

Detection and Classification of Lesions in Mammograms using One-Stage Models

Mohammad Amin Sakha (MSc)¹, Ali Ameri (PhD)^{1*}

¹Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

ABSTRACT

Background: Breast cancer, the most common cancer among women, necessitates early detection. Despite advances in Computer-Aided Diagnosis (CAD), lesion detection in mammograms remains challenging. Artificial Intelligence (AI) in radiology offers significant potential to enhance diagnostic accuracy in medical imaging.

Objective: This study compares object detection methods to identify the most effective model for smart diagnostic systems. This comprehensive study is the first to apply the advanced You Only Look Once version 12 (YOLO-v12) architecture for the automated detection and localization of lesions in mammographic images and to identify their malignancy or benignity status with high precision.

Material and Methods: This comparative experimental study, utilizing retrospective data, also evaluated two state-of-the-art models, the Detection Transformer (DETR) and RetinaNet, for their performance. The models were trained and tested on the publicly available Categorized Digital Database for Low-Energy and Subtracted Contrast-Enhanced Spectral Mammography (CDD-CESM), which contains 1,982 mammograms with 3,720 annotated lesions of various types and sizes.

Results: YOLO-v12 demonstrated excellent diagnostic accuracy (mean Average Precision at an IOU threshold of 0.5 (mAP50)=0.98; Intersection Over Union (IOU)=0.95), significantly outperforming contemporary models and older YOLO versions.

Conclusion: The promising and robust results clearly underscore the remarkable potential of artificial intelligence technologies in effectively assisting radiologists with the early detection and diagnosis of breast cancer. These findings advocate for the implementation of YOLO-v12 in clinical mammography screening applications and suggest that future research should prioritize real-time diagnostic systems to further enhance breast cancer detection capabilities.

Citation: Sakha MA, Ameri A. Detection and Classification of Lesions in Mammograms using One-Stage Models. *J Biomed Phys Eng.* 2025;15(6):543-554. doi: 10.31661/jbpe.v0i0.2506-1941.

Keywords

Artificial Intelligence; Deep Learning; Lesion Detection; Mammography; Breast Cancer Screening

Introduction

Breast cancer stands as the most prevalent malignancy among women worldwide, presenting significant challenges for global healthcare systems. According to the American Cancer Society's statistics, approximately 290,000 women are diagnosed with this disease annually, with nearly 44,000 deaths [1]. The disease process begins with abnormal cellular proliferation within breast tissue, resulting in the formation of either benign or malignant masses. Malignant tumors possess the capability to metastasize to other organs throughout the body.

*Corresponding author:
Ali Ameri

Department of Biomedical Engineering and Medical Physics, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

E-mail:
ameri@sbmu.ac.ir

Received: 23 June 2025
Accepted: 29 September 2025

Early detection of breast cancer plays a crucial role in enhancing survival rates, improving treatment efficacy, and reducing mortality rates. Nevertheless, identifying breast lesions in their initial stages presents considerable difficulties. These challenges arise from substantial variations in the appearance, dimensions, and location of lesions, as well as their resemblance to normal breast tissue. Such factors can lead to diagnostic delays or errors, underscoring the necessity for developing more precise and efficient detection methodologies [2].

Mammography is the most reliable and effective method for screening and identifying suspicious breast lesions, enabling radiologists to detect concerning abnormalities, including masses, microcalcifications (tiny calcium deposits), and structural changes in the breast tissue. During this examination, specialists search for white spots, tissue density patterns, and alterations in breast shape and size to differentiate malignant masses (dangerous with growth and spreading potential) from benign ones (non-cancerous and harmless). Research has demonstrated that regular mammography screening reduces mortality rates through early tumor detection before spread to other tissues occurs. Nevertheless, challenges, such as increasing numbers of mammograms requiring evaluation, specialists' heavy workloads, visual fatigue, and variations in image interpretation can negatively impact diagnostic accuracy [3,4].

In recent years, AI has emerged as a valuable ally for physicians, playing a crucial role in medical image analysis, particularly mammography. Computer-Aided Diagnostic (CAD) systems function as secondary consultants, capable of analyzing vast quantities of mammographic images with remarkable speed and precision, identifying suspicious lesions that warrant further investigation [5-8]. Deep Learning (DL) algorithms have demonstrated the ability to recognize patterns within images that might elude human detection. These sophisticated systems can identify

lesions with high accuracy, precisely delineate their boundaries, and even determine their classification. Recent research indicates that such systems can potentially detect cancer markers approximately six years earlier than conventional diagnostic methods, highlighting their vital contribution to improving treatment outcomes and enhancing patient survival rates [9,10]. Furthermore, AI has shown tremendous potential in assisting doctors with more accurate lesion diagnosis, reducing workload burdens, and eliminating human errors stemming from fatigue or inconsistent interpretations of imaging data.

The primary challenge in breast cancer detection lies in achieving high accuracy when identifying lesions, which is complicated by issues such as false positives (incorrectly diagnosing cancer) and false negatives (failing to detect existing disease). These errors can lead to patient anxiety, unnecessary treatments, or delays in initiating proper care. Since accurate diagnosis depends heavily on radiologists' skill and experience, and different interpretations of the same image may yield contradictory results, utilizing AI systems, i.e., DL models, as supportive tools appears essential [10]. The most advanced DL models for detecting objects (such as lesions) are called one-stage detection models. The most widely used one-stage models include YOLO (You Only Look Once), RetinaNet, and Detection Transformer (DETR). These models have been applied to lesion detection in mammography images in previous studies [11-13].

In 2023, Demirel et al. [11] implemented RetinaNet using a focal loss error function to identify cancerous masses with an accuracy of 0.73, sensitivity of 0.60, and mAP50 of 0.56 on a private dataset. Quiñones-Espín et al. [12] applied a YOLO-v5x model for breast mass detection, attaining a sensitivity of 0.80 and mAP50 of 0.60 on the VinDr-Mammo and MIAS datasets. Duque et al. [13] implemented a DETR model for breast mass detection, attaining an mAP50 of 0.68

on the INbreast dataset. Nevertheless, this model exhibited limitations in detecting small and overlapping masses, particularly in high-density images, which were attributed to insufficient training data and inadequate diversity in training samples. In 2024, Shia and Ku [14] utilized a YOLO-v8 model to detect microcalcifications in mammography images on a private dataset and achieved an accuracy of 0.86, a sensitivity of 0.83, and mAP50 of 0.92, representing substantial improvement over previous research.

In this research, in line with previous work on detecting cancer in mammography data [14], we aimed to improve the performance by applying, for the first time, the latest version of the YOLO algorithm, i.e., v12, on mammography data. Two other one-stage models, including DETR and RetinaNet, were also investigated. A public dataset, specifically the Categorized Digital Database for Low Energy and Subtracted Contrast-Enhanced Spectral Mammography (CDD-CESM), was utilized to train and test the models.

Material and Methods

In this comparative experimental study using retrospective data, three one-stage detection models: DETR, RetinaNet, and YOLO-v12 (medium variant), were trained and evaluated on the CDD-CESM dataset. These one-stage approaches are particularly suitable for real-time clinical applications due to their exceptional speed and accuracy, offering significant potential for rapid and precise lesion identification in mammographic imagery. To improve the performance, transfer learning techniques were employed utilizing pre-trained weights from the Common Objects in Context (COCO) dataset. Key hyperparameters were meticulously calibrated for each model, including input image dimensions, padding techniques, optimization algorithms (such as Adam, AdamW, and Stochastic Gradient Descent (SGD)), and error functions (including Focal Loss for classification and Smooth L1

for regression).

The lesion identification process across all examined models comprises three fundamental automated stages: 1) the extraction of key features from mammography images using Convolutional Neural Networks (CNNs) as the backbone architecture; 2) processing these features in the neck section to enhance spatial and semantic information; and 3) employing the detection head to predict the precise lesions' location (bounding box coordinates) and classification (benign or malignant). The following sections examine the CDD-CESM dataset, the architecture of each model, the implementation of transfer learning, and the evaluation metrics.

CDD-CESM dataset

This study utilized the CDD-CESM dataset comprising mammogram images from 326 female individuals. Each subject typically had 8 images, including 4 images for each breast (low-energy and CESM subtraction images in Cranio-Caudal (CC) and Medio-Lateral Oblique (MLO) views). From the initial collection of 2006 images, 24 were excluded due to incorrect information, resulting in a final set of 1982 images, where 757 images were normal (without lesions), and the remaining 1225 images contained a total of 1744 malignant and 1219 benign lesions. The malignant and benign labels were given to each lesion based on pathology reports. Out of 326 subjects, 62 individuals had images with no lesions, 115 subjects had images with only benign lesions, and the remaining 149 subjects had images with at least one malignant lesion. Figure 1 illustrates the distribution of the dataset [15].

CESM represents an advanced breast imaging technology that demonstrates superior diagnostic accuracy compared to conventional Digital Mammography (DM) and operates through intravenous administration of non-ionic iodinated contrast material (administered at a dosage of 1.5 milliliters per kilogram of body weight). This methodology captures

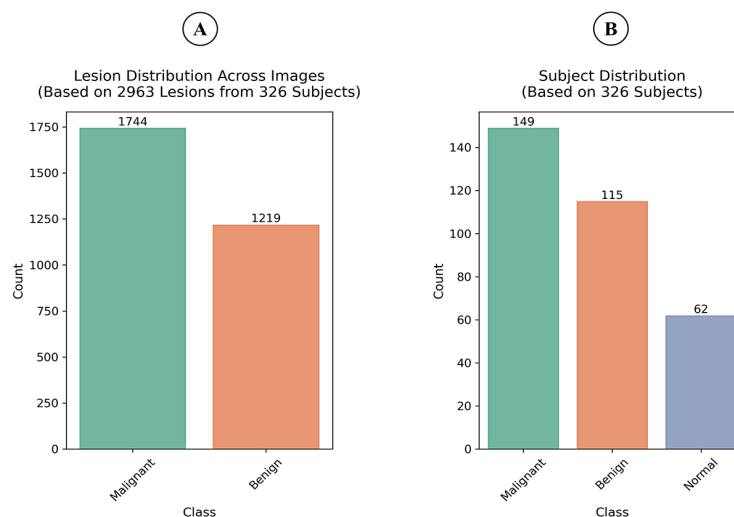


Figure 1: The dataset distribution

two distinct images for each projection (utilizing CC and MLO views): one acquired with low energy (26-31 kV) resembling standard full-field DM, and another obtained with high energy (45-49 kV) that exhibits greater sensitivity to the contrast agent. Then, subtraction of the high-energy images from the low-energy images eliminates normal breast tissue and highlights regions with high contrast uptake, which are typically markers of abnormal vascularization and possible malignancy [15].

DETR

In this study, the DETR model identified lesions in mammography images through three primary stages. Initially, in the backbone section, a ResNet-50 CNN extracted feature maps containing crucial visual information from the input image. Next, in the neck section, these feature maps underwent flattening and merged with positional encoding to preserve spatial information; subsequently, a Transformer Encoder with self-attention mechanisms processed these features to comprehend relationships between different image regions and generated context-aware, enriched features. Finally, in the detection head, the encoder's output, alongside a fixed number of object queries, entered the transformer decoder; the

decoder refined these queries through self-attention and cross-attention, producing output vectors that fed into Feed-Forward Networks. These networks independently predicted lesion classification (malignant or benign) and the corresponding bounding box coordinates. This process was trained using a Hungarian Matching bipartite loss function to ensure accurate predictions without overlaps [16].

RetinaNet

The RetinaNet model used in this study comprised three principal components. The Backbone section employed ResNet-50, resolving the gradient vanishing problem through skip connections while extracting multi-level features from mammography images. In the neck component, the Feature Pyramid Network intelligently combined low-level features (containing precise spatial details) with high-level features (rich in semantic information) via top-down pathways and lateral connections, generating multi-scale feature maps. The feature maps were passed to the detection head, comprising two parallel subnetworks: 1) a classification branch predicting lesion type (benign or malignant) using focal loss, and 2) a regression branch estimating bounding box coordinates. Finally, Non-Maximum

Suppression (NMS) eliminated redundant and overlapping boxes, leading to enhanced detection accuracy. This integrated architecture maintained an optimal balance between speed and precision, demonstrating remarkable capability in detecting lesions of varying dimensions in mammography images [17].

YOLO-v12

The YOLO-v12 model comprised three principal components that worked in concert to accurately detect lesions in mammography images. The backbone section utilized Residual Efficient Layer Aggregation Networks (RELAN), employing skip connections and 7×7 separable (depth-wise) convolutions to efficiently extract multi-scale features while mitigating the gradient vanishing problem. In the neck section, the Path Aggregation Network enriched and integrated the extracted features by establishing bidirectional connections between various layers: downward pathways transmitted semantic information while upward pathways conveyed textural details. Additionally, the area attention mechanism with FlashAttention enabled intelligent focus on critical image areas. Finally, the detection head received the refined feature maps and simultaneously predicted both the precise position of lesions (boundary boxes) and their classification (benign or malignant). It then applied the NMS algorithm to eliminate redundant boxes, thereby delivering high-precision final detection results [18-20].

Transfer Learning

Transfer learning represents an efficient methodology in deep learning, employed in this research to enhance the performance of lesion detection models in mammography images. We utilized pre-trained models on the COCO dataset, which is a large-scale object detection dataset, as a starting point, which we then optimized for lesion detection and classification on our mammography dataset. The transfer learning process encompassed

several fundamental stages: initially, we loaded the pre-trained model weights and converted the CDD-CESM dataset into a standard format compatible with the intended model (for instance, YOLO) to ensure structural compatibility with the models. Subsequently, we replaced the final fully connected and classification layers of the models according to the number of classes. We then proceeded with model fine-tuning. Fine-tuning constitutes a process within transfer learning whereby a pre-trained model undergoes optimization for a specific problem. Two principal fine-tuning methodologies exist: 1) partial fine-tuning, where only the final layers receive training and the weights of other layers are frozen, and 2) full fine-tuning, where the entire network undergoes retraining [21].

In this research, we implemented the latter approach. Given the adequate volume of available data (1982 images from 326 individuals), we retrained the complete network to optimize the extracted features for breast lesion detection. The transfer learning approach not only reduced training time but also significantly enhanced detection accuracy by transferring general object recognition knowledge to the specialized domain of breast lesion identification.

Preprocessing, Dataset Partitioning, and Hyperparameter Settings

A stratified 5-fold cross-validation was employed to ensure balanced class distribution (normal, malignant, and benign), prevent data leakage, and enable accurate evaluation. In each fold, 80% of the images per class were allocated for training, 10% for validation, and 10% for testing, with final performance metrics averaged across all five folds. For data preprocessing, images underwent normalization by subtracting the mean and dividing by the standard deviation to standardize pixel values [15]. Hyperparameter settings included training for 100 epochs with a batch size of 16, using the AdamW optimizer with an

initial learning rate of 0.0001 and a dropout rate of 0.15 to mitigate overfitting. The experiments were conducted on a hardware platform equipped with dual NVIDIA T4 GPUs, 15 GB of dedicated GPU memory, and 29 GB of system RAM.

Performance Metrics

The models' performances of lesion detection and classification were meticulously evaluated using standard metrics, including precision, recall, mean average precision at an Intersection Over Union (IOU) threshold of 0.5 (mAP50), IOU, area under the curve of the receiver operating characteristics (AUC-

ROC), and confusion matrix [22].

Results

Among the three proposed models, YOLO-v12-m achieved the highest performance for both benign and malignant lesion detection, with mAP50=0.98 and IOU=0.95. For malignant lesions, it obtained a precision of 0.92 and a recall of 0.93. RetinaNet also demonstrated robust detection capability (mAP50=0.79), whereas DETR exhibited the lowest overall performance (mAP50=0.65). Detailed performance metrics for each model are provided in Table 1, while Table 2 compares the proposed models with results

Table 1: The lesion detection and classification performance metrics for each model with 95% Confidence Intervals (CI).

Model	Class	Precision	Precision _95%_CI	Recall	Recall _95%_CI	F1- Score	F1-Score _95%_CI	mAP50	IOU
DETR	Benign	0.64	[0.53, 0.74]	0.66	[0.54, 0.76]	0.65	[0.55, 0.73]	0.65	0.80
	Malignant	0.84	[0.77, 0.89]	0.73	[0.66, 0.79]	0.77	[0.73, 0.83]		
RetinaNet	Benign	0.88	[0.78, 0.94]	0.79	[0.69, 0.87]	0.83	[0.76, 0.90]	0.79	0.86
	Malignant	0.92	[0.87, 0.96]	0.80	[0.73, 0.85]	0.86	[0.81, 0.90]		
YOLO-v12-m	Benign	0.94	[0.87, 0.98]	0.93	[0.85, 0.97]	0.94	[0.89, 0.97]	0.98	0.95
	Malignant	0.92	[0.87, 0.95]	0.93	[0.88, 0.96]	0.93	[0.90, 0.96]		

DETR: Detection Transformer, YOLO-v12-m: You Only Look Once-version 12- medium variant, CI: Confidence Intervals, mAP50: mean Average Precision at an IOU threshold of 0.5, IOU: Intersection Over Union

Table 2: Comparison with previous studies

Method	Reference	Year	Database	mAP50	IOU
DETR	[13]	2024	INbreast	0.68	-
RetinaNet	[11]	2023	Private dataset	0.56	-
YOLO-v5	[12]	2023	VinDr-Mammo, MIAS	0.60	-
YOLO-v8	[14]	2024	Private dataset	0.92	-
DETR				0.65	0.80
RetinaNet	Current Study	2025	CDD-CESM	0.79	0.86
YOLO-v12-m				0.98	0.95

DETR: Detection Transformer, YOLO-v5: You Only Look Once-version 5, YOLO-v8: You Only Look Once-version 8, YOLO-v12-m: You Only Look Once-version 12- medium variant, VinDr-Mammo: Vietnamese Clinical Image Database for Radiology – Mammography, MIAS: Mammographic Image Analysis Society, CDD-CESM: Categorized Digital Database for Low-Energy and Subtracted Contrast-Enhanced Spectral Mammography, mAP50: mean Average Precision at an IOU threshold of 0.5, IOU: Intersection Over Union

reported in previous studies. The ROC curves, confusion matrices, and representative image outputs are presented in Figures 2-4.

Statistical analysis confirms the superior performance of YOLO-v12-m (Table 1). We used bootstrap confidence intervals ($n=10,000$) for F1-scores and Wilson confidence intervals for precision and recall. YOLO-v12-m achieved significantly higher F1-scores for both benign (0.94, 95% CI: [0.89, 0.97]) and malignant lesions (0.93, 95% CI: [0.90, 0.96]) compared to RetinaNet and DETR. Independent t-tests demonstrated statistically significant differences ($P\text{-value}<0.001$) for all comparisons, with large effect sizes (t-statistics ranging from 260.21 to 582.94 for F1-scores) indicating substantial clinical relevance. The

non-overlapping confidence intervals validate YOLO-v12-m's consistent superiority across both lesion types, establishing its robustness for clinical lesion detection with 221 true positives, 17 false positives, and only 16 false negatives out of 237 total lesions (malignant and benign).

Discussion

The current study showed that YOLO-v12-m substantially outperformed DETR and RetinaNet in both lesion localization and classification into two classes of malignant and benign; and a comparison of our results with those of previous work that applied DETR, RetinaNet, and older versions of YOLO on mammography data is reported (Table 2).

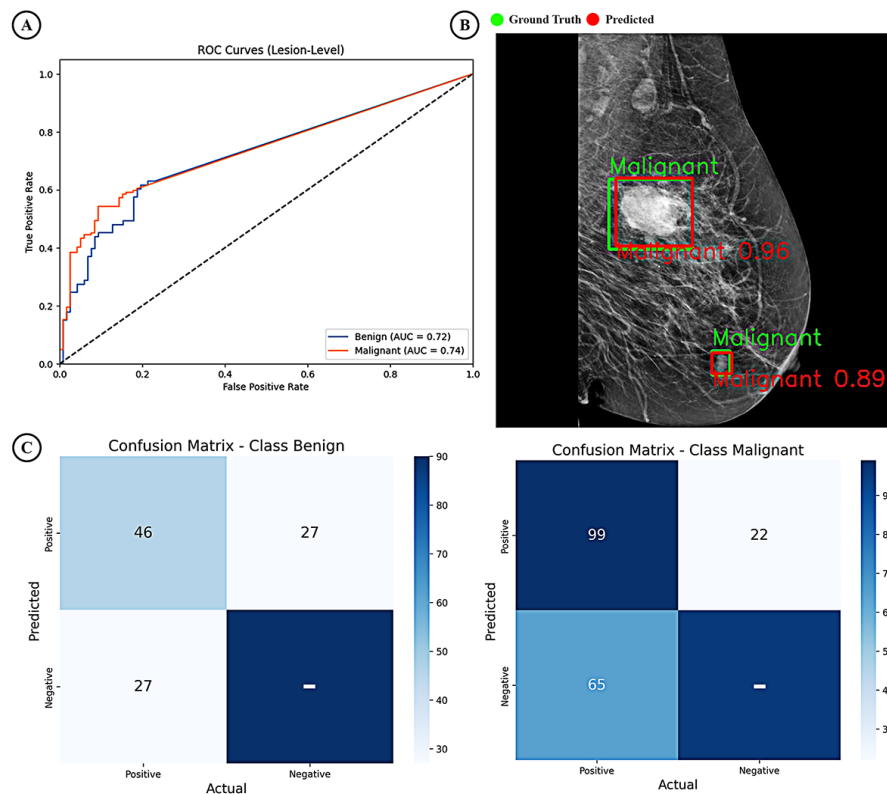


Figure 2: Detection Transformer (DETR): (A) The Receiver Operating Characteristics (ROC) curves demonstrate Area Under Curve (AUC) values for the detection of benign and malignant lesions, indicating the model's discriminative ability between these classes. (B) Representative mammography image demonstrating model predictions with bounding boxes and confidence scores. (C) Confusion matrices for lesion classification (left: benign, right: malignant), highlighting accurate identification of true positives and minimizing false positives.

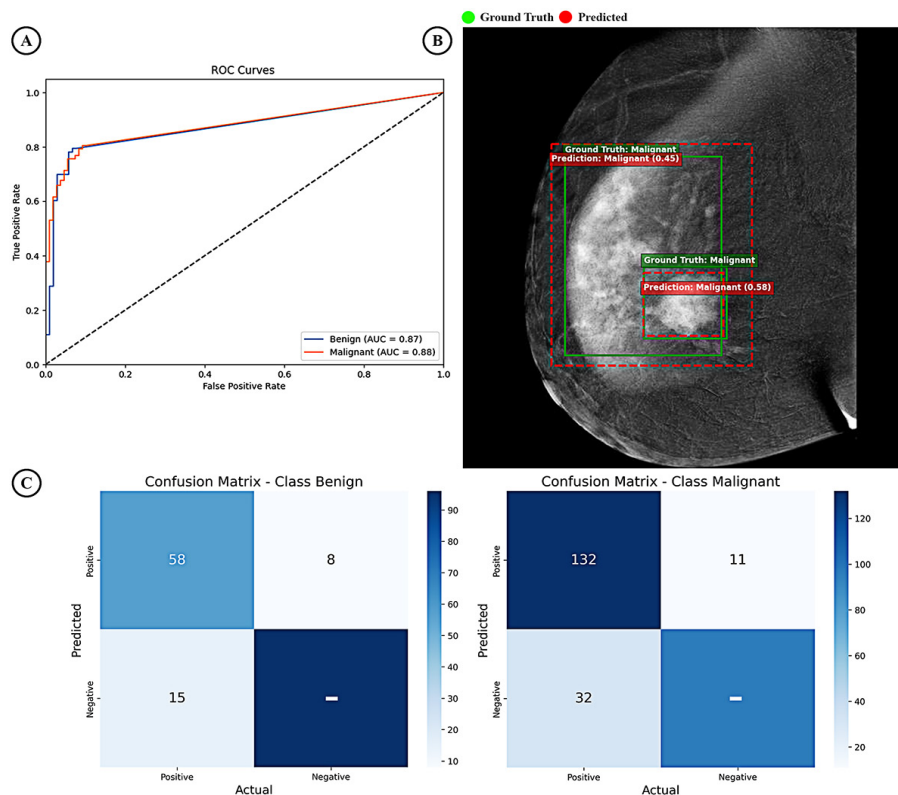


Figure 3: RetinaNet: (A) The Receiver Operating Characteristics (ROC) curves demonstrate Area Under Curve (AUC) values for the detection of benign and malignant lesions, indicating the model's discriminative ability between these classes. (B) Representative mammography image demonstrating model predictions with bounding boxes and confidence scores. (C) Confusion matrices for lesion classification (left: benign, right: malignant), highlighting accurate identification of true positives and minimizing false positives.

Although a direct comparison is not possible due to different datasets, the results show that YOLO-v12 significantly outperformed older versions employed in previous studies.

The YOLO-v12 architecture exhibits fundamental differences compared to versions v5, v8, v10, and v11, enhancing its performance in object detection tasks. Compared with previous iterations (YOLO-v1 to YOLO-v11) [23], YOLO-v12 incorporates R-ELAN within its Backbone component, which optimizes feature aggregation through block-level residual connections while enhancing training stability. Additionally, the Area Attention (A2) mechanism combined with FlashAttention [23] enables the model to intelligently concentrate on critical image regions, resulting in

superior detection of small or overlapping objects, such as benign and malignant lesions in mammography. Furthermore, FlashAttention [23] reduces computational overhead, thereby improving efficiency and consequently allowing YOLO version 12 to deliver higher accuracy and speed in lesion detection compared to earlier versions, leading to a significant reduction in false negative cases and achieving precise differentiation between benign and malignant classes. The implementation of FlashAttention facilitates optimization of memory access and computations related to the attention mechanism. This approach reduces unnecessary operations and improves data management in computational processes, enabling the model to execute faster with

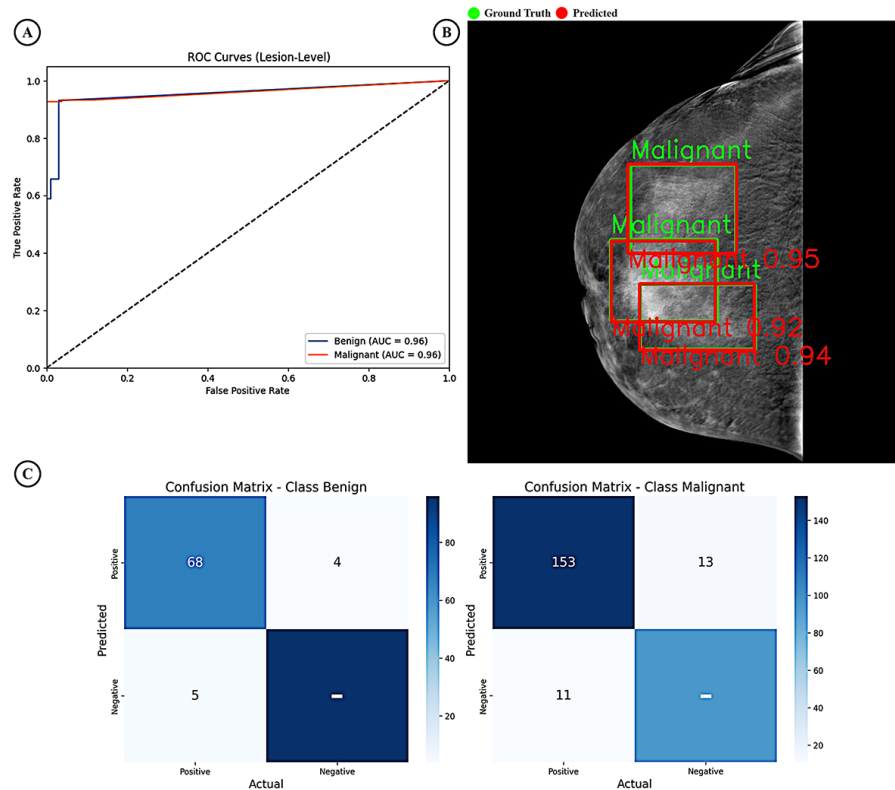


Figure 4: You Only Look Once-version 12- medium variant (YOLO-v12-m): **(A)** The Receiver Operating Characteristics (ROC) curves demonstrate Area Under Curve (AUC) values for the detection of benign and malignant lesions, indicating the model's discriminative ability between these classes. **(B)** Representative mammography image demonstrating model predictions with bounding boxes and confidence scores. **(C)** Confusion matrices for lesion classification (left: benign, right: malignant), highlighting accurate identification of true positives and minimizing false positives.

reduced resource consumption while maintaining accuracy, ultimately resulting in enhanced model performance.

The YOLO-v12-m model, comprising 20.2 million parameters, emerged as an optimal solution by achieving a well-calibrated balance between predictive accuracy and computational efficiency. With an inference time of 4.86 milliseconds on a T4 GPU, it demonstrates both outstanding processing speed and compliance with the stringent latency requirements of real-time clinical applications. This performance advantage stems primarily from the architectural refinements in YOLO-v12—most notably the integration of the FlashAttention mechanism—which optimizes

memory access patterns and eliminates redundant computations. Together, these attributes render YOLO-v12-m not only diagnostically powerful but also exceptionally well-suited for deployment in time-critical diagnostic workflows. FlashAttention substantially enhances processing speed without compromising diagnostic accuracy. Consequently, these findings confirm that the YOLO-v12-m model is not only a diagnostically powerful tool but also a pragmatically viable solution for implementation in real-time diagnostic workflows.

While contemporary deep learning models have achieved diagnostic accuracy levels comparable to those of average radiologists, the primary challenge lies in the fact that most

existing models have been trained exclusively on conventional DM images, as no suitable dataset for CESM images has been available until now. This research addresses this limitation by developing a CDD-CESM that incorporates both DM and CESM data types, aiming to enhance the development of superior medical decision support systems.

Despite the advantages of CESM technology, the development of deep learning models for this imaging modality faces three major limitations: (1) lack of diversity in available training data, which are predominantly collected from a single center and do not cover the full spectrum of imaging protocols, equipment variations, and patient demographic characteristics (such as age, ethnicity, and breast density), thereby limiting the model's generalizability to other methods such as Magnetic Resonance Imaging or to populations with different breast densities and diverse ethnic backgrounds; (2) a general scarcity of breast images for training deep learning models, which can reduce the system's performance and stability; and (3) technical and clinical factors, including the complexity of tumor size, shape, and texture, GPU memory limitations that hinder the use of high-resolution images, and the challenge of detecting lesions in dense breast tissue.

To substantially enhance the generalizability of findings, future investigations will concentrate on improving the model's capability to detect lesions within dense breast tissue through the incorporation of diverse datasets from multiple imaging centers into the current dataset, aiming to advance the training process and achieve more precise evaluation.

Conclusion

This study presents an advanced application of the YOLO-v12 model for detecting and classifying benign and malignant lesions in mammographic images. Evaluations show that this model not only outperforms prominent architectures such as RetinaNet and DETR,

but also surpasses previous versions of YOLO. This superiority underscores the potential of YOLO-v12 as a powerful decision-support tool for radiologists, enhancing both the accuracy and efficiency of diagnostic processes. Nonetheless, future research should focus on expanding the dataset to include images from multiple imaging centers to improve the model's generalizability, and on specialized assessments of its performance in dense breast tissues, which remain a major challenge in mammography.

Acknowledgment

We express our appreciation to the Artificial Intelligence Chatbot whose assistance in refining the manuscript ensured grammatical accuracy and the absence of spelling errors. The authors, nevertheless, bear full responsibility for all concepts, interpretations, and conclusions presented in this work.

Authors' Contribution

MA. Sakha contributed to the conception and design of the study, analysis and interpretation of the results, and drafting of the manuscript. A. Ameri contributed to the conception and design of the study, supervised the project, and reviewed the manuscript. All authors reviewed and approved the final version of the manuscript.

Ethical Approval

This study received ethical approval from the Institutional Review Board of Shahid Beheshti University of Medical Sciences (SBMU). The study was performed using the publicly available and anonymized CDD-CESM dataset. Given that all patient information was de-identified by the original data collectors, the need for informed consent was waived for this secondary analysis.

Conflict of Interest

None

References

- Giaquinto AN, Sung H, Miller KD, Kramer JL, Newman LA, Minihan A, et al. Breast Cancer Statistics, 2022. *CA Cancer J Clin.* 2022;**72**(6):524-41. doi: 10.3322/caac.21754. PubMed PMID: 36190501.
- Chintakunta JP, Lepakshi VA. Quantifying Recent State-of-Arts for Breast Cancer Segmentation, Detection and Classification: A Review. 2nd International Conference on Renewable Energy, Green Computing and Sustainable Development (ICREGCSD 2025); Hyderabad, India: E3S Web of Conferences; 2025. p. 12.
- Haerunisa L, Noviartha D. The Analysis Study of Diagnostic Performance and Accuracy of Mammography as Screening and Diagnostic of Breast Cancer: A Comprehensive Systematic Review. *Int J Med Sci Health Res.* 2025;**8**(2):25-45. doi: 10.70070/8a3ben43.
- Nićiforović D, Nikolić MB, Drvendžija Z, Nikolić O, Mijatović A, Lukač S, Stojanović S. Contrast-enhanced mammography in breast cancer screening: our experiences. *Vojnosanit Pregl.* 2025;**82**(2):86-93. doi: 10.2298/VSP240606005N.
- Dada EG, Oyewola DO, Misra S. Computer-aided diagnosis of breast cancer from mammogram images using deep learning algorithms. *J Electr Syst Inf Technol.* 2024;**11**(1):38. doi: 10.1186/s43067-024-00164-y.
- Rahman MM, Ghasemi Y, Suley E, Zhou Y, Wang S, Rogers J. Machine learning based computer aided diagnosis of breast cancer utilizing anthropometric and clinical features. *IRBM.* 2021;**42**(4):215-26. doi: 10.1016/j.irbm.2020.05.005.
- Ramadan SZ. Methods Used in Computer-Aided Diagnosis for Breast Cancer Detection Using Mammograms: A Review. *J Healthc Eng.* 2020;**2020**:9162464. doi: 10.1155/2020/9162464. PubMed PMID: 32300474. PubMed PMCID: PMC7091549.
- Eltrass AS, Salama MS. Fully automated scheme for computer-aided detection and breast cancer diagnosis using digitised mammograms. *IET Image Process.* 2020;**14**(3):495-505. doi: 10.1049/iet-ipr.2018.5953.
- Yoon JH, Kim EK. Deep Learning-Based Artificial Intelligence for Mammography. *Korean J Radiol.* 2021;**22**(8):1225-39. doi: 10.3348/kjr.2020.1210. PubMed PMID: 33987993. PubMed PMCID: PMC8316774.
- Watanabe AT, Lim V, Vu HX, Chim R, Weise E, Liu J, et al. Improved Cancer Detection Using Artificial Intelligence: a Retrospective Evaluation of Missed Cancers on Mammography. *J Digit Imaging.* 2019;**32**(4):625-37. doi: 10.1007/s10278-019-00192-5. PubMed PMID: 31011956. PubMed PMCID: PMC6646649.
- Demirel S, Urfalı A, Bozkır ÖF, Çelikten A, Budak A, Karataş H. Improving Mass Detection in Mammography Using Focal Loss Based RetinaNet. *Turk J Forecast.* 2023;**7**(1):1-9. doi: 10.34110/forecasting.1326245.
- Quiñones-Espín AE, Perez-Diaz M, Espín-Coto RM, Rodriguez-Linares D, Lopez-Cabrera JD. Automatic detection of breast masses using deep learning with YOLO approach. *Health and Technology.* 2023;**13**(6):915-23. doi: 10.1007/s12553-023-00783-x.
- Duque A, Zambrano C, Pérez-Pérez N, Benítez D, Grijalva F, Baldeon-Calisto M. Exploring the Use of Deformable Detection Transformers for Breast Mass Detection. IEEE Biennial Congress of Argentina (ARGENCON); San Nicolás de los: IEEE; 2024. p. 1-7.
- Shia WC, Ku TH. Enhancing Microcalcification Detection in Mammography with YOLO-v8 Performance and Clinical Implications. *Diagnostics (Basel).* 2024;**14**(24):2875. doi: 10.3390/diagnostics14242875. PubMed PMID: 39767236. PubMed PMCID: PMC11675152.
- Khaled R, Helal M, Alfarghaly O, Mokhtar O, Elkorany A, El Kassas H, Fahmy A. Categorized contrast enhanced mammography dataset for diagnostic and artificial intelligence research. *Sci Data.* 2022;**9**(1):122. doi: 10.1038/s41597-022-01238-0. PubMed PMID: 35354835. PubMed PMCID: PMC8967853.
- Carion N, Massa F, Synnaeve G, Usunier N, Kirillov A, Zagoruyko S. End-to-end object detection with transformers. In European conference on computer vision; Cham: Springer; 2020. p. 213-29.
- Wang M, Liu R, Luttrell Iv J, Zhang C, Xie J. Detection of Masses in Mammogram Images Based on the Enhanced RetinaNet Network With INbreast Dataset. *J Multidiscip Healthc.* 2025;**18**:675-95. doi: 10.2147/JMDH.S493873. PubMed PMID: 39935433. PubMed PMCID: PMC11812562.
- Alif MA, Hussain M. Yolov12: A breakdown of

- the key architectural features [Internet]. arXiv [Preprint]. 2025 [cited 2025 Feb 20]. Available from: <https://arxiv.org/abs/2502.14740>.
19. Sapkota R, Flores-Calero M, Qureshi R, Badgujar C, Nepal U, Poulouse A et al. YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series. *Artif Intell Rev*. 2025;**58**(9):274. doi: 10.1007/s10462-025-11253-3.
 20. Jegham N, Koh CY, Abdelatti M, Hendawi A. Yolo evolution: A comprehensive benchmark and architectural review of yolov12, yolo11, and their previous versions [Internet]. arXiv [Preprint]. 2024 [cited 2024 Oct 31]. Available from: <https://arxiv.org/abs/2411.00201>.
 21. Boudouh SS, Bouakkaz M. Breast cancer: toward an accurate breast tumor detection model in mammography using transfer learning techniques. *Multimed Tools Appl*. 2023;**82**(22):34913-36. doi: 10.1007/s11042-023-14410-4.
 22. Hassan NM, Hamad S, Mahar K. YOLO-based CAD framework with ViT transformer for breast mass detection and classification in CEM and FFDM images. *Neural Comput & Applic*. 2024;**36**(12):6467-96. doi: 10.1007/s00521-023-09364-5.
 23. Tian Y, Ye Q, Doermann D. Yolov12: Attention-centric real-time object detectors [Internet]. arXiv [Preprint]. 2025 [cited 2025 Feb 18]. Available from: <https://arxiv.org/abs/2502.12524>.