Comparative Study and Enhancement of Segmentation Architectures for Breast Tumor Segmentation

Anantajeet Devaraj
Computer Science
Toronto Metropolitan University
Toronto, Canada
anantajeet.devaraj@torontomu.ca

Natasha Narasimhan
Computer Science
Toronto Metropolitan University
Toronto, Canada
natasha.narasimhan@torontomu.ca

***Abstract*—Breast cancer is one of the most common cancers among women and early detection of tumors greatly improves treatment outcomes. Tumor segmentation in breast ultrasound images is an important step in computer aided diagnosis, but it remains challenging because the ultrasound images often have low contrast, noise, and large variations in tumor shapes and appearance. In this work, we compare four widely used segmentation models: U-Net, U-Net++, Attention U-Net, and DeepLabV3. To improve their ability to segment tumors, we also propose a novel module called the Gabor Texture Attention Module (GTAM). GTAM improves the models ability to capture texture patterns that are important in ultrasound images, GTAM uses a set of fixed Gabor filters to highlight texture-rich regions and generate a spatial attention map that is used to guide the models to focus on areas that are more likely to contain tumors. We integrate GTAM into each of the four models and evaluate them on the BUSI and UDIAT breast ultrasound datasets. Our experimental results show that adding GTAM improves the segmentation performance, suggesting that texture-based attention can be useful for breast tumor segmentation. The trained models and the source code of the proposed module can be accessed at:**

https://github.com/Anantajeet23/CPS-843-Final-Project

***Index Terms*—Biomedical imaging, Gabor filters, semantic segmentation.**

I. Introduction

Breast cancer is one of the most commonly diagnosed cancers in women worldwide [1]. Ultrasound imaging is widely used in clinical screening because it is a safe and effective way for identifying abnormalities in breast tissue. However, breast ultrasound images suffer from low contrast, speckle noise and large variations in tumor appearance, all of which make manual interpretation difficult and highly dependent on the radiologists' experience. These challenges have motivated the development of automatic segmentation systems to help radiologists detect tumors more consistently.

Convolutional neural networks have been widely used in medical image segmentation and several architectures have become common baselines. These architectures were not specifically created for breast ultrasound images. However, they are often used as baselines due to their strong performance on medical segmentation tasks. Some of these models include the U-Net [2] which is one of the most popular segmentation architectures because of its skip connections and encoder-decoder design. U-Net++ [3] which modifies the skip connections of the U-Net by the addition of nested dense paths to improve feature refinement. Attention U-Net [4] which adds attention gates to the U-Net skip connections to help the decoder focus on the relevant structures in the image. And DeepLabV3 [5] which uses atrous convolution layers and an Atrous Spatial Pyramid Pooling (ASPP) layout along with post-processing Conditional Random Fields (CRFs) to achieve higher accuracy.

Several recent works propose new architectures that extend or replace these standard baselines. Sulaiman et al. [6] introduced an attention based U-Net that improves focus on the tumor regions. Aslam et al. [7] proposed a hybrid attention model to handle low contrast breast lesions. Shih et al. [8] combined CBAM with a custom DRA-UNet to better capture tumor boundaries. These papers suggest the standard baseline models may not perform effectively on breast ultrasound images, which motivated the development of Gabor Texture Attention Module (GTAM), an enhancement module that extracts important texture patterns in ultrasound images to generate a spatial attention map which can guide the models to focus on regions which are more likely to contain tumors.

GTAM uses Gabor filters [9], to extract oriented edges and texture patterns. Gabor filters have previously been incorporated into convolutional neural networks (CNN) to improve early feature extraction. Luan et al. [10] proposed the Gabor Convolutional Network to reduce the number of filter parameters in deep CNNs. Some earlier works also incorporate Gabor filters into CNN for face detection [11] and handwritten digit recognition [12].

The main contributions of this work are as follows:

- To the best of our knowledge, Gabor filters have not been used as an attention prior inside segmentation architectures.

- We introduce a novel Gabor Texture Attention Module (GTAM) that uses fixed Gabor filters to provide a texture-aware spatial attention prior.

- We compare four baseline models, with and without GTAM, on two breast ultrasound datasets (BUSI and UDIAT) to evaluate both model performance and the effect of the GTAM module.

II. Methods

A. U-Net

U-Net [2] was proposed by Ronneberger et al. as a fully convolutional network for biomedical image segmentation, designed to work well even with only a small number of images. The architecture has a symmetric encoder-decoder structure where the contracting path repeatedly applies 3x3 convolutions with ReLU activation followed by 2x2 max pooling, which reduces the spatial resolution while increasing the number of feature channels capturing high-level features. The expansive path performs up-convolutions to recover spatial resolution, then concatenates the upsampled features with the corresponding high-resolution feature maps from the encoder via skip connections. These skip connections allow the decoder to use both precise low-level details and high-level semantic information to achieve more accurate object boundaries. A final 1x1 convolution is used to map the decoder features to a segmentation map.

U-Net was originally shown to outperform sliding window CNNs on neural structures and cell segmentation benchmarks and has since become a standard baseline for many medical image segmentation tasks.

B. U-Net++

U-Net++ [3] is a variant of U-Net that tries to improve the segmentation quality by redesigning the skip connections between the encoder and decoder. Instead of directly concatenating encoder and decoder feature maps, Zhou et al. proposed the U-Net++ where a series of convolution blocks are inserted along each skip path, arranged in a nested and dense fashion. These nested paths gradually refine the high-resolution encoder features so that by the time they reach the decoder they are more semantically meaningful. This helps the network learn an easier optimization problem and recover finer details in the segmentation.

U-Net++ also adds deep supervision by attaching output heads at multiple decoder layers and combining their losses during training to achieve higher accuracy. Experiments in the original paper show that this design yields higher IoU than standard U-Net on several medical segmentation tasks.

C. Attention U-Net

Oktay et al. proposed Attention U-Net [4] which extends the standard U-Net by adding attention gates on the skip connections between the encoder and decoder. The main idea is that not all spatial locations in a feature map are equally relevant. In a normal U-Net all the encoder feature maps are passed to the decoder unchanged, so the decoder has to learn to ignore background responses on its own.

In Attention U-Net, each skip feature map is filtered by an attention gate that looks at two inputs, the encoder features at that scale and a coarse gating signal coming from the decoder path. The gate computes attention coefficients between 0 and 1 for every spatial position using a small network made of 1x1 convolutions, a ReLU and a sigmoid. These coefficients are then multiplied elementwise with the skip features so the responses in irrelevant regions are suppressed while responses near the target are kept or amplified.

The authors showed that inserting attention gates into U-Net improves segmentation on pancreas and multi-organ CT scans, and the extra computation and parameters are relatively small compared to the base model.

D. DeepLabV3

Originally, DeepLabV1 [13] proposed a Deep Convolutional Neural Network (DCNN) using a VGG16 backbone, it also introduced atrous convolutions, and CRF layers. DeepLabV2 [14] imporved upon V1 by adding Atrous Spatial Pyramid Pooling (ASPP) to handle multiscale objects, replacing downsampling operations with upsampling in the last max pooling layers for denser feature maps, and providing support for other backbone networks such as ResNet.

Then, DeepLabV3 was proposed by Chen et al. [5] to build on the previous versions and improve two main challenges in applying DCNNs for semantic segmentation tasks: reduced feature sampling and multiscale objects.

Atrous convolutions are used to extract dense features and increase feature sampling without reducing spatial resolution. In DeepLabV3 the downsampling operators in the last few layers of the network are removed: standard 3x3 convolutions and pooling operations are replaced with 3x3 atrous convolutions implemented by inserting zeros between filter weights. The atrous rate is adjusted depending on the chosen output stride to preserve the feature map resolution. On top of this atrous backbone, DeepLabV3 employs the ASPP module, which applies multiple atrous convolutions in parallel with different atrous rates to capture information at various spatial scales. Batch normalization is applied within the ASPP module. As the atrous rate increases, kernel weight density and the possible valid filters decrease, so DeepLabV3 also augments ASPP with image-level features by applying global average pooling to the final feature map, the result is passed through a 1×1 convolution with batch normalization, and then bilinearly upsampled to match the spatial resolution of the ASPP outputs.

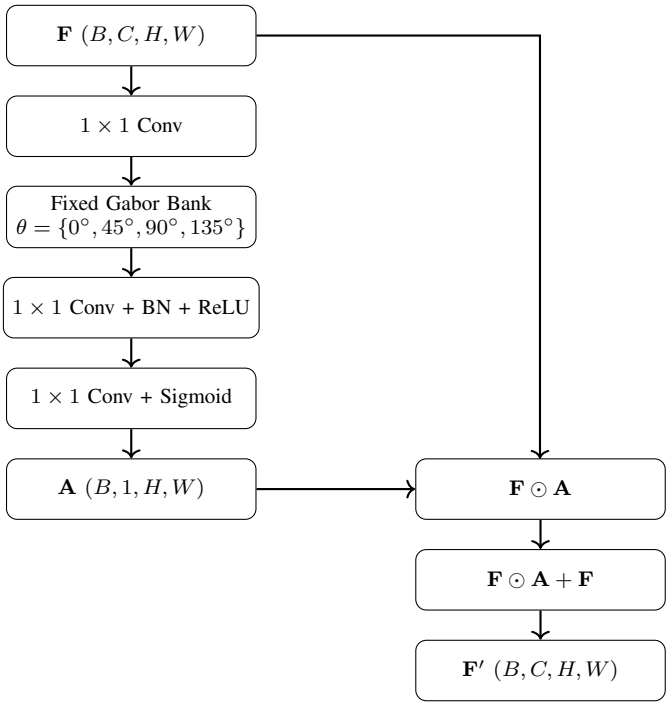


Fig. 1: Architecture of the GTAM. Fixed Gabor filters extract orientation-aware texture features, which are then fused to generate a spatial attention map that refines the input features.

E. Gabor Texture Attention Model

The Gabor Texture Attention Module (GTAM) is a lightweight plug-in designed to highlight texture-rich regions in breast ultrasound images. Unlike standard CNN layers that learn filters from data, GTAM uses fixed Gabor filters, which are well known for capturing orientations, edges and local texture patterns all of which are highly relevant in ultrasound images due to speckle noise, low contrast and irregular tumor boundaries.

GTAM consists of two main stages:

1) Gabor Filter Response Extraction.
2) Spatial Attention Generation.

A high-level overview of the proposed module is illustrated in Fig. 1. Given an input feature map $\mathbf{X} \in \mathbb{R}^{R \times H \times C}$, the proposed Gabor Texture Attention Module produces a spatial attention map $\mathbf{A} \in \mathbb{R}^{1 \times H \times C}$. The attention map $\mathbf{A}$ is then applied to the input feature map $\mathbf{X}$ via element-wise multiplication to enhance texture-relevant spatial locations.

1. Gabor Filter Response Extraction

Gabor filters model the response of simple cells in our eyes and are widely used for detecting texture and oriented surfaces. Each 2-D Gabor filter is defined as such:

$$G(x, y) = \exp\left(-\frac{x'^2 + \gamma^2 y'^2}{2\sigma^2}\right) \cos\left(2\pi \frac{x'}{\lambda} + \psi\right) \quad (1)$$

where $x' = x\cos\theta + y\sin\theta$, $y' = -x\sin\theta + y\cos\theta$, and θ controls the orientation, λ the wavelength, σ the scale, γ the aspect ratio, and ψ the phase offset.

In GATM, we use a fixed bank of Gabor filters with several orientations (0°, 45°, 90° and 135°) which are applied to the input feature map. For each orientation k, we compute a response:

$$\mathbf{R}_k = \mathbf{G}_k * \mathbf{X}_{\text{in}} \quad (2)$$

Where $\mathbf{X}_{\text{in}}$ is the grayscale input image at the first encoder layer. The responses $\{\mathbf{R}_k\}$ capture different directional textures such as vertical edges and slanted edges.

These responses are stacked along the channel dimension:

$$\mathbf{R} = \text{Concat}(\mathbf{R}_1, \mathbf{R}_2, \ldots, \mathbf{R}_K) \quad (3)$$

A 1×1 convolution is then applied to fuse them into a single feature map:

$$\mathbf{F} = \text{Conv}_{1\times1}(\mathbf{R}) \quad (4)$$

This produces a summary of all the important textures across the image.

2. Spatial Attention Generation

The fused texture representation of $\mathbf{F}$ is passed through a small attention head to produce a spatial probability map. First a convolution layer is used to aggregate local information:

$$\mathbf{Z} = \text{Conv}_{3\times3}(\mathbf{F}) \quad (5)$$

Then we apply the sigmoid activation function to generate the map:

$$\mathbf{A} = \text{Sigmoid}(\mathbf{Z}) \quad (6)$$

Where:

$$\text{Sigmoid}(z) = \frac{1}{1 + e^{-z}}. \quad (7)$$

The resulting map $\mathbf{A}$ assigns high values to regions with strong texture cues and low values to smooth background areas.

Finally, the attention map is used to refine the input feature map by element wise multiplication across all channels:

$$\mathbf{X}_{\text{out}} = \mathbf{X} \odot \mathbf{A} \tag{8}$$

This mechanism encourages the model to focus more on regions where tumors are likely to appear while suppressing irrelevant background noise.

III. Experiments

A. Datasets

We evaluated the four baseline models U-Net, U-Net++, Attention U-Net, DeepLabV3 with ResNet-50 backbone and their GTAM-enhanced variants on two publicly available breast ultrasound datasets: BUSI [15] and UDIAT [16]. The BUSI dataset contains 780 ultrasound images categorized into normal, benign, and malignant, and each image is paired with a ground truth mask. The UDIAT dataset is smaller consisting of 133 ultrasound images with corresponding binary masks.

For both datasets, we followed a 70/20/10 split for training, validation, and testing. All images and masks were resized to 256x256, converted to single channel format and normalized before input into the models.

B. Implementation Details

All models were trained using the cross entropy loss function:

$$l_n = -\sum_{c=1}^{C} w_c \log \frac{\exp\left(x_{n,c}\right)}{\sum_{i=1}^{C} \exp\left(x_{n,c}\right)} y_{n,c} \tag{9}$$

$$l(x, y) = \frac{\sum_{n=1}^{N} l_n}{N} \tag{10}$$

For (9) and (10), x is the input, y is the target class, w is the weight, C is the number of classes, and N spans the batch dimensions.

And optimized with the Adam optimizer at a learning rate of 0.0001. We used a batch size of 8 and trained each model for up to 60 epochs. To prevent overfitting, we applied early stopping based on the validation dice score and used several data augmentations such as rotations and flips in both horizontal and vertical directions. For the BUSI dataset the early stopping patience level was set to 10 and for the smaller UDIAT dataset patience was reduced to 4 due to the limited number of training images.

All training was performed on an Apple Mac GPU using the Pytorch framework and the GTAM module was inserted into each architecture after the first encoder block between layer 1 and layer 2, so it could refine early texture patterns before deeper processing.

To evaluate performance, we use the following metrics:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{11}$$

$$Precision = \frac{TP}{TP + FP} \tag{12}$$

$$Recall = \frac{TP}{TP + FN} \tag{13}$$

$$Specificity = \frac{TN}{TN + FP} \tag{14}$$

$$Dice\ Score = \frac{2TP}{2TP + FP + FN} \tag{15}$$

$$IOU = \frac{TP}{TP + FP + FN} \tag{16}$$

For (11), (12), (13), (14), (15), (16), TP, TN, FP, and FN are the number of true positives, true negatives, false positives, and false negatives, respectively.

C. Results

On the BUSI dataset, Attention U-Net is the strongest baseline, with a Dice of 50.36% and IoU of 42.16%. U-Net and U-Net++ perform slightly worse, with Dice scores of 43.41% and 44.29% respectively, while DeeplabV3 reaches a Dice of 47.82%. Adding GTAM generally improves segmentation quality. The Dice of U-Net improves by 3.87% and IoU by 3.72%. U-Net++ with GTAM also increased Dice by 2.48% and IoU by 3.89%. The greatest gain occurs for Attention U-Net, where GTAM raises the Dice by 4.04% and IoU by 3.55%, while also improving recall by 5.49%. DeepLabV3 also shows moderate improvement on BUSI, Dice improves by 2.04% and IoU by 1.78%, although precision and specificity trade off slightly.

On the UDIAT dataset, Attention U-Net again provides the best baseline performance with a Dice of 58.69% and IoU of 45.31%, clearly outperforming U-Net, U-Net++ and DeepLabV3. Adding GTAM improves most architectures here as well. U-Net Dice improves by 3.15% and U-Net++ Dice by 9.81%, with corresponding IoU gains. DeepLabV3 benefits strongly from GTAM on UDIAT, its Dice raises by 14.19% and IoU by 12.48%, along with higher precision and specificity. Attention U-Net with GTAM achieves the best overall performance on this dataset with Dice of 60.97% and IoU of 49.36%, and also maintains a good balance between precision and recall.

UDIAT is a small dataset and training is more sensitive to noise and random initialization, we believe some of the observed improvements with GTAM on this dataset might be slightly exaggerated and should be interpreted with caution.

Overall, the experiments show that GTAM consistently improves or at least maintains performance across multiple architectures and two different datasets. The results show that texture-aware attention prior can help segmentation models focus on tumor regions more effectively.

TABLE I

COMPARISON OF BASELINE MODELS ON BUSI AND UDIAT DATASET.

Dataset	Models	Accuracy	Precision	Recall	Specificity	Dice score	IoU
BUSI	U-Net	92.35	55.48	43.89	**98.05**	43.41	34.41
	U-Net++	92.27	**59.52**	43.97	97.84	44.29	35.41
	Attention U-Net	**93.08**	59.40	**50.30**	98.01	**50.36**	**42.16**
	DeepLabV3	92.35	56.93	49.68	96.98	47.82	38.81
UDIAT	U-Net	89.26	47.28	66.82	94.22	43.87	34.45
	U-Net++	88.53	54.79	53.24	97.53	35.95	27.38
	Attention U-Net	92.37	**65.75**	**76.33**	97.72	**58.69**	**45.31**
	DeepLabV3	**92.55**	54.90	40.44	**98.86**	42.10	33.28

Fig. 2: Segmentation masks produced by the baseline models along with ground truth.

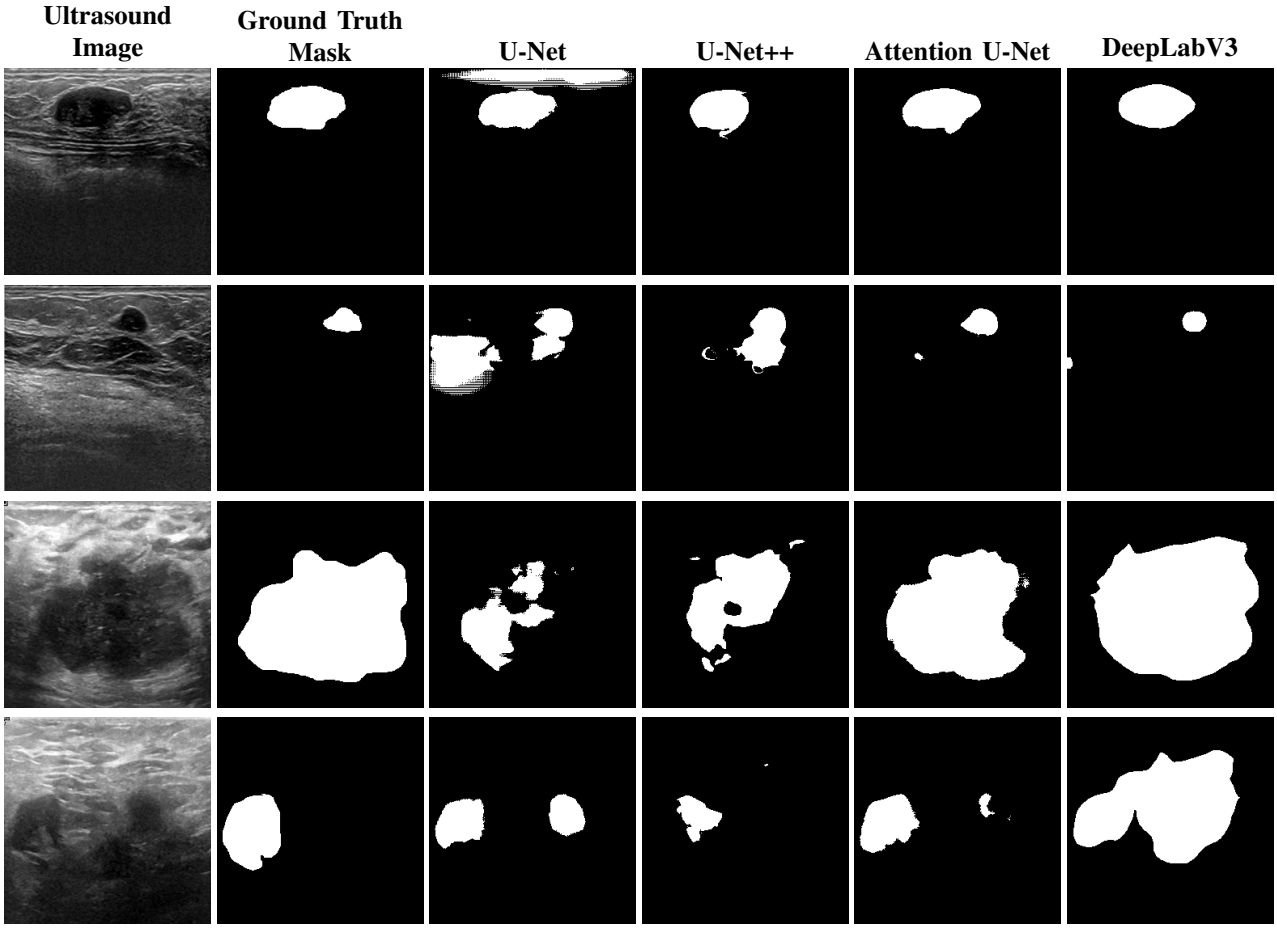


TABLE II

COMPARISON OF U-NET AND U-NET WITH GTAM ON BUSI AND UDIAT DATASET.

Dataset	Models	Accuracy	Precision	Recall	Specificity	Dice score	IoU
BUSI	U-Net	92.35	55.48	43.89	98.05	43.41	34.41
	U-Net + GTAM	92.72	57.63	49.52	96.88	47.28	38.13
	Improvement	0.37	**2.15**	**5.63**	-1.17	**3.87**	**3.72**
UDIAT	U-Net	89.26	47.28	66.82	94.22	43.87	34.45
	U-Net + GTAM	90.75	71.48	50.86	99.19	47.02	37.03
	Improvement	1.49	**24.20**	-15.96	**4.97**	**3.15**	**2.58**

Fig. 3: Segmentation masks produced by U-Net and U-Net with GTAM.

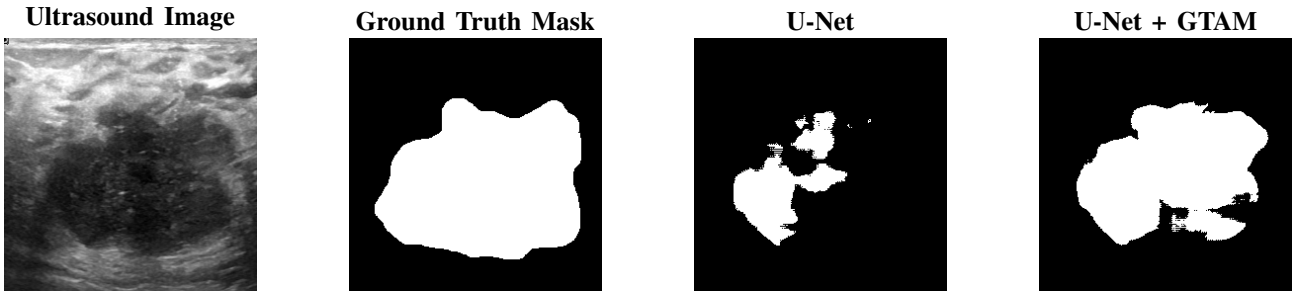


TABLE III

COMPARISON OF U-NET++ AND U-NET++ WITH GTAM ON BUSI AND UDIAT DATASET.

Dataset	Models	Accuracy	Precision	Recall	Specificity	Dice score	IoU
BUSI	U-Net++	92.27	59.52	43.97	97.84	44.29	35.41
	U-Net++ + GTAM	92.73	60.37	46.18	98.40	46.77	38.30
	Improvement	0.46	0.85	**2.21**	0.56	**2.48**	**2.89**
UDIAT	U-Net++	88.53	54.79	53.24	97.53	35.95	27.38
	U-Net++ + GTAM	89.68	54.08	65.88	95.59	45.76	33.48
	Improvement	1.15	-0.71	**12.64**	-1.94	**9.81**	**6.10**

Fig. 4: Segmentation masks produced by U-Net++ and U-Net++ with GTAM.

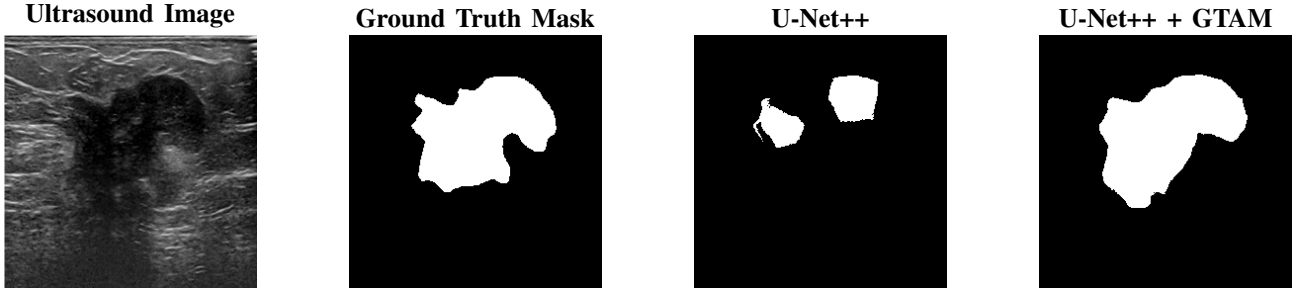


TABLE IV

COMPARISON OF ATTENTION U-NET AND ATTENTION U-NET WITH GTAM ON BUSI AND UDIAT DATASET.

Dataset	Models	Accuracy	Precision	Recall	Specificity	Dice score	IoU
BUSI	Attention U-Net	93.08	59.40	50.30	98.01	50.36	42.16
	Attention U-Net + GTAM	93.77	59.93	55.79	97.19	54.40	45.71
	Improvement	0.69	0.53	**5.49**	-0.82	**4.04**	**3.55**
UDIAT	Attention U-Net	92.37	65.75	76.33	97.72	58.69	45.31
	Attention U-Net + GTAM	92.93	72.60	68.70	98.00	60.97	49.36
	Improvement	0.56	**6.85**	-7.63	0.28	**2.28**	**4.05**

Fig. 5: Segmentation masks produced by Attention U-Net and Attention U-Net with GTAM.

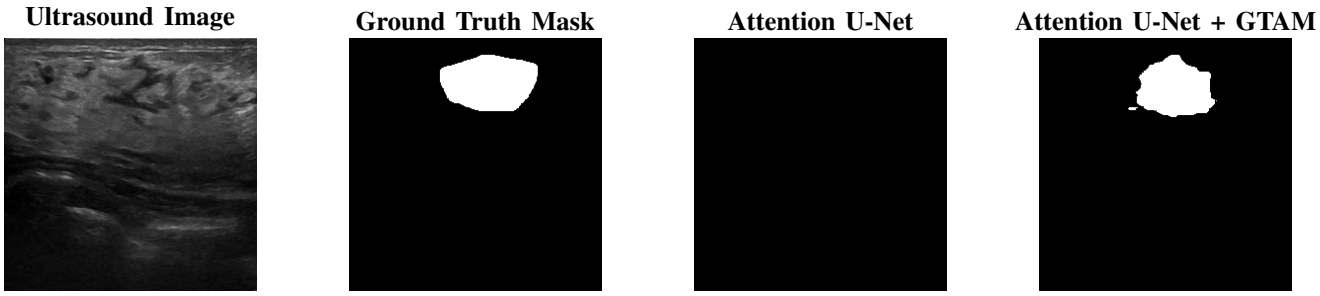
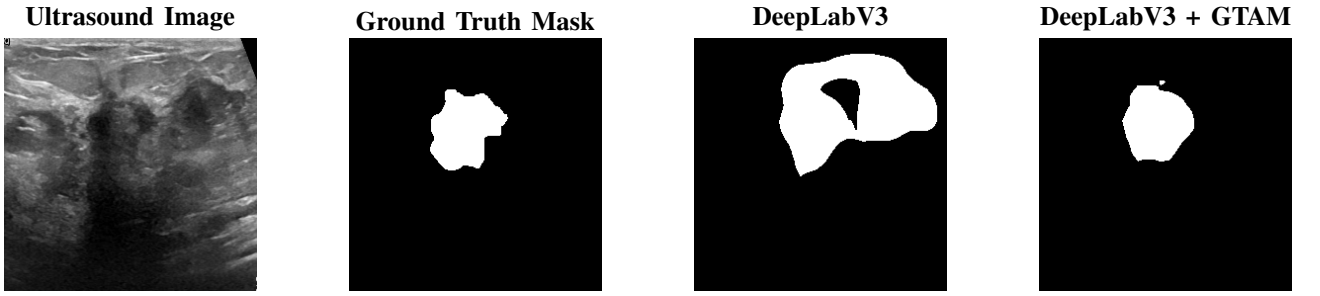


TABLE V

COMPARISON OF DEEPLABV3 AND DEEPLABV3 WITH GTAM ON BUSI AND UDIAT DATASET.

Dataset	Models	Accuracy	Precision	Recall	Specificity	Dice score	IoU
BUSI	DeepLabV3	92.35	56.93	49.68	96.98	47.82	38.81
	DeepLabV3 + GTAM	92.29	52.00	58.58	94.96	49.86	40.59
	Improvement	-0.06	-4.93	**8.90**	-2.02	**2.04**	**1.78**
UDIAT	DeepLabV3	92.55	54.90	40.44	98.86	42.10	33.28
	DeepLabV3 + GTAM	93.02	77.11	54.46	99.06	56.29	45.76
	Improvement	0.47	**22.21**	**14.02**	0.20	**14.19**	**12.48**

Fig. 6: Segmentation masks produced by DeepLabV3 and DeepLabV3 with GTAM.



D. Heatmap Visualization of GTAM

To better understand the effect of GTAM, we visualize the spatial attention maps produced by the module for the BUSI dataset. We overlay the output of the GTAM module as a heatmap on the input ultrasound image. These visualizations allow us to evaluate whether the attention mechanism consistently focuses on meaningful regions such as tumor boundaries and lesion areas.

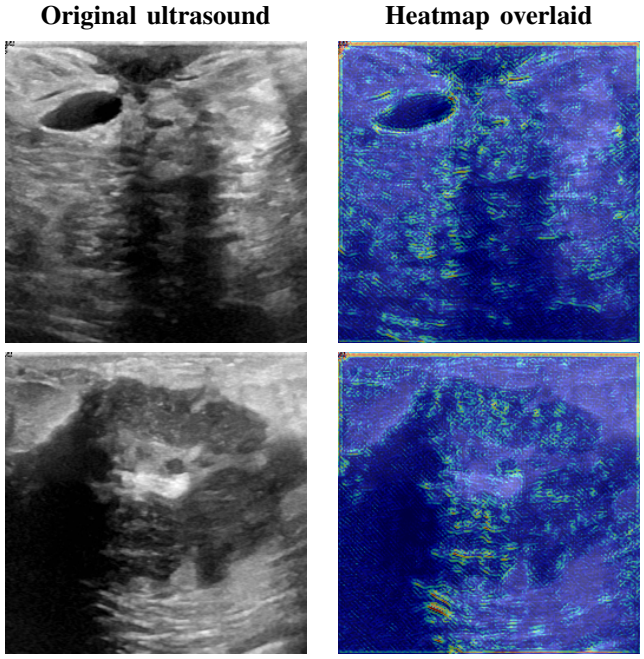


Fig. 7: Original ultrasound and corresponding heatmap overlaid.

The GTAM highlights textured tissue structures and edges including in and around the lesion, while suppressing the background uniformly.

IV. Conclusion

This paper has compared four widely used Segmentation architectures on breast ultrasounds and proposed a novel enhancement module that gives consistent improvements on most metrics, especially Dice and IoU. Our work can be further extended and refined, the next step is to test GTAM on more diverse breast ultrasound datasets, and on other medical modalities such as prostate or liver ultrasounds to see how well the idea generalizes. The GTAM itself can be improved as well by exploring more advanced attention mechanisms.

Member Contributions

Anantajeet developed the GTAM and implemented U-Net, U-Net++, and all the GTAM enhanced models.

Natasha preprocessed the datasets and implemented Attention U-Net, and DeepLabV3.

The rest of the work was shared evenly.

References

[1] WHO, "Breast cancer," World Health Organization, Aug. 14, 2025. https://www.who.int/news-room/fact-sheets/detail/breast-cancer

[2] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," Lecture Notes in Computer Science, vol. 9351, pp. 234–241, 2015, doi: https://doi.org/10.1007/978-3-319-24574-4_28

[3] Z. Zhou, M. M. R. Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: Redesigning Skip Connections to Exploit Multiscale Features in Image Segmentation," IEEE Transactions on Medical Imaging, vol. 39, no. 6, pp. 1–1, 2019, doi: https://doi.org/10.1109/tmi.2019.2959609.

[4] O. Oktay et al., "Attention U-Net: Learning Where to Look for the Pancreas," arXiv.org, 2018. https://arxiv.org/abs/1804.03999

[5] L. -C. Chen, G. Papandreou, F. Schroff, and H. Adam, "Rethinking atrous convolution for semantic image segmentation," arXiv.org, 2017. https://arxiv.org/pdf/1706.05587

[6] A. Sulaiman et al., "Attention based UNet model for breast cancer segmentation using BUSI dataset," Scientific Reports, vol. 14, no. 1, Sep. 2024, doi: https://doi.org/10.1038/s41598-024-72712-5.

[7] M. A. Aslam et al., "A hybrid attention network for accurate breast tumor segmentation in ultrasound images," Scientific Reports, vol. 15, no. 1, pp. 39633–39633, Nov. 2025, doi: https://doi.org/10.1038/s41598-025-23213-6.

[8] Y.-H. Shih, C.-Y. Wu, R.-C. Lin, and C.-T. Huang, "Segmentation of Tumor Regions in BUSI Breast Ultrasound Images Based on DRA-UNet Model with CBAM," pp. 105–113, Mar. 2025, doi: https://doi.org/10.1145/3749859.3749873

[9] D. Gabor, "Theory of communication. Part 1: The analysis of information," J. Inst. Elect. Eng.-III, Radio Commun. Eng., vol. 93, no. 26, pp. 429–441, Jul. 1946.

[10] S. Luan, C. Chen, B. Zhang, J. Han, and J. Liu, "Gabor Convolutional Networks," IEEE Transactions on Image Processing, vol. 27, no. 9, pp. 4357–4366, Sep. 2018, doi: https://doi.org/10.1109/tip.2018.2835143

[11] B. Kwolek, "Face Detection Using Convolutional Neural Networks and Gabor Filters," Lecture notes in computer science, pp. 551–556, Jan. 2005, doi: https://doi.org/10.1007/11550822_86

[12] Calderón, A.; Roa-Valle, S.; Victorino, J. Handwritten digit recognition using convolutional neural networks and Gabor filters. In Proceedings of the 2003 International Conference on Computational Intelligence, Cancun, Mexico, 19–21 May 2003.

[13] L. -C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected CRFs," in ICLR, 2015 [Online]. Available: https://arxiv.org/pdf/1412.7062v3

[14] L. -C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," arXiv.org, 2017. https://arxiv.org/pdf/1606.00915

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

[16] M. H. Yap et al., "Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks," IEEE Journal of Biomedical and Health Informatics, vol. 22, no. 4, pp. 1218–1226, Jul. 2018, doi: https://doi.org/10.1109/jbhi.2017.2731873