Systems and Computing
Volume 1, Issue 1, 1-12, 2025.
https://doi.org/10.64409/sycom.v1.i1.5

Research Article

Open Access

Attention-Gateway U-Net for Mammographic Mass Segmentation: An Empirical Study

Hama Soltani[1*], Yousuf Islam[2], Hazem Fareh[1], Issam Bendib[1], Mohamed Yassine Haouam[1], Mohamed Amroune[1,3]

1. *Laboratory of Mathematics, Informatics and Systems (LAMIS), Echahid Cheikh Larbi Tebessi University, Tebessa, Algeria*
2. *School of Electrnoc Information, Central South University, China*
3. *National School of Nanoscience and Nanotechnology, Algiers, Algeria*

Corresponding author: Hama Soltani (hama.soltani@univ-tebessa.dz)

Manuscript Review Record:

Submitted:
May 13, 2025

Accepted:
June 24, 2025

Published:
July 19, 2025

Cite This:
H. Soltani, Y. Islam, H. Fareh, I. Bendib, M. Y. Haouam, M. Amroune, "Attention-gateway U-Net for mammographic mass segmentation: an empirical study". *Systems and Computing*, Volume 1, Issue 1, 1-12, 2025. https://doi.org/10.64409/sycom.v1.i1.5

Abstract- Context: Accurate segmentation of breast masses in mammographic images is a critical step for early breast cancer detection and effective clinical decision-making. However, this task remains challenging due to variability in mass appearance, low contrast with surrounding tissues, and complex anatomical structures. **Objective:** This study proposes a novel deep learning framework based on an attention-guided U-Net architecture, specifically designed for the segmentation of suspicious masses in mammograms. The primary objective is to investigate the impact of integrating attention gates into the U-Net structure and to evaluate how these mechanisms influence the accuracy and robustness of mass segmentation. **Method:** Attention gates are embedded within the skip connections and function by dynamically highlighting relevant features from the encoder while suppressing less informative regions. This selective focus enables the model to more accurately delineate the often subtle boundaries of breast masses. **Results:** The proposed model is evaluated on the publicly available INbreast mammography dataset, and its performance is compared against the baseline U-Net and other state-of-the-art segmentation methods. Quantitative results demonstrate that the attention-enhanced U-Net significantly outperforms its counterparts in segmentation accuracy, particularly in challenging cases involving dense breast tissue or ill-defined masses. **Conclusions:** This study highlights the effectiveness of attention mechanisms in enhancing mammographic mass segmentation and represents a valuable step toward more intelligent and reliable Computer-Aided Diagnosis (CAD) systems for breast cancer.

Keywords- Attention Gateway U-net, Breast Cancer, Mammography, Segmentation.

Acknowledgement

The authors gratefully acknowledge the financial support provided by the LAMIS Laboratory. They also wish to thank the anonymous reviewers for their valuable comments and constructive suggestions, which helped improve the quality and clarity of this work.

1. Introduction

Breast cancer continues to be a major public health concern, ranking as the most commonly diagnosed cancer among women worldwide and representing a leading cause of cancer-related mortality. According to recent statistics from the World Health Organization (WHO) and the International Agency for Research on Cancer (IARC), the global incidence of breast cancer has been steadily increasing, with over 2.3 million new cases diagnosed annually [1]. Early and accurate detection remains the cornerstone of effective breast cancer management, significantly improving survival rates and enabling less aggressive treatment regimens. Mammography is widely recognized as the gold standard imaging technique for breast cancer screening due to its cost-effectiveness, non-invasiveness, and proven sensitivity in detecting early-stage lesions. Despite its advantages, mammographic image interpretation poses significant challenges, particularly in women with dense breast tissue, where overlapping anatomical structures can obscure masses and lead to diagnostic ambiguity [2]. Masses in mammograms can present with highly variable appearances in terms of size, shape, margin sharpness, and texture, making the accurate identification and segmentation of suspicious lesions a complex task even for experienced radiologists. Computer-Aided Diagnosis (CAD) systems have been developed to support radiologists and reduce diagnostic variability, with mass segmentation constituting a critical step in the CAD pipeline. Precise segmentation enables the delineation of tumor boundaries, facilitating the extraction of morphological and textural features essential for lesion characterization and classification. However, traditional segmentation methods such as region-growing, thresholding, active contour models, and level sets often lack robustness and generalization capability, especially in the presence of imaging artifacts, noise, and diverse lesion characteristics. The emergence of deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized the field of medical image analysis. CNN-based segmentation models have shown remarkable success across a variety of biomedical applications, offering automated, data-driven solutions that adaptively learn complex feature hierarchies from imaging data [3]. Among these models, the U-Net architecture has become a dominant approach in medical image segmentation due to its ability to preserve fine-grained spatial details through symmetric encoder–decoder structures and skip connections. However, a significant limitation of the traditional U-Net lies in its treatment of skip connections, which transfer all encoder features to the decoder without assessing their relevance. This can lead to the inclusion of redundant or misleading features, particularly problematic in mammography where the contrast between abnormal and normal tissue is often subtle. To address this issue, recent advances have introduced attention mechanisms into U-Net-based frameworks, enabling models to selectively focus on informative regions while suppressing irrelevant background [4][5][6].

In this work, we propose an attention-guided U-Net architecture for mass segmentation in mammography images, designed to overcome the limitations of standard segmentation approaches. Our method integrates spatial attention gates into the skip connections of the U-Net, enhancing the model's ability to highlight salient regions associated with suspicious masses and attenuate non-relevant structures. These attention gates operate by leveraging contextual cues from deeper layers to modulate encoder features before fusing them with decoder representations, thereby refining the segmentation process. The contributions of this study are threefold:

- We design a novel deep learning framework based on an attention-enhanced U-Net architecture, specifically tailored for the challenges of breast mass segmentation in mammograms.
- We incorporate attention gates into the skip connections to dynamically weight feature importance, thereby improving boundary localization and segmentation accuracy in dense or low-contrast breast tissue.
- We conduct a comprehensive evaluation of the proposed model on benchmark mammography datasets INbreast, demonstrating significant improvements over baseline U-Net and traditional segmentation approaches in terms of both quantitative metrics and visual accuracy.

This work aims not only to advance the state-of-the-art in automated breast mass segmentation but also to lay the groundwork for future integration into clinically viable CAD systems. By improving segmentation reliability and precision, our approach holds the potential to support radiologists in early cancer detection, risk assessment, and personalized treatment planning. The remainder of the paper is organized as follows: Section 2 reviews related work on breast mass detection and segmentation. Section 3 describes the proposed methodology, including the architectural design and attention mechanism. Section 4 presents the experimental setup, datasets, and evaluation metrics, followed by the results and a comparative analysis. Finally, Section 5 concludes the paper and outlines potential future directions.

2. Related works

Breast mass segmentation is a critical step in CAD systems for breast cancer. Numerous researchers have proposed a wide range of approaches, from traditional image processing to sophisticated hybrid models integrating deep learning, aiming to improve the accuracy and robustness of mass segmentation [7].

Early methods focused primarily on classical image processing techniques, such as thresholding, region-growing, and watershed segmentation. While these approaches laid the groundwork for computer-assisted breast imaging, they often suffered from sensitivity to noise, variability in mass appearance, and dependency on manual parameter tuning. A researcher [8] proposed a novel method based on texture gradient for segmenting the pectoral muscle, involving the Hough transform, polynomial modeling, and Euclidean distance regression. Although effective in many cases, this method struggled when the muscle had low contrast or small size. Similarly, [9] introduced an automated technique combining a genetic algorithm, morphological selection, and polynomial curve fitting, achieving promising results on mini-MIAS, DDSM, and INBreast datasets but encountering difficulties in dense tissue regions and with complex mass boundaries.

The integration of CNNs marked a significant advancement in mass segmentation. Architectures such as U-Net, which combine encoder-decoder structures with skip connections, became particularly influential. [10] developed a U-Net-based model enhanced with preprocessing and data augmentation, achieving high Dice scores and specificity on the DDSM dataset. However, the model's inability to encode object position and orientation limited its effectiveness, particularly on low-resolution images. While U-Net allowed for more precise localization even with limited data, further enhancements were needed to effectively manage complex structures and dense breast tissue.

Other studies explored hybrid pipelines that integrate segmentation with classification. [11] proposed a method that extracts regions of interest (ROIs) using rough set theory, followed by the application of vector field convolution features combined with Random Forest classification. While accurate, the approach was computationally intensive and thus unsuitable for real-time clinical deployment. In [12], the authors addressed misalignment issues in mammograms through an atlas selection method based on clustering. This improved segmentation performance on the DDSM and mini-MIAS datasets but relied on prior image information, limiting its generalizability across diverse datasets.

In parallel, [13] introduced a technique using active contour models and edge detection to segment both the pectoral muscle and breast boundary, achieving high Dice similarity scores. [14] employed a metaheuristic algorithm (EML) for automatic segmentation of suspicious regions, although the method's convergence time remained a limitation, particularly for large datasets or high-resolution images.

Recognizing the challenges of classical methods, researchers began integrating attention mechanisms into deep learning architectures. [15] proposed an enhanced U-Net model incorporating attention gates, leading to significant improvements in segmentation accuracy on the DDSM dataset, albeit with increased computational demands. [16]

introduced a method combining the Spiking Cortical Model (SCM) with an improved Chan-Vese algorithm, achieving precise boundary delineation and high specificity.

Earlier traditional approaches relied heavily on handcrafted features, including wavelet coefficients, texture descriptors, and morphological characteristics (e.g., shape regularity and spiculated margins) to differentiate benign from malignant lesions [17][18]. However, these methods were limited by high computational cost and low feature expressiveness [19].

U-Net, introduced in [20], is particularly notable for its encoder–decoder structure with skip connections, which enables precise localization and contextual understanding even with limited training data. To further improve U-Net's performance, attention mechanisms have been integrated, allowing the model to focus on relevant regions of the image. The Attention U-Net incorporates attention gates that suppress irrelevant regions in the input image while highlighting salient features relevant to the segmentation task. This modification has been shown to enhance accuracy across various medical image segmentation challenges.

Recent studies have proposed various enhancements to the Attention U-Net architecture:

- Adaptive Attention U-Net (AAU-Net): [21] introduced AAU-Net, which integrates a hybrid adaptive attention module comprising channel and spatial self-attention blocks. This design enables the network to capture more discriminative features under different receptive fields, improving segmentation performance on complex ultrasound images.
- Explainable Attention-Based Models: A study published in *Scientific Reports* presented an Attention U-Net model for breast cancer segmentation using the BUSI dataset. The model achieved high accuracy in delineating tumor boundaries, demonstrating the efficacy of attention mechanisms in enhancing model interpretability and performance [6].
- Attention-Enhanced U-Net: [22] proposed an Attention-Enhanced U-Net architecture that integrates attention blocks into the U-Net framework. This models like YOLOv5, YOLOv8 [23] [24] and Mask R-CNN [25] in breast cancer segmentation tasks [22].
- Recurrent Attention U-Net: AlJabri et al. developed a Recurrent Attention U-Net model for the segmentation and quantification of breast arterial calcifications in synthesized model achieved a Dice similarity coefficient of 90.5% on a dataset of 510 images, outperforming other
- 2D mammograms. By incorporating recurrent mechanisms and attention modules, the network's ability to focus on relevant features across different layers was enhanced [26].
- Dual Attention-Aided U-Net (DAU-Net): A novel segmentation method, DAU-Net, was designed to detect tumors in breast ultrasound images. This architecture employs dual attention mechanisms to improve the network's focus on pertinent regions, enhancing segmentation accuracy [5].

These advancements underscore the importance of integrating attention mechanisms into deep learning architectures for medical image segmentation. While attention-enhanced networks have demonstrated superior performance in delineating complex structures like breast tumors, they also introduce additional computational demands, highlighting the ongoing need for optimization strategies to balance accuracy and efficiency.

3. Materials and methodologies

In this study, we propose a deep learning framework for the segmentation of breast masses in mammography images, employing an attention-enhanced U-Net architecture. Accurate segmentation of masses is a critical step in the early diagnosis and characterization of breast cancer, yet it remains a challenging task due to the low contrast between malignant regions and surrounding tissues, the variability in mass shapes and sizes, and the presence of dense breast tissue. Our architecture addresses these challenges by leveraging a symmetric encoder–decoder design, enriched with attention mechanisms that selectively focus on clinically relevant mass regions.

The encoder is responsible for extracting rich hierarchical features while gradually reducing the spatial resolution, capturing both contextual and textural patterns indicative of breast pathology. While skip connections help retain spatial information lost during down-sampling, they often carry redundant or irrelevant background features. To alleviate this, we incorporate attention gates into the skip pathways. These gates dynamically generate attention maps by integrating semantic information from the decoder with spatial features from the encoder, effectively filtering out noise and highlighting salient mass regions. This guided feature selection improves the network's ability to localize and segment masses, particularly in heterogeneous breast tissues. The proposed framework is highly suitable for use in clinical CAD systems, where accurate delineation of breast masses can significantly aid in cancer risk assessment and treatment planning. Moreover, the precise boundary localization achieved by the attention-guided network holds potential for downstream tasks such as malignancy classification and radiomic analysis. In summary, the proposed attention-enhanced U-Net architecture provides a robust and interpretable solution for mass segmentation in mammography, thereby contributing to the development of reliable automated diagnostic tools.

3.1. Network architecture

The proposed architecture consists of three primary components—Encoder Blocks, Decoder Blocks, and Attention Gates—as depicted in Figure 1.

A. Encoder blocks

Each encoder block is designed to capture increasingly abstract representations of the input mammogram. It consists of two consecutive convolutional layers with 3×3 kernels and ReLU activation, followed by a dropout layer to prevent overfitting. Optionally, a max pooling layer is applied to reduce spatial dimensions and capture broader contextual information. Each block outputs a down sampled feature map for deeper layers and a$= \mathrm{MaxPool}(\mathrm{E_i})$.

B. Decoder blocks

The decoder blocks reconstruct the segmentation mask by sequentially up sampling the feature maps and refining them through convolutional layers. Each up sampled feature map is concatenated with its corresponding attended skip connection from the encoder. This fused representation is processed through a convolutional module similar in structure to the encoder, but without pooling, enabling the recovery of high-resolution segmentation masks that preserve both context and anatomical precision.

C. Attention gates

Attention gates play a crucial role in enhancing the focus of the model on relevant breast regions, particularly the masses. Each gate receives a gating signal from a deeper decoder layer and a skip connection from the encoder. These inputs are processed through convolutional layers, summed, and passed through a ReLU activation, followed by a sigmoid-activated convolutional layer to generate an attention map. The resulting map is then upsampled and applied to the original skip connection via element-wise multiplication, producing a filtered, context-aware feature map. This selective focus helps suppress normal tissue while emphasizing abnormal masses.

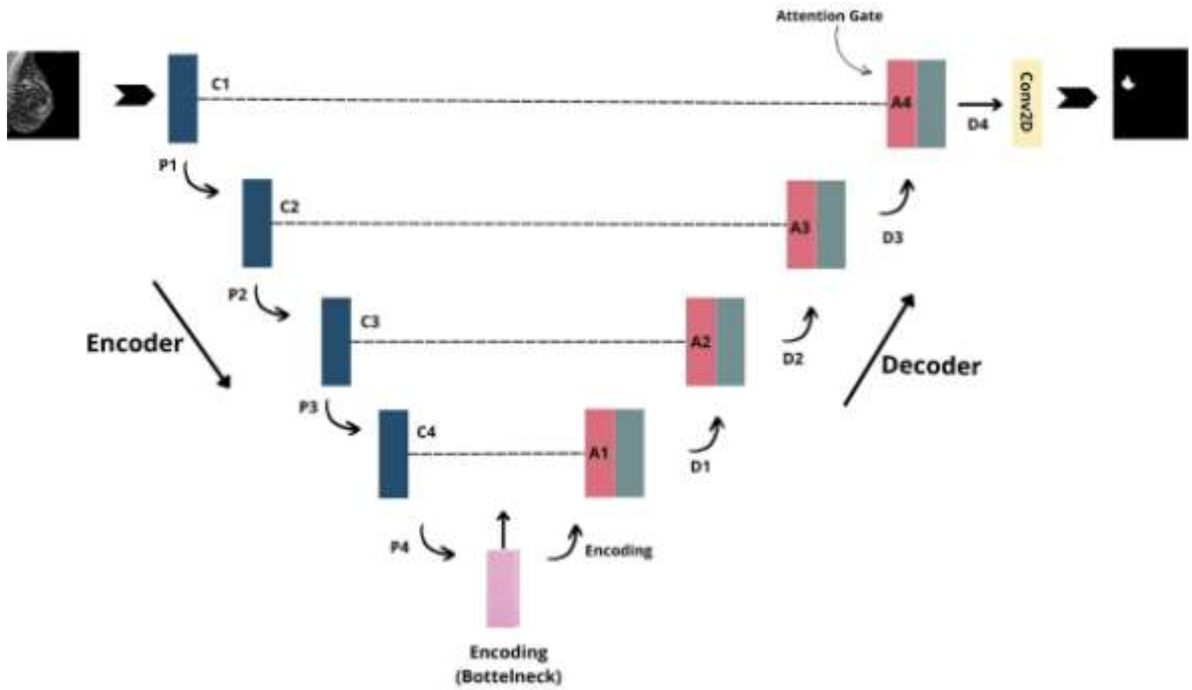

Figure 1. Architecture of proposed method.

3.2. Overall pipeline

The segmentation pipeline begins with the encoding of the input mammogram via a stack of encoder blocks that capture multi-scale contextual features while progressively reducing spatial dimensions. The bottleneck layer at the deepest point of the network represents the most abstracted semantic encoding of the input image. During decoding, each stage includes an Upsampling operation and a concatenation with the corresponding attended encoder features. This combination is processed by the decoder block to refine the segmentation. Finally, a 1×1 convolution with sigmoid activation is applied to generate the segmentation mask, delineating the breast mass with pixel-level precision.

This attention-guided segmentation framework is particularly effective in distinguishing pathological tissue from surrounding structures in dense breast tissues, a known challenge in mammographic analysis. Furthermore, its ability to localize and segment masses with high accuracy lays the foundation for more comprehensive CADe/CADx systems that integrate detection, segmentation, and classification. Future extensions may include multi-view fusion using craniocaudal and mediolateral-oblique views, and comparative evaluations with conventional architectures to assess improvements in generalization and robustness.

4. Experiments and results

4.1. Dataset

For the training and evaluation of our proposed attention-guided U-Net model, we utilized the INbreast dataset, a publicly available and high-quality mammographic database widely recognized in the medical imaging research community. The INbreast dataset consists of 410 full-field digital mammography (FFDM) images collected from 115 different patients. Each case includes standard mammographic views: craniocaudal (CC) and mediolateral oblique (MLO). What distinguishes INbreast is its precise and detailed annotations, manually delineated by expert radiologists.

The dataset provides BI-RADS assessments, breast density categories, and contour-level ground truth annotations of masses and other lesion types, making it highly suitable for supervised segmentation tasks. In this study, we specifically selected the images annotated with mass lesions and corresponding binary masks. Due to its resolution, image clarity, and clinically relevant labeling, the INbreast dataset presents an ideal benchmark for evaluating mass segmentation models. Its diversity in breast density and lesion appearance allows robust testing of deep learning models under realistic diagnostic conditions. Prior to training, all images were preprocessed to normalize intensity values and resized to a consistent spatial resolution to meet the input requirements of our network [27].

4.2 Experimental setup and evaluation

To comprehensively evaluate our proposed attention-guided U-Net architecture for breast mass segmentation, we established a systematic experimental framework. The model was developed using the Keras API with TensorFlow backend in Python and executed on Google Colab Pro, taking advantage of GPU acceleration to expedite training. We employed the INbreast dataset, known for its high-resolution full-field digital mammograms and expert-annotated mass contours. From this dataset, we extracted all cases containing mass lesions and their corresponding binary masks. The dataset was divided into 80% for training and 20% for validation, ensuring a balanced distribution across breast densities and mass types.

To investigate the individual and combined effects of attention mechanisms and data augmentation on breast mass segmentation performance, a structured ablation study was conducted, comprising four experimental configurations. The first configuration, referred to as the Baseline U-Net, involved training the standard U-Net architecture on the original INbreast dataset without any form of attention integration or data augmentation. This setup served as a reference benchmark, providing a foundational understanding of the model’s performance under default conditions. In the second configuration, termed U-Net with Attention, spatial attention gates were incorporated into the skip connections of the U-Net. These mechanisms aimed to enhance the model's ability to focus on the most relevant regions of the input mammograms. No data augmentation was applied, which allowed for an isolated assessment of the contribution made by the attention components. The third experimental setup, U-Net with Data Augmentation, involved applying a variety of augmentation techniques—such as image rotation, horizontal flipping, zooming, intensity normalization, and elastic deformation—to significantly expand the diversity of the training dataset. The model in this configuration did not include attention modules, thereby enabling a focused evaluation of how augmented data influences segmentation performance. Finally, the fourth configuration, combined both attention mechanisms and data augmentation strategies. This model was designed to examine the potential synergistic benefits of combining spatial attention with enriched training diversity. It directly supports the central goal of this study: to explore the impact of attention gateway mechanisms in improving breast mass segmentation from mammography images. In addition to quantitative metrics, we performed qualitative analysis by superimposing predicted segmentation masks on the original mammograms. This visual inspection allowed us to assess boundary accuracy, anatomical coherence, and noise suppression across different experimental setups.

4.3 Performance evaluation metrics

To rigorously evaluate our proposed segmentation framework, we utilized a suite of performance metrics well-established in medical image segmentation research:

- **Dice coefficient (Dice):** This overlap-based metric, defined in Eq (1), quantifies the similarity between predicted and ground truth segmentation masks. A higher Dice score indicates more accurate segmentation, making it particularly useful in class-imbalanced datasets like mammograms.

$$Dice = \frac{2 \mid X \cap Y \mid}{\mid X \mid + \mid Y \mid} \quad (1)$$

Where X and Y denote the sets of predicted and ground truth pixels, respectively.

- **Precision**: Precision, shown in Eq (2), measures the proportion of correctly identified positive pixels relative to all positive predictions. High precision is crucial in medical imaging, as false positives may result in unnecessary clinical follow-up.

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

Where TP and FP are the true and false positives, respectively.

- **Loss function**: We adopted a combination of binary cross-entropy and Dice loss to guide model training. While binary cross-entropy ensures convergence, the Dice loss component directly promotes accurate region overlap.

Throughout training, we monitored these metrics on both training and validation sets. Visual plots of training loss, Dice score, and precision curves across epochs (illustrated in Figures X and Y) provide insight into model convergence and generalization capability. By integrating a robust evaluation protocol and exploring key architectural variations, this study validates the effectiveness of attention-gated U-Net in accurately segmenting breast masses from mammography images and lays the groundwork for future advancements in AI-assisted breast cancer diagnosis.

4.4 Result and discussion

The performance of the proposed segmentation framework was evaluated through a comprehensive ablation study designed to isolate and examine the individual and combined effects of attention gates (AG) and data augmentation (DA) in the context of breast mass segmentation from mammographic images. Table 1 summarizes the quantitative results obtained across four experimental configurations, providing insights into the contributions of each enhancement to segmentation quality. In the baseline configuration, where a standard U-Net was trained without data augmentation or attention mechanisms, the model achieved a Dice similarity coefficient of 62%, a precision of 65%, and a segmentation loss of 0.45.

These results serve as a reference for traditional segmentation performance without architectural or training optimizations. When data augmentation was introduced to this baseline U-Net, a notable performance improvement was observed. The Dice score increased to 69%, precision improved to 73%, and loss decreased to 0.30.

These gains highlight the effectiveness of data augmentation in enhancing model generalization and reducing overfitting by exposing the model to a more diverse set of training samples. Integrating attention gates into the U-Net architecture, without applying data augmentation, led to a significant improvement in segmentation accuracy. The model achieved a Dice score of 81%, precision of 85%, and a reduced loss of 0.20. These results illustrate the capacity of AG to refine feature representations by suppressing irrelevant or noisy information and directing the model’s focus toward the most salient anatomical regions associated with breast masses. The best performance was attained by combining both attention gates and data augmentation. This configuration yielded a Dice score of 85%, a precision of 90%, and the lowest recorded segmentation loss of 0.14. These results reflect a synergistic effect, where AG enhances spatial attention and structural consistency, while DA increases robustness and adaptability to intra-class variations. The joint integration of these components demonstrates their complementary nature in improving segmentation quality.

A qualitative assessment of the segmentation outputs further confirms the quantitative findings. Visual inspection revealed that the inclusion of AG leads to more precise boundary delineation, particularly in regions with complex structures or low contrast. The baseline U-Net exhibited challenges in segmenting such areas, often resulting in over-segmentation or fragmented mass contours. In contrast, the AG-enhanced models demonstrated greater spatial coherence and anatomical fidelity, which is essential for clinical interpretability.

Table 1. Segmentation performance across experimental configurations.

Experiment	**U-net**		**U-net +AG**	
	Without DA	**+DA**	**Without DA**	**+ DA**
Dice	62%	69%	81%	85%
Precision	65%	73%	85%	90%
Loss	0.45	0.30	0.20	0.14

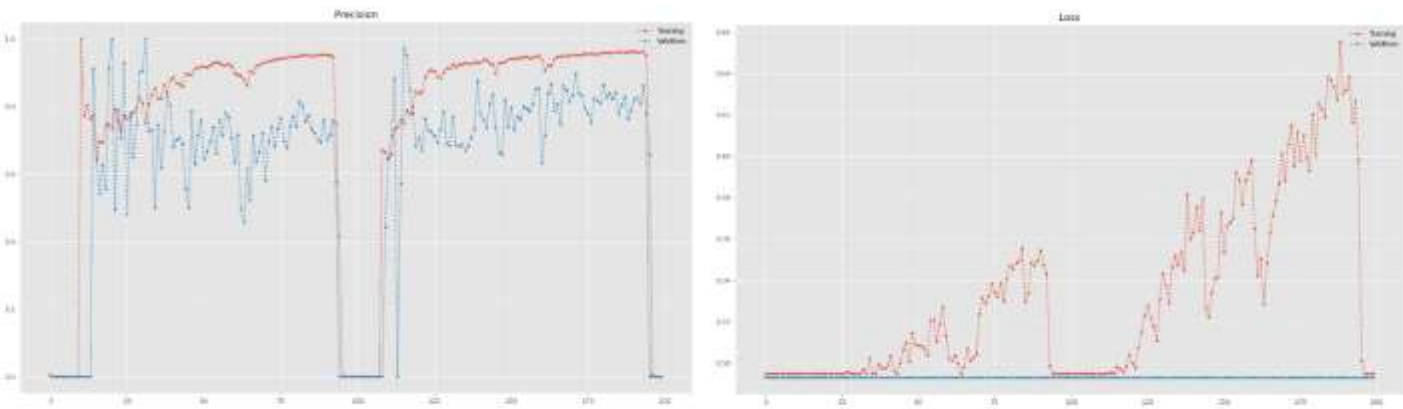

Figure 2. Precision and loss curves for the baseline U-Net model without attention gates

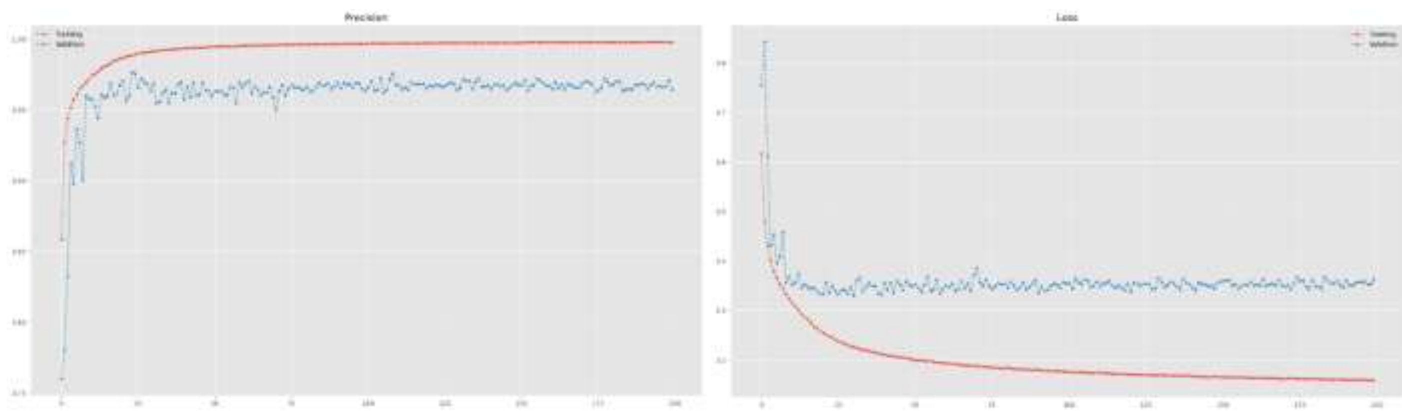

Figure 3. Precision and loss curves for the proposed U-Net model enhanced with attention gates.

Moreover, the performance gap between the baseline and the best-performing model—an increase of 20 percentage points in Dice and 25 in precision—underscores the critical role of attention mechanisms and data augmentation in achieving high-quality, reliable segmentation results.
Importantly, the improved precision achieved with AG indicates a substantial reduction in false positives, a crucial factor in medical image analysis where diagnostic accuracy has direct clinical implications. In summary, the experimental findings validate the effectiveness of incorporating attention gates and data augmentation in U-Net-based architectures for breast mass segmentation. These components collectively enhance the model's ability to generalize, localize, and delineate masses with high fidelity, contributing to more accurate and clinically meaningful outcomes.

Figure 4 (qualitative results) further support these findings by showing that models equipped with AG produced segmentation maps that are much closer to the ground truth. The attention-enhanced model exhibited sharper boundary delineation and reduced noise, particularly in challenging areas such as overlapping bones and low-contrast regions. In contrast, the baseline U-Net tended to generate fragmented or blurred contours, failing to accurately capture finer anatomical details. Finally, the significant performance drop observed in the absence of both AG and DA highlights the necessity of integrating these components for achieving reliable and clinically applicable segmentation outcomes. Attention gates dynamically enhance spatial feature extraction, while data augmentation fortifies the model's resilience against variations in image acquisition, thereby producing segmentation outputs that are both accurate and consistent.
When compared to state-of-the-art methods, the proposed U-Net with attention gates shows clear improvements over the baseline configuration but still falls short of surpassing the best performances reported in the literature. Nonetheless, the enhancement provided by the attention mechanism is evident, and the results suggest that with further experiments—such as applying advanced preprocessing steps, more extensive data augmentation, hyperparameter tuning, and evaluation on more diverse datasets—the proposed model has strong potential to reach or even exceed current state-of-the-art performances in medical image segmentation.

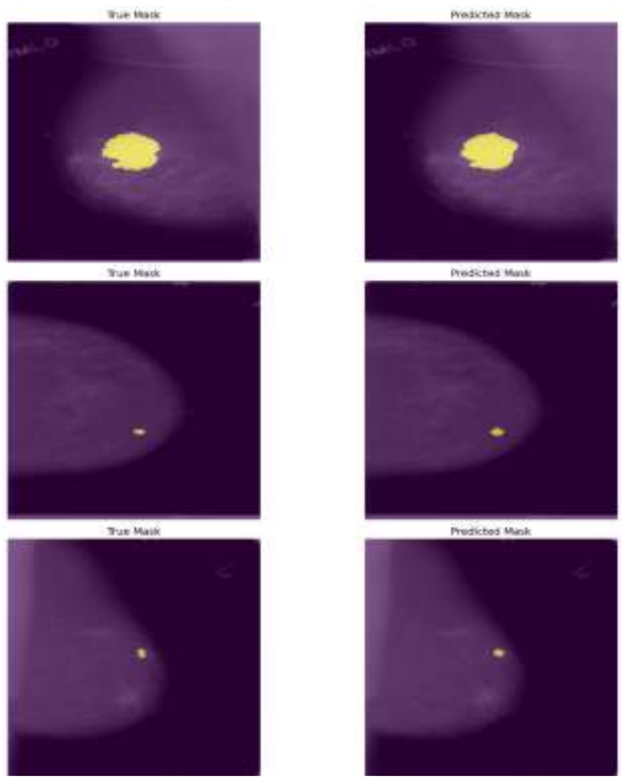

Figure 4. Qualitative comparison of predicted segmentation masks against ground truth annotations.

5. Conclusions and future works

In this study, we investigated the impact of integrating attention gates into a baseline U-Net architecture for medical image segmentation. Extensive quantitative and qualitative evaluations demonstrated that the attention-enhanced U-Net consistently outperformed the baseline configuration.
The baseline U-Net model exhibited significant instability during training, characterized by large fluctuations in precision and a failure to achieve meaningful convergence. In contrast, the proposed model incorporating attention mechanisms demonstrated stable learning behavior, higher precision on the validation set, and a smoother loss curve, indicating improved generalization capability.

Qualitative assessments further supported these findings. As illustrated in Figure 5, the attention-based U-Net generated segmentation masks that better preserved structural details, reduced noise more effectively, and achieved sharper boundary delineation compared to the fragmented and blurred outputs of the baseline model.

Although the proposed model achieved substantial improvements over the baseline, it has not yet surpassed the best-performing state-of-the-art methods reported in the literature. Nevertheless, the promising results indicate that the incorporation of attention mechanisms enhances the model's ability to focus on relevant anatomical features.

We believe that with further refinements—such as incorporating more advanced preprocessing techniques, expanding the training datasets, optimizing hyperparameters, and conducting evaluations on additional benchmark datasets—the performance of the proposed approach could be further improved, potentially matching or even exceeding current state-of-the-art segmentation results. These findings not only validate the effectiveness of attention mechanisms in enhancing segmentation quality but also provide a solid foundation for future research aimed at advancing performance beyond existing benchmarks.

References

[1] E. T. Sedeta, B. Jobre, & B. Avezbakiyev, "Breast cancer: Global patterns of incidence, mortality, and trends". *Journal of Clinical ncology*, 41(16_suppl), 10528–10528, 2023. https://doi.org/10.1200/JCO.2023.41.16_suppl.10528

[2] H. Soltani, I. Bendib, M. Y. Haouam, & M. Amroune, "Advancements in breast cancer diagnosis: A comprehensive review of mammography datasets, preprocessing and classification techniques." *Acta Informatica Pragensia*, 13(2), 308–326, 2024. https://doi.org/10.18267/j.aip.244

[3] S. A. Hassan, M. S. Sayed, M. I. Abdalla, & M. A. Rashwan, "Breast cancer masses classification using deep convolutional neural networks and transfer learning." *Multimedia Tools and Applications*, 79(41–42), 30735–30768, 2020. https://doi.org/10.1007/s11042-020-09518-w

[4] V. M. Tiryaki, "Mass segmentation and classification from film mammograms using cascaded deep transfer learning." *Biomedical Signal Processing and Control*, 84, 104819, 2023. https://doi.org/10.1016/j.bspc.2023.104819

[5] P. Pramanik, A. Roy, E. Cuevas, M. Perez-Cisneros, & R. Sarkar, "DAU-Net: Dual attention-aided U-Net for segmenting tumor in breast ultrasound images." *PLOS ONE*, 19(5), e0303670, 2020. https://doi.org/10.1371/journal.pone.0303670

[6] A. Sulaiman, V. Anand, S. Gupta, A. Rajab, H. Alshahrani, M. S. Al Reshan, A. Shaikh, & M. Hamdi, "Attention based UNet model for breast cancer segmentation using BUSI dataset." *Scientific Reports*, 14(1), 22422, 2024. https://doi.org/10.1038/s41598-024-72712-5

[7] S. Padhi, S. Rup, S. Saxena, & F. Mohanty, "Mammogram segmentation methods: A brief review." *2019 2nd International Conference on Intelligent Communication and Computational Techniques (ICCT)*, 218–223, 2019. https://doi.org/10.1109/ICCT46177.2019.8968781

[8] V. B. Bora, A. G. Kothari, & A. G. Keskar, "Robust automatic pectoral muscle segmentation from mammograms using texture gradient and Euclidean distance regression." *Journal of Digital Imaging*, 29(1), 115–125, 2016. https://doi.org/10.1007/s10278-015-9813-5

[9] S.-C. Tai, Z.-S. Chen, & W.-T. Tsai, "An automatic mass detection system in mammograms based on complex texture features." *IEEE Journal of Biomedical and Health Informatics*, 18(2), 618–627, 2014. https://doi.org/10.1109/JBHI.2013.2279097

[10] F. A. Zeiser, C. A. Costa, T. Zonta, N. Marques, M. C. Roehe, A. V. Moreno, & R. R. Righi, "Segmentation of masses on mammograms using data augmentation and deep learning." *Journal of Digital Imaging*, 33(4), 858–868, 2020. https://doi.org/10.1007/s10278-020-00330-4

[11] M. Dong, X. Lu, Y. Ma, Y. Guo, Y. Ma, & K. Wang, "An efficient approach for automated mass segmentation and classification in mammograms." *Journal of Digital Imaging*, 28(5), 613–625, 2015. https://doi.org/10.1007/s10278-015-9778-4

[12] M. K. Sharma, M. Jas, V. Karale, A. Sadhu, & S. Mukhopadhyay, "Mammogram segmentation using multi-atlas deformable registration." *Computers in Biology and Medicine*, 110, 244–253, 2019. https://doi.org/10.1016/j.compbiomed.2019.06.001

[13] A. Rampun, P. J. Morrow, B. W. Scotney, & J. Winder, "Fully automated breast boundary and pectoral muscle segmentation in mammograms." *Artificial Intelligence in Medicine*, 79, 28–41, 2017. https://doi.org/10.1016/j.artmed.2017.06.001

[14] K. B. Soulami, M. N. Saidi, B. Honnit, C. Anibou, & A. Tamtaoui, "Detection of breast abnormalities in digital mammograms using the electromagnetism-like algorithm." *Multimedia Tools and Applications*, 78(10), 12835–12863, 2019. https://doi.org/10.1007/s11042-018-5934-4

[15] S. Li, M. Dong, G. Du, & X. Mu, "Attention Dense-U-Net for automatic breast mass segmentation in digital mammogram." *IEEE Access*, 7, 59037–59047, 2019. https://doi.org/10.1109/ACCESS.2019.2914873

[16] Y. Guo, X. Gao, Z. Yang, J. Lian, S. Du, H. Zhang, & Y. Ma, "SCM-motivated enhanced CV model for mass segmentation from coarse-to-fine in digital mammography." *Multimedia Tools and Applications*, 77(18), 24333–24352, 2018. https://doi.org/10.1007/s11042-018-5685-2

[17] M. A. Al-antari, M. A. Al-masni, M.-T. Choi, S.-M. Han, & T.-S. Kim, "A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification." *International Journal of Medical Informatics*, 117, 44–54, 2018. https://doi.org/10.1016/j.ijmedinf.2018.06.003

[18] R. Wang, Y. Ma, W. Sun, Y. Guo, W. Wang, Y. Qi, & X. Gong, "Multi-level nested pyramid network for mass segmentation in mammograms." *Neurocomputing*, 363, 313–320, 2019. https://doi.org/10.1016/j.neucom.2019.06.045

[19] M. M. Eltoukhy, M. Elhoseny, K. M. Hosny, & A. K. Singh, "Computer aided detection of mammographic mass using exact Gaussian–Hermite moments." *Journal of Ambient Intelligence and Humanized Computing*, 15, 1139–1147, 2024. https://doi.org/10.1007/s12652-018-0905-1

[20] O. Ronneberger, P. Fischer, & T. Brox, "U-Net: Convolutional networks for biomedical image segmentation." In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Lecture Notes in Computer Science, vol. 9351. Springer, Cham, 2015. https://doi.org/10.1007/978-3-319-24574-4_28

[21] G. Chen, L. Li, Y. Dai, J. Zhang, & M. H. Yap, "AAU-Net: An adaptive attention U-Net for breast lesions segmentation in ultrasound images." *IEEE Transactions on Medical Imaging*, 42(5), 1289–1300, 2023. https://doi.org/10.1109/TMI.2022.3226268

[22] Md. N. Hasan, A. Ishraq, A. A. Emon, J. Shin, & Md. M. Kabir, "Advancing breast cancer diagnosis: Attention-enhanced U-Net for breast cancer segmentation." In: *Data-Driven Clinical Decision-Making Using Deep Learning in Imaging*, Studies in Big Data, vol. 152. Springer, Singapore, 2024. https://doi.org/10.1007/978-981-97-3966-0_11

[23] M. A. Al-masni, M. A. Al-antari, J.-M. Park, G. Gi, T.-Y. Kim, P. Rivera, E. Valarezo, M.-T. Choi, S.-M. Han, & T.-S. Kim, "Simultaneous detection and classification of breast masses in digital mammograms via a deep learning YOLO-based CAD system." *Computer Methods and Programs in Biomedicine*, 157, 85–94, 2018. https://doi.org/10.1016/j.cmpb.2018.01.017

[24] G. H. Aly, M. Marey, S. A. El-Sayed, & M. F. Tolba, "YOLO based breast masses detection and classification in full-field digital mammograms." *Computer Methods and Programs in Biomedicine*, 200, 105823, 2021. https://doi.org/10.1016/j.cmpb.2020.105823

[25] H. Soltani, M. Amroune, I. Bendib, & M. Y. Haouam, "Breast cancer lesion detection and segmentation based on Mask R-CNN." *2021 International Conference on Recent Advances in Mathematics and Informatics (ICRAMI)*, 1–6, 2021. https://doi.org/10.1109/ICRAMI52622.2021.9585913

[26] M. AlJabri, M. Alghamdi, F. Collado-Mesa, & M. Abdel-Mottaleb, "Recurrent attention U-Net for segmentation and quantification of breast arterial calcifications on synthesized 2D mammograms." *PeerJ Computer Science*, 10, e2076, 2024. https://doi.org/10.7717/peerj-cs.2076

[27] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, & J. S. Cardoso, "INbreast." *Academic Radiology*, 19(2), 236–248, 2012. https://doi.org/10.1016/j.acra.2011.09.014