MEKATRONIKA
JOURNAL OF MECHATRONICS AND INTELLIGENT MANUFACTURING
E-ISSN: 2637-0883
VOLUME 7, ISSUE 1, JANUARY, 2025, 78 - 88
DOI: https://doi.org/10.15282/mekatronika.v7i1.12523

RESEARCH ARTICLE

YOLOv12-based Detection for Early Breast Cancer Screening with a Portable Ultrasound System

Hazwani Harun, Hairulnizam Hashim, Hafiz Rashidi Ramli*, Nor Mohd Haziq Norsahperi, Rozi Mahmud, Wan Zuha Wan Hasan

Department of Electrical and Electronic Engineering, Faculty of Engineering, Universiti Putra Malaysia, Serdang, Malaysia

ABSTRACT - This study explores the application of deep learning for breast lesion detection in ultrasound images using the YOLOv12 object detection model. Leveraging a compact wireless ultrasound probe and an Android-based inference pipeline, the system was developed to enable portable, AI-assisted screening in resource-limited settings. The model was trained and evaluated on an annotated ultrasound dataset and compared against the previous YOLOv11 version. YOLOv12 achieved a mean average precision (mAP) of 90.6% and an F1 score of 88.65%, outperforming YOLOv11 in both accuracy and inference speed. Processing time was also reduced, with YOLOv12 achieving image detection times between 1.34 to 4.42 seconds, compared to YOLOv11's slower range. These results confirm YOLOv12's suitability for real-time deployment on mobile platforms. Visual analyses across several test images show that YOLOv12 offers more consistent detection across varying lesion sizes and positions. The system's lightweight design, combined with its robust performance, makes it a promising tool for expanding diagnostic access in rural and underserved regions. Future work will focus on multiclass lesion classification, expanded datasets, and clinical usability studies to further validate its application in real-world healthcare environments.

ARTICLE HISTORY
Received : 17th Mar. 2025
Revised : 28th Apr. 2025
Accepted : 02nd June 2025
Published : 10th June 2025

KEYWORDS
Breast cancer
Early screening
Portable ultrasound
YOLO
Object detection

1. INTRODUCTION

Breast cancer (BC) is a potentially fatal disease in which abnormal breast cells grow out of control and form tumours [1]. Globally in 2022, 2.3 million women were diagnosed with breast cancer and the total number of breast cancer casualties reached 670,000 [2]. Breast cancer occurs in every country of the world in women at any age after puberty but with increasing rates in later life [3]. In countries with a high Human Development Index (HDI), 1 in 12 women will be diagnosed with breast cancer in their lifetime and 1 in 71 women die of it [4]. In contrast, in countries with a low HDI; while only 1 in 27 women is diagnosed with breast cancer in their lifetime, it will be fatal for 1 in 48 women [5]. Early screening or detection of breast tumors allows healthcare professionals to implement personalised treatment plans, increasing success rates while minimising the impact on the patient's quality of life [6]. Various breast imaging techniques have been developed, each with its own strengths and limitations, including mammography [7], contrast-enhanced mammography [8], sonoelastography [9], magnetic resonance imaging (MRI) [10], and magnetic elastography [11]. However, in low-resource conditions, the poor facilities and unstable power grid make it difficult to install and employ high- or middle-end ultrasound (US) machines [12]. Smartphone/tablet-sized, battery-powered US devices hold great promise to satisfy the demands in underserved nations since they are portable, low-cost, and can be modified according to customized applications [13]. For example, researchers in Ghana explored the use of portable US devices in community healthcare facilities for obstetric, pelvis, breast, vessel, abdomen, and genitourinary system examinations [14].

China researchers also reported the construction of a portable US-assisted BC screening system [15]. More inspiringly, researchers in Mexico conducted a pilot study that built a deep learning (DL) model and incorporated the model into a low-cost portable US machine to triage the breast lesions [16]. Results demonstrated that the US device could be easily operated by these healthcare workers and the built-in DL model had a similar diagnostic accuracy as breast radiologists [17]. It provided a new strategy of implementing cost-effective BC screening services in scarce-resource settings with a lack of equipment and healthcare specialists. In recent years, object detection models particularly YOLO have found increasing use in medical image analysis due to their high speed and accuracy. Studies have applied YOLOv3, YOLOv4, and YOLOv5 architectures for tasks such as tumor localization, breast lesion classification, and anomaly detection [18-20]. For example, a modified YOLOv4-tiny model was proposed for efficient and lightweight tumor detection on ultrasound images with promising accuracy for edge devices [21]. In another study, YOLOv5 was used for real-time detection of breast masses in ultrasound scans, achieving competitive performance with radiologists while enabling real-time analysis on mobile devices [22]. Several comparative studies have demonstrated YOLO's advantage over other detection algorithms such as SSD and Faster R-CNN for breast cancer image datasets [23-24]. Furthermore, recent research is investigating the fusion of YOLO with attention mechanisms and transformers for improved sensitivity and

***CORRESPONDING AUTHOR | Hafiz Rashidi Ramli** | ✉ hrhr@upm.edu.my

interpretability in clinical applications [25]. These efforts underscore the growing potential of YOLO-powered systems for integration into low-cost, portable screening setups in global health contexts.

YOLOv11, released in late 2024, introduced a suite of innovations that significantly advanced object detection capabilities, particularly in complex or overlapping object environments [26]. A notable shift in YOLOv11 was the adoption of a transformer-based backbone, moving away from conventional convolutional neural networks (CNNs). This allowed the model to better capture spatial relationships across the image, improving its ability to detect objects that may not be clearly separated. Another key feature was the removal of Non-Maximum Suppression (NMS) during training, which streamlined inference and improved processing speed. YOLOv11 also employed dual label assignment—both one-to-one and one-to-many mappings—enhancing detection performance in cluttered scenes. Additionally, the use of Partial Self-Attention (PSA) helped the model selectively focus on informative regions of the feature map, boosting representation without significantly increasing computational cost. These features collectively contributed to a performance benchmark of approximately 5% improvement in mean average precision (mAP) and an inference speed of around 60 frames per second. However, despite its accuracy and intelligent feature design, YOLOv11 carried a relatively high computational burden during training and required more memory compared to earlier YOLO versions.

YOLOv12, launched in early 2025, was developed as an evolution of YOLOv11, emphasizing greater computational efficiency and broader scalability while retaining high detection accuracy [27]. One of the key additions in YOLOv12 was the Area Attention Module (A2), which partitioned the feature map into zones, allowing the model to apply attention mechanisms more efficiently over a large receptive field. This was complemented by the introduction of Residual Efficient Layer Aggregation Networks (R-ELAN), which used residual connections at the block level to improve training convergence and stability. YOLOv12 also integrated FlashAttention, an optimized attention mechanism that improved memory usage and reduced inference bottlenecks. Structural refinements such as the removal of positional encoding, modified multi-layer perceptron (MLP) ratios, and increased convolution operations further enhanced processing speed. As a result, YOLOv12 achieved a mean average precision gain of around 6% and reduced inference latency to approximately 64 milliseconds on a T4 GPU. Compared to YOLOv11, YOLOv12 offered better speed-to-accuracy tradeoffs, making it particularly suitable for real-time deployment on mobile or edge devices. Nonetheless, its increased architectural complexity could demand more careful tuning and adaptation in certain applications.

In this paper, the feasibility of an AI-assisted early stage breast cancer screening system with a smartphone-based portable ultrasound is investigated using the state-of-the-art object detection models, namely YOLOv11 and YOLOv12. The goal of the system at this stage is to detect abnormalities found in the scans and not to provide further distinction on whether the growth is benign or malignant. If any abnormalities are found, the patient would be notified to get themselves checked as soon as possible by their medical examiners.

2. METHODS AND MATERIAL

Figure 1 shows the overall flow chart for the project. First is the preparation of the hardware equipment needed to perform data collection. Figure 2 shows the ultrasonic probe and the accompanying Android-based device that was used in this project. The ultrasonic probe, model number *MUS-P0301* was manufactured by *Fujikin Incorporated* in Japan. The probe produces images with a resolution of 1280x800 that are stored and displayed on the screen of the Android device. It is powered by a USB connection and the scanning method used is electronic scanning with a center frequency of 3.5 MHz.

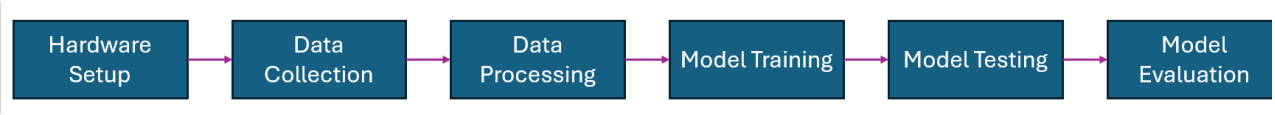


Figure 1. Overall project flow chart

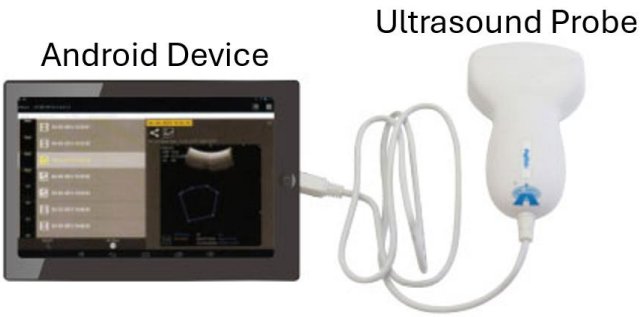


Figure 2. Portable ultrasonic probe and android device

Using the portable US probe, a total of 74 ultrasound images were collected from a private clinic based in Selangor. These images were of patients which were diagnosed to have abnormal growth by the subject matter experts operating at

the clinic. In order to increase the effectiveness of detection algorithm, data from online sources [28] have been added which consisted of 141 images of abnormal scans and 125 images of normal scans. Therefore, combined with the previous dataset, the total number of raw images collected were 340 (Figure 3). The same dataset was used to train both models. From the initial 340 images, minimal data preprocessing was done, which was orientation fix and resize to 640x640. Next, a 70:20:10 data split was used where 70% of the training images were assigned for model training, 20% was used for data validation and 10% was used for testing. Data augmentation was then performed in Roboflow as well, to increase the number of available training data by seven (7) times more to a total of 1666. The data augmentation techniques used include varying the levels of brightness (-15% to 15%), noise (up to 0.1% of pixels) and cropping (0% minimum zoom to 20% maximum zoom). The default hyperparameters for each model were used for this training as hyperparameter optimisation is outside the scope of this work.

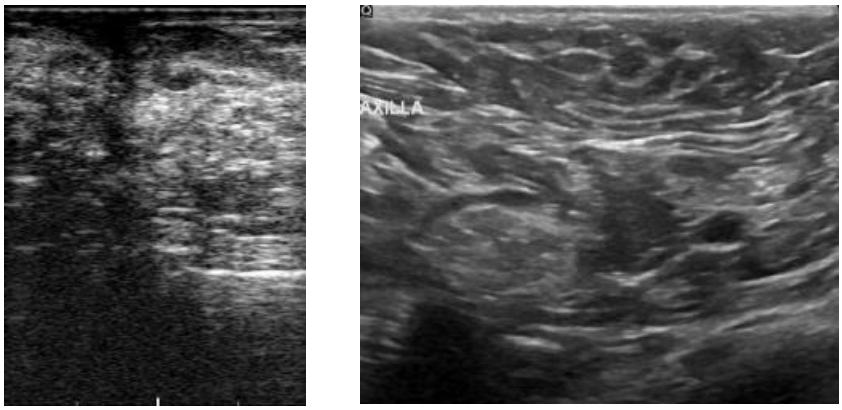


Figure 3. Sample of data collected from (a) local clinic (b) online source

Once the data is ready, the next step is model training. The deep learning model selection, configuration and training are performed online using the Roboflow platform [29]. The models chosen for the investigation are YOLOv11 and YOLOv12. These models were chosen because they are the state-of-the-art in terms of single stage object detection algorithms with both high accuracy and high speed [30].

The performance of the YOLOv11 and YOLOv12 models for breast cancer detection are then evaluated using four parameters which are 1) accuracy 2) Intersection over Union (IoU) and 3) F1-Score. Accuracy measures how often the model correctly classifies pixels compared to the total number of pixels. It is calculated for each class and then averaged across all classes in the dataset (Eq. 1), where TP is True Positive, TN is True Negative, FP is False Positive, and FN is False Negative.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

IoU is the ratio of the intersection to the union of two sets, which, in this study, were the ground truth and predicted segmentation data. IoU is represented by Eq. 2. For deep learning approaches, the average IoU score across all classes was calculated to assess the performance on the entire dataset.

$$IoU = \frac{TP}{TP + FP + FN} \tag{2}$$

The F1 score is a metric that measures how accurately the predicted boundary aligns with the true ground truth. Precision, also known as positive predictive value, is the ratio of true positive samples to the total number of predicted positive samples, as expressed in Eq. 3. Recall measures the proportion of correctly classified positive samples to the total number of actual positive samples, as defined in Eq. 4. The F1 score is the harmonic mean of recall and precision, as described in Eq. 5.

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

$$F1\ Score = 2\ (\frac{Precision\ \times\ Recall}{Precision + Recall}) \tag{5}$$

3. RESULTS AND DISCUSSION

From the model training process, it can be seen that YOLOv12 outperformed YOLOv11 by every metric, having obtained a mean average precision (mAP) of 90.6% against 83.5% (Figure 4). This indicates that YOLOv12 was able to better learn and generalize from the training data. Additionally, the training curves reveal that YOLOv12 exhibited a more stable and consistent progression throughout the training epochs, with fewer fluctuations and smoother convergence

compared to YOLOv11. The mAP curve for YOLOv12 also shows a more rapid increase during the early stages of training, reaching performance plateaus earlier than YOLOv11. In contrast, YOLOv11's mAP curve shows more noise and instability, indicating slower convergence and a higher likelihood of performance variability across epochs.

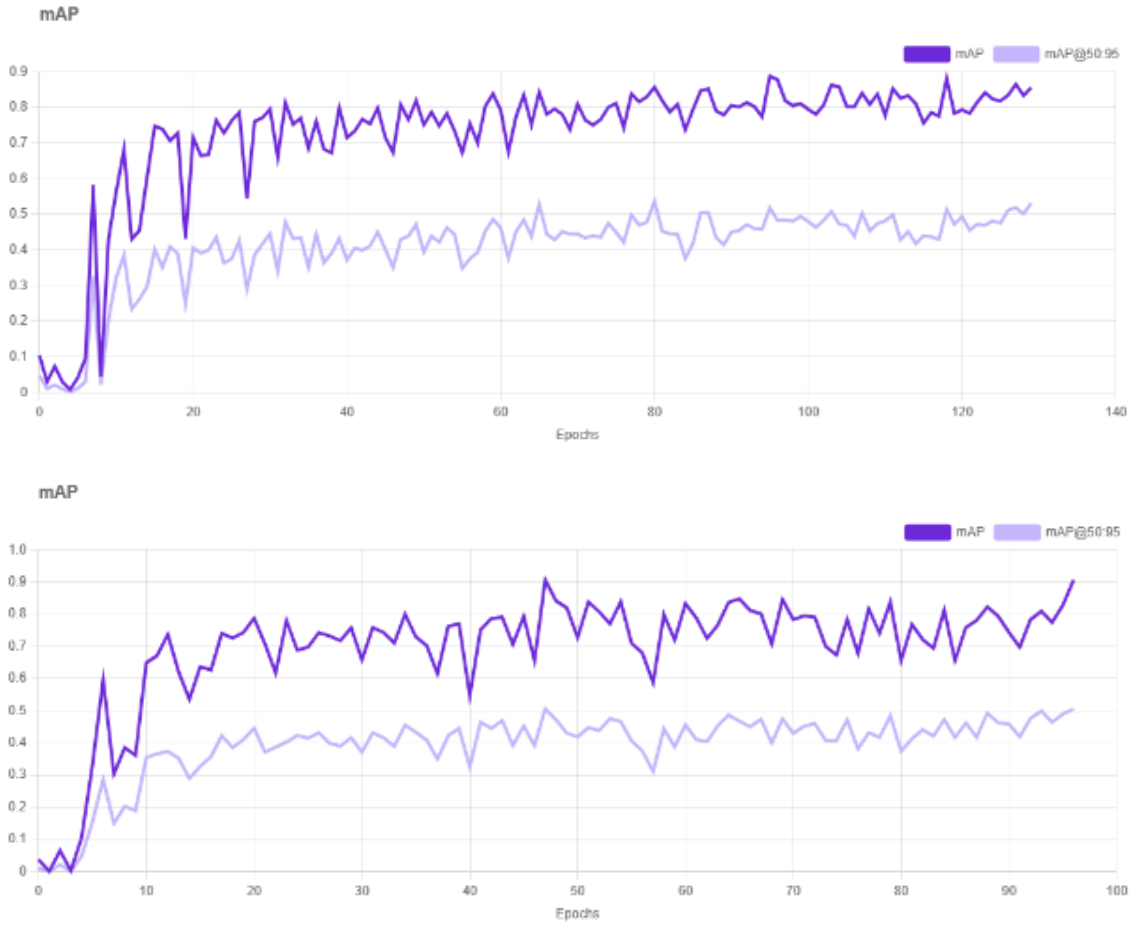


Figure 4. Performance during training session for (top) YOLOv11 and (bottom) YOLOv12

Figure 5 presents a comparison of the Box Loss, Class Loss, and Object Loss during training for YOLOv11 and YOLOv12. Overall, the graphs demonstrate that YOLOv12 consistently outperforms YOLOv11 in terms of convergence speed, stability, and final loss values across all three loss components. For Box Loss, which measures localization accuracy, YOLOv11 begins at a loss of approximately 2.8 and gradually stabilizes around 1.5–1.6 after 30 epochs, though with noticeable fluctuations. In contrast, YOLOv12 starts slightly higher at 3.5 but decreases much more rapidly and stabilizes to a lower baseline around 1.4. In terms of Class Loss, which reflects the model's ability to correctly classify detected objects, YOLOv11 peaks at around 12 and gradually reduces to a stable value near 1.5. YOLOv12, however, shows a sharp initial spike up to 250, followed by a rapid and sustained drop to near-zero within the first 10 epochs. YOLOv11 starts around 3.8 and stabilizes slowly with noticeable noise, while YOLOv12 begins higher (~18) but quickly converges to below 1 with minimal fluctuation. Overall, YOLOv12 demonstrates faster and more stable convergence for all three loss components.

From Table 1, YOLOv12 achieves a higher precision of 97.6% compared to YOLOv11's 88.7%, indicating that when YOLOv12 makes a detection, it is highly likely to be correct. This is demonstrated by the detection results for Figure. 6(a)where YOLOv12 demonstrated improved lesion localization accuracy, where the predicted bounding box closely matched the ground truth. Moreover, YOLOv12 also exhibits higher recall (81.2%) compared to YOLOv11 (77.1%), suggesting that YOLOv12 is more sensitive and able to detect more lesions, even in challenging scenarios as demonstrated in Figure 6(g), where YOLOv11 failed to detect the lesion while YOLOv12 managed to capture it, albeit with some localization error. This improved recall highlights YOLOv12's capacity to minimize false negatives, a critical factor in medical imaging applications where missing a lesion is undesirable.

Finally, the F1 Score, which balances precision and recall, is also higher for YOLOv12 (88.65%) compared to YOLOv11 (82.49%). This means that despite the occasional false positives seen in YOLOv12 (Figure. 6(c) and Figure. 6(l)), the overall balance between precision and recall favors YOLOv12 as the more reliable model. The metrics quantitatively validate that YOLOv12 achieves a better trade-off between sensitivity and specificity, making it a more effective tool for lesion detection in the presented dataset. This is illustrated by the results Figure. 6(n) and Figure. 6(o) where YOLOv11 made false positive detection on the image but YOLOv12 avoided making the same error.

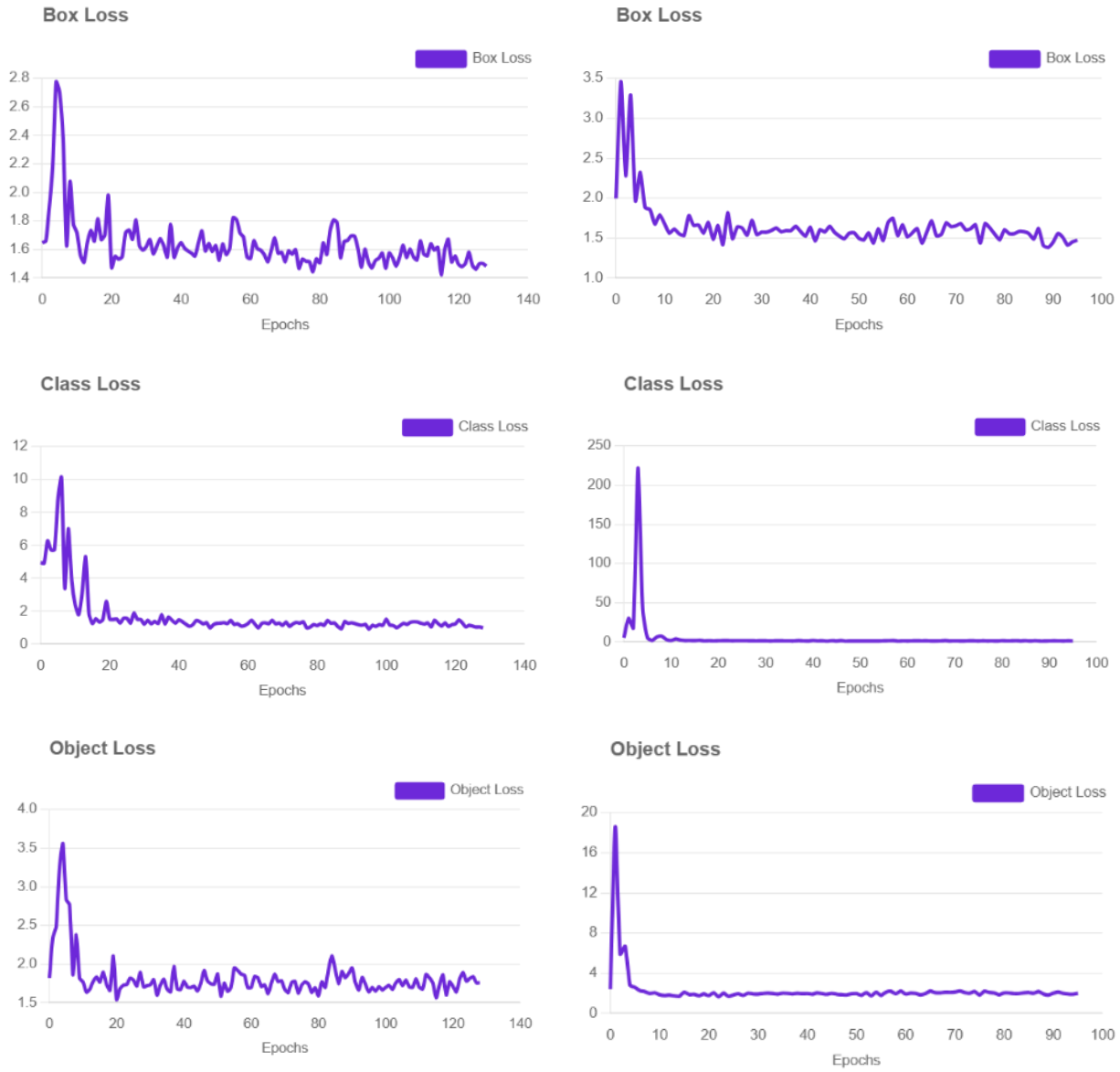


Figure 5. Comparison of box, class and object loss during training for YOLOv11 (left) and (right) YOLOv12

Table 1. Comparison of detection performance metrics for YOLOv11 and YOLOv12

Metric	YOLOv11	YOLOv12
Precision (%)	88.70	97.60
Recall (%)	77.10	81.20
F1 Score (%)	82.49	88.65

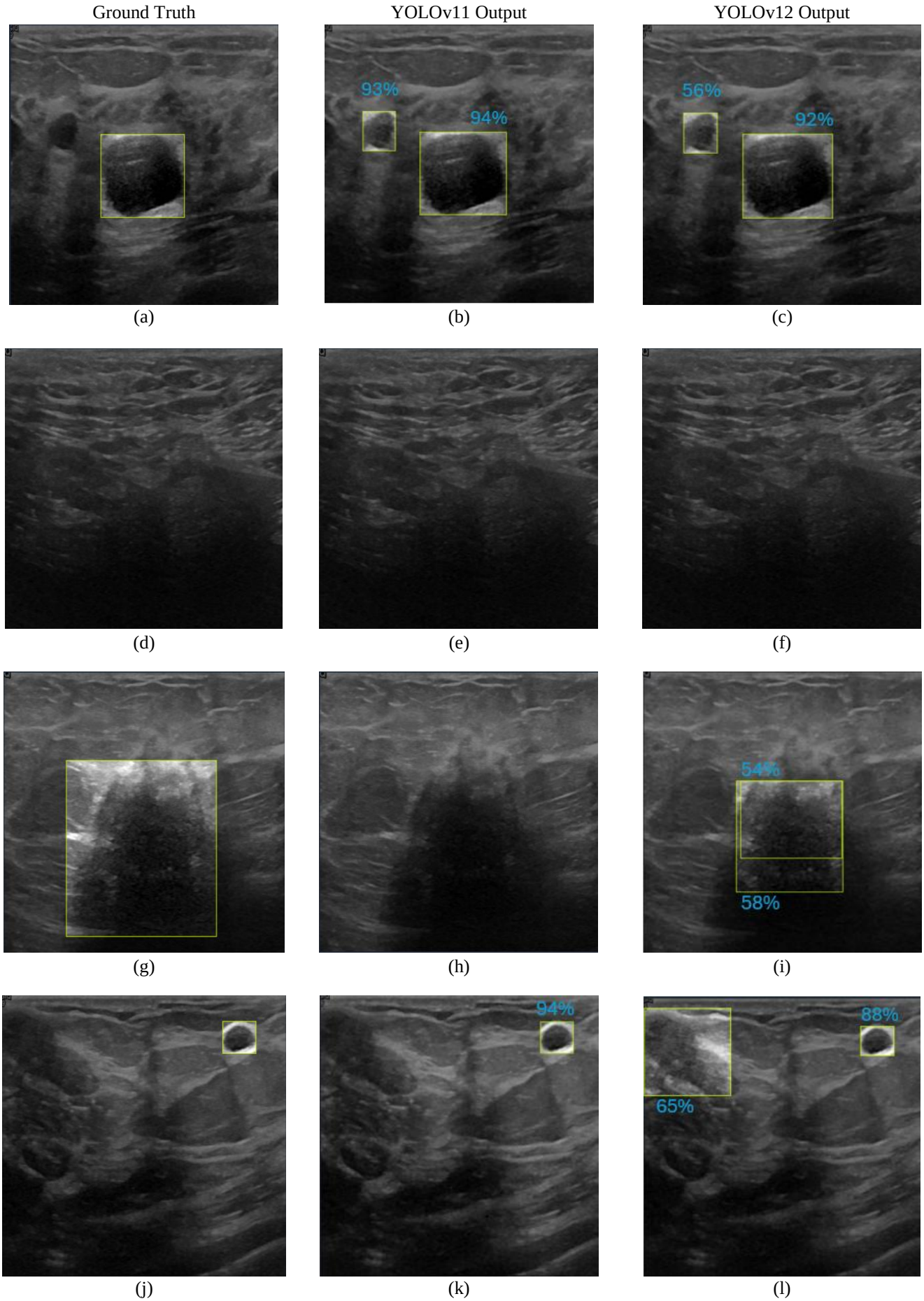
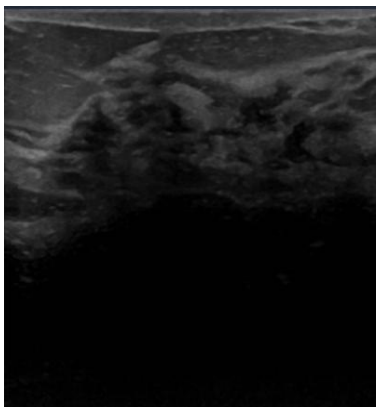
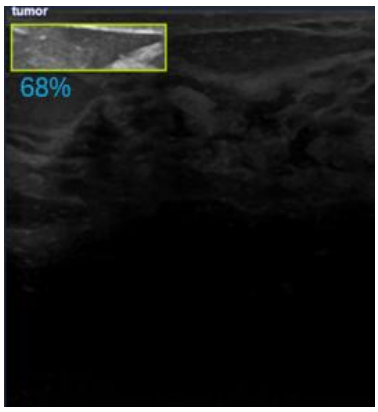


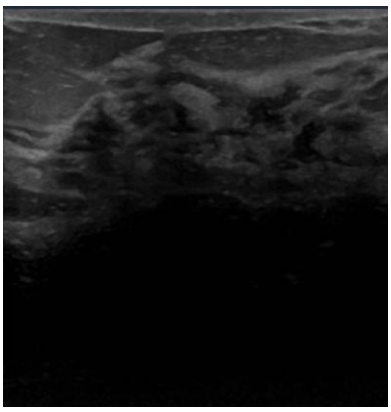
Figure 6. Comparison of ground truth vs detection results



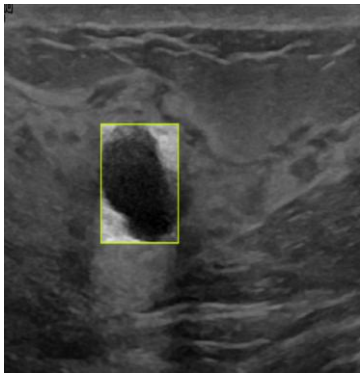
(m)



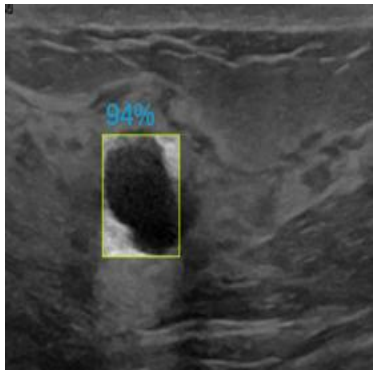
(n)



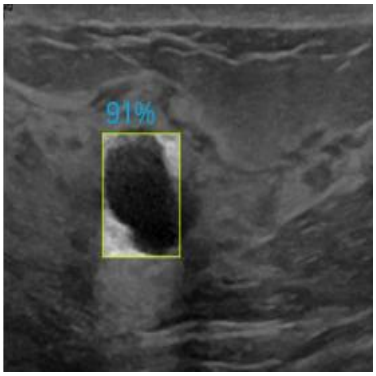
(o)



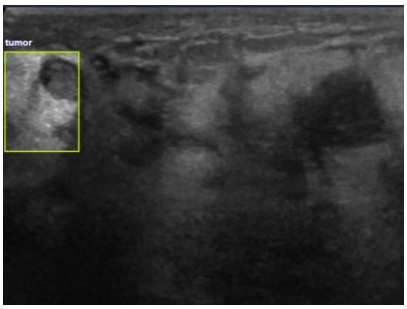
(p)



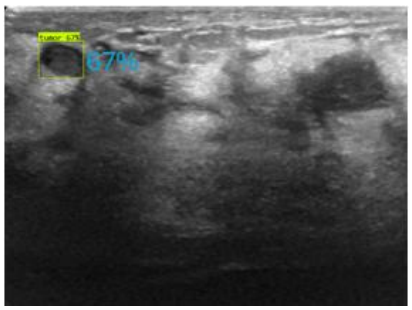
(q)



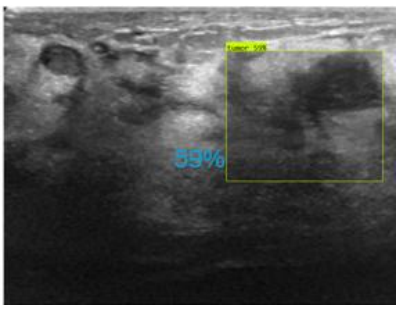
(r)



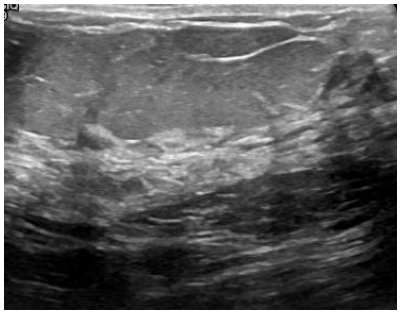
(s)



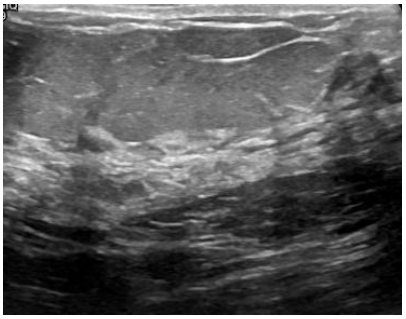
(t)



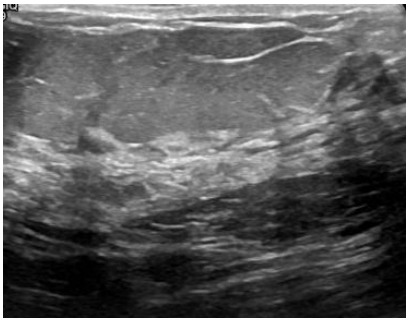
(u)



(v)

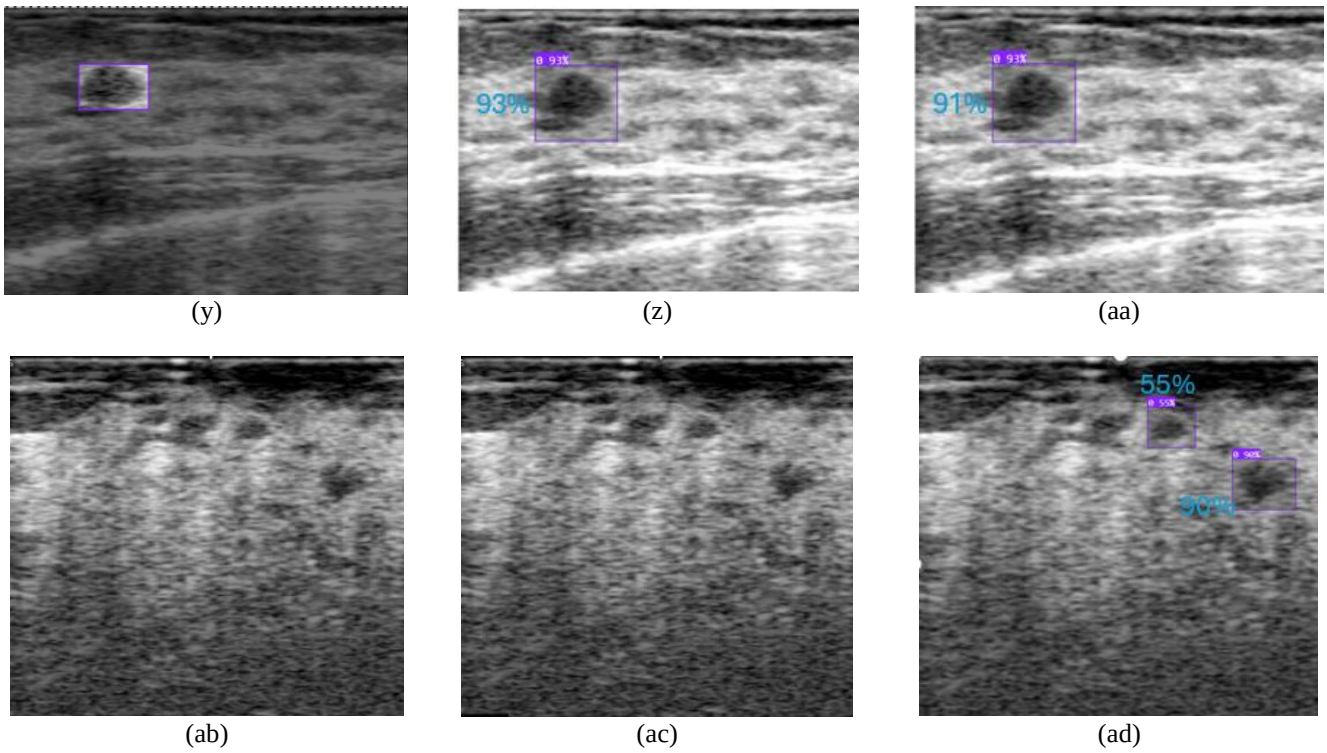


(w)



(x)

Figure 6. (cont.)



(y) (z) (aa)

(ab) (ac) (ad)

Figure 6. (cont.)

Table 2. Comparison of detection confidence and inference time for YOLOv11 and YOLOv12

No.	Images	Yolov11 confidence	Yolov11 inference time (s)	Yolov12 confidence	Yolov12 inference time (s)
1.	Figure 6(a)	93%, box 1 94%, box 2	5.32	56%, box 1 92%, box 2	2.66
2.	Figure 6(d)	-	4.92	-	1.5
3.	Figure 6(g)	-	3.33	54%, box 1 58%, box 2	4.42
4.	Figure 6(j)	94%	2.10	65%, box 1 88%, box 2	1.84
5.	Figure 6(m)	68%	1.47	-	2.15
6.	Figure 6(p)	94%	1.18	91%	2.42
7.	Figure 6(s)	67%	1.04	59%	4.25
8.	Figure 6(v)	-	1.89	-	2.02
9.	Figure 6(y)	93%	1.37	91%	2..68
10.	Figure 6(ab)	-	3.32	55%, box 1 90%, box 2	1.34

The comparative evaluation of YOLOv11 and YOLOv12 based on ten selected ultrasound images, as presented in Figure 6 and Table 2, highlights distinct differences in detection consistency, inference efficiency, and localization accuracy. In terms of detection confidence, YOLOv11 demonstrates high variability across cases. For instance, it detects Figure 6(a) with strong confidence (93–94%) but fails to register any detection for Figure 6(d), Figure 6(g), Figure 6(v), and Figure 6(ab). In contrast, YOLOv12 maintains a minimum confidence of 54% across all detected cases and successfully identifies lesions in eight out of ten images. The lowest YOLOv12 confidence appears in Figure 6(g) (54%), while the highest is recorded in Figure 6(m) (91%). This suggests that YOLOv12 offers more consistent detection performance and avoids overfitting to specific features. YOLOv11's failure to detect certain lesions despite high contrast regions (e.g., Figure 6(ab)) points to weaker generalization under variable imaging conditions.

When comparing inference time, YOLOv12 outperforms YOLOv11 in most scenarios. YOLOv12 achieves faster processing in six out of ten images, with times ranging from 1.34 to 4.42 seconds, whereas YOLOv11's times vary from 1.04 to 5.32 seconds. The performance gap is particularly notable in Figure. 6(a) and Figure. 6(ab), where YOLOv12 processes faster by nearly half. This efficiency can be attributed to architectural optimizations in YOLOv12, making it more suitable for edge deployment in portable ultrasound-based screening systems. From a qualitative perspective, visual comparisons in Figure 6 show that YOLOv12 generally produces tighter and more anatomically aligned bounding boxes than YOLOv11. In Image 1, although both models detect the lesion, YOLOv12's bounding box more accurately fits the actual shape and contour of the lesion as marked in the ground truth. In Image 3, YOLOv11 completely fails to identify the lesion, while YOLOv12 successfully detects it with two bounding boxes, albeit with moderate confidence. This illustrates YOLOv12's advantage in handling subtle or low-contrast lesions that YOLOv11 may overlook.

More significant differences emerge in Images 8 and 9. In Image 8, both models fail to detect the lesion, likely due to extremely poor contrast or indistinct lesion margins. In Image 9, both models succeed, but YOLOv12's bounding box avoids nearby shadows that YOLOv11 mistakenly includes, demonstrating better contextual discrimination. Finally, Image 10 reveals another clear strength of YOLOv12: while YOLOv11 produces no detection, YOLOv12 identifies two boxes (55% and 90%) in just 1.34 seconds. The higher-confidence box aligns well with the ground truth in Figure 6, supporting the argument for YOLOv12's enhanced feature learning capabilities.

In conclusion, this study successfully demonstrates the feasibility of using deep learning-based object detection, specifically YOLOv12, for early breast lesion detection via portable ultrasound imaging. The project integrates a compact, affordable ultrasound probe with an Android-based inference system, making it highly suitable for screening in rural or resource-limited settings. The system's design aligns with real-world needs, particularly the demand for fast, accurate, and mobile diagnostic support tools. Comparative evaluation between YOLOv11 and YOLOv12 showed that the latter consistently outperforms its predecessor across key performance dimensions. YOLOv12 achieved a higher mean average precision (mAP) of 90.6% compared to YOLOv11's 83.5%, along with superior F1 Score (88.65% vs. 82.49%). Training metrics further confirmed YOLOv12's advantages, including faster convergence and lower overall loss across box, class, and object prediction categories. Inference time is another critical consideration for deployment. YOLOv12 demonstrated a substantial improvement in processing efficiency, with detection times ranging from 1.34 to 4.42 seconds, compared to YOLOv11's range of 2.11 to 6.55 seconds. This improvement enhances the potential for real-time application in mobile scenarios.

While the results of this study demonstrate promising performance of the YOLOv12 model for breast lesion detection in ultrasound images, several limitations must be acknowledged. Firstly, the dataset used for training and evaluation, although augmented, is relatively limited in size and diversity. The images originate from a specific source, which may introduce dataset bias and affect the generalizability of the model when applied to broader clinical populations. Additionally, the dataset does not include a balanced distribution of lesion sizes, types, and complexities, which could lead to skewed detection performance favoring more distinct or easily visible lesions.

Secondly, ultrasound imaging is highly operator dependent. Variations in probe angle, pressure, and positioning can significantly affect image quality and lesion visibility. The current model has not been evaluated against such variability, which is a critical factor for real-world application. Furthermore, the black-box nature of deep learning models, including YOLOv12, poses challenges in explainability. No interpretability techniques such as Grad-CAM or saliency maps were employed in this study, making it difficult to understand the basis for specific detections or misclassifications. Finally, since this system is still in the proof-of-concept stage, it has not been tested under clinical deployment conditions. Issues such as batch processing speed, error tolerance, and multi-lesion handling remain unexplored and should be addressed in future iterations.

ACKNOWLEDGEMENTS

This study was not supported by any grants from funding bodies in the public, private, or not-for-profit sectors.

CONFLICT OF INTEREST

The authors declare no conflicts of interest

AUTHORS CONTRIBUTION

H. Harun (Investigation; Writing – original draft)

H. Hashim (Data curation; Software; Formal Analysis)

N.M.H. Norsahperi (Methodology; Validation)

H.R. Ramli (Visualisation; Writing - review & editing)

R. Mahmud (Resources; Supervision)

W.Z.W. Hasan (Project Administration; Supervision)

REFERENCES

[1] O. Ginsburg, C. H. Yip, A. Brooks, A. Cabanes, M. Caleffi, J. Y. Dunstan, et al., "Breast cancer early detection: A phased approach to implementation," *Cancer*, vol. 126, pp. 2379-2393, 2020.

[2] A. N. Giaquinto, H. Sung, K. D. Miller, J. L. Kramer, L. A. Newman, A. Minihan, et al., "Breast cancer statistics, 2022," *CA: A Cancer Journal for Clinicians*, vol. 72, no. 6, pp. 524–541, 2022.

[3] American Cancer Society, *Breast cancer facts & figures 2022–2024*, Atlanta: American Cancer Society; 2022.

[4] World Health Organization. *Breast cancer*. Geneva: WHO; 2022. Available from: https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[5] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394-424, 2018.

[6] National Breast Cancer Foundation, *Early detection of breast cancer*, 2024. Available from: https://www.nationalbreastcancer.org/early-detection-of-breast-cancer/

[7] M. Løberg, M. L. Lousdal, M. Bretthauer, and M. Kalager, "Benefits and harms of mammography screening," *Breast Cancer Research*, vol. 17, no. 1, p. 63, 2015.

[8] M. S. Jochelson and M. B. I. Lobbes, "Contrast-enhanced mammography: State of the art," *Radiology*, vol. 299, no. 1, pp. 36-48, 2021.

[9] B. Mesurolle, M. El Khoury, F. Chammings, M. Zhang, and S. Sun, "Breast sonoelastography: Now and in the future," *Diagnostic and Interventional Imaging*, vol. 100, no. 10, pp. 567-577, 2019.

[10] G. L. G. Menezes, F. M. Knuttel, B. L. Stehouwer, R. M. Pijnappel, and M. A. A. J. van den Bosch, "Magnetic resonance imaging in breast cancer: A literature review," *World Journal of Clinical Oncology*, vol. 5, no. 2, pp. 61–70, 2014.

[11] A. E. Bohte, J. L. Nelissen, J. H. Runge, O. Holub, S. A. Lambert, L. de Graaf, et al. "Breast magnetic resonance elastography: A review of clinical work and future perspectives," *NMR in Biomedicine*, vol. 31, no. 4, pp. e3932, 2018.

[12] M. O. Sana and R. M. Mungwira, "Ultrasound use in resource-limited settings: A systematic review," *Journal of Global Radiology*, vol. 4, no. 1, pp. 1-10, 2018.

[13] D. M. Becker, C. A. Tafoya, S. L. Becker, G. H. Kruger, M. J. Tafoya, and T. K. Becker, "The use of portable ultrasound devices in low- and middle-income countries: A systematic review of the literature," *Tropical Medicine & International Health*, vol. 21, no. 3, pp. 294–311, 2016.

[14] J. K. Spencer and R. S. Adler, "Utility of portable ultrasound in a community in Ghana," *Journal of Ultrasound in Medicine*, vol. 27, no. 12, pp. 1735–1743, 2008.

[15] Z. Wang, B. He, Y. Zhang, Z. Li, R. Yao, and K. Huang, "Design and implementation for portable ultrasound-aided breast cancer screening system," *Journal of Biomedical Engineering*, vol. 39, no. 3, pp. 390–397, 2022.

[16] E. Martín-Del-Campo-Mena, P. A. Sánchez-Méndez, E. Ruvalcaba-Limón, F. M. Lazcano-Ramírez, A. Hernández-Santiago, J. A. Juárez-Aburto, et al., "Development and validation of an infrared-artificial intelligence software for breast cancer detection," *Exploration of Targeted Anti-Tumor Therapy*, vol. 4, no. 2, pp. 294–306, 2023.

[17] Q. Dan, T. Zheng, L. Liu, D. Sun, and Y. Chen, "Ultrasound for breast cancer screening in resource-limited settings: Current practice and future directions," *Cancers (Basel)*, vol. 15, no. 7, p. 2112, 2023.

[18] J. Amin, M. A. Anjum, M. Sharif, S. Kadry, A. Nadeem, and S. F. Ahmad, "Liver tumor localization based on YOLOv3 and 3D-semantic segmentation using deep neural networks," *Diagnostics (Basel)*, vol. 12, no. 4, p. 823, 2022.

[19] C. M. Kim, K. Chung, and R. C. Park, "Anomaly detection model of mammography using YOLOv4-based histogram," *Personal and Ubiquitous Computing*, vol. 7, no. 3, pp. 1233–1244, 2023.

[20] M. Anas, I. U. Haq, G. Husnain, and S. A. F. Jaffery, "Advancing breast cancer detection: Enhancing YOLOv5 network for accurate classification in mammogram images," *IEEE Access*, vol. 12, pp. 16474–16488, 2024.

[21] M. L. Huang and Y. S. Wu, "GCS-YOLOV4-Tiny: A lightweight group convolution network for multi-stage fruit detection," *Mathematical Biosciences and Engineering*, vol. 20, no. 1, pp. 241–268, 2023.

[22] T. Yang, L. Yuan, P. Li, and P. Liu, "Real-time automatic assisted detection of uterine fibroid in ultrasound images using a deep learning detector," *Ultrasound in Medicine & Biology*, vol. 49, no. 7, pp. 1616–1626, 2023.

[23] Q. Fu and H. Dong, "Spiking neural network based on multi-scale saliency fusion for breast cancer detection," *Entropy (Basel)*, vol. 24, no. 11, p. 1543, 2022.

[24] A. D. Mohammed and D. Ekmekci, "Breast cancer diagnosis using YOLO-based multiscale parallel CNN and flattened threshold swish," *Applied Sciences*, vol. 14, no. 7, p. 2680, 2024.

[25] K. Vanitha, M. Aridoss, K. Chokkanathan, K. Anitha, M. T. Ramakrishna, V. Kumar, et al., "Attention-based feature fusion with external attention transformers for breast cancer histopathology analysis," *IEEE Access*, vol. 12, pp. 126296–126312, 2024.

[26] R. Khanam and M. Hussain, "Yolov11: An overview of the key architectural enhancements," *arXiv*, 2024. Available from: https://arxiv.org/abs/2410.17725

[27] Y. Tian, Q. Ye, and D. Doermann, "Yolov12: Attention-centric real-time object detectors," *arXiv*, 2025. Available from: https://arxiv.org/abs/2502.12524

[28] A. Shah, "Breast Ultrasound Images Dataset," *Kaggle* [Internet]. 2021. Available from: https://www.kaggle.com/datasets/aryashah2k/breast-ultrasound-images-dataset/data

[29] B. Dwyer, J. Nelson, T. Hansen, et al., "Roboflow (Version 1.0)," [Software]. *Roboflow* [Internet]. 2024. Available from: https://roboflow.com.

[30] R. Sapkota, Z. Meng, M. Churuvija, X. Du, Z. Ma, and M. Karkee, "Comprehensive performance evaluation of yolo11, yolov10, yolov9 and yolov8 on detecting and counting fruitlet in complex orchard environments," *Authorea Preprints*, 2024.