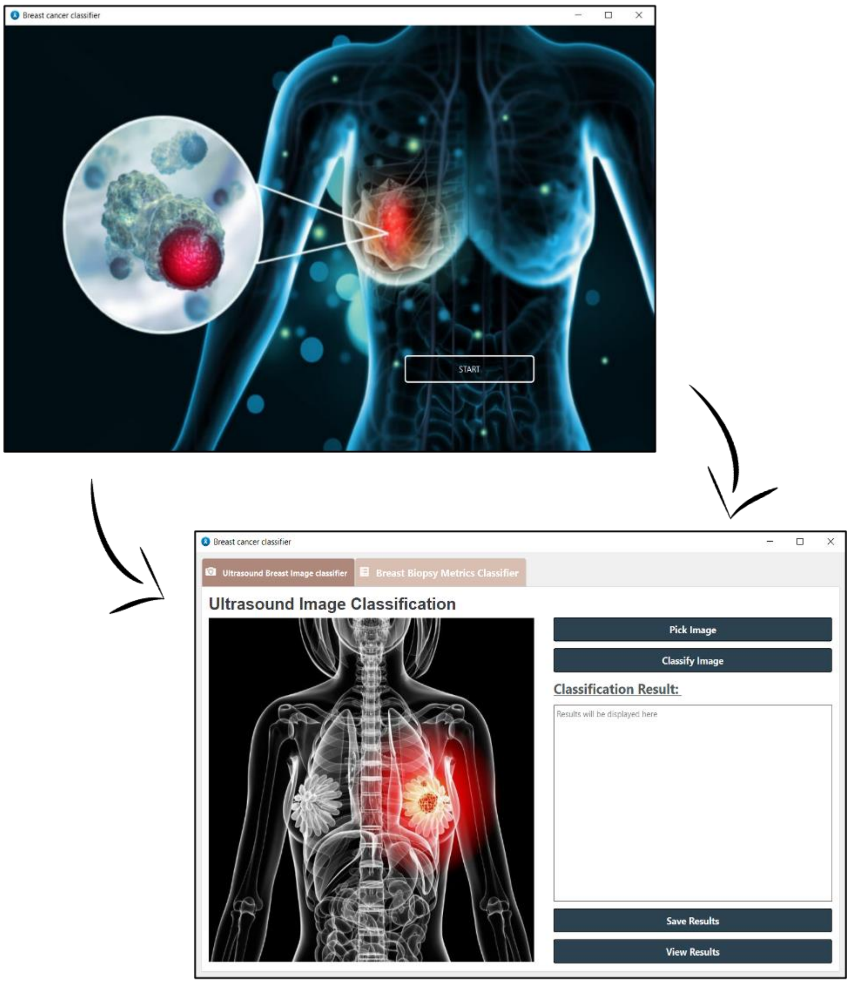


Figure 10. App UI

An example of the classification process in the desktop application is provided in Figure 11.

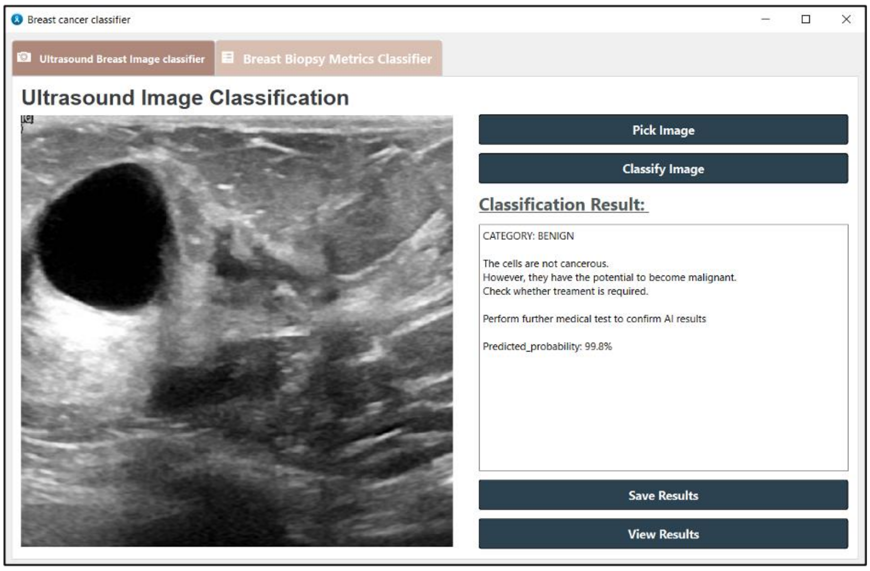


Figure 11. Example of image-based classification

As indicated in Figure 5, the best-performing model was selected after training various image classification models. This model achieved an accuracy of 90%. Comparing the results with existing related works, different research works obtained a range of accuracies. The lowest accuracy achieved when employing a machine learning algorithm was 62.3%, while the highest reached 99.8%. Therefore, it is evident that the obtained result falls within the higher end of the accuracy spectrum.

5- Conclusion

Breast Cancer has become a prevalent disease among women, underscoring the significance of early detection and classification for disease prevention. Integrating machine learning minimizes the cost and time of disease diagnosis, thus facilitating the prediction process. Similarly, in this research work, we aimed to construct a VGG-16 CNN model with Transfer Learning using the BUSI dataset from Kaggle. Before training, the dataset underwent essential preprocessing stages, including image resizing, denoising, and contrast enhancement. Various test-size ratios were employed during training, and the most successful configuration yielded an accuracy, loss, validation accuracy, and validation loss of 0.9012, 0.2641, 0.9015, and 0.3115, respectively. The trained model was then locally downloaded for integration into our desktop app. These results are indeed promising, but our approach has potential for improvement. It is also important to note that even the classification process performed by doctors is only sometimes 100% accurate. Combining doctors' expertise with the application's automated capabilities can enhance the diagnosis, increasing time efficiency and reducing subjectivity during the classification process. In future research, additional data types such as mammography, MRI scans, or patient medical history shall be combined with ultrasound images through multi-modal CNN architectures to enhance diagnostic accuracy.

6- Declarations

6-1- Author Contributions

Conceptualization, S.A., K.M., D.A.D., T.B.K., and J.K.; methodology, S.A. and K.M.; software, K.M., S.A., and D.A.D.; validation, S.A., K.M., and J.K.; formal analysis, T.B.K., D.A.D., and K.M.; investigation, D.A.D., T.B.K., and J.K.; resources, S.A.; data curation, S.A., K.M., T.B.K., and J.K.; writing—original draft preparation, K.M., T.B.K., and J.K.; writing—review and editing, S.A., T.B.K. and D.A.D.; visualization, K.M., T.B.K. and J.K.; supervision, S.A. and D.A.D.; project administration, S.A., D.A.D., T.B.K., and J.K.; funding acquisition, D.A.D. All authors have read and agreed to the published version of the manuscript.

6-2- Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6-3- Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

6-4- Institutional Review Board Statement

Not applicable.

6-5- Informed Consent Statement

Not applicable.

6-6- Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

7- References

[1] W.H.O. (2023). Global Breast Cancer Initiative Implementation Framework: Assessing, Strengthening and Scaling up of Services for the Early Detection and Management of Breast Cancer. World Health Organization (WHO), Genève, Switzerland. Available online: https://www.who.int/publications/i/item/9789240065987 (accessed on November 2024).

[2] MNCR. (2024). Cancer in the Republic of Mauritius Incidence and Mortality Study for 2022 Report of the National Cancer Registry. Mauritius National Cancer Registry, Port Louis, Mauritius. Available online: https://health.govmu.org/health/wp-content/uploads/2024/02/MNCR-2022-REPORT-final-with-cover-updated-3.pdf (accessed on November 2024).

[3] Liu, L., Feng, W., Chen, C., Liu, M., Qu, Y., & Yang, J. (2022). Classification of breast cancer histology images using MSMV-PFENet. Scientific Reports, 12(1), 17447. doi:10.1038/s41598-022-22358-y.

[4] Yan, F., Huang, H., Pedrycz, W., & Hirota, K. (2023). Automated breast cancer detection in mammography using ensemble classifier and feature weighting algorithms. Expert Systems with Applications, 227. doi:10.1016/j.eswa.2023.120282.

[5] Reshan, M. S. Al, Amin, S., Zeb, M. A., Sulaiman, A., Alshahrani, H., Azar, A. T., & Shaikh, A. (2023). Enhancing Breast Cancer Detection and Classification Using Advanced Multi-Model Features and Ensemble Machine Learning Techniques. Life, 13(10), 2093. doi:10.3390/life13102093.

[6] Huang, J., Mei, L., Long, M., Liu, Y., Sun, W., Li, X., Shen, H., Zhou, F., Ruan, X., Wang, D., Wang, S., Hu, T., & Lei, C. (2022). BM-Net: CNN-Based MobileNet-V3 and Bilinear Structure for Breast Cancer Detection in Whole Slide Images. Bioengineering, 9(6), 261. doi:10.3390/bioengineering9060261.

[7] Samee, N. A., Alhussan, A. A., Ghoneim, V. F., Atteia, G., Alkanhel, R., Al-Antari, M. A., & Kadah, Y. M. (2022). A Hybrid Deep Transfer Learning of CNN-Based LR-PCA for Breast Lesion Diagnosis via Medical Breast Mammograms. Sensors, 22(13), 4938. doi:10.3390/s22134938.

[8] Law, D., Najm, A. A., Chong, J. X., K'ng, J. Z. Y., Amran, M., Ching, H. L., Wong, R. R., Leong, M. H., Mahdi, I. M., & Fazry, S. (2023). In silico identification and in vitro assessment of a potential anti-breast cancer activity of antimicrobial peptide retrieved from the ATMP1 Anabas testudineus fish peptide. PeerJ, 11, 15651. doi:10.7717/peerj.15651.

[9] Chaudhury, S., & Sau, K. (2023). Classification of Breast Masses Using Ultrasound Images by Approaching GAN, Transfer Learning, and Deep Learning Techniques. Journal of Artificial Intelligence and Technology, 3(4), 142–153. doi:10.37965/jait.2023.0175.

[10] Pavithra, S., Vanithamani, R., & Justin, J. (2020). Computer aided breast cancer detection using ultrasound images. Materials Today: Proceedings, 33, 4802–4807. doi:10.1016/j.matpr.2020.08.381.

[11] Hijab, A., Rushdi, M. A., Gomaa, M. M., & Eldeib, A. (2019). Breast Cancer Classification in Ultrasound Images using Transfer Learning. International Conference on Advances in Biomedical Engineering, ICABME, 2019-October, 1–4. doi:10.1109/ICABME47164.2019.8940291.

[12] Uysal, F., & Köse, M. M. (2022). Classification of Breast Cancer Ultrasound Images with Deep Learning-Based Models †. Engineering Proceedings, 31(1), 8. doi:10.3390/ASEC2022-13791.

[13] Arooj, S., Atta-ur-Rahman, Zubair, M., Khan, M. F., Alissa, K., Khan, M. A., & Mosavi, A. (2022). Breast Cancer Detection and Classification Empowered With Transfer Learning. Frontiers in Public Health, 10. doi:10.3389/fpubh.2022.924432.

[14] Zhuang, Z., Kang, Y., Joseph Raj, A. N., Yuan, Y., Ding, W., & Qiu, S. (2020). Breast ultrasound lesion classification based on image decomposition and transfer learning. Medical Physics, 47(12), 6257–6269. doi:10.1002/mp.14510.

[15] Zeimarani, B., Costa, M. G. F., Nurani, N. Z., Bianco, S. R., De Albuquerque Pereira, W. C., & Filho, C. F. F. C. (2020). Breast Lesion Classification in Ultrasound Images Using Deep Convolutional Neural Network. IEEE Access, 8, 133349–133359. doi:10.1109/ACCESS.2020.3010863.

[16] Hajipour Khire Masjidi, B., Bahmani, S., Sharifi, F., Peivandi, M., Khosravani, M., & Hussein Mohammed, A. (2022). CT-ML: Diagnosis of Breast Cancer Based on Ultrasound Images and Time-Dependent Feature Extraction Methods Using Contourlet Transformation and Machine Learning. Computational Intelligence and Neuroscience, 2022, 1–15. doi:10.1155/2022/1493847.

[17] Raza, A., Ullah, N., Khan, J. A., Assam, M., Guzzo, A., & Aljuaid, H. (2023). DeepBreastCancerNet: A Novel Deep Learning Model for Breast Cancer Detection Using Ultrasound Images. Applied Sciences (Switzerland), 13(4). doi:10.3390/app13042082.

[18] Dan, Q., Zheng, T., Liu, L., Sun, D., & Chen, Y. (2023). Ultrasound for Breast Cancer Screening in Resource-Limited Settings: Current Practice and Future Directions. Cancers, 15(7), 2112. doi:10.3390/cancers15072112.

[19] Ashurov, A., Chelloug, S. A., Tselykh, A., Muthanna, M. S. A., Muthanna, A., & Al-Gaashani, M. S. A. M. (2023). Improved Breast Cancer Classification through Combining Transfer Learning and Attention Mechanism. Life, 13(9), 1945. doi:10.3390/life13091945.

[20] Al-Dhabyani, W., Gomaa, M., Khaled, H., & Fahmy, A. (2020). Dataset of breast ultrasound images. Data in Brief, 28, 104863. doi:10.1016/j.dib.2019.104863.

[21] Adjobo, E. C., Mahama, A. T. S., Gouton, P., & Tossa, J. (2022). Towards Accurate Skin Lesion Classification across All Skin Categories Using a PCNN Fusion- Based Data Augmentation Approach. Computers, 11(3), 44–44. doi:10.3390/computers11030044.

[22] Saponara, S., & Elhanashi, A. (2022). Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance. Lecture Notes in Electrical Engineering, 866 LNEE, 10–17. doi:10.1007/978-3-030-95498-7_2.

[23] Cheng, P. M., & Malhi, H. S. (2017). Transfer Learning with Convolutional Neural Networks for Classification of Abdominal Ultrasound Images. Journal of Digital Imaging, 30(2), 234–243. doi:10.1007/s10278-016-9929-2.

[24] Yu, Y., Lin, H., Meng, J., Wei, X., Guo, H., & Zhao, Z. (2017). Deep transfer learning for modality classification of medical images. Information (Switzerland), 8(3), 91. doi:10.3390/info8030091.

[25] Manna, A., Kundu, R., Kaplun, D., Sinitca, A., & Sarkar, R. (2021). A fuzzy rank-based ensemble of CNN models for classification of cervical cytology. Scientific Reports, 11(1). doi:10.1038/s41598-021-93783-8.

[26] Wu, N., Phang, J., Park, J., Shen, Y., Huang, Z., Zorin, M., Jastrzebski, S., Fevry, T., Katsnelson, J., Kim, E., Wolfson, S., Parikh, U., Gaddam, S., Lin, L. L. Y., Ho, K., Weinstein, J. D., Reig, B., Gao, Y., Toth, H., … Geras, K. J. (2020). Deep Neural Networks Improve Radiologists' Performance in Breast Cancer Screening. IEEE Transactions on Medical Imaging, 39(4), 1184–1194. doi:10.1109/TMI.2019.2945514.