Breast Masses Detection Using YOLOv8

Abdul Moiz, Hikmat Ullah, Hannia Naseem, Umar Sadique, Muhammad Abeer Irfan, Amaad Khalil
Dept. of CSE, UET Peshawar, Pakistan
*Correspondence: moizmansoor68@gmail.com, hikmatcse1919@gmail.com, hanninaseem836@gmail.com, umar.sadique@uetpeshawar.edu.pk, abeer.irfan@uetpeshawar.edu.pk, amaadkhalil@uetpeshawar.edu.pk.

Citation | Moiz. A, Ullah. H, Naseem. H, Sadique. U, Irfan. M. A, Khalil. A, "Breast Masses Detection Using YOLOv8", IJIST, Special Issue. pp 198-206, May 2024
Received | May 05, 2024 **Revised |** May 13, 2024 **Accepted |** May 19, 2024 **Published |** May 27, 2024.

Breast cancer stands as a formidable global health challenge, necessitating swift and precise diagnostic measures to combat its devastating impact. In this study, we delve into the efficacy of YOLOv8, a cutting-edge artificial intelligence model, for the precise detection and localizing of breast masses in digital mammography images. YOLOv8's inherent capability to simultaneously detect and localize masses showcases accurate pinpointing of the exact locations of abnormalities within mammographic scans. Our comprehensive evaluation reveals compelling performance metrics, including an F1 score of 0.91 and a mean Average Precision (mAP) of 0.942. These results depict the robustness of the YOLOv8 in mass detection but also show better results than the conventional clinical methods, offering higher accuracy and efficiency in the diagnostic process. This study explains the transformative potential of YOLOv8 in revolutionizing breast cancer detection paradigms, presenting a promising pathway toward enhancing early detection rates and ultimately improving patient outcomes.

Keywords: YOLOv8, Mass Detection and Localization, Digital Mammography

Introduction:

Breast cancer poses a significant global health challenge, underscoring the urgent need for advancements in diagnostic techniques to ensure early identification and improved patient outcomes [1]. While traditional methods like mammography have been pivotal in breast cancer screening, their limitations in sensitivity and specificity have prompted increased interest in leveraging technological breakthroughs, particularly the application of YOLO (You Only Look Once) and its latest iteration, YOLO v8, in medical imaging for enhanced breast cancer detection.

The intricate development of malignant breast masses stems from aberrant cell division within human tissues, leading to the emergence of benign and malignant masses. Benign masses, non-cancerous in nature, exhibit localized growth without aggressive tendencies. Conversely, malignant breast masses, driven by cancerous cells, possess an uncontrolled propensity to multiply and potentially spread to different body parts and adjacent tissues [2]. YOLO v8, as a cutting-edge object detection system, has emerged as a transformative force in medical imaging, heralding improved capabilities for disease detection [3].

In the realm of deep learning for breast cancer detection, the primary focus is on utilizing a diverse dataset to train YOLO v8 with representations from various mammography views. This strategic approach aims to augment the model's practi- cal performance by fostering a comprehensive understanding of breast cancer lesions. Overcoming challenges associated with traditional diagnostic techniques is crucial for achieving greater accuracy and efficacy [4]. The deliberate inclusion of mediolateral oblique and craniocaudal views is deemed essential, enhancing the model's ability to detect subtle pat- terns indicative of malignant growth. This comprehensive strategy elevates sensitivity and equips YOLO v8 to navigate complexities in identifying and classifying breast cancer [2].

This study is different from conventional approaches by taking leverage of a carefully curated breast mass dataset obtained from Roboflow to accurately annotate them which leads to achieving better results; further, we also conducted a validation process to ensure the dataset's quality and accuracy. Furthermore, collaborating with a radiologist for the result validation strengthens the clinical relevance of our findings. By employing YOLOv8 on this dataset and incorporating expert validation, our research offers valuable insights into the efficacy of YOLOv8 for breast mass detection. This investigation paves the way for further exploration of deep learning in breast cancer screening, potentially leading to more accurate diagnoses and improved patient care.

Literature Review:

Each year, the American Cancer Society estimates the numbers of new cancer cases and deaths in the United States and compiles the most recent data on population-based cancer occurrence and outcomes using incidence data collected by central cancer registries (through 2020) and mortality data collected by the National Center for Health Statistics (through 2021). In 2024, 2,001,140 new cancer cases and 611,720 cancer deaths are projected to occur in the United States [5]. The im- portance of mammography images in the diagnosis of breast cancer has led to a thorough investigation of developments in detection and classification.

Breast cancer impacts more than one in ten women globally, but it is particularly prominent—across all racial and ethnic groups—in the United States. The need for focused diagnostic efforts is highlighted by differences in incidence rates amongst ethnic groups [2][6]. Breast lesions are complex, three-dimensional anomalies that reflect a variety of radio- logically defined illnesses. The distinction between benign and malignant lesions must be made early to improve the prognosis of patients with this cancer, which is the most common in women and the second largest cause of cancer-related fatalities [7][8]. The use of deep learning techniques in computer vision, segmentation, detection, and image identification has increased dramatically in recent years, overcoming the drawbacks of conventional computer-aided diagnosis (CAD)

methods [9][10][11]. Despite these developments, problems in manually identifying lesions and controlling memory complexity during training still exist.

Surprisingly, when compared to shallower models, deep learning models like Alex Net, Res Net, VGG16, Inception, Google Net, and Dense Net have shown improved classification performance. The classification accuracies attained by VGG16, ResNet50, and Inception-V3 were 95%, 92.5%, and 95.5%, respectively. Although these deep learning techniques perform better than shallow models, problems with memory complexity during training and manual detection still exist. Building upon the landscape of breast cancer detection, recent advancements by Mahoro and Akhloufi [12] showcase the potential of YOLOv7 and YOLOv8 in breast mass detection. The utilization of the Vin Dr-Mammo dataset, coupled with innovative image enhancement techniques, positions YOLOv8 as a superior model, outperforming its predecessor YOLOv7 and contributing to enhanced breast cancer diagnostics.

Al-antari et al.'s [13] important study involved estimating a Full Resolution Convolutional Network (FrCN) using a Computer-Aided Diagnosis (CAD) model. Their method, which used X-ray mammography and a four-fold cross- validation, showed a high accuracy of 95.96%. One approach to identifying breast cancer is called Diverse Features-Based Detection (DFeBCD), proposed by Chouhan and colleagues [14]. They assessed their method on the IRMA mammography dataset, and it attained an accuracy rate of 80.30% by combining an emotion learning-inspired integrated classifier (ELiEC) with the Support Vector Machine (SVM). Through the application of the Lifting Wavelet Transform (LWT) for feature extraction from breast images, Muduli et al. [15] made a substantial contribution to the field. With the use of the Extreme Learning Machine (ELM) and moth flame optimization methodology, their method, which combined Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA), produced remarkable accuracy rates of 98.76% and 95.80% on the DDSM and MIAS databases, respectively.

A technique for detecting breast cancer based on diversity analysis, geostatistics, and alpha form was presented by Junior et al. [16]. They achieved a 96.30% detection accuracy using the Support Vector Machine (SVM) classifier on the DDSM and MIAS datasets. An approach for segmentation using a Mult granulation rough set and intuitionistic fuzzy soft set was presented by Ghosh et al. [17]. Their method distinguished between malignant and unaffected tissue in mammograms, hence addressing ambiguity in pixels. An effective Adaboost deep-learning technique for identifying breast cancer was created by Zheng et al. [18]. Their strategy, which combined multiple deep learning techniques with feature extraction and selection, produced a noteworthy 97.2% accuracy. Mahoro and Akhloufi investigated sophisticated deep-learning methods for breast mass identification, particularly the YOLO (You Only Look Once) framework. The promise of YOLO- based techniques in improving breast cancer diagnostics was demonstrated by the researchers by utilizing the Vin Dr- Mammo dataset and incorporating the YOLOv7 and YOLOv8 architectures.

Methodology:

In this study, a dataset containing breast mass mammograms was obtained from Rob flow publically available [19]. The dataset was carefully observed, ensuring that each image was annotated with bounding boxes around the identified masses. Block diagram of overall system is shown in figure 1.

Preprocessing:

Mammograms were preprocessed using functionality such as Auto-Orient and resizing the dataset. An auto-orient operation ensures that images are oriented correctly, and resizing is applied to stretch images to a constant size of 640x640 pixels. CLAHE (Contrast Limited Adaptive Histogram Equalization) was applied to enhance mammogram quality and visual- ize masses. Preprocessed mammograms can be seen in figure 2.

Data Splitting:

The previously developed dataset consists of a total of 1025 mammographic images. This dataset is divided into three groups: training, validation, and testing, to ensure robust model training, performance validation, and unbiased evaluation. The dataset is split as follows:

Train Set: 80% of the dataset, totaling 823 images, is allocated for training the model.

Valid Set: 13% of the dataset, totaling 135 images, is reserved for validating the model's performance during training.

Test Set: 7% of the dataset, totaling 67 images, is kept separate for final evaluation on unseen data.

Furthermore, each training example underwent data augmentation to enhance the model's robustness. The augmentation process included horizontal flipping, resulting in three augmented outputs per training example. Regarding the categories of the dataset, it comprises two classes:" mass" and" null." The" mass" class represents cases where a mass is present in the mammogram, indicating a potential abnormality, while the" null" class denotes mammograms without any detectable masses, indicating normal cases.

Model Selection and Training:

The YOLOv8 model was selected for its superior ability in object detection tasks. The YOLOv8 architecture allows for the simultaneous detection of several objects in one image, which is convenient for efficient breast mass detection, so the pre- processed data set for model training is transferred to Google Colab and the YOLOv8 model is trained on the training set.

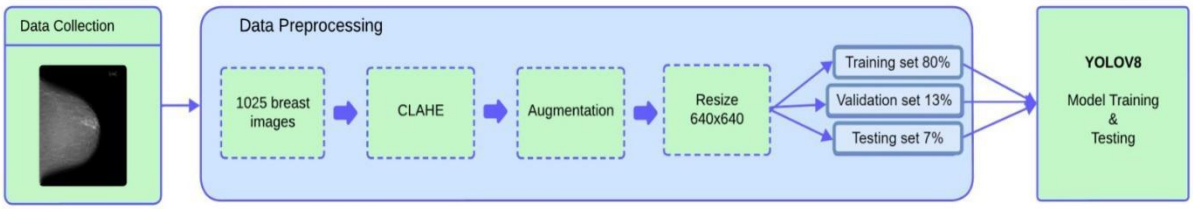

Figure 1: Block Diagram of YOLOv8-based Breast Mass Detection System

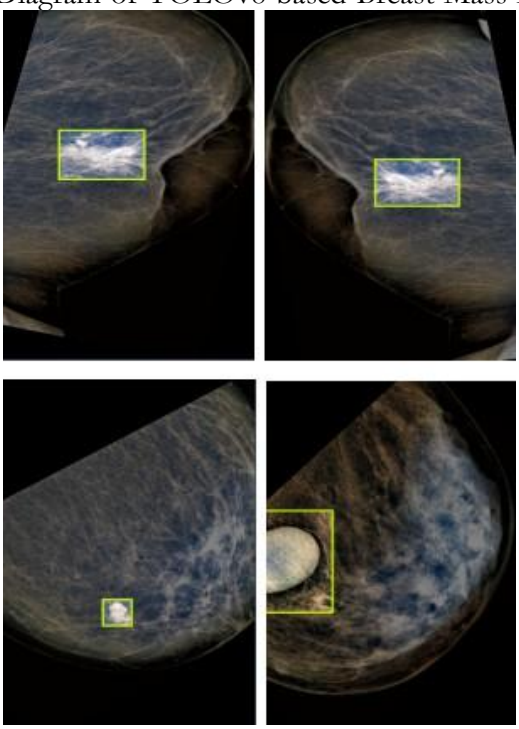

Figure 2: Preprocessed Mammograms

Testing Unseen Mammograms:

To test the generalization ability of the model, it was tested on images that were not used

during training or validation. This step was taken to measure and visualize the ability of the model to detect breast mass accurately in real-world and unseen scenarios.

Result and Discussion:

The breast mass detection system using YOLOv8 was evaluated accurately and comprehensively, and the results showed its effectiveness in detecting and localizing masses in mammographic images.

Performance Metrics:

To test the robustness and generalization of the system, the model is evaluated using matrices such as F1-scores, maps (average accuracy) and PR curves.

F1-Score: F1-score is measured as the harmonic mean of precision and recall and is a valuable metric for assessing the balance between false positives and false negatives. It is calculated by the formula as shown in equation 1.

$$F1 = \frac{2 \times P \times R}{P + R} \qquad (1)$$

The F1 score for the breast mass detection system is 0.91. This score is an important indicator of the model's ability to achieve high accuracy and recall in breast mass detection, as shown in Figure 3.

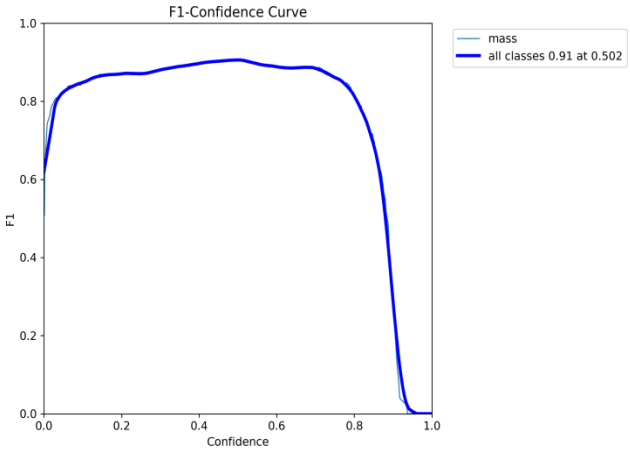

Figure 3: F1-Score Variation with Confidence Threshold

Figure 3 also shows the variation of F1-score with different confidence limits. It provides insight into model performance at different confidence points, showing robustness in balancing accuracy and recall at different operating points.

MAP (Mean Average Precision):

The map is calculated because it is a comprehensive measure that takes into account accuracy at different levels of confidence. The system achieved a commendable map of 0.942 as shown in Figure 4, confirming its high retention ability.

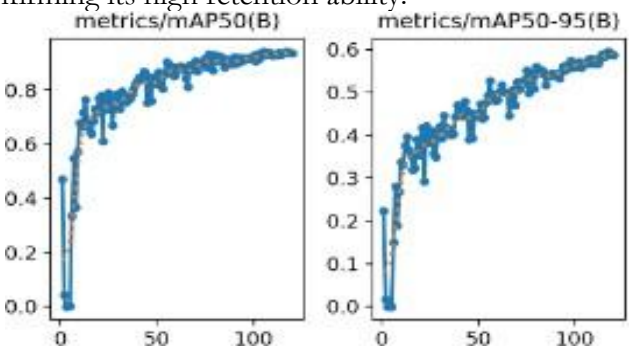

Figure 4: Mean Average Precision with IoU Threshold on the x-axis and Mean Average Precision on the y-axis

Recall and Precision Confidence Curves:

Recall and precision confidence curves provide a detailed analysis of the model's

behavior at different confidence limits, showing how precision and recall vary. This curve is necessary to under- stand the trade-off between precision and recall at different confidence levels. Figure 5 shows the confidence Recall curve for the breast mass detection system. This curve represents the relationship between the return (R) and the confidence limit, helping to determine the optimal return point for the model without compromising accuracy.

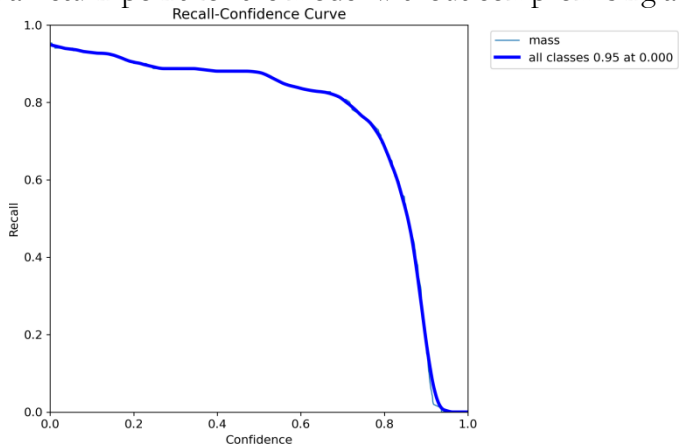

Figure 5: Recall Confidence Curve for Breast Masses Detection From equation 2 Recall (R) is defined as:

$$R = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Negatives (FN)}} \quad (2)$$

Similarly, Figure 6 shows the precision confidence curve showing how precision (P) varies with different confidence limits. This curve helps to choose an appropriate operating point based on the desired balance between accuracy and recall. The formula for Precision (P) is given in equation 3 as:

$$P = \frac{\text{True Positives (TP)}}{\text{True Positives (TP)} + \text{False Positives (FP)}} \quad (3)$$

PR Curve:

Likewise, the Precision-Recall (PR) curve is an important visual tool that shows the model's performance across various precision-recall tradeoffs. Precision is produced by plotting against recall at different confidence limits. The PR curve of this system is shown in Figure 7, which is equal to 1.00, indicating its ability to provide high accuracy even at high recall levels.

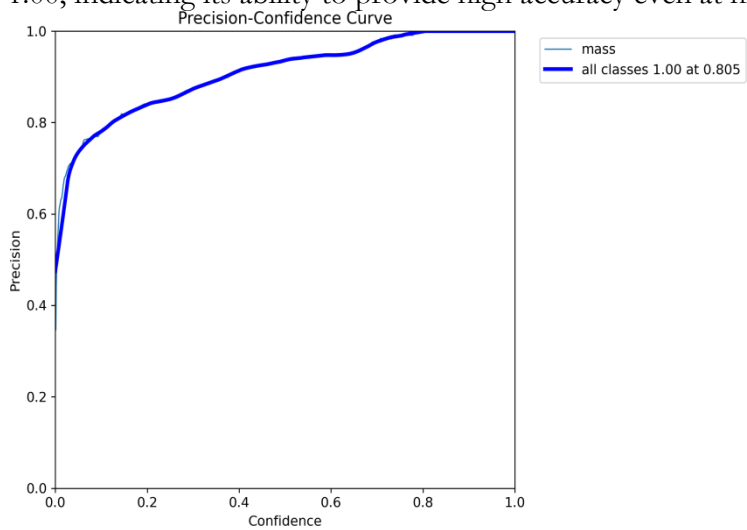

Figure 6: Precision Confidence Curve for Breast Masses Detection

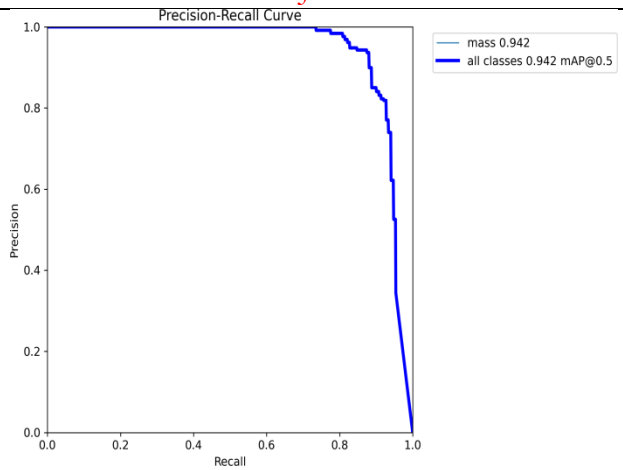

Figure 7: Precision-Recall Curve for Breast Masses Detection

Overall System Performance:

The overall performance of the system is carefully evaluated through a combination of quantitative measures, and various learning and evaluation metrics are considered to gain a deeper understanding of system performance. Figure 8 shows the plot of training and validation loss (Lbox, Lcls, Ldfl), recall and map (average accuracy) score during training.

Box-Loss (Bounding Box Loss):

Measures the localization accuracy of predicted bounding boxes and it is defined by the formula mentioned in equation 4.

$$L_{\text{box}} = \frac{1}{N \cdot B} \sum_{i=1}^{N} \sum_{j=1}^{B} \Big[(x_{i,j} - x^*_{i,j})^2 + (y_{i,j} - y^*_{i,j})^2 + (w_{i,j} - w^*_{i,j})^2 + (h_{i,j} - h^*_{i,j})^2 \Big] \tag{4}$$

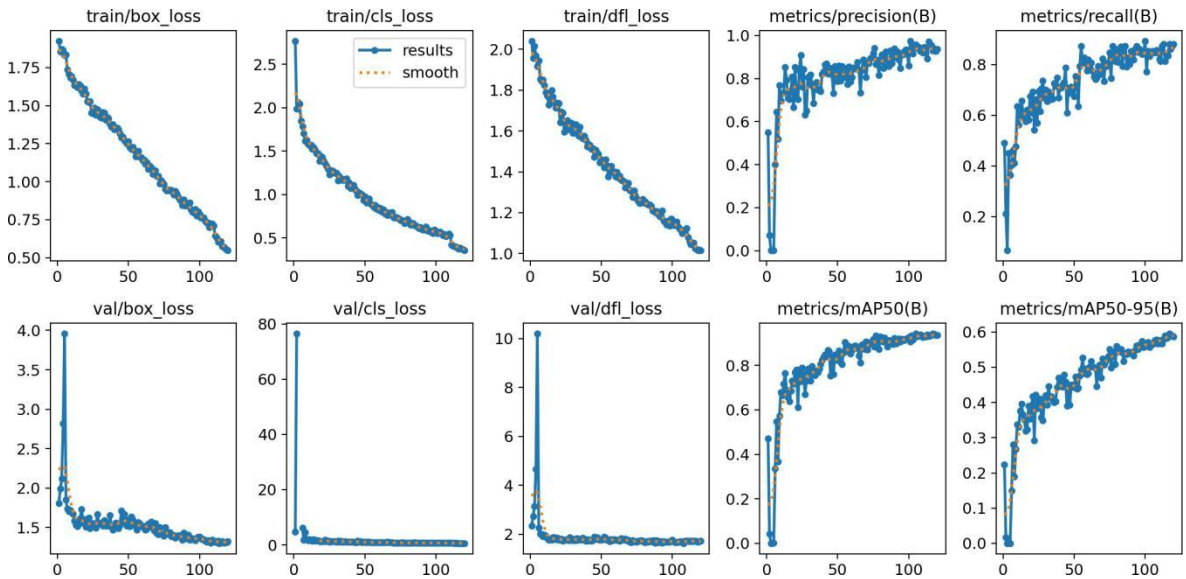

Figure 8: Overall Results of Breast Masses Detection

N: Number of samples.

B: Number of bounding boxes.

$x_{i,j}, y_{i,j}, w_{i,j}, h_{i,j}$: Predicted coordinates.

$x^*_{i,j}, y^*_{i,j}, w^*_{i,j}, h^*_{i,j}$: Ground truth coordinates.

Cls-Loss (Classification Loss):

Evaluates the accuracy of object class predictions and it is calculated as in equation 5.

$$L_{\text{cls}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}\left[y_{i,c}\cdot\log(p_{i,c}) + (1-y_{i,c})\cdot \log(1-p_{i,c})\right] \quad (5)$$

N: Number of samples.
C: Number of classes.
$Y_{i,c}$: Ground truth class.
$P_{i,c}$: Predicted probability.

Dfl-Loss (Detection Focal Loss):

Represents the overall detection performance, combining localization and classification and its formula can be seen in equation 6.

$$L_{\text{dfl}} = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C}\left[y_{i,c}\cdot(1-p_{i,c})^{\gamma}\cdot\log(p_{i,c}) + (1-y_{i,c})\cdot p_{i,c}^{\gamma}\cdot\log(1-p_{i,c})\right] \quad (6)$$

N: Number of samples.
C: Number of classes.
$Y_{i,c}$: Ground truth class.
$P_{i,c}$: Predicted probability.
P: Tunable parameter.

MAP (Mean Average Precision):

Quantifies the precision-recall trade-off across different confidence thresholds. (Computed numerically using algorithms like the trapezoidal rule)

Conclusion:

Our results demonstrate the effectiveness of the advanced YOLOv8-based breast mass detection system. This model shows promising performance in mass localization in breast mammograms showing the potential to help in the diagnosis of breast cancer. Validation of system results with radiologists strengthens its clinical utility and reliability. However, it is important to acknowledge that the quality and quantity of descriptive images play an important role in model performance. In Rob flow, a more comprehensive and annotated dataset can improve the accuracy and reliability of the model. Our collab- oration with radiologists in the annotation process ensures that the database faithfully reflects real-world scenarios.

Future Directions:

For the future direction, it is recommended to expand interpretation efforts to include supplementary mammographic views like axillary tail, and tangential views. Adding these additional views to existing mammograms can significantly improve the model's ability to detect breast masses in a wider range of scenarios. In addition, the study of advanced imaging modalities such as tomosynthesis and ultrasound may contribute to a more comprehensive and multimodal breast cancer detection system.

References:

[1] G. Hamed, M. A. E. R. Marey, S. E. S. Amin, and M. F. Tolba, "Deep Learning in Breast Cancer Detection and Classification," Adv. Intell. Syst. Comput., vol. 1153 AISC, pp. 322–333, 2020, doi: 10.1007/978-3-030-44289-7_30.

[2] "World Health Organization," 2020, [Online]. Available: https://www.who.int/news-room/fact- sheets/detail/cancer

[3] J. Li, Q., Smith, "Deep Learning in Medical Imaging: A Comprehensive Review," J. Comput. Assist. Tomogr., vol. 42, no. 2, pp. 223–233, 2018.

[4] et al Johnson, M., "Dataset Diversity in Deep Learning for Medical Image Analysis," Int. J.

Comput. Vis., vol. 128, no. 3, pp. 777–795, 2020.
[5] "Rebecca L. Siegel, Surveillance Research, American Cancer Society, 270 Peachtree Street, Atlanta, GA 30303, USA".
[6] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet- Tieulent, and A. Jemal, "Global cancer statistics, 2012," CA. Cancer J. Clin., vol. 65, no. 2, pp. 87–108, Mar. 2015, doi: 10.3322/CAAC.21262.
[7] E. Ward et al., "Cancer Disparities by Race/Ethnicity and Socioeconomic Status," CA. Cancer J. Clin., vol. 54, no. 2, pp. 78–93, Mar. 2004, doi: 10.3322/CANJCLIN.54.2.78.
[8] et al Dalaker, "Bureau of the Census, current population report, series P60–210. Poverty in the United States; 1997. U.S," Gov. Print. Off. Wash- ington, DC, 1999.
[9] "Cancer in North America: 2008–2012. Volume one: combined Cancer incidence for the United States, Canada and North America." Accessed: May 08, 2024. [Online]. Available: https://www.naaccr.org/wp-content/uploads/2016/11/Cina2015.v1.combined-incidence.pdf
[10] H. D. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," Pattern Recognit., vol. 43, no. 1, pp. 299–317, Jan. 2010, doi: 10.1016/J.PATCOG.2009.05.012.
[11] M. A. Al-antari, M. A. Al-masni, M. T. Choi, S. M. Han, and T. S. Kim, "A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification," Int. J. Med. Inform., vol. 117, pp. 44–54, Sep. 2018, doi: 10.1016/J.IJMEDINF.2018.06.003.
[12] E. Mahoro and M. A. Akhloufi, "Breast masses detection on mammograms using recent one-shot deep object detectors," BioSMART 2023 - Proc. 5th Int. Conf. Bio-Engineering Smart Technol., 2023, doi: 10.1109/BIOSMART58455.2023.10162036.
[13] M. A. Al-antari, M. A. Al-masni, and T. S. Kim, "Deep Learning Computer-Aided Diagnosis for Breast Lesion in Digital Mammogram," Adv. Exp. Med. Biol., vol. 1213, pp. 59–72, 2020, doi: 10.1007/978-3-030-33128-3_4.
[14] N. Chouhan, A. Khan, J. Z. Shah, M. Hussnain, and M. W. Khan, "Deep convolutional neural network and emotional learning based breast cancer detection using digital mammography," Comput. Biol. Med., vol. 132, p. 104318, May 2021, doi: 10.1016/J.COMPBIOMED.2021.104318.
[15] D. Muduli, R. Dash, and B. Majhi, "Automated diagnosis of breast cancer using multi-modal datasets: A deep convolution neural network based approach," Biomed. Signal Process. Control, vol. 71, Jan. 2022, doi: 10.1016/J.BSPC.2021.102825.
[16] G. Braz Junior, S. V. Da Rocha, M. Gattass, A. C. Silva, and A. C. De Paiva, "A mass classification using spatial diversity approaches in mammography images for false positive reduction," Expert Syst. Appl., vol. 40, no. 18, pp. 7534–7543, Dec. 2013, doi: 10.1016/J.ESWA.2013.07.034.
[17] P. Ghosh, M. Mitchell, J. A. Tanyi, and A. Y. Hung, "Incorporating priors for medical image segmentation using a genetic algorithm," Neurocomputing, vol. 195, pp. 181–194, Jun. 2016, doi: 10.1016/J.NEUCOM.2015.09.123.
[18] "Correction to 'Deep Learning Assisted Efficient AdaBoost Algorithm for Breast Cancer Detection and Early Diagnosis' | IEEE Journals & Magazine | IEEE Xplore." Accessed: May 08, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/9285266
[19] "Breast Cancer. "Breast Cancer Dataset." Open Source Dataset. Roboflow Universe, Roboflow", [Online]. Available: https://universe.roboflow.com/breast- cancer-ce1zx/breast-cancer-jtuaz