Critical Review on Deep Learning-Based Breast Cancer Detection and Segmentation: Challenges, Gaps, and Future Directions

Hassan Mahichi[1], Vahid Ghods[2,], Mohammad Karim Sohrabi[3], Arash Sabbaghi[4]

Department of Electrical and Computer Engineering, Se.C., Islamic Azad University, Semnan, Iran.
Email: V.ghods@iau.ac.ir (Corresponding author)

ABSTRACT:
Breast cancer remains a leading cause of cancer-related mortality among women worldwide, where early and accurate detection is vital for effective intervention and prognosis. Deep learning has emerged as a cornerstone in the automated analysis of breast cancer imaging, offering substantial improvements in tumor detection, segmentation, and classification across modalities such as mammography, ultrasound, and MRI. Despite notable progress, clinical integration remains constrained by challenges including limited dataset availability, suboptimal generalization, lack of interpretability, high computational complexity, and insufficient multi-task learning optimization. This critical review synthesizes findings from 70 peer-reviewed studies published between 2018 and 2025, encompassing convolutional neural networks, U-Net derivatives, Vision Transformers, instance segmentation models, and hybrid frameworks. Comparative evaluation highlights architectural strengths, modality-specific adaptations, and diagnostic performance metrics. Emphasis is placed on the comparative analysis of single-task versus multi-task frameworks, the integration of handcrafted features, transfer learning, and optimization strategies to improve model generalizability and robustness. Key limitations are identified in areas such as cross-domain robustness, real-time applicability, interpretability, and standardized benchmarking. Emerging solutions are examined, including self-supervised and semi-supervised learning strategies, lightweight and explainable architectures, adaptive loss balancing for MTL, cross-modal fusion techniques, and unified end-to-end pipelines.

KEYWORDS: Breast cancer detection, Segmentation, Medical Image Analysis, Deep learning, Convolutional neural network.

1. INTRODUCTION

Breast cancer remains the most commonly diagnosed cancer and the leading cause of cancer-related death among women worldwide, accounting for a substantial burden on global healthcare systems [1]. Early detection and accurate localization of breast tumors play a crucial role in improving prognosis, enabling timely intervention, and guiding therapeutic strategies. Medical imaging modalities—such as mammography, ultrasound (US), and magnetic resonance imaging (MRI)—are integral to breast cancer screening and diagnosis[2], [3]. However, the manual interpretation of these images by radiologists is subject to inter-observer variability, fatigue, and diagnostic error, especially in complex cases or dense breast tissues [4], [5]. In recent years, deep learning (DL) has emerged as a transformative technology in medical image analysis, particularly for tasks such as tumor detection, segmentation, classification, and localization. Architectures such as CNNs, transformers, and multi-task learning (MTL) frameworks have achieved remarkable success in various computer vision domains and are increasingly applied to breast imaging due to their ability to learn hierarchical features from large datasets [6]-[8]. Despite significant advancements, several critical challenges remain.

Paper type: Research paper
https://doi.org/xxx
Received: 25 February 2025, Revised: 5 April 2025, Accepted: 2 May 2025, Published: 1 June 2025
How to cite this paper: H. Mahichi, V. Ghods, M. K. Sohrabi, A. Sabbaghi, **"Critical Review on Deep Learning-Based Breast Cancer Detection and Segmentation: Challenges, Gaps, and Future Directions"**, *Majlesi Journal of Telecommunication Devices*, Vol. 14, No. 2, pp. 105-129, 2025.

These include the scarcity of annotated datasets, poor cross-dataset generalization, lack of model interpretability, and computational demands that limit real-time clinical deployment. Furthermore, many models are designed in isolation for either detection or segmentation, lacking the integration required for practical, end-to-end diagnosis workflows [6]-[9]. Moreover, while transformer-based models and hybrid architectures improve contextual understanding, their real-world applicability is often constrained by their complexity and high resource requirements [10], [11].

While several reviews have surveyed deep learning techniques in medical imaging, few have delivered a critical, task-specific, and end-to-end analysis tailored for breast cancer detection and segmentation across modalities. This critical review aims to systematically evaluate the current landscape of deep learning-based breast cancer detection and segmentation models. The primary objectives are to:

- Analyze recent methods across different imaging modalities and model types;
- Identify technical gaps, methodological inconsistencies, and clinical challenges;
- Highlight the strengths and weaknesses of existing detection, segmentation, and unified models;
- And propose future research directions that address outstanding issues and move toward more robust, interpretable, and deployable solutions.

Through an in-depth synthesis of over 70 peer-reviewed studies from 2018 to 2025, this review provides a comprehensive knowledge base for researchers, clinicians, and developers aiming to enhance AI-driven breast cancer diagnosis systems. By framing the discussion around methodological innovation, clinical utility, and scalability, this work contributes a forward-looking perspective on how deep learning can evolve into clinically reliable tools for breast cancer care.

2. CONCEPTUAL AND THEORETICAL FRAMEWORK

Deep learning-based breast cancer detection and segmentation are rooted in the intersection of medical imaging, artificial intelligence (AI), and pattern recognition. The theoretical framework for this research domain encompasses several key components: the nature of breast imaging modalities, the principles of deep neural networks, the role of encoder-decoder architectures for segmentation, attention mechanisms, transformers, and the emerging trend of multi-task learning. This section provides a detailed overview of these concepts and explains how they form the scientific foundation of modern breast cancer analysis systems.

2.1. Imaging Modalities in Breast Cancer Diagnosis

Breast cancer diagnosis relies on various imaging techniques, each providing distinct structural and textural information. The three most widely used modalities in clinical and research settings include [4]:

Mammography: This is the gold standard for breast cancer screening. Mammograms are grayscale 2D X-ray images that reveal calcifications, asymmetries, and masses. Their high resolution is beneficial, but the modality often suffers from reduced sensitivity in dense breast tissues [12].

Ultrasound (US): Ultrasound imaging is non-invasive, inexpensive, and safe. It is frequently used as an adjunct to mammography, particularly for younger women or patients with dense breast tissues. However, ultrasound images are often noisy, operator-dependent, and may exhibit lower spatial resolution, making automated analysis more challenging [13].

Magnetic Resonance Imaging (MRI): Breast MRI offers high sensitivity and 3D anatomical detail. It is especially useful for high-risk patients and for evaluating tumor extent. Despite its benefits, it is expensive, time-consuming, and requires expert interpretation [14]

Each modality introduces unique challenges in image interpretation and thus influences the architecture and preprocessing strategies of deep learning models (Fig. 1).

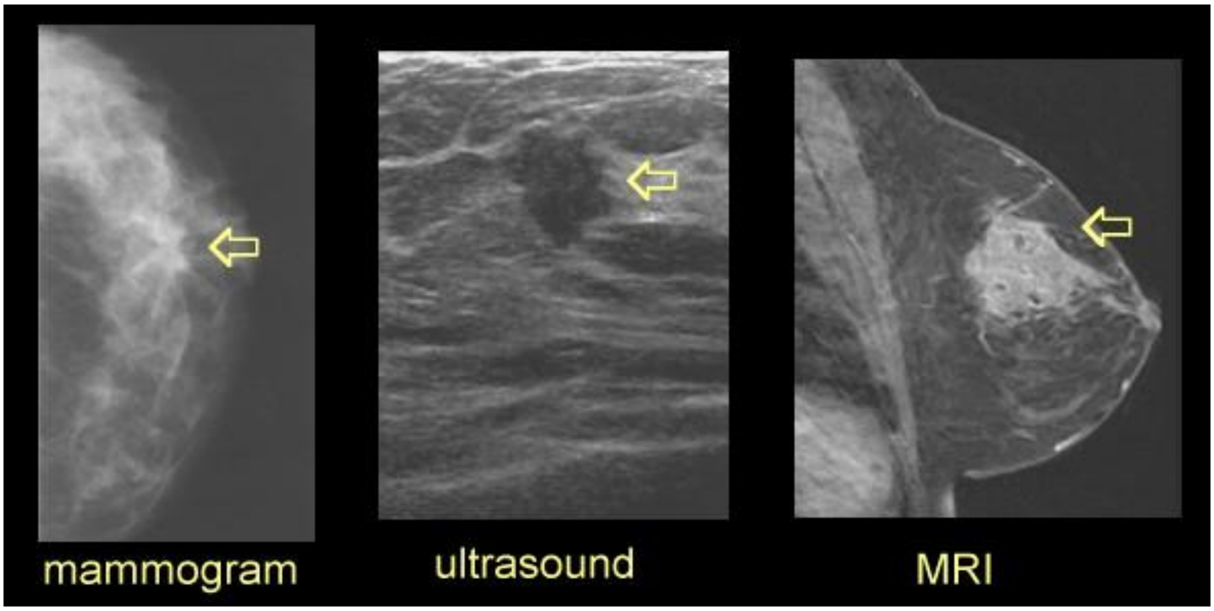

Fig. 1. breast cancer imaging [15].

2.2. Fundamentals of Deep Learning in Medical Imaging

Deep learning is a subset of machine learning that automatically learns hierarchical representations from data. It has revolutionized computer vision tasks and shown exceptional promise in the medical imaging domain [16]. Convolutional Neural Networks (CNNs) are the most widely used architecture in breast cancer diagnosis tasks due to their ability to capture spatial features, detect edges, and model patterns in images with varying levels of abstraction [17]. CNNs consist of layers such as convolution, pooling, and fully connected layers [18], [19].

Deeper CNNs like ResNet (with residual skip connections) [19], DenseNet [20] (which encourages feature reuse), and InceptionNet [21] (multi-scale feature extraction) have improved image classification performance by addressing issues like vanishing gradients and overfitting [22]. These architectures are commonly used for tumor detection and classification tasks and have been further fine-tuned using transfer learning approaches in medical contexts [23].

However, standard CNNs are typically limited to image-level predictions. For tasks such as delineating tumors pixel by pixel, specialized architectures like encoder-decoder networks are required.

2.3. Encoder-Decoder Architectures for Tumor Segmentation

Segmentation, which involves delineating tumor boundaries at the pixel level, is a critical task in medical imaging. U-Net [24], introduced by Ronneberger et al. [25], is the foundational architecture for biomedical segmentation (Fig. 2). It follows an encoder-decoder scheme, where the encoder progressively downsamples the input to extract features, and the decoder upsamples these features to the original resolution, enabling fine-grained segmentation [25]. Key enhancements to the U-Net architecture (Fig. 2) have been proposed to address limitations such as class imbalance, low contrast, and small lesion size:

- **U-Net^{++}** introduces dense skip pathways to bridge semantic gaps between encoder and decoder features [26].
- **Attention U-Net** incorporates attention gates that allow the model to focus on relevant tumor regions while suppressing irrelevant background [27].
- **ResUNet**, **V-Net**, and **DenseUNet** adapt popular classification backbones into segmentation frameworks to enhance feature richness [28].

These models have shown superior performance on various breast imaging datasets; however, their generalization across datasets and imaging conditions remains a challenge.

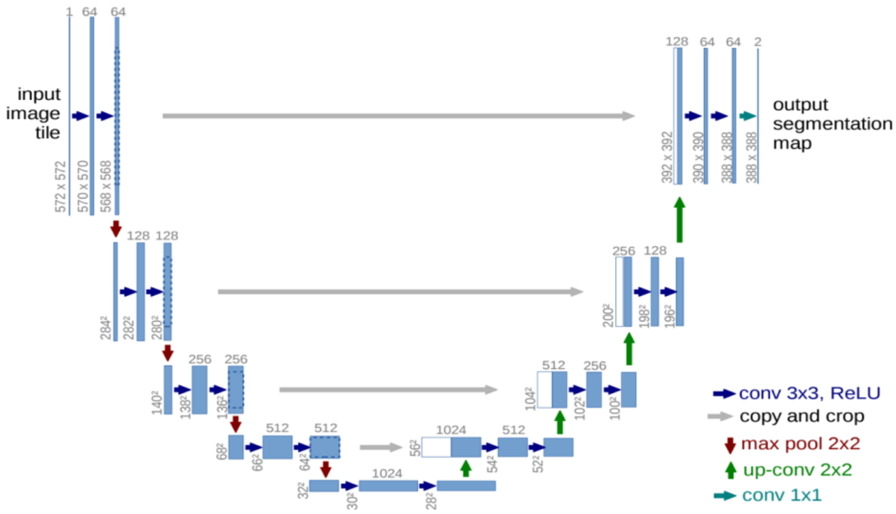

Fig. 2. U-net architecture [29].

2.4. Role of Attention Mechanisms

Inspired by human visual attention, attention mechanisms enable neural networks to focus selectively on important regions of an image[30]. In breast cancer segmentation, attention modules help improve model robustness by suppressing irrelevant background and emphasizing tumor boundaries. The most common types include channel attention, spatial attention, and self-attention mechanisms [30]. Attention U-Net (Fig. 3) and its variants improve segmentation quality, particularly when tumors are small or embedded in complex tissue structures. Moreover, coordinate attention and non-local attention blocks have also been applied to improve feature representation in noisy and low-contrast ultrasound and mammogram images [14].

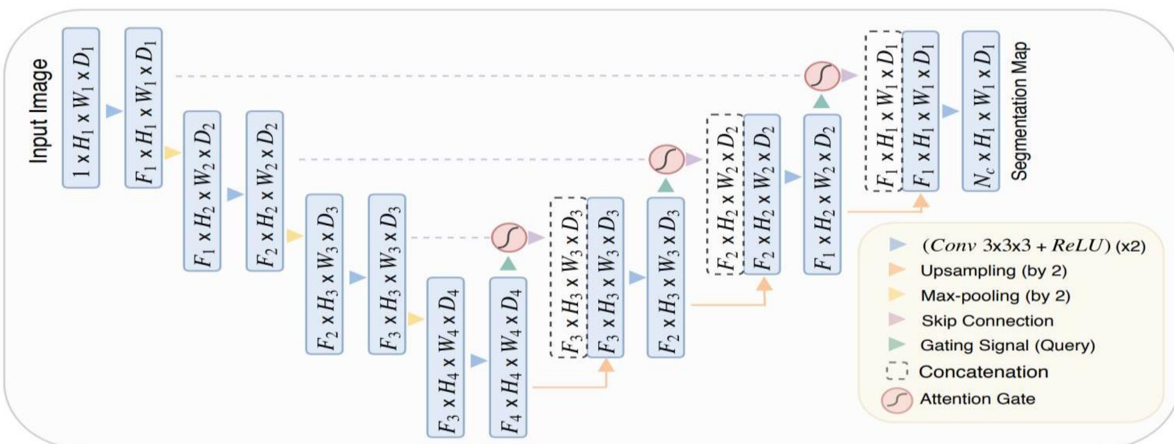

Fig. 3. Attention U-Net [14].

2.5. Transformers in Medical Imaging

While CNNs are effective for local feature extraction, they often struggle with modeling long-range dependencies. Vision Transformers (ViTs) [31], introduced in 2020, overcome this limitation by treating images as sequences of patches and applying self-attention across the entire image (Fig. 4). Although initially developed for natural images, transformers have been successfully adapted for medical tasks such as tumor segmentation and detection [31].

TransUNet, one of the first hybrid models, combines a transformer-based encoder with a U-Net decoder, achieving state-of-the-art results in several biomedical segmentation benchmarks [32]. Transformers provide better global context modeling, which is especially useful in breast cancer detection where the tumor may be diffuse, poorly defined, or appear in multiple regions [31], [32].

However, transformer-based models typically require large datasets and significant computational resources. This poses limitations in medical applications where annotated datasets are often limited. To address this, recent studies have explored lightweight transformer variants and semi-supervised training techniques to make these models more practical [33], (Fig. 4).

(a) standard Transformer architecture (b) Transformer for tumor classification

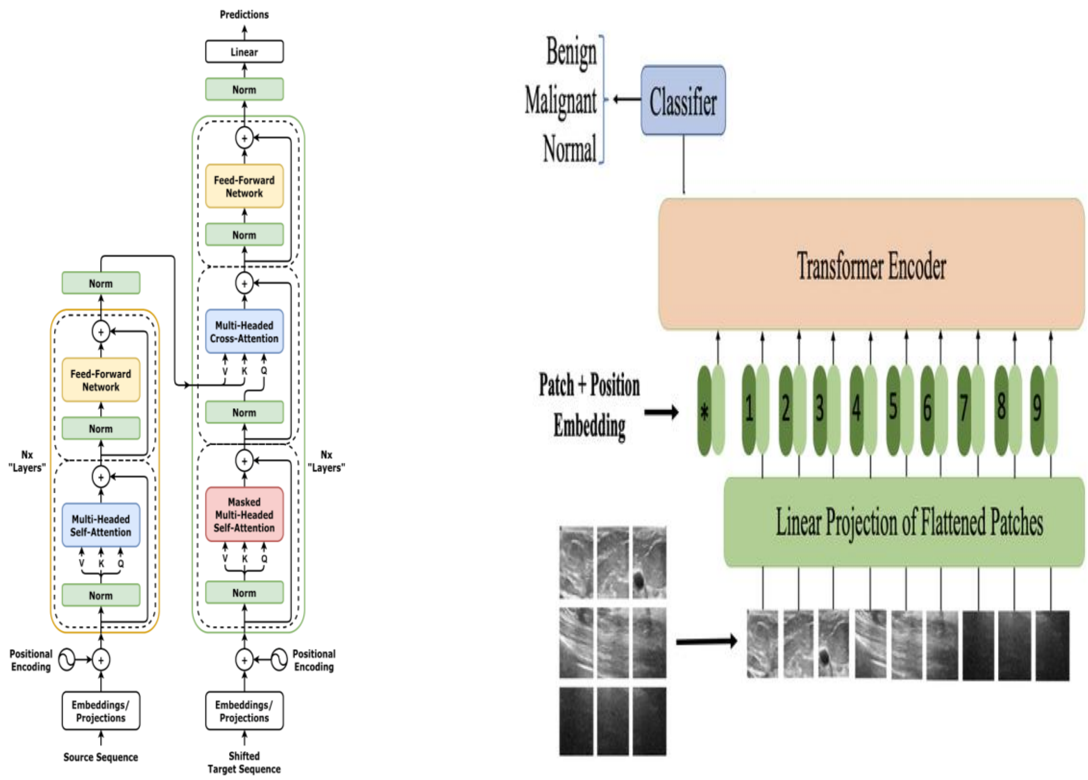

Fig. 4. Transformers [31]: (a): simple standard Transformer architecture, and (b): example of using transformer encoder for breast tumor classification.

2.6. Multi-Task Learning Paradigm

Traditional models perform either detection or segmentation as independent tasks. However, breast cancer diagnosis in clinical practice often involves both tasks simultaneously. Multi-task learning (MTL) addresses this by training a single model to optimize multiple objectives (e.g., classification and segmentation) concurrently, encouraging shared feature learning and improving model efficiency [8]. Shared encoder-decoder structures are commonly used in MTL, with task-specific heads for segmentation and classification. While MTL (Fig. 5) can improve performance and reduce training time, balancing the learning of each task is a challenge. Improper loss weighting or architectural imbalance can lead to dominance of one task over the other [34].

Recent works have introduced dynamic loss balancing, task-specific attention modules, and cross-task consistency learning to enhance MTL performance in breast cancer imaging [17]. These frameworks align well with real-world clinical workflows and represent an important trend in future model development [31], [34].

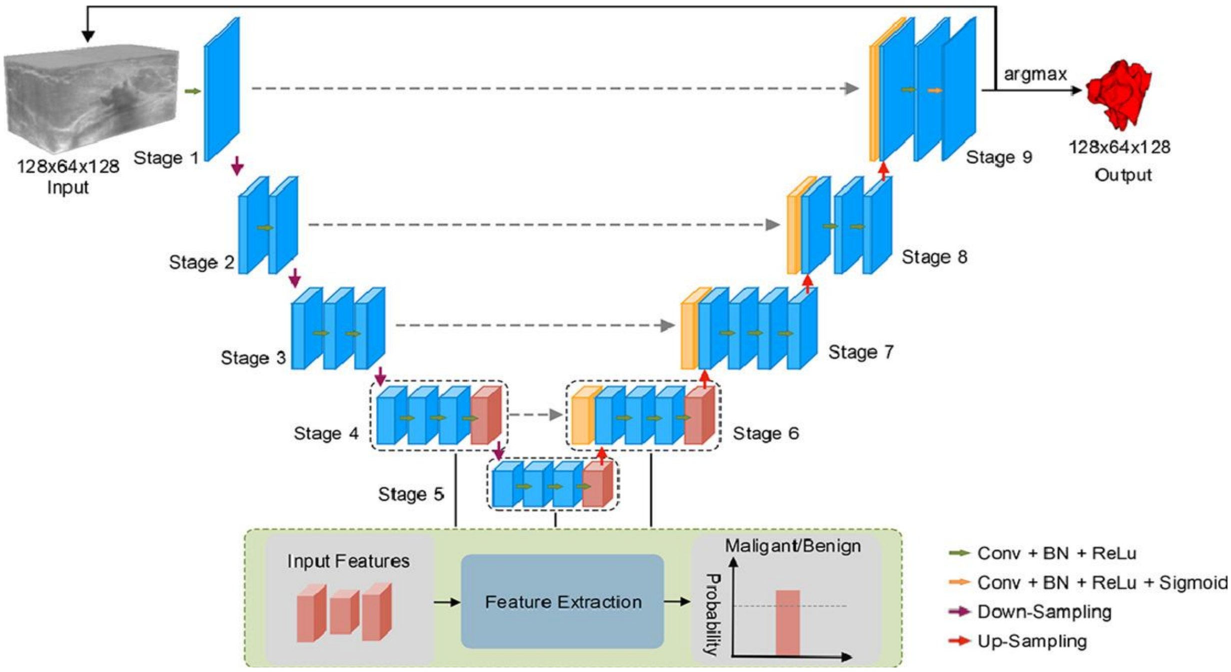

Fig. 5. Multi-Task Learning [35].

3. LITERATURE SELECTION APPROACH

To ensure the integrity, relevance, and scientific rigor of this critical review, a systematic and reproducible literature selection strategy was adopted. The focus was to identify high-quality, peer-reviewed research works that employed deep learning approaches for breast cancer detection, classification, and segmentation across different imaging modalities, including mammography, ultrasound, and MRI. The literature selection methodology adhered to widely accepted guidelines such as PRISMA and was guided by a multi-phase screening process encompassing database querying, inclusion/exclusion filtering, and quality assessment.

Initially, a comprehensive search was conducted across five major academic databases—IEEE Xplore, Scopus, PubMed, Web of Science, and ScienceDirect—owing to their extensive coverage of medical imaging, computer vision, and AI research. The search queries were designed using a combination of targeted keywords and Boolean operators. Typical queries included phrases such as: (“breast cancer” OR “mammography” OR “ultrasound” OR “MRI”) AND (“deep learning” OR “CNN” OR “transformer” OR “multi-task learning” OR “semantic segmentation”) AND (“detection” OR “classification” OR “segmentation” OR “localization”). To enhance the specificity of the search, filters were applied to restrict the publication period from January 2015 to July 2025, including only peer-reviewed articles, and exclude non-English publications and preprints.

The literature selection process proceeded in multiple stages. In the first stage, all retrieved articles were screened based on titles and abstracts. Articles that did not clearly indicate the use of deep learning for breast cancer analysis or that focused exclusively on traditional machine learning, radiomics without DL components, or non-imaging modalities were excluded. In the second stage, the remaining papers were subjected to full-text review to determine methodological depth, dataset transparency, and the presence of quantitative performance metrics. Only those papers which provided sufficient implementation detail, comparative evaluation with state-of-the-art methods, and performance results on benchmark datasets (e.g., INbreast, DDSM, BUSI, CBIS-DDSM)[36] were retained.

To maintain the scientific quality of this review, an internal quality appraisal framework was used. This framework evaluated each article against criteria such as clarity in problem formulation, innovation in methodology, use of public or multi-institutional datasets, inclusion of ablation studies, reproducibility of results, and relevance to breast imaging. Articles were rated on a 5-point scale for each criterion, and only those achieving a minimum threshold across all dimensions were included. This step ensured that the review captures robust, and impactful studies rather than merely reiterating superficial trends.

In total, 70 primary studies were selected for detailed analysis. These works covered a diverse range of deep learning techniques, from standard CNNs to advanced architectures such as Vision Transformers [4], hybrid attention mechanisms [12], U-Net variants [25], and generative adversarial networks (GANs) used for data augmentation or unsupervised segmentation [37]. The selected studies also addressed different challenges such as multi-task learning [8],

[34]-[35], class imbalance [5], [7], [38]-[40], model generalization across domains [37], [41]-[42], and interpretability [4], [20], [31], [43]-[46], reflecting the breadth and complexity of the current research landscape. Special attention was paid to recent innovations such as lightweight DL models for edge deployment [47], multimodal learning frameworks integrating clinical metadata [48], [49], and large-scale self-supervised pretraining [50].

Furthermore, priority was given to studies published in high-impact journals such as IEEE Transactions on Medical Imaging, Medical Image Analysis, Nature Scientific Reports, Computer Methods and Programs in Biomedicine, and *Pattern Recognition*. Inclusion of recent systematic reviews and meta-analyses published between 2020 and 2025 helped validate the selection strategy and ensured thematic consistency.

4. CRITICAL ANALYSIS OF DETECTION METHODS

Over the past decade, CNNs have become the dominant framework for breast cancer detection due to their ability to extract high-level semantic features from medical images[5], [13], [17], [51]-[55]. Despite considerable improvements in detection accuracy, these methods exhibit diverse performance characteristics depending on architecture, dataset, optimization technique, and modality.

Traditional CNN Approaches: Early CNN-based models for breast cancer detection, such as those by Mikhailov et al.[56], focused on relatively shallow architectures applied to histopathological images, achieving moderate accuracy (~85%) but struggling with generalization across magnification levels and datasets [56]. Later studies incorporated deeper networks and transfer learning. For instance, Gonçalves et al. utilized ResNet50, DenseNet201, and VGG16 with transfer learning, significantly improving classification accuracy to 91.67% on the DMR-IR dataset [57]. However, these models often relied on pre-trained weights from natural image datasets, introducing domain mismatch issues that limit their performance in breast imaging.

Optimization-Based CNN Enhancements: To address the limitations of conventional architectures, optimization techniques such as evolutionary algorithms have been integrated. The BreastCNet model [57], for instance, introduced a dual-optimizer approach using the Grey Wolf Optimizer (GWO) [58] to fine-tune neuron counts in dense layers and the Parrot Optimizer (PO) [59] to adjust the learning rate dynamically. These strategies mitigated overfitting, reduced convergence time, and improved generalization, yielding 98.10% validation accuracy and an AUC of 0.995 on the BUSI dataset. The integration of multi-task learning for simultaneous classification and bounding box regression further distinguished BreastCNet from traditional single-task CNNs, enabling efficient, real-time diagnostic inference [5].

Transfer Learning and Feature Fusion: Numerous studies have leveraged pre-trained models such as ResNet, EfficientNet, and DenseNet to transfer learned features from large-scale datasets like ImageNet to the breast imaging domain [60], [61], [62]. While EfficientNet-based models have shown good trade-offs between accuracy and parameter count, they are typically optimized for classification only and require architectural adaptation for localization. For example, Petrini et al.'s EfficientNet-based CAD system achieved 0.9344 AUC on CBIS-DDSM, yet lacked integrated lesion localization capabilities [9]. Similarly, YOLO-based detectors (YOLOv4, YOLOv5) demonstrated real-time performance but required additional classification branches for dual-task capability, adding complexity and memory overhead [63], [9].

Multi-Task Learning (MTL) for Joint Detection and Localization: MTL-based models have gained attention for their ability to share feature representations across related tasks. Ding et al. introduced a ResNet-GAP architecture that jointly performed classification and localization in ultrasound images, though with limited optimization and modest accuracy (88.6%)[5], [64]. In contrast, the BreastCNet architecture offered significant innovation by integrating MTL with hybrid optimization, allowing efficient learning of both tasks in a unified framework [57]. Compared to previous models, BreastCNet's dual-branch output achieved higher F1-scores (0.98) and Intersection over Union (IoU) scores (0.96), indicating superior localization precision [5].

Lightweight Models and Clinical Feasibility: While heavy models like YOLOv4 (~60M parameters) offer state-of-the-art detection accuracy, their computational cost poses barriers for clinical deployment. Lightweight architectures such as MobileNetV2 [41], SqueezeNet [42], and EfficientNet-B0 [21] offer fewer parameters (~3–5M) and reduced inference latency, but require additional modules for localization and lack specialization for medical image characteristics. BreastCNet, with ~1.12M parameters and ~1.58 GFLOPs, achieves a practical balance between performance and efficiency, making it suitable for edge deployment in clinical workflows [57].

Generalization and Cross-Modality Validation: Many CNN models suffer from limited generalization due to overfitting on specific datasets[65]. BreastCNet's training on BUSI and validation on DDSM and INbreast datasets confirmed its robust cross-modality performance. It achieved 97.50% accuracy on DDSM and 96.20% on INbreast, outperforming other models that showed performance degradation when exposed to unseen data [57], [23], [35].

Interpretability in Detection Models: Interpretability is essential for clinical adoption. Grad-CAM was used in BreastCNet to visualize the decision-making process, improving trust and transparency [57]. However, the most detection models in the literature do not incorporate sufficient interpretability tools, reducing their acceptance in clinical settings[5], [57].

In conclusion, current detection methods in breast cancer imaging show a clear progression from basic CNN architectures to sophisticated multi-task, optimization-enhanced, and lightweight models. While high accuracy has been achieved in many studies, key limitations persist in terms of computational cost, lack of real-time integration, generalization across imaging modalities, and explainability.

5. CRITICAL ANALYSIS OF SEGMENTATION METHODS

Segmentation of breast tumors from medical images plays a vital role in identifying the extent and location of the malignancy, aiding in treatment decisions. With advances in deep learning techniques, CNN based methods have gained attention for their ability to achieve an accurate and automated segmentation [66]. Early and accurate segmentation of breast tumors is essential for precise diagnosis, treatment planning, and monitoring. CNN methods have emerged as powerful tools for medical image analyses, including breast cancer segmentation [67]. Segmentation methods, such as U-Net[1], YOLO[2], Mask R-CNN, and contour-based methods, have been extensively studied and developed for various applications in computer vision and medical imaging [68]-[73]. However, instance segmentation goes beyond semantic segmentation by not only assigning class labels, but also distinguishing between individual instances of objects within a given class [74]. It involves identifying and segmenting each distinct object instance separately, providing precise localization and differentiation of objects. It is typically built upon object detection algorithms, and extends them to include pixel-level segmentation masks for each detected object instance [74].

Other segmentation methods, such as graph-cut algorithms, region-based approaches, and deep learning-based architectures, have also been explored and developed for various segmentation tasks [72], [73]. These methods aim to improve segmentation accuracy, efficiency, and generalization capabilities. However, the main problem is that a highly accurate and efficient instance segmentation of breast cancer tumors in images should be reached, a challenging task due to varying tumor shapes, sizes, and appearances. Methods range from basic CNN architectures to more advanced models, such as U-Net, Mask R-CNN, and DenseNet for medical image segmentation tasks. While these methods are likely to be promising, they often struggle with complex variations in tumor presentation and suffer from lower generalization across diverse datasets[12]. Currently, models like Mask R-CNN and U-Net are considered among the best for medical image segmentation tasks due to their ability to handle detailed pixel-wise segmentation and detection. However, even the most advanced models have limitations in terms of generalizability, especially when applied to different datasets like BUSI, DDSM, and INbreast. Breast cancer segmentation methods like UNet, VGG16, and similar CNN-based models face multiple limitations[12], [75]. These include poor generalization across different imaging datasets due to overfitting, difficulty in accurately detecting small, irregular or diffuse lesions, and high sensitivity to noise and artifacts in medical images. Additionally, the computational cost and resource requirements of these models make real-time clinical application challenging, particularly in resource-limited settings. The models also lack interpretability, limiting clinical trust, and fail to effectively capture multiscale information, which is critical for segmenting tumors of varying shapes and sizes. These methods also lack full automation for clinical workflows and are unable to perform precise instance segmentation, which could otherwise distinguish between multiple, adjacent tumor instances [12].

Despite considerable advancements, current segmentation methods face key limitations, particularly in generalization, precision, and clinical applicability.

5.1. Traditional CNN-Based Segmentation: From U-Net to Its Variants

The seminal U-Net architecture by Ronneberger et al. [25] revolutionized biomedical segmentation through its encoder-decoder framework with skip connections, enabling pixel-level precision with relatively few annotated samples. However, baseline U-Net often underperforms in segmenting small or low-contrast lesions, a common issue in mammography and ultrasound imaging. To address this, U-Net++ introduced nested dense skip pathways, reducing the semantic gap between encoder and decoder features [76]. Meanwhile, Attention U-Net incorporated attention gates to emphasize tumor-relevant features [77].

Despite these innovations, several studies have highlighted U-Net's limitations. Its rigid architecture struggles with significant shape variability and inter-patient heterogeneity. For example, experiments on the BUSI and DDSM datasets revealed sensitivity drops when segmenting irregular or low-contrast tumors, particularly in ultrasound and dense breast tissues[78], [79]. Moreover, performance often degrades when models trained on one dataset are applied to another, due to domain shift and overfitting [12], [80]-[83].

---

[1] U-shaped encoder-decoder network
[2] You only look once

5.2. Instance Segmentation: Mask R-CNN and Beyond

Instance segmentation, unlike semantic segmentation, aims to identify each object instance separately. Mask R-CNN has emerged as a leading architecture for this task by extending Faster R-CNN with an additional segmentation head, allowing simultaneous object detection and pixel-level mask generation [13], [55], [84]-[85]. In breast imaging, Mask R-CNN has shown superior results on datasets like INbreast and CBIS-DDSM, especially for well-defined lesions [86]-[87]. However, its reliance on bounding box proposals makes it less effective for diffuse or small-scale tumors. Additionally, the model's computational complexity (~140M parameters) hinders real-time clinical application, particularly in low-resource settings [13].

To reduce computational burden, lightweight adaptations like Lite-Mask R-CNN and EfficientDet+Mask modules have been proposed. While these models maintain reasonable segmentation performance (IoU ~0.82 on BUSI), they often sacrifice fine-grained accuracy and are prone to false positives in noisy modalities like ultrasound [53].

5.3. Transformer-Based and Hybrid Architectures

Transformers have recently been incorporated into segmentation pipelines to model long-range dependencies. TransUNet, Swin-UNet, and MedT combine CNN backbones with transformer encoders, yielding better contextual awareness and improved Dice coefficients, particularly in MRI segmentation [88]-[89]. However, their reliance on large-scale datasets and high GPU memory limits widespread adoption in clinical practice.

Hybrid models like CNN-ViT and Res-TransUNet have been proposed to balance spatial feature extraction and global context modeling [90]. For instance, Res-TransUNet achieved a 3–5% gain in Dice and Jaccard scores over pure CNN models on the BUSI dataset, especially for multi-region tumors [90] .However, these models require careful tuning of transformer depth and attention granularity to avoid overfitting in small datasets.

5.4. Graph-Based and Edge-Aware Segmentation

To overcome boundary ambiguity, several studies have explored graph-cut methods and edge-aware segmentation. Deep contour-aware networks (e.g., DCAN) and GCN-integrated U-Nets leverage boundary supervision to improve precision around lesion edges[91]-[92]. This is especially beneficial in ultrasound, where tumors often lack clear demarcation. However, integrating these modules increases architectural complexity and training time.

Recent efforts using graph neural networks (GNNs) in post-processing stages have enhanced connectivity preservation and reduced artifacts [93]. For example, GNN-refined U-Nets improved boundary recall by 6% on BUSI but required additional annotations for edge maps, limiting scalability [93].

5.5. Generalizability, Robustness, and Cross-Modality Performance

A recurring challenge in breast cancer segmentation is the generalization of models across imaging modalities and datasets. Many studies report high Dice coefficients (>0.90) on training datasets but observe significant performance drops on external validation sets due to domain shifts in image quality, resolution, and annotation styles [31,33,80]. Few-shot learning, domain adaptation (e.g., CycleGAN-based style transfer), and data augmentation using generative adversarial networks (GANs) have shown promise in mitigating this issue [37], [41], [94].

However, GAN-based augmentation, while beneficial, can introduce unrealistic samples that skew model learning. Unsupervised domain adaptation techniques remain underexplored in breast cancer imaging compared to other fields like retinal or brain imaging [36].

5.6. Computational Constraints and Clinical Integration

Many of the top-performing segmentation models—such as Transformer-based or hybrid networks—demand significant memory, long inference times, and are not optimized for real-time application. Lightweight segmentation architectures like Mobile-UNet and Squeeze-UNet have addressed this by reducing parameter counts while maintaining reasonable accuracy (Dice ~0.85) [95]. Nonetheless, a trade-off persists between efficiency and performance.

Moreover, clinical adoption is hindered by the lack of model interpretability. Few segmentation frameworks integrate explainable AI (XAI) tools like Grad-CAM or saliency maps, which are vital for clinician trust[4], [47], [65], [88]. Embedding such mechanisms in future models is essential for transparency in segmentation decisions.

Despite the considerable progress in deep learning-based segmentation methods for breast cancer, several unresolved issues remain. One major challenge is the poor generalization of models across different imaging modalities and institutions, often caused by dataset bias and overfitting. Additionally, many models struggle with accurately detecting small, irregular, or overlapping tumors, particularly in low-resolution or noisy images such as those produced by ultrasound. The high computational complexity of advanced architectures further limits their feasibility for real-time deployment on edge devices, especially in resource-constrained clinical environments. Moreover, the lack of interpretability and clinical explainability in most segmentation models diminishes their acceptance and trust among radiologists. Lastly, there is limited integration of multi-modal information, such as patient metadata and radiomics features, which could otherwise enhance diagnostic precision and model robustness.

6. CRITICAL ANALYSIS OF HYBRID AND MULTI-TASK LEARNING MODELS

Hybrid models and multi-task learning (MTL) frameworks have recently emerged as promising paradigms in breast cancer analysis due to their capacity to jointly address related objectives—such as classification, detection, and segmentation—within a single, integrated architecture. These frameworks aim to enhance computational efficiency, promote feature sharing, and improve predictive consistency across tasks. However, several fundamental challenges still hinder their full potential in clinical applications.

6.1. Efficiency vs. Specialization Trade-off

MTL models are increasingly adopted in breast cancer analysis due to their ability to jointly handle tasks such as tumor detection and segmentation within a shared computational framework. A central advantage of these models lies in their computational efficiency: by employing a shared encoder with dedicated decoders for each task, MTL architectures significantly reduce model size, training time, and inference latency. This shared representation facilitates the transfer of common features, which can enhance generalization, particularly when data resources are limited [96].

However, this shared structure also introduces a fundamental trade-off between efficiency and task-specific specialization. While shared layers capture general features beneficial to multiple tasks, they may struggle to model the fine-grained distinctions needed for each individual objective. This can result in task interference, where gradients from competing tasks update shared parameters in conflicting directions, ultimately hindering convergence or degrading performance on one or more tasks [97].

To address these challenges, recent studies have explored adaptive sharing strategies, such as soft-parameter sharing or modular encoders with attention-based routing, which aim to maintain efficiency while allowing task-specific specialization [96], [98]. Nevertheless, determining the optimal balance between shared and task-specific components remains a significant design challenge in MTL architectures, particularly in complex medical imaging scenarios where tasks demand differing levels of semantic abstraction.

6.2. Architectural Complexity and Optimization Challenges

Architectural design plays a central role in the effectiveness of hybrid models. Simple hard parameter sharing (i.e., full encoder sharing) often fails to capture task-specific nuances. More sophisticated designs, such as partially shared encoders, gated multi-branch networks, and task-aware attention mechanisms, aim to separate shared and private representations more effectively. For example, Task-Aware Attention Networks (TAAN) have shown notable improvements by dynamically routing features to task-specific decoders based on learned attention maps [2], [52].

Nonetheless, determining the optimal depth, width, and connection schema for shared versus private layers remains an open research problem. There is no one-size-fits-all solution, especially for medical imaging tasks that differ in spatial precision and semantic interpretation.

6.3. Joint Loss Function Formulation

The joint optimization of multiple tasks relies heavily on well-designed loss functions. Most hybrid models combine individual losses such as Dice loss (for segmentation), focal loss (for detection), and cross-entropy loss (for classification). A fixed weighting of these losses can cause training imbalance, where one task dominates the learning process. Adaptive approaches, like Uncertainty Weighting [99], Gradient Normalization (GradNorm) [75], [100] and Adaptive Loss Balancing (ALB), offer dynamic weighting schemes to stabilize training and improve convergence across all tasks.

However, these strategies are computationally intensive and may still be sensitive to hyperparameter tuning, making real-world implementation more complex.

6.4. Evaluation Metrics and Benchmarking Inconsistencies

Evaluating hybrid models is challenging due to the diversity of metrics involved. Typically, segmentation is assessed using Dice coefficient and IoU [101], while detection/classification uses AUC [102] accuracy, sensitivity, and specificity. Reporting these metrics independently makes it difficult to compare models holistically. Furthermore, a model that excels in one task might underperform in another, which is not acceptable in clinical workflows.

Some researchers have proposed composite evaluation metrics or task-weighted scoring systems to provide a balanced perspective [35], [96]-[97]. However, such methods are not yet standardized, leading to inconsistencies across published studies.

6.5. Data and Annotation Bottlenecks

High-quality labeled datasets are essential for effective multi-task learning. Unfortunately, datasets that provide aligned annotations for all tasks (e.g., segmentation masks along with class/detection labels) are rare in the medical imaging domain[12], [103]. Many publicly available datasets offer either classification labels or segmentation masks, but not both.

To mitigate this, hybrid training schemes have been proposed, including semi-supervised learning, weak supervision, and pseudo-label generation. For instance, self-supervised pre-training followed by fine-tuning on limited labeled data has shown promise [50]. Nevertheless, the lack of comprehensive, multi-task annotated datasets remains a significant bottleneck.

6.6 Generalization and Clinical Translation Challenges

Despite their academic success, hybrid models often face performance degradation when applied to real-world clinical data due to domain shifts—variations in imaging modality, scanner settings, and patient demographics. Models trained on public datasets like BUSI or CBIS-DDSM may fail to generalize to images from other institutions or populations [12], [82], [104].

Transfer learning and domain adaptation techniques have been employed to mitigate this issue, but the robustness and interpretability of these models are still under scrutiny [39]. Clinical translation also requires transparency and reproducibility—qualities that many current MTL models lack due to their complex architecture and black-box nature.

hybrid and multi-task learning frameworks present a compelling solution for unified breast cancer diagnosis, offering the potential to perform classification, detection, and segmentation in a single model. While these models show significant promise in reducing computational redundancy and improving diagnostic coherence, several challenges persist. Future research must address task interference, architectural optimization, adaptive loss balancing, and data scarcity. Moreover, cross-domain generalization and clinical interpretability should be prioritized to ensure successful real-world deployment of these advanced AI systems.

7. COMPARATIVE ANALYSIS AND SUMMARY

This section presents a critical comparative analysis of recent deep learning approaches for breast cancer detection, segmentation, and unified end-to-end diagnosis frameworks. By systematically evaluating performance metrics, architectural design, computational demands, and clinical relevance, we identify the strengths, limitations, and gaps across state-of-the-art models. The methods are categorized into three core areas: (1) tumor detection, (2) tumor segmentation, and (3) integrated models capable of simultaneous detection and segmentation. Within each category, related models are grouped and contrasted based on shared methodologies or design goals, enabling a more cohesive understanding of trends and challenges in the field. This structured comparison facilitates the identification of promising directions for future research and practical deployment in clinical settings.

7.1. Breast cancer detection

The landscape of breast cancer detection using deep learning is rich with diverse approaches, ranging from classical CNNs to transformer-based and optimization-enhanced architectures. Traditional models, like those used by Das & Rana [100] and Kumar et al. [103], employ deep CNNs or ResNet variants but primarily focus on classification without localization. This limits their clinical utility, as they fail to offer tumor-specific visual guidance. While they achieve moderate to high accuracy (~88–97%), these methods generally lack interpretability and task diversity (e.g., no segmentation or bounding box regression).

In contrast, models like those by Petrini et al. [56] and Sait & Nagaraj [104] explore transfer learning with advanced networks like EfficientNet, demonstrating improved accuracy (~85–99%) on large datasets. However, they still fall short on localization and lack multi-task capabilities, often requiring high computational resources without offering real-time or clinical insight.

Optimization-enhanced methods such as BreastCNet [30] mark a substantial advancement. By integrating hybrid optimizers (GWO and PO) and a multi-task learning (MTL) structure, BreastCNet not only achieves superior performance (AUC: 0.995, IOU: 0.96) but also addresses critical limitations such as task separation, overfitting, and interpretability. This is in contrast to YOLO-based detectors [57], which offer real-time performance but suffer from high false positives and architectural complexity.

Transformer-based models like BUViTNet [55] show promising global context awareness but are limited by their data and compute requirements, as well as their lack of bounding box outputs. Meanwhile, deformable-attention approaches (e.g., ACSNet [43]) offer innovative gating mechanisms yet still lack full integration of classification and localization. Table 1 provides a comprehensive comparison of several methods in recent years (Table 1).

Table 1. Comparison of breast cancer detection models

Author/Model	Year	Technique	Dataset	Accuracy / AUC / F1 / IOU	Gaps/Shortcomings
Das & Rana [105]	2021	ResNet variants	BUSI	Accuracy: 88.89%	Accuracy decreased with deeper models; no localization; lacks optimization and interpretability.
Ding et al. [64]	2022	ResNet-GAP + elastography	BUSI	Accuracy: 88.6%	High complexity; lacks efficient optimization; no cross-dataset generalization shown.
Petrini et al. [9]	2022	EfficientNet + transfer learning	CBIS-DDSM	AUC: 0.9344, Acc: 85.13%	No localization; no MTL; weak generalization; lacks hybrid optimization.
Huynh et al. [106]	2023	YOLOX + EfficientNet	Private	Accuracy: 92%	No mention of optimization or interpretability; only classification; lacks cross-dataset validation.
Ayana et al. [62](BUViTNet)	2022	Vision Transformer	BUSI	AUC: 0.968, Kappa: 0.959	No localization; lacks optimization; limited to ultrasound only.
Sahu et al. [107] (HADLCM)	2023	Hybrid deep-layer cascade	BUSI	Accuracy: 95.0%	No bounding box regression; lacks optimizer usage; evaluated on one dataset.
Kumar et al.[108]	2022	CNN (3 conv layers)	DDSM, CBIS-DDSM	Accuracy: 97.2%	Only classification; no MTL or localization; lacks interpretability framework.
Sait & Nagaraj [109]	2024	EfficientNetB7 + LightGBM	CMMD	Accuracy: 99.9%	Very high complexity; not tested on ultrasound; no localization or MTL.
Prinzi et al.[63]	2023	YOLOv5 + transfer learning	Proprietary	Accuracy: 95.3%, mAP: 0.621	High false positives; lacks classification branch; no dual-task support.
Yu H et al. (ACSNet)[50]	2024	Deformable attention + gated CNN	BUSI	Accuracy: 94.44%	No bounding box output; no optimization; interpretability limited.
Mishra et al.[110] (MultiRUSNet)	2024	UNet-ResNet hybrid	BUSI	Accuracy: 95.2%, Dice: 0.741	No localization; no optimizer tuning; lacks cross-modality testing.
Mahichi et al.[5] BreastCNet	2025	CNN + GWO + PO (MTL)	BUSI, DDSM, INbreast	Accuracy: 98.10%, AUC: 0.995, F1: 0.98, IOU: 0.96	Balanced performance, MTL, hybrid optimization, interpretable with Grad-CAM.

7.2. Breast cancer segmentation

Breast tumor segmentation methods (Table 2) primarily rely on U-Net variants, instance segmentation models like Mask R-CNN, and more recently, transformer hybrids. Classical U-Net and its extensions (e.g., U-Net++, Attention U-Net) [71]-[72] are favored for their pixel-wise segmentation capabilities. However, they struggle with small or irregular lesions, particularly in ultrasound images where contrast is low. Studies such as those by Vakanski et al. [101] and Zhao & Dai [103] demonstrate strong performance (DSC: ~90%) but often lack robustness across datasets.

Multi-scale and attention-based networks like MDF-Net [104] and CV-VAE [106] introduce advanced feature fusion and spatial disentanglement, achieving better generalization and Dice scores (up to 93.70%). However, they are often limited by training complexity and computational demands, making real-time deployment difficult. Similarly, hybrid and ensemble models like those by Bobowicz et al. [105] and Karunanayake et al. [82] offer high performance but require complex setups that hinder scalability and interpretability.

Instance segmentation models (e.g., Mask R-CNN) provide finer object-level delineation, yet their reliance on bounding box proposals often limits performance on diffused or small tumors. They also introduce high computational overhead (~140M parameters), making them impractical for edge deployment.

Table 2. Comparison of Breast tumor segmentation methods

Author	Year	Dataset	Method	Results	Gaps/Shortcomings
Vogl et al. [111]	2019	Multi-modal data (34 Patients)	Random Forest Classifier with mpPET/MRI	Segmentation Dice Coeff: 0.665	small dataset size, feature predictiveness, less modalitation between classification and segmentation
Vakanski et al. [112]	2020	510 breast ultrasound images	U-Net with attention blocks incorporating visual saliency	DSC of 90.5%	Lower accuracy compared to other advanced models, Reliance on visual saliency may not fully capture variability in medical images, Limited dataset scope, Need for further validation on diverse datasets, Potential overfitting, Scalability concerns
Irfan et al. [113]	2021	Breast ultrasonic image	Dilated Semantic Segmentation Network (Di-CNN) + DenseNet201 + 24-layer CNN + SVM	Accuracy: 98.9%, Mean-IoU: 52.89%, Mean Accuracy: 79.61%, Weighted-IoU: 73.83%, Mean-BF Score: 0.18218	Potential overfitting due to high complexity, Time-consuming training with 500 epochs, Limited explanation on performance with varied dataset sizes,
Zhao and Dai [114]	2022	Ultrasound images	U-Net + Residual Block + Attention Mechanism	Dice Index: 0.921	Evaluation metrics limited to Dice Index, IoU, and HD Index, Dataset details not clearly specified
Qi W et al. [115]	2023	Dataset A: Breast Ultrasound Lesions Dataset Dataset B: Baheya Hospital Ultrasound Dataset	Multi-scale Dynamic Fusion Network (MDF-Net)	Dataset A: DSC: 83.63%, IoU: 75.83%, Sensitivity: 85.07%, Specificity: 99.42% Dataset B: DSC: 78.20%, IoU: 70.22%, Sensitivity: 79.44%, Specificity: 98.22%	- Sensitivity slightly lower on Dataset B - No explicit reporting on overall accuracy, AUC, F1 Score, and mAP, 500 epoch cause overfitting.

Zhang S et al. [77]	2023	External datasets for BUS imaging	Dual-branch model (classification & segmentation)	AUC: 0.991 (classification); DSC: 0.898 (segmentation)	Slightly lower segmentation precision compared to specialized models; Increased complexity due to dual-branch architecture; Synchronization challenges between classification and segmentation tasks; Complex loss function design needed.
Karunanayake N et al. [88]	2024	Three distinct ultrasound datasets	AI-based hybrid model combining deep learning and multi-agent artificial life	Dice coefficients: 0.96 (easy), 0.91 (medium), 0.90 (hard); Relative Hausdorff distance: H3 = 0.26 (easy), H3 = 0.82 (medium), H3 = 0.84 (hard)	Reliance on initial DL segmentation quality; Lack of direct comparison to instance segmentation methods, Computational Complexity
Bobowicz M et al. [116]	2024	BUS B, OASBUD, BUSI, UCC BUS	PraNet, CaraNet, FCBFormer (ensemble classifiers)	IoU: 0.81, 0.80, 0.73; Dice: 0.89, 0.87, 0.82	Lower segmentation accuracy compared to ISRFE-DO; Requires complex ensemble setup for classification, Computational complexity, scalability issues, complexity in implementation.
Ma Yet al. [117]	2024	CBIS-DDSM, INbreast	Cross-View Variational Autoencoder (CV-VAE) with spatial hidden factor disentanglement, FPN-based classifier, and U-Net-like decoder	DSC: 92.46%, 93.70%. F1-score: 0.864	Requires large labeled datasets and high computational power, limiting use in small clinics.

7.3. End to end breast cancer detection and segmentation Simultaneously

End-to-end models that simultaneously detect and segment tumors represent a promising, holistic approach for breast cancer diagnosis. Most such models attempt to unify classification and segmentation, although often at the expense of performance trade-offs in one domain.

For example, EDCNN by Islam et al. [14] combines MobileNet and Xception for improved detection, yet it falls short in segmentation accuracy (IoU: 0.77). Models like DDA-AttResUNet [96] and DAU-Net [107] emphasize segmentation using attention-enhanced U-Net variants, achieving Dice scores above 92%, but do not support classification or localization, reducing their diagnostic completeness.

UCapsNet [108] and CoAtUNet [109] incorporate capsule networks and transformer modules, respectively, offering strong classification and segmentation performance (~99% accuracy and Dice ~95%). However, their high computational cost, lack of real-time capability, and evaluation on limited datasets weaken their generalizability and clinical viability. Table 3 summarizes several methods for breast cancer classification, localization and tumor segmentation, (Table 3).

Table 3. Comparison of breast cancer detection and segmentation Simultaneously methods

Author (Year)	Dataset	Model Architecture	Performance Metrics	Strengths	Gaps/Shortcomings
Islam M. et al. [24](2024)	BUSI, UDIAT	EDCNN (Ensemble Deep CNN - MobileNet + Xception)	Accuracy: 87.82%, AUC: 0.91, F1-score: 86.00%, IoU: 0.77	Combined MobileNet and Xception for enhanced detection, interpretability with Grad-CAM	Limited segmentation effectiveness, lower accuracy than BreaVisioNet, no classification component
Hekal A. et al.[101] (2024)	BUSI	DDA-AttResUNet (Dual Decoder Attention ResUNet)	Dice: 92.92%, IoU: 87.39%, Accuracy: 98.82%, Sensitivity: 92.16%, Precision: 93.90%	Focuses on segmentation accuracy, Dual Decoder Attention enhances tumor region emphasis	Restricted to segmentation without classification, image resizing reduces fine tumor detail, limited generalizability
Carriero A. et al. [118](2024)	BUSI	Unet3+ (with FCN-32s, Unet, SegNet, DeepLabV3+, PSPNet)	Accuracy: 82.53%, IoU: 52.57%, Weighted IoU: 89.14%, F1-score: N/A	Unet3+ showed high accuracy and was compared with various models	Struggles with small/irregular tumors, poor generalization (training vs. eval IoU gap), dataset size is insufficient
Pramanik et al. [119](2024)	BUSI, UDIAT	DAU-Net (Dual Attention U-Net with PCBAM & SWA)	Dice: 74.23% (BUSI), 78.58% (UDIAT)	Enhanced feature extraction with dual attention mechanisms, hybrid loss function	Lacks tumor classification, limited to segmentation, no real-time processing capabilities, dataset-specific evaluation
Madhu et al.[120] (2024)	BUSI	UCapsNet (Hybrid U-Net + Capsule Network)	Classification Accuracy: 99.22%, Segmentation Accuracy: 99.07%, Dice Score: 95.14%	Strong segmentation, capsule network improves classification accuracy	High computational cost, not optimized for real-time deployment, limited dataset evaluation (only BUSI dataset)
Zaidkilani et al. [11] (2025)	BUSI, UDIAT	CoAtUNet (Convolutional Transformer U-Net)	(Results not specified yet)	Strong transformer-based architecture, attention mechanisms to enhance feature extraction	Potential computational complexity, may require further optimization for clinical deployment
Schutte et al. [121] (2024)	Private Dataset	CNN-based Model (Multi-stage CNN)	Accuracy: 86.4%; Sensitivity: 90.2%; Specificity: 85.1%	Uses multi-stage processing for improved sensitivity	Dataset limited to a private source, restricting generalizability

8. IDENTIFIED RESEARCH GAPS AND CHALLENGES

Despite the rapid advancement of deep learning in breast cancer detection and segmentation, numerous unresolved challenges continue to hinder real-world clinical integration. These challenges span across data availability, model generalization, interpretability, efficiency, and evaluation consistency. Based on critical analyses throughout this review, the following research gaps have been identified.

8.1. Data Limitations and Annotation Scarcity

The development of effective deep learning models is fundamentally limited by the scarcity of large, annotated, and diverse breast cancer imaging datasets. Most existing datasets (e.g., BUSI, INbreast, CBIS-DDSM) either lack segmentation masks, classification labels, or are heavily imbalanced in terms of lesion types and sizes [5], [36]-[37] . Particularly, datasets with aligned annotations for multi-task learning (e.g., joint classification, detection, and segmentation) are rare, hindering unified model training [5], [8], [12]. Moreover, manual annotation of breast lesions requires expert radiologists, making the process costly and time-consuming [103]. Although weakly supervised, semi-supervised, and self-supervised methods have been introduced to reduce dependency on labels [51], their use in breast imaging remains under-investigated and not yet mature for deployment.

8.2. Generalization and Domain Shift Challenges

A significant number of models perform well on internal validation sets but struggle to generalize across external datasets or modalities. Domain shift caused by differences in imaging protocols, scanner types, or patient demographics often leads to performance degradation [31], [41], [82], [122]. Techniques such as domain adaptation using CycleGANs [37], few-shot learning, and cross-modality training[42] have been explored, but robust cross-institutional generalization remains limited, especially for segmentation tasks [80], [83].

8.3. Limited Interpretability and Clinical Trust

Deep learning models are often viewed as "black-box" systems, which reduces trust and hinders clinical acceptance. Many high-performing models lack mechanisms to explain their decision-making process, making them unsuitable for clinical workflows where transparency is essential [4], [20], [43]. While methods like Grad-CAM, saliency maps, or attention visualization have been adopted in some frameworks (e.g., BreastCNet [5]), few models embed interpretability natively within their architectures, particularly in segmentation pipelines [57], [65], [88].

8.4. Trade-off Between Model Complexity and Real-Time Feasibility

Transformer-based models, capsule networks, and ensemble frameworks often offer superior accuracy at the cost of massive computational overhead, making them unsuitable for real-time diagnosis or use in low-resource clinical environments[62], [88], [90]. On the other hand, lightweight models such as MobileNet [95]and EfficientNet-B0 [9] offer reduced complexity but typically suffer from performance trade-offs in segmentation accuracy or lesion localization. Striking an optimal balance between model performance and deployment feasibility remains a key challenge for breast imaging systems [95].

8.5. Multi-Task Learning Optimization and Interference

Multi-task learning (MTL) frameworks enable simultaneous learning of classification, detection, and segmentation, promoting efficiency and feature sharing [34], [35], [50], [123]. However, MTL introduces task interference, where one task (e.g., classification) may dominate training and suppress others (e.g., segmentation), especially under fixed loss weights [8], [99]. Although adaptive loss weighting strategies like GradNorm [75], Uncertainty Weighting [99], and ALB have shown promise, MTL optimization is still highly sensitive to architecture design and hyperparameter tuning, limiting stability and reproducibility [98], [100].

8.6. Inconsistent Evaluation Protocols and Benchmarking

Current studies adopt diverse datasets, metrics, and evaluation criteria, making direct comparisons difficult and impeding progress [24], [101]-[102]. Some report only segmentation metrics (Dice, IoU), others emphasize classification (accuracy, AUC), while very few attempt a comprehensive multi-task evaluation. The lack of standardized benchmarks or leaderboards for breast cancer imaging (as seen in fields like natural image segmentation) hampers reproducibility and model comparison, leading to fragmented progress[8], [35].

8.7. Limited Integration of Multimodal and Clinical Metadata

Most models rely solely on imaging data, overlooking valuable clinical context such as patient history, genetic markers, and radiomics features [48]-[49]. This represents a missed opportunity for improving diagnostic accuracy and robustness. While some recent models integrate structured metadata with imaging via multimodal fusion techniques, this remains rare, technically complex, and poorly standardized. Effective fusion of non-image data with imaging pipelines demands further exploration to enhance diagnostic precision and personalization [50].

Table 4 summarizes the major research gaps and challenges identified in recent deep learning-based breast cancer imaging studies, including issues related to data scarcity, model generalization, interpretability, computational efficiency, multi-task learning optimization, evaluation inconsistency, and multimodal integration.

Table 4. Gaps and challenges identified in recent deep learning-based breast cancer imaging studies

Research Gap / Challenge	Description
Limited Annotated Datasets	Public datasets (e.g., BUSI, CBIS-DDSM, INbreast) often lack multi-task annotations (classification + detection + segmentation). Manual labeling is time-consuming and expert-dependent. Multi-task datasets are rare, limiting unified model development.
Poor Generalization and Domain Shift	Models trained on specific datasets often fail on unseen domains due to variations in image quality, acquisition devices, and patient demographics. Generalization across modalities is not guaranteed.

Research Gap / Challenge	Description
Lack of Interpretability and Clinical Trust	Most models are black-box systems. Few integrate interpretable tools like Grad-CAM or saliency maps natively. Clinicians require transparent decision support.
High Computational Complexity	Transformer-based, capsule, and ensemble models achieve high accuracy but are resource-intensive. Lightweight models reduce complexity but often compromise segmentation accuracy or localization ability.
Multi-Task Learning (MTL) Interference	MTL frameworks often face imbalanced task learning where one objective dominates. Optimizing joint loss functions is difficult. Adaptive balancing methods (e.g., GradNorm) are promising but computationally demanding.
Inconsistent Evaluation Protocols	Studies use varied datasets, metrics (e.g., AUC, Dice, IoU), and lack statistical validation. Absence of unified benchmarks prevents direct model comparisons.
Limited Multimodal Integration	Most models use imaging data alone. Clinical metadata (e.g., age, family history, radiomics) is underutilized. Integration techniques are complex and poorly standardized.

9. FUTURE RESEARCH DIRECTIONS

Building on the identified challenges and gaps outlined in the preceding sections, future research in deep learning-based breast cancer diagnosis must prioritize several critical areas to ensure greater clinical applicability, robustness, and interpretability of AI systems. Despite the rapid evolution of deep neural network architectures and their promising diagnostic capabilities, translating these systems into real-world healthcare environments requires overcoming persistent limitations related to data scarcity, generalization, transparency, computational efficiency, and workflow integration.

While our review highlights significant progress in the field, a critical evaluation reveals a dichotomy in the development of DL models for breast cancer detection and segmentation. On one hand, complex, state-of-the-art models like YOLOv4 demonstrate impressive accuracy but often come with a substantial computational cost and latency, making them less suitable for real-time clinical applications. On the other hand, lightweight models such as BreastCNet offer faster inference times but may sacrifice a degree of diagnostic precision. This trade-off between performance and efficiency presents a key challenge for clinical translation. Furthermore, the limited interpretability of many black-box models, despite their high-performance metrics, remains a significant barrier to their adoption by clinicians who require confidence and transparency in AI-driven diagnoses. This highlights a critical need for future research to not only pursue higher accuracy but to do so in a manner that is both computationally efficient and inherently interpretable.

A foundational priority is addressing the heavy reliance on large-scale, fully annotated datasets, which continues to hinder the development of robust and scalable models. Annotating breast cancer imaging data—especially for pixel-wise segmentation or instance-level detection—is time-consuming, costly, and requires expert radiologists. As such, future work should explore the development and integration of self-supervised and semi-supervised learning strategies that can leverage unlabeled or partially labeled data to enhance model learning. Techniques such as contrastive learning, pseudo-label generation, and consistency regularization have demonstrated success in other domains and hold significant promise in medical imaging contexts. Their application in breast cancer analysis, however, remains underutilized and demands greater research attention [43]. Implementing such methods can alleviate the annotation burden, enable more efficient use of available datasets, and ultimately lead to models that generalize better across patient populations and imaging devices.

Interpretability is another crucial direction for future exploration. Many of the high-performing deep learning models currently used in breast cancer imaging function as "black boxes," which limits their acceptability in clinical settings. Although post hoc techniques such as Grad-CAM or saliency maps offer visual cues about model attention, they do not provide inherently explainable reasoning mechanisms. Future architectures should incorporate interpretability directly into their structure, such as through attention-guided feature selection, decision-aware modules, or explicit representation learning. By embedding transparency into the model pipeline, researchers can improve clinical trust and facilitate model validation, a necessary step for regulatory approval and physician acceptance [1], [36], [50], [59].

Another key area for advancement lies in the optimization of MTL frameworks. These models are increasingly used for joint prediction tasks, such as simultaneous tumor classification, detection, and segmentation, as they improve computational efficiency and promote knowledge sharing across related tasks. However, one common challenge with MTL approaches is task interference, where learning one objective may negatively affect performance on another. To address this, dynamic loss-balancing methods such as Uncertainty Weighting, GradNorm, or Adaptive Loss Balancing (ALB) have been proposed, offering better stability and task convergence during training [25], [26], [94]-[95]. In addition, neural architecture search (NAS) and automated hyperparameter tuning may be employed to adaptively optimize MTL configurations for diverse imaging modalities and objectives.

From a deployment perspective, the creation of lightweight, resource-efficient deep learning models is imperative, particularly for use in point-of-care or low-resource environments. Current state-of-the-art models, particularly transformer-based or ensemble networks, often demand high computational resources and large memory footprints, which hinder their usability in real-time clinical workflows. Designing and optimizing architectures such as MobileNet, EfficientNet-lite, or SqueezeNet—alongside model compression techniques including pruning, quantization, and knowledge distillation—can yield smaller, faster, and more efficient models without significantly compromising diagnostic accuracy. These models would be particularly valuable in settings with limited access to high-performance computing infrastructure, such as community hospitals or remote clinics [9], [95], [124].

In parallel, future research must explore multimodal fusion strategies that incorporate both imaging data and complementary clinical metadata. Most current deep learning models are limited to analyzing image pixels, disregarding rich contextual data such as patient age, genetic risk factors, hormonal status, and prior imaging history. Integrating such non-image data into diagnostic models via multimodal fusion—using strategies like cross-modal transformers, graph-based neural networks, or attention-based late fusion—can enhance prediction robustness, enable personalized diagnosis, and improve the interpretability of model outputs. This direction also aligns well with precision medicine objectives, which aim to tailor interventions based on both phenotypic and clinical profiles [41]-[43], [125].

A persistent challenge in current research is the lack of model generalization across domains, institutions, and imaging protocols. Many models perform well on internal validation sets but exhibit significant degradation when applied to external datasets due to domain shift. Future efforts should focus on improving domain adaptation and transfer learning strategies. Techniques such as unsupervised domain adaptation, few-shot learning, and image-to-image translation using generative adversarial networks (e.g., CycleGAN) have shown potential in mitigating domain variability and improving robustness [28], [34], [77], [126]-[127]. However, their systematic application to breast cancer imaging remains limited. Moreover, constructing large-scale, diverse, multi-institutional datasets would provide the necessary benchmark for fair and reproducible evaluation of generalization performance across populations and imaging devices.

Equally important is the design of unified, end-to-end architectures that consolidate multiple diagnostic tasks—such as classification, localization, and segmentation—within a single streamlined model. At present, many models are developed and trained separately for each task, leading to redundant computation and lack of clinical integration. A unified approach, typically based on a shared encoder and multiple task-specific decoders, would significantly improve system efficiency, enable joint learning of diagnostic cues, and better reflect real-world diagnostic workflows. Furthermore, these models should be modular and extensible, supporting plug-and-play configurations for varied imaging modalities and clinical contexts [13], [30], [128].

Finally, the lack of standardized evaluation protocols and benchmarking practices poses a substantial barrier to progress. Current studies often employ disparate datasets, inconsistent evaluation metrics, and variable training-validation splits, which limits the reproducibility and comparability of results. The field would benefit greatly from the establishment of common benchmarking datasets with fixed training/testing splits and the inclusion of multi-task performance metrics that evaluate models holistically. The development of open-access leaderboards and community challenges—similar to those established for natural image classification and segmentation—could foster healthy competition, increase transparency, and drive methodological advancement [24], [26], [104], [129].

future research in deep learning for breast cancer imaging must not only pursue higher accuracy but also address fundamental issues related to data efficiency, interpretability, generalization, and deployment feasibility. Only by tackling these multidimensional challenges can researchers create AI systems that are not only powerful and accurate, but also trustworthy, efficient, and ready for real-world clinical impact.

Table 5 presents a consolidated summary of the key future research directions identified in this review, outlining their core objectives and associated references to guide ongoing advancements in deep learning-based breast cancer detection and segmentation, (Table 5).

Table 5. Future Research Directions in Deep Learning-Based Breast Cancer Detection and Segmentation

Research Direction	Description and Goals
Self-Supervised and Semi-Supervised Learning	Address data scarcity by enabling models to learn from unlabeled or partially labeled data. Techniques like contrastive learning, pseudo-labeling, and consistency training reduce reliance on expert annotations and improve generalization.
Integrated Explainable AI (XAI)	Improve clinical trust and regulatory approval through native interpretability mechanisms. Move beyond Grad-CAM to attention-based modules and decision-aware architectures that explain predictions in real time.

Research Direction	Description and Goals
Optimizing Multi-Task Learning (MTL)	Enhance learning of classification, detection, and segmentation in unified frameworks. Tackle task interference via dynamic loss weighting (e.g., GradNorm, ALB) and explore automated hyperparameter tuning for stable training.
Lightweight and Edge-Deployable Models	Design efficient architectures (e.g., MobileNet, EfficientNet-lite) with low computational overhead suitable for real-time inference on portable or embedded devices. Use pruning, quantization, and distillation for compression.
Multimodal and Clinical Metadata Integration	Fuse imaging data with patient metadata (e.g., age, family history, genetics) and radiomics features to enhance diagnostic accuracy and personalization. Apply cross-modal attention or graph-based fusion.
Domain Adaptation and Generalization	Develop models that generalize across institutions, scanners, and demographics using techniques like few-shot learning, CycleGANs, and unsupervised domain adaptation. Encourage cross-dataset evaluations.
Unified End-to-End Architectures	Integrate classification, detection, and segmentation into a single pipeline using shared encoders and task-specific decoders. Improve efficiency, coherence, and clinical usability.
Benchmarking and Evaluation Standardization	Establish shared benchmarks, multi-task leaderboards, and standardized datasets to promote transparency and reproducibility. Ensure fair comparison of models across different tasks and datasets.

10. CONCLUSION

This review provides a critical and multi-faceted synthesis of deep learning-based methodologies for breast cancer detection and segmentation, encompassing single-task, multi-task, and hybrid approaches across mammography, ultrasound, and MRI modalities. While significant progress has been made in algorithmic development—marked by the adoption of CNNs, transformers, attention mechanisms, and optimization techniques—several structural, methodological, and translational gaps remain unresolved.

One of the most pressing issues is the scarcity of large, annotated, and multi-task-compatible datasets, which limits both the training of robust models and the reproducibility of results. Despite the emergence of models with exceptional accuracy and segmentation performance, generalization across domains, modalities, and populations remains inadequate, frequently leading to performance degradation in real-world settings. Moreover, the widespread use of black-box architectures continues to challenge clinical trust, with many models lacking inherent interpretability or transparency in decision-making. Segmentation frameworks, although mature in architecture, often fail when faced with noisy, low-contrast, or irregular lesions—particularly in ultrasound imaging. Detection models frequently lack integrated localization capabilities or interpretability tools, reducing their diagnostic value. While end-to-end and multi-task learning frameworks offer a compelling route toward unified diagnostic pipelines, they are currently hindered by task interference, complex optimization demands, and insufficient benchmarking protocols. From a deployment perspective, current transformer-based and ensemble architectures, though accurate, are computationally expensive and unsuitable for point-of-care environments. Conversely, lightweight models often sacrifice accuracy or diagnostic completeness. Balancing computational feasibility with diagnostic precision is a critical yet underexplored challenge. Looking ahead, the field must prioritize several transformative directions: developing semi-supervised and self-supervised training pipelines to mitigate annotation bottlenecks; embedding explainability directly into model design to enhance transparency; optimizing multi-task architectures for stability and scalability; and integrating non-image clinical metadata through multimodal learning.

The future directions identified—such as the adoption of self-supervised learning for robust feature extraction from limited datasets, the integration of explainable AI for enhanced clinical trust, and the development of multi-modal fusion and domain adaptation techniques—are not merely incremental improvements. They are essential pathways for overcoming the existing barriers to create trustworthy, efficient, and clinically-impactful AI systems. By focusing on these areas, the next generation of deep learning models can be developed to provide a reliable and indispensable tool for breast cancer screening and diagnosis. In addition, a move toward standardized datasets, shared benchmarks, and clinically validated evaluation protocols is essential to promote reproducibility and fair comparison. Ultimately, the future of deep learning in breast cancer diagnosis depends not only on architectural innovation but also on cross-disciplinary collaboration—where AI researchers, radiologists, and clinical stakeholders co-develop solutions that are interpretable, generalizable, efficient, and ethically sound. Only then can deep learning systems transition from promising research tools to trusted, real-world clinical allies in the fight against breast cancer.

REFERENCES

[1] M., Leonardi, P., Martelletti, R., Burstein, A., Fornari, L., Grazzi, A., Guekht, and et al., "**The World Health Organization Intersectoral Global Action Plan on Epilepsy and Other Neurological Disorders and the Headache Revolution: From Headache Burden to A Global Action Plan for Headache Disorders**," *J. Headache Pain*, Vol. 25, No. 4, https://doi.org/10.1186/s10194-023-01700-3, 2024.

[2] H., Zhao, H., Li, and D., Cheng, Data "**augmentation for medical image analysis**," *Biomedical Image Synthesis and Simulation*, Elsevier, pp. 279–302, https://doi.org/10.1016/B978-0-12-824349-7.00021-9, 2022.

[3] G., Litjens, T., Kooi, B. E., Bejnordi, A. A. A., Setio, F., Ciompi, M., Ghafoorian, and et al., "**A Survey on Deep Learning in Medical Image Analysis,"** *Med Image Anal*, Vol. 42, https://doi.org/10.1016/j.media.2017.07.005, 2017.

[4] E., Abu Abeelh, and Z., AbuAbeileh, "**Comparative Effectiveness of Mammography, Ultrasound, and MRI in the Detection of Breast Carcinoma in Dense Breast Tissue: A Systematic Review**," *Cureus*, https://doi.org/10.7759/cureus.59054, 2024.

[5] H., Mahichi, V., Ghods, M. K., Sohrabi, and A., Sabbaghi, "**BreastCNet: Breast Cancer Detection, Classification, and Localization Convolutional Neural Network with Advanced Optimization Techniques**," *IEEE Access*, Vol. 13, pp. 87386–400, https://doi.org/10.1109/ACCESS.2025.3570364, 2025.

[6] P., Arbeláez, M., Maire, C., Fowlkes, and J., Malik, "**Contour Detection and Hierarchical Image Segmentation**," *IEEE Trans Pattern Anal Mach Intell*, Vol. 33, pp. 898–916. https://doi.org/10.1109/TPAMI.2010.161, 2011.

[7] O., Rainio, and R., Klén, "**Comparison of Simple Augmentation Transformations for A Convolutional Neural Network Classifying Medical Images**," *Signal Image Video Process*, Vol. 18, pp. 3353–60, https://doi.org/10.1007/s11760-024-02998-5, 2024.

[8] C., Aumente-Maestro, J., Díez, and B., Remeseiro, "**A multi-task framework for breast cancer segmentation and classification in ultrasound imaging**," *Comput Methods Programs Biomed*, Vol. 260, pp. 108540, https://doi.org/10.1016/j.cmpb.2024.108540, 2025.

[9] D. G. P., Petrini, C., Shimizu, R. A., Roela, G. V., Valente, M. A. A. K., Folgueira, and H. Y., Kim, "**Breast Cancer Diagnosis in Two-View Mammography Using End-to-End Trained EfficientNet-Based Convolutional Network**," *IEEE Access*, Vol. 10, pp. 77723–31, https://doi.org/10.1109/ACCESS.2022.3193250, 2022.

[10] M., Kaddes, Y. M., Ayid, A. M., Elshewey, and Y., Fouad, "**Breast Cancer Classification Based on Hybrid CNN with LSTM Model**," *Sci Rep*, Vol. 15, pp. 4409, https://doi.org/10.1038/s41598-025-88459-6, 2025.

[11] N., Zaidkilani, M. A., Garcia, and D., Puig, "**CoAtUNet: A symmetric encoder-decoder with hybrid transformers for semantic segmentation of breast ultrasound images**," *Neurocomputing*, Vol. 629, pp. 129660, https://doi.org/10.1016/j.neucom.2025.129660, 2025.

[12] J., Logan, P. J., Kennedy, and D., Catchpoole, "**A Review of The Machine Learning Datasets in Mammography, Their Adherence to The Fair Principles and The Outlook for The Future**," *Sci Data*, Vol. 10, pp. 595, https://doi.org/10.1038/s41597-023-02430-6, 2023.

[13] A., Anand, S., Jung, and S., Lee, "**Breast Lesion Detection for Ultrasound Images Using MaskFormer**," *Sensors*, Vol. 24, pp. 6890, https://doi.org/10.3390/s24216890, 2024.

[14] K., Yang, J., Bin, M. J., Lee, and J., Yang, "**Multi-Class Semantic Segmentation of Breast Tissues from MRI Images Using U-Net Based on Haar wavelet Pooling**," *Sci Rep*, Vol. 13, pp. 11704, https://doi.org/10.1038/s41598-023-38557-0, 2023.

[15] Y. A., Chen, L. J., Grimm, M., Nedrud, and H., Rahbar, "MRI Characteristics of Ductal Carcinoma in Situ," pp. 145–56, https://doi.org/10.1016/B978-0-12-822729-9.00026-6, 2022.

[16] Y., Le Cun, Y., Bengio, and G., Hinton, "**Deep learning**," *Nature*, Vol. 521, pp. 436–44, https://doi.org/10.1038/nature14539, 2015.

[17] N., Carion, F., Massa, G., Synnaeve, N., Usunier, A., Kirillov, and S., Zagoruyko, "**End-to-End Object Detection with Transformers,**" 2020.

[18] H., Chan, L. M., Hadjiiski, and R. K., Samala, "**Computer-Aided Diagnosis in The Era of Deep Learning**", *Med Phys*, Vol. 47, https://doi.org/10.1002/mp.13764, 2020.

[19] Y., Yu, J., Huang, L., Wang, S., Liang, "**A 1D-Inception-Resnet Based Global Detection Model for Thin-Skinned Multifruit Spectral Quantitative Analysis**," *Food Control*, Vol. 167, pp. 110823, https://doi.org/10.1016/j.foodcont.2024.110823, 2025.

[20] Y., Hou, Z., Wu, X., Cai, and T., Zhu, "**The Application of Improved Densenet Algorithm in Accurate Image Recognition**," *Sci Rep*, Vol. 14, pp. 8645, https://doi.org/10.1038/s41598-024-58421-z, 2024.

[21] A., Qayyum, M., Mazher, T., Khan, and I., Razzak, "**Semi-Supervised 3D-InceptionNet for Segmentation and Survival Prediction of Head and Neck Primary Cancers**," *Eng Appl Artif Intell*, Vol. 117, pp. 105590, https://doi.org/10.1016/j.engappai.2022.105590, 2023.

[22] L., Hamnett, M., Adewunmi, M., Abayomi, K., Raheem, and F., Ahmed, "**Enhancing Transformer-Based Segmentation for Breast Cancer Diagnosis using Auto-Augmentation and Search Optimisation Techniques**," 2023.

[23] F., Manigrasso, R.,, Milazzo, A. S.,, Russo, F., Lamberti, F., Strand, A., Pagnani, and et al., "**Mammography Classification with Multi-View Deep Learning Techniques: Investigating Graph and Transformer-Based Architectures**," *Med Image Anal*, Vol. 99, pp. 103320, https://doi.org/10.1016/j.media.2024.103320, 2025.

[24] M. R., Islam, M. M., Rahman, M. S., Ali, A. A. N., Nafi, M. S., Alam, T. K., Godder, and et al., "**Enhancing Breast Cancer Segmentation and Classification: An Ensemble Deep Convolutional Neural Network And U-Net Approach on**

Ultrasound Images," *Machine Learning with Applications*, Vol. 16, pp. 100555, https://doi.org/10.1016/j.mlwa.2024.100555, 2024.

[25] O., Ronneberger, P., Fischer, and T., Brox, "**U-Net: Convolutional Networks for Biomedical Image Segmentation**," 2015.

[26] X., Wu, D., Hong, and J., Chanussot, "**UIU-Net: U-Net in U-Net for Infrared Small Object Detection**," *IEEE Transactions on Image Processing*, Vol. 32, pp. 364–76, https://doi.org/10.1109/TIP.2022.3228497, 2023.

[27] R., Maqsood, F., Abid, J., Rasheed, O., Osman, S., Alsubai, "**Optimal Res-UNET architecture with deep supervision for tumor segmentation**," *Front Med (Lausanne)*, Vol. 12, https://doi.org/10.3389/fmed.2025.1593016, 2025.

[28] C., Qin, Y., Wu, J., Zeng, L., Tian, Y., Zhai, F., Li, and et al., "**Joint Transformer and Multi-Scale CNN for DCE-MRI Breast Cancer Segmentation**," *Soft Comput*, Vol. 26, pp. 8317–34, https://doi.org/10.1007/s00500-022-07235-0, 2022.

[29] O., Ronneberger, P., Fischer, and T., Brox, "**U-Net: Convolutional Networks for Biomedical Image Segmentation**," 2015.

[30] X., Li, M., Li, P., Yan, G., Li, Y., Jiang, H, Luo, and et al., "**Deep Learning Attention Mechanism in Medical Image Analysis: Basics and Beyonds**," *International Journal of Network Dynamics and Intelligence*, pp. 93–116, https://doi.org/10.53941/ijndi0201006, 2023.

[31] S., Jamil, M. D., Jalil Piran, and O. J., Kwon, "**A Comprehensive Survey of Transformers for Computer Vision**," *Drones*, Vol. 7, pp. 287, https://doi.org/10.3390/drones7050287, 2023.

[32] J., Chen, J., Mei, X., Li, Y., Lu, Q., Yu, Q., Wei, and et al., "**Trans UNet: Rethinking the U-Net Architecture Design For Medical Image Segmentation Through the Lens of Transformers**," *Med Image Anal*, Vol. 97, pp. 103280, https://doi.org/10.1016/j.media.2024.103280, 2024.

[33] H., Afrin, N. B., Larson, M., Fatemi, and A., Alizad, "**Deep Learning in Different Ultrasound Methods for Breast Cancer, from Diagnosis to Prognosis: Current Trends, Challenges, and an Analysis**," *Cancers (Basel)*, Vol. 15, pp. 3139, https://doi.org/10.3390/cancers15123139, 2023.

[34] Q., He, Q., Yang, H., Su, and Y., Wang, "**Multi-Task Learning for Segmentation and Classification of Breast Tumors from Ultrasound Images,**" *Comput Biol Med*, Vol. 173, pp. 108319, https://doi.org/10.1016/j.compbiomed.2024.108319, 2024.

[35] Y., Zhou, H., Chen, Y., Li, Q., Liu, X., Xu, S., Wang, and et al., "Multi-task Learning for Segmentation and Classification of Tumors In 3d Automated Breast Ultrasound Images," *Med Image Anal*, Vol. 70, pp. 101918, https://doi.org/10.1016/j.media.2020.101918, 2021.

[36] A., Oliver, J., Freixenet, J., Martí, E., Pérez, J., Pont, E. R. E., Denton, and, et al., "**A Review of Automatic Mass Detection and Segmentation in Mammographic Images**," *Med Image Anal*, Vol. 14, https://doi.org/10.1016/j.media.2009.12.005, 2010.

[37] T., Shen, K., Hao, C., Gou, and F. Y., Wang, "Mass Image Synthesis in Mammogram with Contextual Information Based on GANs," *Comput Methods Programs Biomed*, Vol. 202, pp. 106019, https://doi.org/10.1016/j.cmpb.2021.106019, 2021.

[38] R. C., Joshi, D., Singh, V., Tiwari, and M. K., Dutta, "**An Efficient Deep Neural Network Based Abnormality Detection and Multi-Class Breast Tumor Classification**," *Multimed Tools Appl*, Vol. 81, pp. 13691–711, https://doi.org/10.1007/s11042-021-11240-0, 2022.

[39] R., Maalej, and A., "**Mezghani, Transfer Learning and Data Augmentation for Improved Breast Cancer Histopathological Images Classifier**," *International Journal of Computer Information Systems and Industrial Management Applications*, Vol. 15, No. 10, 2023.

[40] A. M., Carrington, D. G., Manuel, P. W., Fieguth, T., Ramsay, V., Osmani, B., Wernly, and et al., "**Deep ROC Analysis and AUC as Balanced Average Accuracy, for Improved Classifier Selection, Audit and Explanation**," *IEEE Trans Pattern Anal Mach Intell*, Vol. 45, pp. 329–41, https://doi.org/10.1109/TPAMI.2022.3145392, 2023.

[41] M., Frid-Adar, I., Diamant, E., Klang, M., Amitai, J., Goldberger, and H., "**Greenspan, GAN-Based Synthetic Medical Image Augmentation for Increased CNN Performance in Liver Lesion Classification**," *Neurocomputing*, Vol. 321, ppp 321–31. https://doi.org/10.1016/j.neucom.2018.09.013, 2018.

[42] S. D., S. M., S. K., S. S., and R. M., "**GAN Based Data Augmentation for Enhanced Tumor Classification. 2020 4th International Conference on Computer, Communication and Signal Processing (ICCCSP)**," *IEEE*, pp. 1–5, https://doi.org/10.1109/ICCCSP49186.2020.9315189, 2020.

[43] C., Shorten, and T. M., Khoshgoftaar, "**A Survey on Image Data Augmentation for Deep Learning**," *J. Big Data*, Vol. 6, pp. 60, https://doi.org/10.1186/s40537-019-0197-0, 2019.

[44] C., Shorten, and T. M., Khoshgoftaar, "**A Survey on Image Data Augmentation for Deep Learning**," *J. Big Data*, Vol. 6, https://doi.org/10.1186/s40537-019-0197-0, 2019.

[45] W., Scharff, W., Rethfeldt, L., McNeilly, M., Laasonen, N., Meir, H., Abutbul-Oz, S., Smolander, and et al., "**Assessment of Developmental Language Disorder in Multilingual Children: Results from an International Survey**," *Folia Phoniatrica et Logopaedica*, Vol. 76, pp. 127–50, https://doi.org/10.1159/000533139, 2024.

[46] J., Rony, S., Belharbi, J., Dolz, I., Ben Ayed, L., McCaffrey, and E., Granger, "**Deep Weakly-Supervised Learning Methods for Classification and Localization in Histology Images: A Survey**," *Machine Learning for Biomedical Imaging*, Vol. 2, https://doi.org/10.59275/j.melba.2023-5g54, 2023.

[47] C. H., Wang, K. Y., Huang, Y., Yao, J. C., Chen, H. H., Shuai, and W. H., Cheng, "**Lightweight Deep Learning: An Overview**," *IEEE Consumer Electronics Magazine*, Vol. 13, pp. 51–64, https://doi.org/10.1109/MCE.2022.3181759, 2024.

[48] F. Z., Nakach, A., Idri, and E., Goceri, "**A Comprehensive Investigation of Multimodal Deep Learning Fusion Strategies for Breast Cancer Classification**," *Artif Intell Rev*, Vol. 57, pp. 327, https://doi.org/10.1007/s10462-024-10984-z, 2024.

[49] A., Iqbal, and M., Sharif, "**BTS-ST: Swin Transformer Network for Segmentation and Classification of Multimodality Breast Cancer Images**," *Knowl Based Syst*, Vol. 267, pp. 110393, https://doi.org/10.1016/j.knosys.2023.110393, 2023.

[50] Yu, H., and Dai, Q., "**Self-Supervised Multi-Task Learning for Medical Image Analysis**," *Pattern Recognit*, Vol. 150, pp. 110327, https://doi.org/10.1016/j.patcog.2024.110327, 2024.

[51] O., Díaz, A., Rodríguez-Ruíz, and I., Sechopoulos, "**Artificial Intelligence for Breast Cancer Detection: Technology, Challenges, and Prospects**," *Eur J Radiol*, Vol. 175, pp. 111457, https://doi.org/10.1016/j.ejrad.2024.111457, 2024.

[52] H. S., Das, A., Das, A., Neog, S., Mallik, K., Bora, and Z., Zhao, "**Breast Cancer Detection: Shallow Convolutional Neural Network Against Deep Convolutional Neural Networks Based Approach**," *Front Genet*, Vol. 13, https://doi.org/10.3389/fgene.2022.1097207, 2023.

[53] M., Tan, R., Pang, and Q. V., Le, "**Efficient Det: Scalable and efficient object detection**," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, https://doi.org/10.1109/CVPR42600.2020.01079, 2020.

[54] A., Bochkovskiy, C. Y., Wang, and H. Y. M., Liao, "**YOLOv4: Optimal Speed and Accuracy of Object Detection**" 2020.

[55] A., Anand, S., Jung, and S., Lee, "**Breast Lesion Detection for Ultrasound Images Using MaskFormer**," *Sensors*, Vol. 24, pp. 6890. https://doi.org/10.3390/s24216890, 2024.

[56] N., Mikhailov, M., Shakeel, A., Urmanov, M. H., Lee, and M. F., "**Demirci, Optimization of CNN Model for Breast Cancer Classification**", *2021 16th International Conference on Electronics Computer and Computation (ICECCO), IEEE*, p p. 1–3, 2021.

[57] K. J., Geras, R. M., Mann, and L., Moy, "**Artificial Intelligence for Mammography and Digital Breast Tomosynthesis: Current Concepts and Future Perspectives**," *Radiology*, Vol. 293, pp. 246–59, https://doi.org/10.1148/radiol.2019182627, 2019.

[58] Y., Hou, H., Gao, Z., Wang, and C., Du, "**Improved Grey Wolf Optimization Algorithm and Application**", *Sensors*, Vol. 22, pp. 3810, https://doi.org/10.3390/s22103810, 2022.

[59] J., Lian, G., Hui, L., Ma, T., Zhu, X., Wu, A. A., Heidari, and et al., "**Parrot Optimizer: Algorithm and Applications to Medical Problems**," *Comput Biol Med*, Vol. 172, pp. 108064. https://doi.org/10.1016/j.compbiomed.2024.108064, 2024.

[60] D. G. P., Petrini, C., Shimizu, R. A., Roela, G. V., Valente, M. A. A. K., Folgueira, and H. Y., Kim, "**Breast Cancer Diagnosis in Two-View Mammography Using End-to-End Trained EfficientNet-Based Convolutional Network**," *IEEE Access*, Vol. 10, pp. 77723–31, https://doi.org/10.1109/ACCESS.2022.3193250, 2022.

[61] H., Rahman, T. F., Naik Bukht, R., Ahmad, A., Almadhor, and A. R., Javed, "**Efficient Breast Cancer Diagnosis from Complex Mammographic Images Using Deep Convolutional Neural Network**," *Comput Intell Neurosci*, pp. 1–11. https://doi.org/10.1155/2023/7717712, 2023.

[62] G., Ayana, and S., Choe, "**BUViTNet: Breast Ultrasound Detection via Vision Transformers**," *Diagnostics*, Vol. 12, pp. 2654. https://doi.org/10.3390/diagnostics12112654, 2022.

[63] F., Prinzi, M., Insalaco, A., Orlando, S., Gaglio, and S., Vitabile, "**A Yolo-Based Model for Breast Cancer Detection in Mammograms**," *Cognit Comput*, https://doi.org/10.1007/s12559-023-10189-6, 2023.

[64] W., Ding, J., Wang, W., Zhou, S., Zhou, C., Chang, and J., Shi, "**Joint Localization and Classification of Breast Cancer in B-Mode Ultrasound Imaging via Collaborative Learning with Elastography**," *IEEE J. Biomed Health Inform*, Vol. 26, pp. 4474–85. https://doi.org/10.1109/JBHI.2022.3186933, 2022.

[65] I. Z., Yao, M., Dong, and W. Y. K., Hwang, "**Deep Learning Applications in Clinical Cancer Detection: A Review of Implementation Challenges and Solutions**," *Mayo Clinic Proceedings: Digital Health*, Vol. 100253, https://doi.org/10.1016/j.mcpdig.2025.100253, 2025.

[66] R. L., Siegel, K. D., Miller, and A., Jemal, "**Cancer statistics**", *CA Cancer J. Clin*, Vol. 70, pp. 7–30, https://doi.org/10.3322/caac.21590, 2020.

[67] F., Bray, J., Ferlay, I., Soerjomataram, R. L., Siegel, L. A., Torre, and A., Jemal, "**Global Cancer Statistics**", *Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries*, CA Cancer J Clin, Vol. 68, pp. 394–424, https://doi.org/10.3322/caac.21492, 2018.

[68] O., Ronneberger, P., Fischer, and T., Brox, "**U-net: Convolutional Networks for Biomedical Image Segmentation**," *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference*, Munich, Germany, Proceedings, Part III, 18, Springer, pp. 234–41, 2015.

[69] J., Redmon, and A., Farhadi, "**YOLOv3: An Incremental Improvement**", 2018.

[70] K., He, G., Gkioxari, P., Dollar, and R., Girshick, "**Mask R-CNN**," *IEEE International Conference on Computer Vision (ICCV), IEEE*, 2017, pp. 2980–8, https://doi.org/10.1109/ICCV.2017.322, 2017.

[71] P., Arbeláez, M., Maire, C., Fowlkes, and J., Malik, "**Contour Detection and Hierarchical Image Segmentation**," IEEE Trans Pattern Anal Mach Intell, Vol. 33, pp. 898–916, https://doi.org/10.1109/TPAMI.2010.161, 2011.

[72] Y., Boykov, and G., Funka-Lea, "**Graph Cuts and Efficient N-D Image Segmentation**," *Int J. Comput Vis*, Vol. 70, pp. 109–31, https://doi.org/10.1007/s11263-006-7934-5, 2006.

[73] L. C., Chen, G., Papandreou, I., Kokkinos, K., Murphy, and A. L., Yuille, "**Deep Lab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs**," *IEEE Trans Pattern Anal Mach Intell,* Vol. 40, pp. 834–48, https://doi.org/10.1109/TPAMI.2017.2699184, 2018.

[74] K., He, G., Gkioxari, P., Dollar, and R., Girshick, "**Mask R-CNN**," *IEEE International Conference on Computer Vision (ICCV), IEEE*, pp. 2980–8, https://doi.org/10.1109/ICCV.2017.322, 2017.

[75] M., Liu, D., Yao, Z., Liu, J., Guo, and J., Chen, "**An Improved Adam Optimization Algorithm Combining Adaptive Coefficients and Composite Gradients Based on Randomized Block Coordinate Descent**," *Comput Intell Neurosci*, https://doi.org/10.1155/2023/4765891, 2023.

[76] M. J., Umer, M. I., Sharif, and J., Kim, "**Breast Cancer Segmentation from Ultrasound Images Using Multiscale Cascaded Convolution with Residual Attention-Based Double Decoder Network**," *IEEE Access*, Vol. 12, pp. 107888–902, https://doi.org/10.1109/ACCESS.2024.3429386, 2024.

[77] S., Zhang, M., Liao, J., Wang, Y., Zhu, Y., Zhang, J., Zhang, and et al., "**Fully Automatic Tumor Segmentation of Breast Ultrasound Images with Deep Learning**," *J. Appl Clin Med Phys*, Vol. 24, https://doi.org/10.1002/acm2.13863, 2023.

[78] J., Redmon, S., Divvala, R., Girshick, and A., Farhadi, "**You Only Look Once: Unified, Real-Time Object Detection**," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, https://doi.org/10.1109/CVPR.2016.91, 2016.

[79] S., Goudarzi, J., Whyte, M., Boily, A., Towers, R. D., Kilgour, and H., Rivaz, "**Segmentation of Arm Ultrasound Images in Breast Cancer-Related Lymphedema: A Database and Deep Learning Algorithm**," *IEEE Trans Biomed Eng*, Vol. 70, pp. 2552–63, https://doi.org/10.1109/TBME.2023.3253646, 2023.

[80] F. A., Spanhol, L. S., Oliveira, C., Petitjean, and L., Heutte, "**A Dataset for Breast Cancer Histopathological Image Classification**," *IEEE Trans Biomed Eng*, Vol. 63, pp. 1455–62, https://doi.org/10.1109/TBME.2015.2496264, 2016.

[81] University of Nottingham. Nottingham Prognostic Index (NPI) dataset. https://www.nottingham.ac.uk/pathology/protocols/npi/npi.aspx. Accessed May 12, 2023.

[82] M., Tafavvoghi, L. A., Bongo, and N., Shvetsov, "**Busund L-TR, Møllersen K. Publicly available datasets of breast histopathology H & amp; E whole-slide images: A scoping review**," *J. Pathol Inform*, Vol. 15, pp. 100363, https://doi.org/10.1016/j.jpi.2024.100363, 2024.

[83] University of California Irvine Machine Learning Repository. Wisconsin Diagnostic Breast Cancer (WDBC) dataset. https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Diagnostic). Accessed May 12, 2023.

[84] H., Soltani, M., Amroune, I., Bendib, M. Y., Haouam, E., Benkhelifa, and M. M., Fraz, "**Breast Lesions Segmentation and Classification in A Two-Stage Process Based on Mask-RCNN and Transfer Learning**," *Multimed Tools Appl*, pp. 1–18, 2023.

[85] F. S., Khan, M. N. H., Mohd, M. D., Khan, and S., Bagchi, "**Breast Cancer Histological Images Nuclei Segmentation using Mask Regional Convolutional Neural Network**," *IEEE Student Conference on Research and Development (SCOReD), IEEE*; pp. 1–6, https://doi.org/10.1109/SCOReD50371.2020.9383186, 2020.

[86] H. M. A., Bhatti, J., Li, S., Siddeeq, A., Rehman, and A., Manzoor, "**Multi-Detection and Segmentation of Breast Lesions Based on Mask RCNN-FPN. 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM),**" *IEEE*, pp. 2698–704, https://doi.org/10.1109/BIBM49941.2020.9313170, 2020.

[87] Y., Zhang, S., Chan, V. Y., Park, K. T., Chang, S., Mehta, M. J., Kim, and et al., "**Automatic Detection and Segmentation of Breast Cancer on MRI Using Mask R-CNN Trained on Non–Fat-Sat Images and Tested on Fat-Sat Images**," *Acad Radiol*, Vol. 29, pp. 135–44. https://doi.org/10.1016/j.acra.2020.12.001, 2022.

[88] N., Karunanayake, and S. S., Makhanov, "**When Deep Learning Is Not Enough: Artificial Life as A Supplementary Tool for Segmentation of Ultrasound Images of Breast Cancer**," *Med Biol Eng Comput*, https://doi.org/10.1007/s11517-024-03026-x, 2024.

[89] D., Abdelhafiz, J., Bi, R., Ammar, C., Yang, and S., Nabavi, "**Convolutional Neural Network for Automated Mass Segmentation in Mammography**," *BMC Bioinformatics*, Vol. 21, pp. 192, https://doi.org/10.1186/s12859-020-3521-y, 2020.

[90] A., Baccouche, B., Garcia-Zapirain, C., Castillo Olea, and A. S., Elmaghraby, "**Connected-UNets: A Deep Learning Architecture for Breast Mass Segmentation**," *NPJ Breast Cancer*, Vol. 7, pp. 151, https://doi.org/10.1038/s41523-021-00358-x, 2021.

[91] T., Siriapisith, W., Kusakunniran, and P., Haddawy, "**Pyramid graph cut: Integrating intensity and gradient information for grayscale medical image segmentation**," *Comput Biol Med*, Vol. 126, pp. 103997, https://doi.org/10.1016/j.compbiomed.2020.103997, 2020.

[92] H., Chen, X., Qi, L., Yu, and P. A., Heng, "**DCAN: Deep Contour-Aware Networks for Accurate Gland Segmentation**," *IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE*, pp. 2487–96, https://doi.org/10.1109/CVPR.2016.273, 2016.

[93] P., Coupeau, J. B., Fasquel, and M., "**Dinomais, On the use of GNN-Based Structural Information to Improve CNN-Based Semantic Image Segmentation,**" *J. Vis Commun Image Represent,* Vol. 101, pp. 104167, https://doi.org/10.1016/j.jvcir.2024.104167, 2024.

[94] Y., Gao, B., Liu, Y., Zhu, L., Chen, M., Tan, X., Xiao, and et al., "**Detection and Recognition of Ultrasound Breast Nodules Based on Semi-Supervised Deep Learning: A Powerful Alternative Strategy**," *Quant Imaging Med Surg*, Vol. 11, pp. 2265–78, https://doi.org/10.21037/qims-20-12B, 2021.

[95] J., Dafni Rose, K., VijayaKumar, L., Singh, and S. K., Sharma, “**Computer-Aided Diagnosis for Breast Cancer Detection and Classification Using Optimal Region Growing Segmentation with Mobile Net Model**,” *Concurrent Engineering*, Vol. 30, pp. 181–9. https://doi.org/10.1177/1063293X221080518, 2022.

[96] C., Aumente-Maestro, J., Díez, and B., Remeseiro, “**A Multi-Task Framework for Breast Cancer Segmentation and Classification in Ultrasound Imaging**,” *Comput Methods Programs Biomed*, Vol. 260, pp. 108540, https://doi.org/10.1016/j.cmpb.2024.108540, 2025.

[97] C., Aumente-Maestro, J., Díez, and B., Remeseiro, “**A Multi-Task Framework for Breast Cancer Segmentation and Classification in Ultrasound Imaging**,” *Comput Methods Programs Biomed*, Vol. 260, pp. 108540, https://doi.org/10.1016/j.cmpb.2024.108540, 2025.

[98] Z., Cheng, M., Liu, C., Yan, and S., Wang, “**Dynamic Domain Generalization for Medical Image Segmentation**,” *Neural Networks*, Vol. 184, pp. 107073, https://doi.org/10.1016/j.neunet.2024.107073, 2025.

[99] M., Aslam, “**Cochran’s Q Test for Analyzing Categorical Data Under Uncertainty**,” **J. Big Data**, Vol. 10, pp. 147, https://doi.org/10.1186/s40537-023-00823-3, 2023.

[100] Y., Lecun, L., Bottou, Y., Bengio, and P., Haffner, “**Gradient-Based Learning Applied to Document Recognition**,” *Proceedings of the IEEE*, Vol. 86, pp. 2278–324, https://doi.org/10.1109/5.726791, 1998.

[101] A. A., Hekal, A., Elnakib, H. E. D., Moustafa, and H. M., Amer, “**Breast Cancer Segmentation from Ultrasound Images Using Deep Dual-Decoder Technology with Attention Network**,” *IEEE Access*, Vol. 12, pp. 10087–101, https://doi.org/10.1109/ACCESS.2024.3351564, 2024.

[102] T., Alam, W. C., Shia, F. R., Hsu, and T., Hassan, “**Improving Breast Cancer Detection and Diagnosis through Semantic Segmentation Using the Unet3+ Deep Learning Framework**,” *Biomedicines*, Vol. 11, pp. 1536, https://doi.org/10.3390/biomedicines11061536, 2023.

[103] 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, pp. 102825, https://doi.org/10.1016/j.bspc.2021.102825, 2022.

[104] C., Thomas, M., Byra, R., Marti, M. H., Yap, and R., Zwiggelaar, “**Bus-Set: A Benchmark for Quantitative Evaluation of Breast Ultrasound Segmentation Networks with Public Datasets**,” *Med Phys*, 2023.

[105] A., Das, and S., Rana, “**Exploring Residual Networks for Breast Cancer Detection from Ultrasound Images**” *12th International Conference on Computing Communication and Networking Technologies (ICCCNT), IEEE*, pp. 1–6, https://doi.org/10.1109/ICCCNT51525.2021.9580160, 2021.

[106] H., Huynh, A. T., Tran, and T. N., “**Tran, Region-of-Interest Optimization for Deep-Learning-Based Breast Cancer Detection in Mammograms**,” *Applied Sciences*, Vol. 13, pp. 6894, https://doi.org/10.3390/app13126894, 2023.

[107] A., Sahu, P. K., Das, and S., Meher, “**An Automatic Sparse-Based Deep Cascade Framework with Multilayer Representation for Detecting Breast Cancer**,” *Measurement*, Vol. 228, pp. 114375, https://doi.org/10.1016/j.measurement.2024.114375, 2024.

[108] P., Kumar, S., Srivastava, R. K., Mishra, and Y. P., Sai, “**End-to-end improved convolutional neural network model for breast cancer detection using mammographic data**,” *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, Vol. 19, pp. 375–84, https://doi.org/10.1177/1548512920973268, 2022.

[109] A. R. W., Sait, and R., Nagaraj, “**An Enhanced Light GBM-Based Breast Cancer Detection Technique Using Mammography Images**,” Diagnostics, Vol. 14, pp. 227, https://doi.org/10.3390/diagnostics14020227, 2024.

[110] A. K., Mishra, P., Roy, S., Bandyopadhyay, and S. K., Das, “**A multi-task learning based approach for efficient breast cancer detection and classification**,” Expert Syst, Vol. 39, https://doi.org/10.1111/exsy.13047, 2022.

[111] W. D., Vogl, K., Pinker, T. H., Helbich, H., Bickel, G., Grabner, W., Bogner, and et al., “**Automatic Segmentation and Classification of Breast Lesions Through Identification of Informative Multiparametric PET/MRI Features**,” *Eur Radiol Exp*, Vol. 3, pp. 18, https://doi.org/10.1186/s41747-019-0096-3, 2019.

[112] A., Vakanski, M., Xian, and P. E., Freer, “**Attention-Enriched Deep Learning Model for Breast Tumor Segmentation in Ultrasound Images**,” *Ultrasound Med Biol*, Vol. 46, pp. 2819–33, https://doi.org/10.1016/j.ultrasmedbio.2020.06.015, 2020.

[113] R., Irfan, A. A., Almazroi, H. T., Rauf, R., Damaševičius, E. A., Nasr, and A. E., Abdelgawad, “**Dilated Semantic Segmentation for Breast Ultrasonic Lesion Detection Using Parallel Feature Fusion**,” *Diagnostics*, Vol. 11, pp. 1212, https://doi.org/10.3390/diagnostics11071212, 2021.

[114] T., Zhao, and H., Dai, “**Breast Tumor Ultrasound Image Segmentation Method Based on Improved Residual U-Net Network**”, *Comput Intell Neurosci*, pp. 1–9, https://doi.org/10.1155/2022/3905998, 2022.

[115] M., Bobowicz, M., Badocha, K., Gwozdziewicz, M., Rygusik, P., Kalinowska, E., Szurowska, and et al., “**Segmentation-based BI-RADS ensemble classification of breast tumours in ultrasound images**,” *Int J Med Inform*, Vol. 189, pp. 105522, https://doi.org/10.1016/j.ijmedinf.2024.105522, 2024.

[116] P., Hurtik, V., Molek, J., Hula, M., Vajgl, P., Vlasanek, and T., Nejezchleba, “**Poly-YOLO: Higher Speed, More Precise Detection and Instance Segmentation for YOLOv3**,” 2020.

[117] Y., Ma, and Y., Peng, “**Mammogram Mass Segmentation and Classification Based on Cross-View VAE and Spatial Hidden Factor Disentanglement**,” *Phys Eng Sci Med*, Vol. 47, pp. 223–38, https://doi.org/10.1007/s13246-023-01359-9, 2024.

[118] A., Carriero, L., Groenhoff, E., Vologina, P., Basile, and M., Albera, "**Deep Learning in Breast Cancer Imaging: State of the Art and Recent Advancements in Early 2024**," *Diagnostics*, Vol. 14, pp. 848, https://doi.org/10.3390/diagnostics14080848, 2024.

[119] P., Pramanik, A., Roy, E., Cuevas, M., Perez-Cisneros, and R., Sarkar, "**DAU-Net: Dual Attention-Aided U-Net for Segmenting Tumor in Breast Ultrasound Images**," *PLoS One*, Vol. 19, pp. 0303670, https://doi.org/10.1371/journal.pone.0303670, 2024.

[120] G., Madhu, A., Meher Bonasi, S., Kautish, A. S., Almazyad, A. W., Mohamed, F., Werner, and et al., "**U Caps Net: A Two-Stage Deep Learning Model Using U-Net and Capsule Network for Breast Cancer Segmentation and Classification in Ultrasound Imaging**," *Cancers (Basel)*, Vol. 16, pp. 3777, https://doi.org/10.3390/cancers16223777, 2024.

[121] S., Schutte, and J., Uddin, "**Deep Segmentation Techniques for Breast Cancer Diagnosis,**" *Bio Med Informatics*, Vol. 4, pp. 921–45, https://doi.org/10.3390/biomedinformatics4020052, 2024.

[122] H., Afrin, N. B., Larson, M., Fatemi, and A., Alizad, "**Deep Learning in Different Ultrasound Methods for Breast Cancer, from Diagnosis to Prognosis: Current Trends, Challenges, and an Analysis**," *Cancers (Basel)*, Vol. 15, pp. 3139, https://doi.org/10.3390/cancers15123139, 2023.

[123] Z., Huang, X., Zhang, Y., Ju, G., Zhang, W., Chang, H., Song, and et al., "**Explainable breast cancer molecular expression prediction using multi-task deep-learning based on 3D whole breast ultrasound**," *Insights Imaging*, Vol. 15, pp. 227, https://doi.org/10.1186/s13244-024-01810-9, 2024.

[124] J., Peta, and S., Koppu, "**Explainable Soft Attentive Efficient Net for Breast Cancer Classification in Histopathological Images**," *Biomed Signal Process Control*, Vol. 90, pp. 105828, https://doi.org/10.1016/j.bspc.2023.105828, 2024.

[125] M., Frid-Adar, E., Klang, M., Amitai, J., Goldberger, and H., Greenspan, "**Synthetic Data Augmentation using GAN for Improved Liver Lesion Classification**," 2018.

[126] K., He, G., Gkioxari, P., Dollár, and R., Girshick, "**Mask R-CNN**," *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2961–9, 2017.

[127] Y., Zhang, S., Chan, V. Y., Park, K. T., Chang, S., Mehta, M. J., Kim, and et al., "**Automatic Detection and Segmentation of Breast Cancer on MRI Using Mask R-CNN Trained on Non–Fat-Sat Images and Tested on Fat-Sat Images**," *Acad Radiol*, Vol. 29, pp. 135–44. https://doi.org/10.1016/j.acra.2020.12.001, 2022.

[128] R., Irfan, A. A., Almazroi, H. T., Rauf, R., Damaševičius, E. A., Nasr, and A. E., "**Abdelgawad, Dilated Semantic Segmentation for Breast Ultrasonic Lesion Detection Using Parallel Feature Fusion,**" *Diagnostics*, Vol. 11, pp. 1212, https://doi.org/10.3390/diagnostics11071212, 2021.

[129] F., Di Salvo, S., Doerrich, and C., Ledig, "**MedMNIST-C: Comprehensive Benchmark and Improved Classifier Robustness by Simulating Realistic Image Corruptions**," 2024.