Segmentation-Guided Lesion-Centric Deep Learning Framework for Interpretable Breast Ultrasound Classification

Ishita Rana[1], Jaimeel Shah[2], Hardik Variya[3]

[1]Ph.D. Scholar, Computer Science and Engineering, Parul University, Vadodara, India

[2]Associate Professor, Computer Science and Engineering, Parul University, Vadodara, India

[3]Associate Professor, G.J. Patel Institute of Ayurvedic Studies and Research, CVM University, Vallabh Vidyanagar, Anand, India

Abstract

Breast ultrasound imaging is commonly employed for the assessment of breast lesions, especially in women with dense breast tissue and in circumstances where mammography sensitivity is constrained. Nonetheless, interpreting ultrasound images continues to be difficult because of speckle noise, uneven backgrounds, and differences between operators. This study proposes a segmentation-guided, lesion-centric deep learning framework for the automated classification of breast ultrasound images. The framework combines U-Net-based lesion segmentation, region-of-interest (ROI) extraction, EfficientNetB0-based classification, and Gradient-weighted Class Activation Mapping (Grad-CAM) to make the model easier to understand. We tested the method on the Breast Ultrasound Images (BUSI) dataset, which had normal, benign, and malignant images. The training dataset went through preprocessing that included bilateral filtering, uniform resizing, and class-balanced augmentation. Before classification, segmentation outputs were used to find the areas of lesions. This reduced background noise and made it easier to learn features. The EfficientNetB0 classifier got an overall classification accuracy of 79.6% on the independent test set, with performance that was balanced across classes. Grad-CAM visualizations showed that the network's predictions were mostly based on lesion-relevant areas and not on artifacts in the surrounding tissue. The results show that segmentation-guided lesion-centric pipelines make it easier to understand and more reliable to diagnose breast cancer using ultrasound. They also lay the groundwork for computer-aided diagnostic systems that can be used in real life.

Keywords: Breast ultrasound, deep learning, U-Net segmentation, EfficientNet, explainable AI, medical image analysis

How to cite this article: Rana I, Shah J, Variya H. Segmentation-Guided Lesion-Centric Deep Learning Framework for Interpretable Breast Ultrasound Classification. Int J Drug Deliv Technol. 2026;16(12s): 323-327. DOI: 10.25258/ijddt.16.12s.35

1. Introduction

Breast cancer is still the most common type of cancer diagnosed in women around the world, and it is a major public health problem. Recent epidemiological estimates say that there were more than 2.3 million new cases of breast cancer around the world in 2022. The disease caused a lot of deaths, especially in low- and middle-income countries. Early detection is a key factor that affects survival rates, so breast cancer screening and diagnosis must use imaging technologies that are both accurate and easy to get to mammography, ultrasound, and magnetic resonance imaging (MRI) are all types of medical imaging that are commonly used to find breast cancer. Mammography is the most common way to screen a large group of people, but it doesn't work as well in younger people and women with dense breast tissue. Ultrasound imaging is an important additional diagnostic tool because it doesn't use ionizing radiation, is widely available, and is relatively inexpensive. As a result, breast ultrasound is often used to find lesions, screen for them in dense breasts, and follow up on imaging. Even with these benefits, interpreting ultrasound images is very hard. Ultrasound images frequently exhibit speckle noise, heterogeneous background structures, and variability due to discrepancies in acquisition devices and operator techniques. These factors can cause a lot of differences between radiologists and within the same radiologist, which could lead to missed cancers or unnecessary biopsies.

Computer-aided diagnosis (CAD) systems have been created in the past to help radiologists understand images. Early CAD methods used a mix of hand-made feature extraction and traditional machine learning algorithms. But these methods were limited because they depended on features that were manually engineered and weren't very strong in different imaging settings.

Recent improvements in deep learning have changed the way medical images are analyzed by allowing models to learn hierarchical features directly from imaging data. Convolutional neural networks (CNNs) have shown that they can find breast cancer better than traditional machine learning methods. But there are still some problems to solve. A lot of deep learning studies look at whole ultrasound images without clearly marking the areas where lesions are. This can lead to model predictions being affected by background information that isn't relevant. The "black-box" nature of deep neural networks can also make it hard for doctors to trust and use them.

To overcome these constraints, lesion-centric frameworks have been introduced wherein segmentation occurs prior to classification. These pipelines mimic the diagnostic reasoning of radiologists by focusing on the lesion region of interest and reducing background noise. Moreover, explainable artificial intelligence methods, like Gradient-weighted Class Activation Mapping (Grad-CAM), can give visual explanations of what a model predicts.

The current study suggests a segmentation-directed lesion-focused deep learning framework for the classification of breast ultrasound images. The framework combines U-Net-based lesion segmentation, region-of-interest extraction, EfficientNet-based classification, and Grad-CAM visualization. The goal is to build a pipeline that makes it easier to understand and more reliable for diagnosis while still being fast enough to be used in clinical settings.

2. Related Work

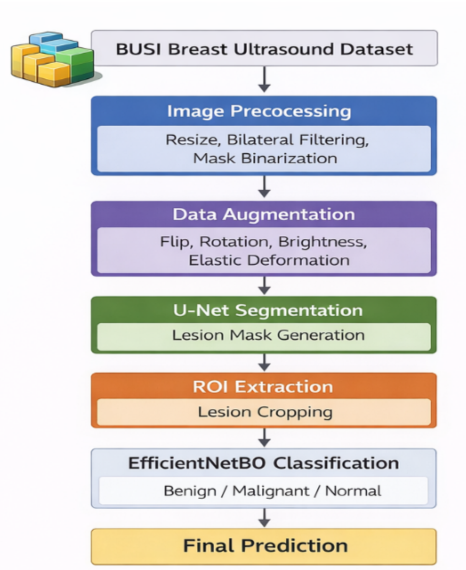
Deep learning is becoming more common in the analysis of medical images, especially when it comes to finding, segmenting, and classifying tumors in oncology imaging. VGG, ResNet, DenseNet, and EfficientNet are all CNN-based architectures that are often used to look at mammography, ultrasound, and histopathology images of breast cancer.

Early machine learning techniques employed handcrafted features such as texture descriptors, gray-level co-occurrence matrices (GLCM), and morphological measurements derived from lesion regions. After that, we used algorithms like support vector machines, random forests, and k-nearest neighbors to put these features into groups. These methods worked to some extent, but they weren't very good because the features that were made by hand couldn't fully represent the data.

Deep learning enabled the acquisition of features from raw imaging data in a comprehensive manner. Transfer learning methods that use networks that have already been trained on large natural image datasets have worked well for classifying breast ultrasound images. ResNet and Inception are two examples of architectures that have been used a lot and often lead to better diagnostic accuracy than traditional machine learning methods. But the current deep learning methods still have some problems. Many studies categorize complete ultrasound images without distinctly delineating the regions of lesions. Ultrasound images frequently contain substantial background tissue and artifacts, which may introduce noise into the learned representations during whole-image classification. Also, small datasets in medical imaging can make complex neural network architectures fit too well.

Segmentation-based pipelines are a good choice because they first find the area of the lesion and then sort it. U-Net and its variations are the most popular ways to segment medical images because they have an encoder-decoder structure and skip connections that keep spatial information. Segmentation-guided pipelines help classifiers focus on structures that are important for diagnosis by separating the area of the lesion.

Another big problem for medical AI is that it is hard to understand how the models work. AI systems must be clear and easy to understand for doctors to trust them in the clinic. Grad-CAM is one of the most common ways to explain CNN models. It makes heatmaps that show which parts of the image had the most effect on the model's guess.Not many studies have put segmentation, classification, and explainability together into one framework for breast ultrasound analysis, even with these improvements. The present study aims to address this gap by developing a segmentation-driven, lesion-focused pipeline that integrates U-Net segmentation, EfficientNet classification, and Grad-CAM explainability.



3. Dataset and Preprocessing

3.1 BUSI Dataset

The proposed framework was evaluated using the Breast Ultrasound Images (BUSI) dataset, a publicly available dataset widely used in research on breast lesion classification. The dataset contains ultrasound images categorized into three diagnostic classes:

- Benign lesions
- Malignant lesions
- Normal tissue

Each image is accompanied by segmentation masks outlining lesion boundaries for abnormal cases.

The dataset contains:

- 437 benign images
- 210 malignant images
- 133 normal images

Total images: 780 ultrasound images.

3.2 Data Preprocessing

Speckle noise and changes in intensity can affect ultrasound images, which can make segmentation and classification less accurate. To fix these problems, a number of preprocessing steps were taken. To begin, bilateral filtering was used to get rid of speckle noise while keeping the edges of the lesions. This filtering method keeps edge information while making smooth areas. Second, all images and their masks were resized to 256 × 256 pixels so that the segmentation network would always get the same size input.

Third, segmentation masks were turned into binary images to make sure that the edges of the lesions were clear. When several masks were linked to one image, they were combined to make a single lesion annotation.

3.3 Dataset Split and Class Balancing

Prior to data augmentation, the dataset was divided into training and testing subsets using an 80:20 split. This resulted in:

Training set

- 349 benign
- 168 malignant
- 106 normal

Test set

- 88 benign
- 42 malignant
- 27 normal

To address class imbalance, data augmentation was applied exclusively to the training dataset. Augmentation techniques included:

1. Horizontal and vertical flips
2. Random rotations
3. Brightness and contrast adjustments
4. Elastic deformation
5. Spatial transformations (shift, scale, rotate)

4. Segmentation Network (U-Net)

Lesion segmentation was performed using a U-Net architecture designed for biomedical image segmentation tasks.

The network consists of an encoder–decoder architecture with skip connections that allow spatial information from early layers to be preserved during upsampling.

Encoder

The encoder contains four convolutional blocks with increasing channel depths:

32 → 64 → 128 → 256 filters.

Each block contains:

1. Two convolutional layers
2. ReLU activation
3. Max-pooling layer

Bottleneck

The bottleneck layer connects the encoder and decoder and consists of two convolutional layers with 512 filters.

Decoder

The decoder performs progressive upsampling using transposed convolution layers. Skip connections from the encoder allow the decoder to recover spatial resolution and maintain precise lesion boundaries.

Training Setup

Loss function: Dice Loss

Optimizer: Adam

Learning rate: 1×10^{-4}

Batch size: 4

Segmentation performance was evaluated using:

- Dice Similarity Coefficient (DSC)
- Intersection over Union (IoU)

The network produces a binary lesion mask of size 256 × 256.

5. ROI Extraction Strategy

The segmentation mask generated by the U-Net model was used to identify the lesion region within the ultrasound image.

The following steps were applied:

1. Non-zero pixels in the segmentation mask were detected.
2. A bounding box was computed around the lesion region.
3. Margin padding was added to include contextual tissue information.

4. The lesion region was cropped from the original image.
5. The cropped ROI was resized to 128 × 128 pixels.

In cases where no lesion was detected, a fallback strategy was applied using either a center crop or the full image to ensure consistent input for classification.
This ROI extraction strategy ensures that the classifier focuses on diagnostically relevant regions rather than background tissue.

6. Classification Model (EfficientNetB0)

Lesion classification was performed using the EfficientNetB0 architecture. EfficientNet models are designed to achieve high performance while maintaining computational efficiency through compound scaling of network depth, width, and resolution.
The EfficientNetB0 backbone was initialized using pretrained ImageNet weights to leverage transfer learning.
The classification architecture consists of:
EfficientNetB0 backbone
→ Global average pooling
→ Dense layer (256 units)
→ Dropout layers for regularization
→ Softmax output layer (3 classes)
Input size: 128 × 128 × 3
Loss function: Sparse categorical cross-entropy
Optimizer: Adam
Batch size: 16
Class weights were applied during training to further mitigate class imbalance.
The classifier outputs both the predicted class label and the associated confidence score.

7. Explainable AI Using Grad-CAM

To enhance model interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) was applied to the classification network.
Grad-CAM generates heatmaps highlighting regions of the image that contribute most strongly to the model's prediction.
The Grad-CAM process involves:

1. Computing gradients of the predicted class score with respect to feature maps of the final convolutional layer.
2. Weighting these feature maps using the computed gradients.
3. Aggregating the weighted feature maps to produce a localization heatmap.

The heatmap is then superimposed onto the original ROI image, allowing visual interpretation of the model's decision-making process.

8. Results

The proposed segmentation-guided classification framework was evaluated on the BUSI dataset test set.
Classification Performance

Metric	Benign	Malignant	Normal
Precision	0.88	0.78	0.65
Recall	0.8	0.74	0.89
F1-Score	0.84	0.76	0.75

Overall accuracy: 79.6%
Macro-average F1 score: 0.78
Weighted F1 score: 0.80
The model demonstrated balanced performance across classes, with satisfactory detection of malignant lesions.
Grad-CAM visualizations indicated that the network's predictions were primarily driven by lesion regions identified during segmentation, supporting the interpretability of the lesion-centric pipeline.

9. Discussion

The results show that using segmentation to find lesions makes deep learning models for classifying breast ultrasounds more reliable. The model is less affected by background tissue and imaging artifacts when the lesion area is separated before classification. The EfficientNetB0 classifier did a good job of classifying even though the dataset was not very big. Transfer learning from ImageNet helped the model learn useful visual features and cut down on the amount of training it needed.
Adding Grad-CAM makes it even easier for doctors to understand by showing areas that affect classification decisions. This openness is important for making doctors trust AI-assisted diagnostic systems. But there are some problems that need to be pointed out. First, the BUSI dataset only has a small number of images, which could make it hard to apply the findings to larger groups of patients. Second, ultrasound images can look very different depending on the imaging device and the way they were taken. Finally, we need to test the proposed framework on datasets from more than one institution to make sure it works well.
To make the model even better at generalizing, future research should look into combining multi-modal imaging data with bigger clinical datasets.

10. Conclusion

This study put forward a segmentation-guided, lesion-centric deep learning framework for classifying breast ultrasound images. The framework combines U-Net-based lesion segmentation, ROI extraction, EfficientNet-based classification, and Grad-CAM explainability. Testing on the BUSI dataset showed that segmentation-guided classification makes the results easier to understand and keeps the classification performance balanced for benign, malignant, and normal cases. Grad-CAM visualizations showed that the model's predictions were based on areas that were important to the lesions.

The suggested framework offers a solid and understandable method for detecting breast cancer using ultrasound, and it lays the groundwork for future testing on larger clinical datasets from multiple centers.

11. Reference:

[1]World Health Organization. Breast cancer. 2024. Available online: https://www.who.int/news-room/fact-sheets/detail/breast-cancer (accessed on 29 December 2024).

[2] Arnold M, Morgan E, Rumgay H, Mafra A, Singh D, Laversanne M, et al. Current and future burden of breast cancer: Global statistics for 2020 and 2040. The Breast. 2022;66:15–23.

[3] Cheng JZ, Ni D, Chou YH, Qin J, Tiu CM, Chang YC, et al. Computer-aided diagnosis with deep learning architecture: Applications to breast lesions in ultrasound images and pulmonary nodules in CT scans. Scientific Reports. 2016;6:24454.

[4] Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. Data in Brief. 2020;28:104863. https://doi.org/10.1016/j.dib.2019.104863

[5] Rouhi R, Jafari M, Kasaei S, Keshavarzian P. Benign and malignant breast tumors classification based on region growing and CNN segmentation. Expert Systems with Applications. 2015;42(3):990–1002. https://doi.org/10.1016/j.eswa.2014.09.020

[6] Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, et al. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE Transactions on Medical Imaging. 2016;35(5):1299–1312.

[7] Kooi T, Litjens G, van Ginneken B, Gubern-Mérida A, Sánchez CI, Mann R, et al. Large scale deep learning for computer aided detection of mammographic lesions. Medical Image Analysis. 2017;35:303–312. https://doi.org/10.1016/j.media.2016.07.007

[8] Mendel K, Li H, Sheth D, Giger ML. Transfer learning from convolutional neural networks for computer-aided diagnosis: A comparison of digital breast tomosynthesis and full-field digital mammography. Academic Radiology. 2019;26(6):735–743. https://doi.org/10.1016/j.acra.2018.06.019

[9] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional networks for biomedical image segmentation. In: Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI). 2015;234–241.