**April 18, 2025**

**EXPLORING DEEP LEARNING AI ULTRASOUND AS A PRIMARY BREAST CANCER SCREENING TEST: AN INFORMATICS APPROACH TO CANCER PREVENTION**

Frederick Damptey & Benjamin Odoom Asomaning

**Abstract**

**Introduction:**  
Breast cancer remains a leading cause of cancer-related mortality among women globally. Early and accurate detection is essential for effective treatment and improved patient outcomes. While ultrasound imaging is widely used due to its accessibility and safety, interpretation variability and dependence on specialist expertise pose challenges. This study explores the potential use of deep learning ultrasound scan (USS) AI as a primary breast cancer tool in resource-constrained areas.

**Methods:**  
This study employed convolutional neural networks (CNNs) with transfer learning to classify breast ultrasound images into normal, benign, and malignant categories. Among the models evaluated, EfficientNetB0, pre-trained on the ImageNet dataset, achieved the best performance. The dataset underwent preprocessing including image resizing, normalization, and augmentation. The model was fine-tuned on the training set and evaluated on a separate test set using accuracy, AUC, sensitivity, and specificity metrics.

**Results:**  
EfficientNetB0 achieved the highest classification performance with an accuracy of 95%, an AUC of 0.99, sensitivity of 93%, and specificity of 97.0%. The model showed strong ability to distinguish between malignant and non-malignant lesions and outperformed findings from most meta-analytic studies.

**Conclusion:**  
The results validate that USS deep learning models in general and our model in particular provide very efficient breast cancer detection with ultrasound images. Deep-learning USS models, either standalone or in conjunction with human expertise, hold promising potential to be adopted as a first-line screening tool for breast cancer, especially in regions with limited access to radiological experience. Further studies and clinical validation are needed to enable their use in actual healthcare settings.

**INTRODUCTION**

Breast cancer occurs when malignant cells grow within the breast tissue. Though it affects both sexes, it predominantly impacts women, with men accounting for only 0.5 - 1% of cases. The global burden of breast cancer is immense. As of 2022, the World Health Organization (WHO) estimated that 1 in 12 women in countries with high Human Development Index (HDI) and 1 in 27 women in low-HDI countries will develop breast cancer in their lifetime. However, mortality is disproportionately higher in low-HDI countries (2.1%) compared to high-HDI countries (1.4%) (WHO, 2022).

In the United States, the American Cancer Society (2024) reports that 1 in 8 women (13%) are at lifetime risk of breast cancer. It is the second most common type of cancer among American women and remains the leading cause of cancer-related deaths in women. Risk factors include age, late parity, obesity and family history, among others.

Early detection of breast cancer is known to significantly reduce mortality by allowing timely intervention and reducing the chances of malignancy because treatment interventions such as surgical excision, mastectomy and radiotherapy are more effective when cancer is detected early. The American Cancer Society (2024) emphasizes that when detected at the localized stage, breast cancer has a 99% 5-year survival rate.

**Screening and Diagnostic Modalities**

Standard screening methods for breast cancer include clinical breast exams and mammography. Mammography remains the gold standard for breast cancer screening, especially in high-income countries (HICs). It is widely accessible and affordable in these settings, supported by trained radiologists who interpret results (Miglioretti et al., 2015). However, in many low- and middle-income countries (LMICs), mammography remains out of reach for most women due to cost, availability, and infrastructure. A study in Ghana, a resource-constrained country which found that only 21 of 328 surveyed hospitals provided mammography services, and just one performed over 100 exams per month (Mathew et al., 2024) illustrates this accessibility challenge.

MRI and ultrasound are additional diagnostic tools. But while MRI is typically used in high-risk populations and quite expensive, ultrasound scan (USS) is more commonly employed in LMICs due to its affordability, portability, absence of ionizing radiation (Mirabito, 2023) and high diagnostic accuracy in dense breasts. Moreover, USS does not always require radiologists, as sonographers can perform breast ultrasounds. Despite these advantages, ultrasound's effectiveness is limited by operator dependency, inter-observer variability and inconsistent imaging quality.

**Emerging Technologies and the Role of AI**

Recent advancements in ultrasound technology have significantly improved its diagnostic performance (Rana et al, 2024). Among them, AI breast ultrasound is perhaps the most promising for detecting and classifying breast lesions. Several studies support the potential of AI to match or even outperform radiologists in diagnostic accuracy. For instance, a model had an accuracy of 0.82, compared to 0.72 for radiologists; sensitivity was 0.81 versus 0.79, and specificity was 0.83 versus 0.67 (Zhou et al., 2023). Furthermore, in evaluating dense breasts, ultrasound - especially when AI assisted - outperformed mammography, with combined methods reaching a sensitivity of 98.7% and specificity of 84.5% (Zhang et al., 2023).

**Comparative Effectiveness: Ultrasound vs. Mammography**

In a Chinese prospective trial, ultrasound demonstrated significantly higher sensitivity than mammography (95.7% vs. 78.9%) and was more accurate in breast cancer detection (76.8% vs. 71.3%) (Wang et al., 2022). Although mammography remains the only screening tool proven to reduce mortality, as shown in the Swedish Two-County trial (27–31% mortality reduction) (Tabár et al., 2011), ultrasound is proving to be a viable adjunct or alternative in settings where mammography is less feasible (Dan et al., 2023). Many systematic literature reviews using metanalysis have found that USS DL AI model performance evaluation are either comparable to or better than digital mammography AI models and usually better than human readers (Hanis eta al , 2022), Xue et al (2022), Li eta al (2024) and Hickman et al (2021).

**Performance Benchmarks and AI Validation**

A study by Miglioretti et al (2015) to update mammography performance benchmarks for radiologists established that the minimum threshold sensitivity was ≥80% while specificity was ≥85%. However, only 51% of the radiologists involved in the study met the sensitivity benchmark with 62% meeting the specificity requirement. Various meta-analyses have shown that standalone AI systems can meet or exceed these benchmarks. Yoon et al. (2023) found AI sensitivity of 81% vs. 74% for radiologists and specificity at 86% vs. 90% for mammography in breast cancer screening. Similarly, Rodriguez-Ruiz et al. (2019) found AI achieved an AUC of 0.84 compared to the average of 101 radiologists (0.814), demonstrating its potential in screening workflows. These findings were confirmed by Hickman et al (2021).

Meta-analyses of the performance of USS AI have also shown enhanced performance to human readers and mammography AI. On the premise that mammography AI is generally superior to human readers, this presents a promising potential for the early detection of breast cancer by screening, particularly in LMICs. In Li et al (2024) meta-analysis of USS AI, sensitivity of 93, specificity of 0.90 and AUC of 0.732 were attained. Hijab et al (2019) study on deep learning models also reported an AUC of 0.97 and accuracy of 0.95 for their CNN model and respective pooled values of 0.98 and 0.97 for meta-analysis of USS AI diagnostic models.

By overcoming the problem associated with accessing mammography services, and addressing the issues of operator variability, and diagnostic accuracy, AI-enhanced ultrasound could serve as a viable primary screening tool. However, despite the growing body of evidence, breast ultrasound, either AI-assisted or standalone, is yet to be formally endorsed as an alternative to mammography in global breast cancer screening guidelines. Future research and policy development must address this gap.

**Research question:** Given the ongoing advancements, can deep learning USS AI improve breast cancer detection? And can USS AI be adopted as an alternative breast cancer screening test, especially in resource-constrained environments?

**Problem statement:** Breast cancer remains a leading cause of cancer-related mortality among women worldwide, especially in resource-constrained regions where access to standard mammography screening is limited due to financial, infrastructural and operational barriers. USS, though cost-effective and widely available, is understudied and underutilized due to concerns over diagnostic efficacy, hampering its use as a standard screening tool and limiting its impact on the global fight against breast cancer.

**Aim:** The aim of this project is to develop an effective CNN deep learning model and determine whether deep learning AI can enhance the diagnostic performance of ultrasound for breast cancer detection and justify its use as an alternative to mammography in screening.

**METHODS**

This project will employ standard methods of deep learning modeling and performance evaluation. The results will be compared with previous studies to achieve meaningful conclusions. The procedure for CNN-based deep learning modeling is as follows:

**Dataset Acquisition:** The dataset named, Breast Ultrasound Images Dataset (Dataset\_BUSI\_with\_GT) used in this project was sourced from Kaggle, a publicly available open data platform. It comprises a total of 780 USS images, grouped in 3 folders: normal, benign and malignant according to their class. Each diagnostic label was assigned by a radiologist, following a careful comparison with corresponding mammograms and confirmed by histopathology diagnosis. A semi-automated segmentation was done through collaboration between an IT professional and radiologists, ensuring high-quality ground truth labels for model training.

**Preprocessing:** The model was developed using Google Colab Python 3.x. Initial data exploration was performed to understand dataset characteristics, followed by a robust preprocessing pipeline which included image resizing, data splitting into training and test sets and normalization. The training set size was augmented and subjected to K-Fold cross-validation to promote robustness and mitigate overfitting. This comprehensive training strategy ensured strong model generalization, which was then evaluated on the untouched test set.

**Model Implementation:** This project employed deep learning to build a diagnostic classification model for breast cancer detection from ultrasound images. Four pretrained CNN-based models - ResNet50, EfficientNet, InceptionV3, and VGG16 - were chosen to leverage transfer learning capabilities. These models were trained on the large-scale ImageNet dataset (Deng et al, 2009) and have demonstrated strong performance in image classification, object detection, and segmentation tasks, including medical imaging (Shen et al., 2017).

To address class imbalance, the dataset was split using stratified K-Fold Cross-Validation, ensuring that each fold preserved the distribution of classes for robust model training and validation. This technique has been shown to be effective in handling imbalanced datasets. Each pretrained model was fine-tuned using identical hyperparameters for a fair and consistent comparison.

**Evaluation:** The metrics used for evaluation are accuracy, AUC, sensitivity and specificity which are the recommended metrics for clinical multi-class machine learning models. Confusion matrices and AUC\_ROC curves were also plotted to visualize and gain insights into models’ predictions. These metrics are recommended for supervised learning tasks. The obtained values were evaluated against findings from other USS AI, mammography AI and human readers studies.

**RESULTS**

Among the four fine-tuned pretrained models, EfficientNetB0 had the best performance with an accuracy of 95%, AUC of 0.99, sensitivity of 93.3% and specificity of 97%. All the other models performed quite well with minimum accuracy, AUC, sensitivity and specificity values of 93.3%, 0.964, 0.881 and 0.945. Considering their performances, ResNet50 was the least performing.

The results for the performance evaluation are displayed below:

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  |  |  | | --- | --- | --- | --- | --- | | **Performance Of The 4 Different CNN Models** | | | | | |  | **Test Accuracy** | **AUROC** | **Sensitivity** | **Specificity** | | **ResNet50** | 0.92 | 0.964 | 0.881 | 0.945 | | **EfficientNetB0** | 0.95 | 0.99 | 0.933 | 0.97 | | **InceptionV3** | 0.941 | 0.982 | 0.934 | 0.962 | | **VGG16** | 0.933 | 0.982 | 0.909 | 0.951 | |  |

A screenshot of a graph

AI-generated content may be incorrect.

**DISCUSSION**

This entails a discussion of our best performing finetuned model, EfficientNetB0 against findings from other USS AI, MM AI and human readers studies. A emphasis on the implications of the deep learning model for screening in resource-constrained settings will be discussed as well.

**Model Architecture and Rationale**

One of the primary objectives of this project was to investigate whether an AI-driven system could enhance the diagnostic accuracy of breast cancer detection using ultrasound imaging. A key motivation was the potential of such a solution to serve as a primary screening tool in LMICs where access to advanced radiological infrastructure is limited by employing Convolutional Neural Network (CNN) model for the multi-class disease classification task. Each model was tuned using identical hyperparameters to ensure a fair comparison. All parameters for pixel size, normalization technique, data split, data augmentation (rotation, width and height shifts, horizontal flip, zoom, brightness and shear), epochs (50), K-Fold Cross-Validation (K = 5) were maintained for consistency. L1 regularization value, dropout and learning rate were also constant. All hidden layers were frozen except the last two to maximize learning from pretrained models. Two dense layers with ReLU and Softmax activations were created for the final multi-class output.

**Model Selection and Evaluation**

To achieve high diagnostic performance, CNN-based pre-trained models were evaluated:

* ResNet50
* EfficientNet
* InceptionV3
* VGG16

These models, pre-trained on over 14 million images from ImageNet (including medical images), have demonstrated outstanding capabilities in image classification, object detection, and segmentation tasks.

Among them, EfficientNet demonstrated the most consistent and superior performance across all key evaluation metrics - accuracy, sensitivity, specificity, and AUROC - making it the optimal model for this task. It also exhibited smoother and more stable learning curves, reinforcing its reliability in clinical applications. Python libraries used included TensorFlow, Scikit-learn, OpenCV (CV2), Seaborn, Matplotlib, Pandas, NumPy, OS, etc. were utilized.

The final model achieved:

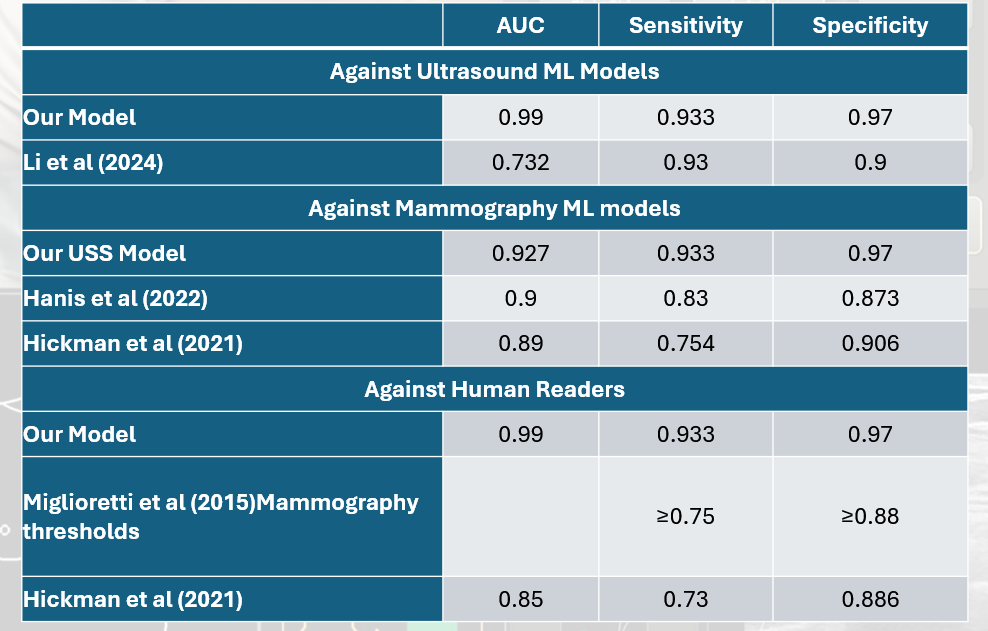
* Predictive Accuracy: 95%
* Sensitivity (Recall): 93%
* Specificity: 97%
* AUC: 0.99

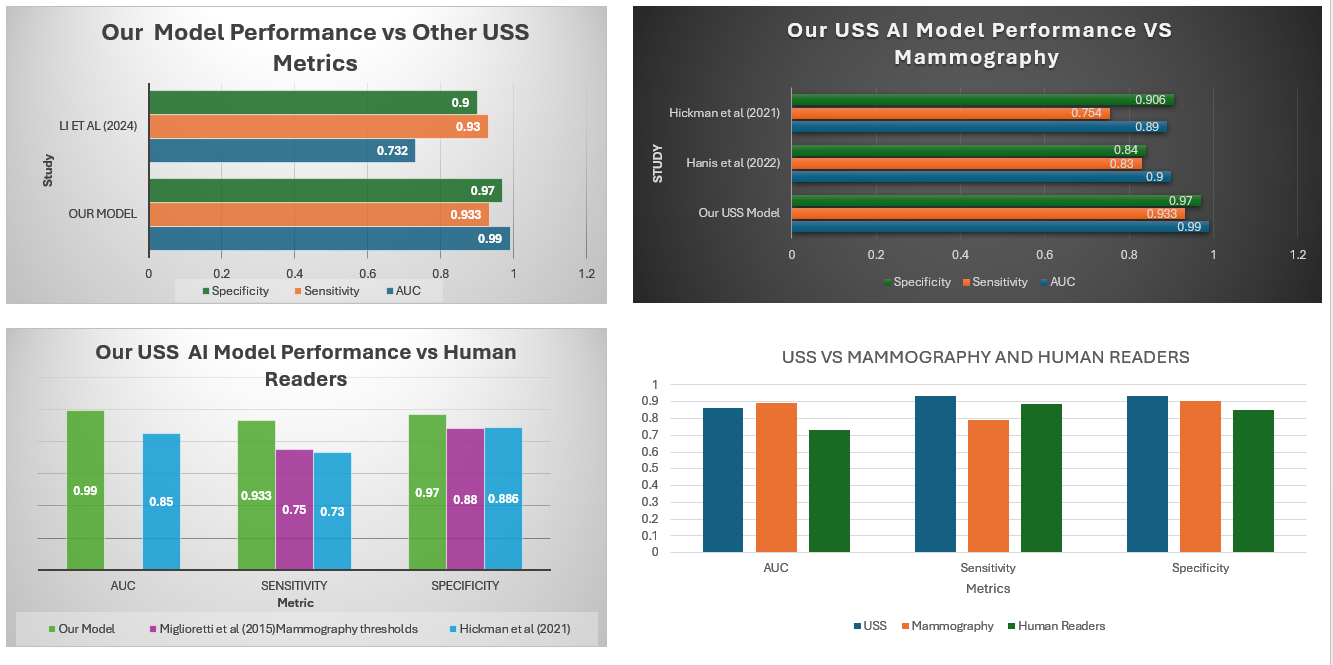
These results exceed benchmark radiologist performance (80% sensitivity, 85% specificity, Miglioretti et al., 2015) and are above the suggested minimum thresholds for AI-based mammography screening (81% sensitivity, 90% specificity, Yoon et al., 2023). The model also outperformed meta-analytic findings for USS AI systems, such as Li et al. (2024), with an AUC of 0.732, sensitivity of 0.93, and specificity of 0.90, and Hijab et al. (2019), with respective accuracy and AUC values of 0.973 and 0.97.

Additionally, the performance metrics of our CNN model outperform the results of Hanis et al. (2022) and Hickman et al. (2021). Both of these studies had lower AUC (0.9 and 0.89, respectively) and lower sensitivity (0.83 and 0.754) and specificity (0.873 and 0.906). Compared to the average human reader performance (AUC: 0.85, sensitivity: 0.73, specificity: 0.886), our model has a clear advantage. In addition, when AUC values were averaged over imaging modalities and compared, our ultrasound-based AI approach performed as well as mammography and outperformed human radiologists.

The higher performance metrics of our model, including its accuracy (95%), sensitivity (93%), specificity (97%), and AUR (0.99), attest to its reliability for clinical application. This demonstrates the potential of deep learning models in the ultrasound-based detection of breast cancer, and it presents a strong case for the incorporation of such models into diagnostic practices, particularly in regions where access to advanced radiological tools is limited.

A display of the comparative performance of our model against other studies is found below.





**Interpretation of Results**

Another notable finding is the model's strong performance on positive (malignant) cases, a pattern consistent with the existing literature wherein sensitivity performs better than specificity in detecting breast cancer. In our research, though, specificity surpasses sensitivity (97% vs. 93%), illustrating the model is particularly effective at correctly identifying non-cancer cases, and thus, reducing false positives. This finding is significant in that a test with high specificity for cancer screening will decrease patient anxiety unnecessarily and minimize the emotional and psychological burden of false cancer diagnoses.

In comparison, Wang et al. (2022), in a China-based study, found that single ultrasound-based AI was more sensitive to the detection of breast cancer than mammography (95.7% vs. 78.9%) but less specific (42.9% vs. 62.3%). Our results have a different and positive narrative to share, illustrating that AI-based ultrasound can achieve both high sensitivity and specificity when properly trained and fine-tuned.

**Population Health Implications**

From a population health perspective, the implications of these findings are profound. Breast cancer remains one of the leading causes of cancer-related deaths among women worldwide, with disparities in early detection contributing to poorer outcomes in low-resource or rural settings. An AI model that demonstrates high specificity and sensitivity offers a scalable, low-cost screening solution, particularly in areas where access to mammography is limited or unavailable. Reducing false positives at a population level can also prevent overburdening already constrained healthcare systems with unnecessary follow-up procedures and biopsies, freeing up resources for patients who truly need further diagnostic workup. Moreover, fewer false alarms help preserve public trust in screening programs, potentially increasing participation rates and improving early cancer detection and survival outcomes across populations.

**CONCLUSION**

Many studies that have reported that stand-alone AI ultrasound and AI-supported ultrasound are highly effective in the detection of breast cancer. Despite all these promising results, ultrasound remains largely an adjunct to mammography. There have frequently been concerns about the diagnostic accuracy, ethical concerns, and biases of AI algorithms. But with the rapid advancements in medical imaging and artificial intelligence technologies, all of these problems are now being addressed by innovation, transparency and robust evaluation frameworks.

From the public informatics and population perspective, it is an inevitability. In the scarce availability of mammography equipment in resource-poor settings, such as LMICs, delayed diagnosis and elevated mortality due to breast cancer are consequences of limited utilization of mammography services.

Our study offers evidence for the roll-out of AI-enhanced ultrasound as a first-line screening tool for breast cancer, especially where traditional screening infrastructure is lacking. With appropriate validation, such technology can enable early detection, reduce mortality and improve health equity. This would ultimately enhance survivors' quality of life, decrease healthcare spending and contribute significantly to public health goals like reducing non-communicable disease burdens.

Although it is stated, broader rollout awaits more validation tests on larger, heterogeneous sets, and addressing regulatory, ethical, and infrastructural preparedness. These are necessary in order to facilitate safe, effective, and equitable application of AI tools within national health systems.

**LIMITATIONS**

Several limitations were encountered in this study, which may affect the generalizability and scalability of the results:

1. Sample size constraints: One of the primary challenges was acquiring large enough labelled dataset. While the model demonstrated strong performance metrics, a larger and more diverse dataset is essential to support external validation and broader generalization across populations, imaging devices and clinical settings.
2. Computational resource demands: Training our CNN deep learning models was resource-intensive, requiring substantial computing power and memory. These limitations restricted the extent of hyperparameter optimization, cross-model experimentation, and retraining, thereby impacting the efficiency of the development cycle.
3. Time constraints hindered rigorous literature review.

**FUTURE WORK AND RECOMMENDATIONS**

1. We will further develop our model with a larger, diverse dataset to improve performance and allow generalization. We will also expand our model to include segmentation and incorporate explainable AI to make it more relevant for clinical applications.
2. A rigorous meta-analytic systematic review will be performed, especially on recent studies to reflect current advancements.
3. We recommend that there should be a careful assessment of global and national policies on breast cancer screening programs to inform possible adoption of USS AI as a primary screening tool.

**DATASOURCE**

Primary dataset - <https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset/data>

Secondary - <https://www.cancerimagingarchive.net/collection/breast-lesions-usg/#:~:text=This%20dataset%20consists%20of%20256,tumor%20classification%20and%20histopathological%20diagnosis>.

Link to imported of data: <https://www.cancerimagingarchive.net/collection/breast-lesions-usg/>

**REFERENCES**

American Cancer Society. (2025). Breast cancer. <https://www.cancer.org/cancer/types/breast-cancer.html>

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. <https://doi.org/10.3390/cancers15072112>

Hanis, T. M., Islam, M. A., & Musa, K. I. (2022). Diagnostic accuracy of machine learning models on mammography in breast cancer classification: A meta-analysis. *Diagnostics (Basel), 12*(7), 1643. <https://doi.org/10.3390/diagnostics12071643> (Link: PMC)

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770–778. <https://doi.org/10.1109/CVPR.2016.90>

Hickman, S. E., Woitek, R., Le, E. P. V., Im, Y. R., Luxhøj, C. M., Aviles-Rivero, A. I., Baxter, G. C., MacKay, J. W., & Gilbert, F. J. (2021). Machine learning for workflow applications in screening mammography: Systematic review and meta-analysis. *Radiology*. <https://doi.org/10.1148/radiol.2021210391> (Link: RSNA)

Hijab, A., Rushdi, M. A., Gomaa, M. M., & Eldeib, A. (2019). Breast cancer classification in ultrasound images using transfer learning. *Proceedings of the Fifth International Conference on Advances in Biomedical Engineering (ICABME)*, Tripoli, Lebanon, 17–19 October 2019, pp. 1–4. [DOI](https://doi.org/) | [Google Scholar](https://scholar.google.com/)

Li, H., Zhao, J., & Jiang, Z. (2024). Deep learning-based computer-aided detection of ultrasound in breast cancer diagnosis: A systematic review and meta-analysis. *Clinical Radiology, 79*(11), e1403-e1413. <https://doi.org/10.1016/j.crad.2024.08.002> (Link: ScienceDirect)

Mathew, D., Mali, M. E., Dedey, F., Dzefi-Tettey, K., Li, Y., Almeida, C., Brownson, K. E., Price, R. R., & Sutherland, E. K. (2024). Mammography and breast ultrasonography services in Ghana: Availability and geographic access. *JCO Global Oncology*, *10*, e2400218. <https://doi.org/10.1200/GO.24.00218>

Miglioretti, D. L., Ichikawa, L., Smith, R. A., Bassett, L. W., Feig, S. A., Monsees, B., Parikh, J. R., Rosenberg, R. D., Sickles, E. A., & Carney, P. A. (2015). Criteria for identifying radiologists with acceptable screening mammography interpretive performance on basis of multiple performance measures. *AJR American Journal of Roentgenology, 204*(4), W486–W491. <https://doi.org/10.2214/AJR.13.12313>

Pawłowska, A., Ćwierz-Pieńkowska, A., Domalik, A., Jaguś, D., Kasprzak, P., Matkowski, R., Fura, Ł., Nowicki, A., & Zolek, N. (2024). A curated benchmark dataset for ultrasound-based breast lesion analysis (Breast-Lesions-USG) (Version 1) [Dataset]. *The Cancer Imaging Archive*. <https://doi.org/10.7937/9WKK-Q141>

Pawłowska, A., Ćwierz-Pieńkowska, A., Domalik, A., Jaguś, D., Kasprzak, P., Matkowski, R., Fura, Ł., Nowicki, A., & Zolek, N. (2024). A curated benchmark dataset for ultrasound-based breast lesion analysis. *Scientific Data, 11*, 148. <https://doi.org/10.1038/s41597-024-02984-z>

Price, M. D., Mali, M. E., Dedey, F., Dzefi-Tettey, K., Li, Y., Almeida, C., Brownson, K. E., Price, R. R., & Sutherland, E. K. (2024). Mammography and breast ultrasonography services in Ghana, availability, and geographic access. *JCO Global Oncology, 10*, e2400218. <https://doi.org/10.1200/GO.24.00218>

Rana, A. S., Rafique, J., & Riffat, H. (2024). Advances in breast ultrasound imaging: Enhancing diagnostic precision and clinical utility. *IntechOpen*. <https://doi.org/10.5772/intechopen.1005373>

Rodriguez-Ruiz, A., Lång, K., Gubern-Merida, A., Broeders, M., Gennaro, G., Clauser, P., Helbich, T. H., Chevalier, M., Tan, T., Mertelmeier, T., Wallis, M. G., Andersson, I., Zackrisson, S., Mann, R. M., & Sechopoulos, I. (2019). Stand-alone artificial intelligence for breast cancer detection in mammography: Comparison with 101 radiologists. *Journal of the National Cancer Institute, 111*(9), 916–922. <https://doi.org/10.1093/jnci/djy222>

Szeghalmy, S., & Fazekas, A. (2023). A comparative study of the use of stratified cross-validation and distribution-balanced stratified cross-validation in imbalanced learning. *Sensors, 23*(4), 2333. <https://doi.org/10.3390/s23042333>. Available at <https://www.mdpi.com/1424-8220/23/4/2333>.

Suh, Y. J., Jung, J., & Cho, B. J. (2020). Automated breast cancer detection in digital mammograms of various densities via deep learning. *Journal of Personalized Medicine, 10*(4), 211. <https://doi.org/10.3390/jpm10040211> | PubMed | PMC

Verywell Health. (2023). Breast ultrasound vs. mammography: Which is best? By S. Mirabito, RN, OCN. <https://www.verywellhealth.com/why-not-annual-ultrasounds-instead-of-mammograms-430185>

Wang, Y., Li, Y., Song, Y., Chen, C., Wang, Z., Li, L., Liu, M., Liu, G., Xu, Y., Zhou, Y., Sun, Q., & Shen, S. (2022). Comparison of ultrasound and mammography for early diagnosis of breast cancer among Chinese women with suspected breast lesions: A prospective trial. *Thoracic Cancer, 13*(22), 3145–3151. <https://doi.org/10.1111/1759-7714.14666> | PubMed | PMC

World Health Organization. (2024). Breast cancer: Fact sheet. <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>

Xue, P., Wang, J., Qin, D., & et al. (2022). Deep learning in image-based breast and cervical cancer detection: A systematic review and meta-analysis. *npj Digital Medicine, 5*(19). <https://doi.org/10.1038/s41746-022-00559-z> (Link: Nature)

Yoon, J. H., Strand, F., Baltzer, P. A. T., Conant, E. F., Gilbert, F. J., Lehman, C. D., Morris, E. A., Mullen, L. A., Nishikawa, R. M., Sharma, N., Vejborg, I., Moy, L., Mann, R. M. (2023). Standalone AI for breast cancer detection at screening digital mammography and digital breast tomosynthesis: A systematic review and meta-analysis. *Radiology, 307*(5), e222639. <https://doi.org/10.1148/radiol.222639>