

Efficient Deep Networks for Tumor Segmentation in Breast Ultrasound Imaging

Sharmila Meinam¹, Kishorjit Nongmeikapam² and N. Basanta Singh³

^{1,3}Department of Electronics and Communication Engineering, Manipur Institute of Technology, Manipur University, Manipur, India

Email: sharmila_meinam@yahoo.com, basanta_n@rediffmail.com

²Department of Computer Science and Engineering, Indian Institute of Information Technology Manipur, Manipur, India

Email: kishorjit@iiitmanipur.ac.in

Abstract— Breast cancer is regarded as one of the major causes of cancer-related deaths among women in today's world. Accurate detection at an early stage plays a crucial role in the improvement of survival rates. In such a scenario, ultrasound imaging has emerged as a safe and cost-effective screening modality, particularly in resource-constrained settings. However, automated tumor segmentation in ultrasound images is a challenging task due to low contrast, speckle noise and heterogeneous lesion boundaries. This study presents a comparative evaluation of two lightweight deep learning architectures, MobileNetV2–UNet and MobileNetV2–DeepLabV3, for automated breast tumor segmentation using the BUSI dataset. The study focuses exclusively on benign and malignant categories. Each model is assessed based on segmentation accuracy (Dice coefficient, IoU, precision, recall) and computational efficiency (inference time and model size). Experimental results demonstrate that MobileNetV2–UNet achieved 79.50% IoU, 86.14% Dice coefficient and an overall accuracy of 97.95%, with an inference speed of 6.54 FPS. In comparison, MobileNetV2–DeepLabV3 attained 78.61% IoU, 86.06% Dice coefficient and 98.02% accuracy, but with a slower inference speed of 2.02 FPS. Additionally, inference benchmarking demonstrates that MobileNetV2–UNet delivers 3× faster inference speed, making it better suited for edge deployment in mobile health screening applications. Both models demonstrated robust tumor localization performance and edge deployability. This work contributes an efficient, scalable solution for automated breast lesion segmentation, relevant for deployment in clinical settings with limited computational resources.

Index Terms— breast cancer, tumor segmentation, ultrasound images, MobileNetV2, edge deployable.

I. INTRODUCTION

Breast cancer continues to be one of the most common and life-threatening malignancies among women globally. According to recent reports by the World Health Organization, breast cancer affects an estimated 2.3 million people annually, making it the most diagnosed cancer worldwide [1]. As per the latest findings, 1 in every 8 women is expected to develop breast cancer in her life. Breast cancer, if detected at an early stage along with timely intervention, can significantly improve prognosis and survival rates. This highlights the critical need for reliable and accessible diagnostic tools.

Ultrasound imaging is a widely popular, safe and cost-effective modality. Ultrasound imaging is non-invasive in nature and possesses the capability of capturing high-resolution images of breast tissue [2] for screening breast cancer. Not like the ionizing techniques, ultrasound provides real-time visualization and is particularly suitable for dense breast tissues. Nevertheless, interpretation of ultrasound images manually is highly operator-dependent and prone to variability [3, 4]. Eventually, the need arises for developing automated computer-aided diagnostic (CAD) systems that can assist clinicians in tumor localization and segmentation [5].

In the last few years, deep convolutional neural networks (CNNs) have been extensively applied to breast tumor segmentation. Though there has been great advancement in deep learning-based medical image analysis, automated segmentation of breast tumors in ultrasound images still remains as a challenging task [6]. This can be attributed to issues such as low contrast, speckle noise and irregular tumor boundaries in sonographic images [7]. Moreover, many state-of-the-art segmentation networks are computationally intensive. Hence they are not suitable for real-time use on low-resource environments such as rural clinics or portable medical devices [8].

In this research, we design two deep networks, MobileNetV2–UNet and MobileNetV2–DeepLabV3 for semantic segmentation of breast tumors on the ultrasound images of BUSI Dataset. The models have been designed to offer a balance between accuracy and computational cost. These models have been chosen for their proven capability in real-time semantic segmentation tasks while maintaining lightweight architectures. Additionally, the study aims to analyze the trade-offs between performance and computational efficiency, particularly in the context of real-time applications. The major contributions of this work include a benchmark evaluation of the selected lightweight deep learning architectures using a consistent experimental setup designed for breast tumor segmentation. This study provides both quantitative and qualitative analyses to highlight the trade-offs between segmentation accuracy, model size and inference time. Furthermore, the findings demonstrate the feasibility of using these models for real-time and mobile deployment. This supports the development of scalable, cost-effective solutions for breast cancer screening, particularly in under-resourced healthcare settings.

The remainder of the manuscript has been organized as follows. Section II reviews the related work. The workflow of the proposed methodology is described in Section III. Section IV discusses the experimental results. Finally, Section V concludes the study and highlights the directions for future research.

II. RELATED WORK

In medical imaging, breast tumor segmentation has long been acknowledged as a critical step for computer-aided diagnosis (CAD) systems to improve early cancer detection [9]. Early research focused on classical machine learning approaches. These approaches relied on handcrafted features such as texture, intensity, shape descriptors which are extracted from ultrasound or mammographic images. Horsch et al. [10] pioneered one of the first automated methods to segment breast lesions in ultrasound imagery, laying the groundwork for CAD in breast ultrasound. Subsequent studies engineered various texture features. As for example, Chen et al. [11] used nearly setting-independent sonographic features with an ANN classifier to distinguish benign from malignant lesions. Traditional classifiers were applied to identify tumor regions. Chang et al. [12] notably improved ultrasound tumor discrimination by combining speckle-emphasizing texture features with a Support Vector Machine (SVM) classifier. Across the literature, researchers explored a range of supervised classifiers for tumor segmentation and classification, including SVMs, k-Nearest Neighbors (k-NN), Decision Trees and ensemble methods like Random Forests [13, 14]. For instance, Al-Fahaidy et al. [15] confirmed that using wavelet-transform coefficients and gray-level co-occurrence matrix (GLCM) texture features alongside an SVM yielded robust performance in mammogram tumor detection. The author in [16] reported that using k-NN with GLCM texture features slightly outperformed SVM for classifying breast lesion regions. However, traditional methods often faced problems due to low contrast and speckle noise in ultrasound images. These issues made it hard for the models to generalize well. Xian et al. [17] also pointed out that speckle noise and a poor signal-to-noise ratio in ultrasound make it very difficult to accurately detect tumor boundaries. Because of this, classical approaches needed a lot of preprocessing, like speckle noise filtering and careful feature selection.

Recent advances in deep learning have led to the adoption of deep learning based methods for breast tumor segmentation. These methods perform better than the traditional machine learning approaches. Badawy et al. [18] conducted a pioneering study in breast ultrasound tumor segmentation by combining fuzzy logic preprocessing with CNNs. This hybrid approach significantly improved segmentation performance, thereby achieving a mean Intersection over Union (IoU) of about 78.70% compared to 49.61% without fuzzy preprocessing. In later years, Hesamian et al. [19] presented an extensive review of deep learning based techniques for segmentation of medical images. Through their research they concluded that modern CNNs generally achieve higher accuracy and robustness than earlier handcrafted-feature methods. Ronneberger et al.

[20], in the year 2015, first introduced the U-Net architecture. This architecture revolutionized biomedical image segmentation by using symmetric downsampling and upsampling paths with lateral skip connections to preserve high-resolution information. Chen et al. [21] designed HAAM, which is a hybrid adaptive attention module for segmenting nodules of breast from ultrasound images. Experimental results demonstrated that their method performed better compared with many state-of-the-art deep learning segmentation approaches on three public breast ultrasound image datasets. Over the years, the focus has shifted to efficient deep networks that can achieve faster inference without significant loss of accuracy. Paszke et al. [22] introduced ENet, one of the first ultra-lightweight segmentation CNNs. The model runs in real time, up to $18\times$ faster than conventional models, by drastically reducing parameter counts while still maintaining competitive accuracy. Zhou et al. [23] developed LAEDNet, a Lightweight Attention Encoder–Decoder Network for segmentation of ultrasound images. Their model could achieve a trade-off between segmentation accuracy and inference efficiency for the complex speckled imagery of ultrasound. Guo et al. [24] proposed LightBTSeg which is a dual-path knowledge distillation framework for segmentation of breast tumor on ultrasound images. Experimental results demonstrated that LightBTSeg can outperform many heavier segmentation models and hence deliver improved performance for breast tumor segmentation with much lower computational complexity. Anari et al. [25] presented an attention-augmented U-Net approach for breast ultrasound tumor segmentation that emphasizes both efficiency and interpretability. Their approach could achieve significant improvement in Dice scores on the BUSI ultrasound dataset. Following these trends, recent studies have proposed segmentation models built on efficient backbones such as MobileNetV2. Such models aim to balance high precision with low computational cost. Such lightweight architectures are particularly attractive for clinical use in real-time and deployment on portable devices. Inspired by these developments, our research evaluates two such architectures, MobileNetV2–UNet and MobileNetV2–DeepLabV3, for segmentation of breast tumor in ultrasound images. Our study also assesses whether the models can deliver accurate segmentation results with the speed and resource economy required in practical healthcare settings.

III. METHODOLOGY

The proposed workflow for Breast Tumor Segmentation has been presented in Fig. 1. The figure illustrates the complete framework starting from input ultrasound images through preprocessing, dataset splitting, model training using MobileNetV2–UNet and MobileNetV2–DeepLabV3 architectures. This is followed by evaluation using both quantitative and qualitative metrics. The final output includes segmented tumor masks and comparative performance analysis for potential deployment in real-time clinical settings.

A. Data Preprocessing

In this research, the experiments were conducted using the Breast Ultrasound Images (BUSI) dataset, a publicly available collection of annotated ultrasound scans categorized into three subfolders: benign, malignant and normal. Although the BUSI dataset includes a subset of normal images without tumor presence, this study focuses exclusively on the segmentation of benign and malignant lesions. Normal cases were excluded as their masks contain no region of interest relevant for tumor boundary detection. Hence, only the benign and malignant images were used for model training, validation and evaluation. To maintain consistency and to enable efficient training across all models, all images and their corresponding masks are resized to 256×256 pixels. Pixel intensities are normalized to the range $[0, 1]$ to ensure stable gradient flow during training.

B. Data Augmentation

A common problem faced by deep learning models is the problem of overfitting due to limited number of the training samples. To expand the dataset size, augmentation techniques were applied for enhanced model generalization and to mitigate overfitting. The augmentation techniques adopted included random rotations, horizontal and vertical flips and brightness adjustments. Because of these transformations, the dataset size was significantly increased from 647 original images to a total of 3235 images. This expansion provided a richer and more diverse training distribution thus supporting the development of robust segmentation models. In medical imaging, dataset size is generally limited. Such augmentations simulate variations commonly observed in clinical ultrasound imaging and help reduce model overfitting.

C. Model Architectures

To evaluate the effectiveness of lightweight segmentation models suitable for real-time applications and edge deployment, we benchmarked two prominent encoder–decoder based architectures. They are MobileNetV2–UNet and MobileNetV2–DeepLabV3 for the task of breast tumor segmentation in ultrasound images. We chose

these models for their ability to deliver an optimum balance between segmentation accuracy and computational efficiency. These features make them suitable for real-time deployment in clinical settings.

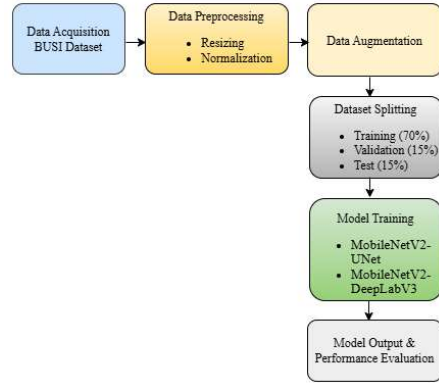


Fig. 1. Workflow diagram for Breast Tumor Segmentation.

Here, in this study, we have proposed a lightweight semantic segmentation architecture, MobileNetV2–UNet, designed for efficient and accurate segmentation of breast tumor in ultrasound images. It combines the computational efficiency of MobileNetV2 as an encoder with the spatial localization strength of the U-Net decoder. The encoder is composed of depthwise separable convolutions along with inverted residual blocks employing linear bottlenecks. Compared to traditional encoders like ResNet or VGG, MobileNetV2 substantially decreases both the model parameters and computational complexity. This characteristic facilitates its deployment on resource constrained devices such as mobile or portable diagnostic units. MobileNetV2 progressively reduces spatial dimensions while increasing feature complexity, enabling robust encoding of multi-scale lesion patterns. In the decoder, the model follows the symmetric U-Net architecture, consisting of upsampling blocks paired with skip connections from the encoder. These skip connections play a crucial role in restoring fine-grained spatial features that are typically lost during downsampling. This is particularly beneficial for delineating tumor boundaries in low-contrast ultrasound images.

The combination of MobileNetV2 with U-Net provides a good balance between speed, memory efficiency and segmentation accuracy. The model performs well without the need for much pre-processing or complex post-processing. This makes it suitable for use in real-time and deployment on edge devices. It combines strong feature extraction with precise spatial reconstruction, making it effective for automatic breast tumor segmentation. Its lightweight design makes it especially suitable for deployment in resource-constrained clinical settings. This balance of efficiency and precision enhances its practical applicability in real-world medical scenarios. Also, it demonstrates strong robustness against speckle noise, which is a common challenge in ultrasound imaging. Ben Ahmed et al. [26] developed a hybrid U-Net model using MobileNetV2 and VGG16 encoders for breast ultrasound segmentation. They fine-tuned these pre-trained encoders on a public dataset. Their approach could achieve dice scores and segmentation accuracy consistently above 80% on benign and malignant cases in the BUSI dataset. Wang et al. [27] designed the CSAU Net, a cross-scale attention-guided U Net model that combines CNN and transformer modules with gated dilated convolution and lightweight channel self-attention. Their research demonstrated enhanced performance on multiple public ultrasound datasets by effectively handling blurring and noise artifacts.

Next, we have explored another high-performing yet computationally efficient semantic segmentation architecture, MobileNetV2–DeepLabV3. This hybrid model integrates the MobileNetV2 encoder with the DeepLabV3 segmentation head. MobileNetV2 is a lightweight convolutional network optimized with depthwise separable convolutions and inverted residual connections. The use of MobileNetV2 as the encoder enables efficient extraction of hierarchical features from ultrasound images. Its low parameter count and fast inference speed makes it well-suited for real-time deployment in clinical environments. The DeepLabV3 decoder uses Atrous (dilated) convolutions to improve segmentation accuracy. These convolutions expand the receptive field while keeping the number of parameters low and preserving spatial resolution. Most importantly, it includes an Atrous Spatial Pyramid Pooling (ASPP) module. This module extracts multi-scale features by using parallel atrous convolutions with varying dilation rates. This design helps the model accurately detect tumor boundaries of different shapes and sizes. It is especially useful in breast ultrasound images, where lesions often vary and have low contrast. Recent studies have shown that DeepLab-based models work well for analyzing breast

lesions. They are especially effective in handling varied lesion appearances in ultrasound and mammogram images when combined with lightweight encoders. For instance, Chen et al. [28] originally introduced DeepLabV3 for semantic segmentation in natural scenes, featuring ASPP module to capture multi scale context in dense prediction tasks.

Subsequent adaptations have demonstrated enhanced performance in lesion segmentation with improved generalization. Shia et al. [29] applied DeepLabV3+ to breast ultrasound images with a lightweight backbone and achieved solid segmentation quality in clinical lexicon extraction tasks. Zhang et al. [30] employed an automatic DeepLabV3 based architecture in breast ultrasound tumor segmentation using MobileNetV2, balancing high accuracy with computational efficiency.

Furthermore, MobileNetV2–DeepLabV3 avoids the use of fully connected layers. This enables faster inference and lower memory consumption, thus making it highly deployable on mobile or edge devices. This architecture, therefore, presents a compelling solution for scenarios where both segmentation precision and runtime efficiency are paramount.

D. Loss Function

In our study, we employ a hybrid loss function [31] which is a combination of Dice Loss and Binary Cross Entropy, BCE Loss. This is to optimize the performance of the segmentation models. Dice Loss is particularly effective for medical image segmentation tasks. It directly maximizes the overlap between the predicted mask and the ground truth mask.

This makes it well-suited for handling class imbalance which is a common issue in breast ultrasound images where lesion regions occupy a small portion of the image. To complement this, BCE Loss penalizes individual pixel-wise misclassifications, encouraging the model to correctly classify both foreground and background pixels. The Dice loss is given by (1) as

$$L_{\text{Dice}} = 1 - \frac{2 \sum_{i=1}^N p_i g_i}{\sum_{i=1}^N p_i + \sum_{i=1}^N g_i + \varepsilon}. \quad (1)$$

The Binary Cross-Entropy, BCE loss is defined by (2) as

$$L_{\text{BCE}} = -\frac{1}{N} \sum_{i=1}^N [g_i \log(p_i) + (1 - g_i) \log(1 - p_i)]. \quad (2)$$

where p_i is the predicted probability of pixel i

g_i is the corresponding ground truth label

N is the total number of pixels in the image

ε is a small constant to avoid division by zero.

The combined loss function, denoted as L_{total} which is a hybrid loss function is defined by (3) as

$$L_{\text{total}} = \alpha L_{\text{Dice}} + \beta L_{\text{CrossEntropy}}. \quad (3)$$

where $\alpha = 0.7$ and $\beta = 0.3$ are weight factors that control the contribution of each component to the total loss. This hybrid formulation leverages the strengths of both region-based and pixel-wise learning. This ultimately leads to improved convergence and enhanced segmentation accuracy, particularly in images with low contrast and complex lesion boundaries.

D. Experimental Setup

All experiments in this study were conducted using the BUSI dataset. Only the benign and malignant categories were considered for our research. The normal class was excluded due to the absence of meaningful lesion regions for segmentation.

To ensure a balanced and representative evaluation, the augmented dataset was randomly split into 70% training, 15% validation and 15% testing subsets. Model training and evaluation were performed using the PyTorch deep learning framework on a system configured with an Intel Core i7 CPU and 16 GB RAM. No GPU acceleration was used to emphasize the suitability of the selected models, MobileNetV2–UNet and MobileNetV2–DeepLabV3, for low-resource, CPU-only environments.

A consistent set of hyperparameters were used for training both the models. The Adam optimizer was employed with a learning rate of 1×10^{-4} . The training of the models were conducted for 20 epochs with a batch size of 8. A hybrid loss function comprising $0.7 \times$ Dice Loss and $0.3 \times$ Binary Cross Entropy was applied during training. Table I list the hyperparameters used to train the models.

IV. RESULTS AND DISCUSSION

A. Quantitative Metrics

Both the models were evaluated on the test set using standard segmentation metrics. These metrics include Intersection over Union (IoU), F1-score, precision, Dice coefficient, recall, and accuracy. Table II summarizes the evaluation results of the two models. Equations (4) – (8) define the evaluation metrics used in this study. Here, TP stands for True Positives, TN means True Negatives, FP for False Positives and FN denotes False Negatives respectively.

$$\text{Accuracy} = \frac{(TP+TN)}{(TP+TN+FP+FN)} \cdot \quad (4)$$

$$\text{Precision} = \frac{TP}{(TP+FP)} \cdot \quad (5)$$

$$\text{Recall} = \frac{TP}{(TP+FN)} \cdot \quad (6)$$

$$\text{F1 - Score} = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \cdot \quad (7)$$

$$\text{IoU (Jaccard Score)} = \frac{TP}{(TP+FP+FN)} \cdot \quad (8)$$

TABLE I. HYPERPARAMETERS USED TO TRAIN MOBILENETV2-UNET AND MOBILENETV2-DEEPLABV3 MODELS

Model Architecture	Optimizer	Loss function	Learning Rate	Batch size	Number of epochs	Activation function
MobileNetV2-UNet/ MobileNetV2- DeepLabV3	Adam optimizer	Weighted Dice Loss + Binary Cross- Entropy Loss	1×10^{-4}	8	20	Sigmoid (applied during inference for binary segmentation)

TABLE II. EVALUATION RESULTS

Model	Accuracy (%)	IoU (%)	Recall (%)	Precision (%)	F1-Score (%)	Dice Coefficient (%)
MobileNetV2-UNet	97.95	79.50	86.57	87.38	86.14	86.14
MobileNetV2-DeepLabV3	98.02	78.61	85.79	88.47	86.06	86.06

Both MobileNetV2-UNet and MobileNetV2-DeepLabV3 delivered strong results on the BUSI breast tumor segmentation test set. As observed from Table II, MobileNetV2-UNet achieved a higher IoU of 79.50% and Dice coefficient of 86.14%, reflecting better overlap with the ground truth masks and more consistent region-level segmentation. While MobileNetV2-DeepLabV3 yielded a slightly higher overall accuracy of 98.02% and precision of 88.47%, its recall value, 85.79% was notably lower than MobileNetV2-UNet, indicating a tendency to miss tumor pixels. On the other hand, MobileNetV2-UNet exhibited better balance between precision and recall, accurately delineating tumor boundaries with fewer misclassifications. Overall, MobileNetV2-UNet balances efficiency and performance, aligning better with the lightweight deployment objectives of this research. Since the Dice coefficient and F1-score are mathematically equivalent in binary segmentation tasks, both metrics yielded identical values in our evaluation results.

B. Visual Results

Fig. 2 presents the visual comparison of segmentation outputs from the MobileNetV2-UNet and MobileNetV2-DeepLabV3 models on three representative samples from the BUSI test dataset. The first column (a) displays the original images followed by their corresponding ground truth masks in column (b). The third column (c) and the fourth column (d) presents the predicted masks generated by the MobileNetV2-UNet and MobileNetV2-DeepLabV3 models respectively. The qualitative comparison of predicted segmentation masks reveals that MobileNetV2-UNet consistently produces more precise and compact masks with well-defined lesion boundaries. The model is thus more suitable for fine-grained segmentation. MobileNetV2-DeepLabV3, on the other hand, tends to over-segment by capturing a wider area around the lesion with slightly blurred edges. This can help with rough lesion detection but may also cause small errors and false positives. As depicted in the last row and column (d) of the figure, the prediction by MobileNetV2-DeepLabV3 fails to capture the lesion region accurately, resulting in fragmented and incomplete segmentation. A significant portion of the lesion is either misclassified as background or missed entirely. Thus we can conclude that though both models can detect and

segment lesions, MobileNetV2–UNet gives more accurate shapes and cleaner boundaries. This makes it better for clinical use where precise localization is important. On the other hand, MobileNetV2–DeepLabV3 is more useful for early-stage detection where capturing the full lesion area matters more than fine detail.

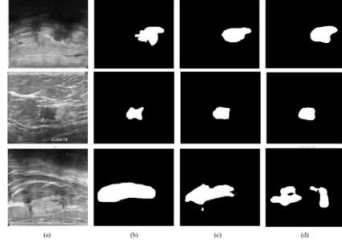


Fig. 2. Visual comparison of segmentation results for MobileNetV2–UNet and MobileNetV2–DeepLabV3.

C. Training And Validation Curves

The training and validation curves for MobileNetV2–UNet and MobileNetV2–DeepLabV3 are illustrated in Fig. 3 and Fig. 4 respectively. The comparative analysis of the curves reveals that both models exhibit stable and consistent learning behavior throughout the training process. For the MobileNetV2–UNet architecture, the IoU curve in Fig. 3(a) and the Dice coefficient curve in Fig. 3(b) demonstrates a consistent upward trend over epochs, indicating effective learning and segmentation improvement. The loss curve in Fig. 3(c) shows a steady decline, with a slight gap between training and validation losses, suggesting mild overfitting. In comparison, the MobileNetV2–DeepLabV3 model exhibits slightly lower but stable performance, as shown in the IoU curve in Fig. 4(a) and Dice curve in Fig. 4(b). Notably, the training and validation losses in Fig. 4(c) are more closely aligned, reflecting better generalization and reduced overfitting. Although MobileNetV2–DeepLabV3 achieves a slightly higher overall accuracy of 98.02% than MobileNetV2–UNet with 97.95% accuracy, the latter exhibits more pronounced segmentation improvements during training. This comparative trend highlights the trade-off between segmentation precision and generalization capability across the two lightweight architectures.

D. Computational Efficiency And Model Complexity Analysis

The comparison table presented in Table III highlights the computational efficiency and model complexity of the two lightweight segmentation architectures, MobileNetV2–UNet and MobileNetV2–DeepLabV3, evaluated on the BUSI breast tumor test dataset. As highlighted in Table III, MobileNetV2–UNet achieves an average inference time of 0.1530 seconds per image and 6.54 FPS indicating its superior inference performance. This makes it highly suitable for real-time and edge deployment scenarios. On the contrary, MobileNetV2–DeepLabV3 incurs a higher computational cost, with an average inference time of 0.4940 seconds per image and only 2.02 FPS. MobileNetV2–UNet has a smaller number of parameters of 6.6 million and a smaller model size of 25.55 MB. In comparison, MobileNetV2–DeepLabV3 has 12.6 million parameters and a model size of 48.51 MB. These results show a clear trade-off between accuracy and efficiency. MobileNetV2–UNet is more suitable for devices with limited resources like mobile phones or embedded systems. The total inference time for MobileNetV2–UNet is 74.34 seconds for 486 images, while MobileNetV2–DeepLabV3 takes 240.06 seconds. This means MobileNetV2–UNet is over three times faster and better suited for real-time use in resource-limited settings.

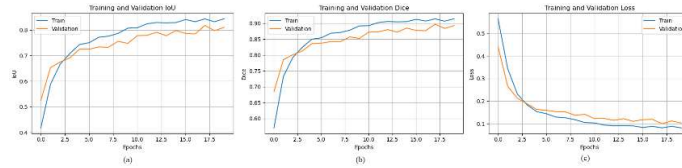


Fig. 3. Training and validation curves for MobileNetV2–UNet

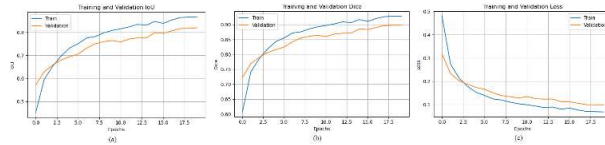


Fig. 4. Training and validation curves for MobileNetV2–DeepLabV3

TABLE III. COMPARISON OF INFERENCE PERFORMANCE BETWEEN MOBILENetV2-UNET AND MOBILENetV2-DEEPLABV3

Model	Total parameters	Model Size	Inference Time/ Image (sec)	FPS (Frames Per Second)	Total Inference Time (486 images)
MobileNetV2-UNet	6.6M	25.55 MB	0.1530 sec	6.54	74.3422 sec
MobileNetV2-DeepLabV3	12.6M	48.51 MB	0.4940 sec	2.02	240.0642 sec

E. Comparison With Existing Work On BUSI Dataset

The publicly available dataset, the BUSI dataset has been widely adopted in recent years for developing and evaluating deep learning models aimed at breast tumor segmentation. Using this dataset, many studies have explored a range of architectures from conventional U-Net variants to more advanced attention-based and transformer-integrated networks. The results obtained from these studies have demonstrated strong segmentation performance on this dataset. Most of these models often focus on improving accuracy and generalization, sometimes at the cost of computational efficiency. However, till date, very few works explicitly evaluate the performance of their models in terms of inference time, model size, or suitability for deployment on real-world, resource-constrained settings. Our work, on the other hand, emphasizes the balance between segmentation accuracy and lightweight design. Our proposed models could achieve competitive performance while significantly reducing model size and inference time. This makes them potentially more suitable for real-time applications on mobile or embedded devices. A brief summary of selected recent works utilizing the BUSI dataset is presented in Table IV for comparative context. This comparative context helps position our research within the broader landscape of BUSI-based segmentation studies. Moreover, it also highlights the importance of assessing both segmentation performance and computational efficiency in medical imaging applications.

V. CONCLUSION AND FUTURE WORK

This study presented a comparative evaluation of two lightweight deep segmentation architectures, MobileNetV2-UNet and MobileNetV2-DeepLabV3, for segmentation of breast tumor using ultrasound images from the BUSI dataset. The models were trained and validated on augmented benign and malignant classes, omitting the normal category due to its lack of meaningful annotations. This targeted approach enhanced the capability of the model to accurately localize lesions and generalize well within the domain of tumor detection and segmentation.

Comprehensive quantitative and qualitative analyses demonstrated that MobileNetV2-UNet strikes a favorable balance between segmentation accuracy and computational efficiency, achieving a Dice coefficient of 86.14% and an average inference speed of 6.54 FPS. Its performance, coupled with a lower model size and faster inference time, makes it potentially suitable for deployment in resource-constrained and real-time clinical environments such as mobile health screening applications.

While MobileNetV2-DeepLabV3 achieved slightly higher precision, it incurred significantly higher computational overhead, suggesting its suitability for offline or high-resource settings. In future work, the current segmentation framework will be extended to include lesion classification into benign and malignant categories, enabling a full end-to-end diagnostic framework.

Further experiments will also evaluate the model's cross-domain robustness using external datasets such as the Mendeley Breast Ultrasound BUSC Dataset and OASBUD Dataset. These datasets differ in acquisition settings and patient demographics, allowing assessment of the model's generalizability across diverse clinical environments.

TABLE IV. COMPARATIVE SUMMARY OF RECENT DEEP LEARNING MODELS USING THE BUSI DATASET FOR BREAST TUMOR SEGMENTATION

Author	Model/Method	Dataset	Segmentation Performance	Edge Deployability
Pramanik et al.[32]	Dual-Branch U-Net (DBU-Net)	BUSI (benign+malignant)	Dice=85.28%; IoU =74.34%	Not Specified
Pramanik et al.[33]	Residual U-Net with Dual Attention (DAU-Net)	BUSI (benign+malignant)	Dice = 74.23%	Not Specified
Rahmani et al.[34]	U-Net with Residual, Attention & Resolution Enhancement blocks	BUSI (benign+malignant)	Dice = 84.72% IoU = 85.95%, (benign), Dice=82.14% IoU=82.21% (malignant)	Not Specified
Pan et al.[35]	Multi-Scale & Multi-Edge U Net(M ² U-Net)	BUSI (benign+malignant)	Mean Dice = 79.43%; IoU = 83.00%; Recall = 87.17%	Not Specified

Ours (Proposed)	MobileNetV2–UNet	BUSI (benign+malignant)	Dice=86.14%, IoU=79.50%	Potentially suitable for edge deployment
	MobileNetV2–DeepLabV3		Dice=86.06%, IoU=78.61%	Potentially suitable for edge deployment (with efficiency trade-offs)

REFERENCES

- [1] Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., & Bray, F., “Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries,” *CA: a cancer journal for clinicians*, vol. 71, no. 3, pp. 209-249, 2021.
- [2] Pramanik, P., Mukhopadhyay, S., Kaplun, D., & Sarkar, “A Deep Feature Selection Method for Tumor Classification in Breast Ultrasound Images,” In *International conference on mathematics and its applications in new computer systems*, pp. 241-252, 2021.
- [3] Iacob, R., Iacob, E. R., Stoicescu, E. R., Ghenciu, D. M., Cocolea, D. M., Constantinescu, A., ... & Manolescu, D. L., “Evaluating the Role of Breast Ultrasound in Early Detection of Breast Cancer in Low-and Middle-Income Countries: A Comprehensive Narrative Review,” *Bioengineering*, vol. 11, no. 3, p. 262, 2024.
- [4] Vega, R., Dehghan, M., Nagdev, A., Buchanan, B., Kapur, J., Jaremko, J. L., & Zonoobi, D., “Overcoming Barriers in the Use of Artificial Intelligence in Point of Care Ultrasound,” *npj Digital Medicine*, vol. 8, no. 1, p. 213, 2025.
- [5] Jalalian, A., Mashohor, S., Mahmud, R., Karasfi, B., Saripan, M. I. B., & Ramli, A. R. B., “Foundation and Methodologies in Computer-Aided Diagnosis Systems for Breast Cancer Detection,” *EXCLI journal*, vol. 16, p. 113, 2017.
- [6] Huang, Q., Luo, Y., & Zhang, Q., “Breast Ultrasound Image Segmentation: A Survey,” *International journal of computer assisted radiology and surgery*, vol. 12, no. 3, pp. 493-507, 2017.
- [7] Sun, Y., Huang, Z., Cai, G., Su, J., & Gong, Z., “NSBR-Net: A Novel Noise Suppression and Boundary Refinement Network for Breast Tumor Segmentation in Ultrasound Images,” *Algorithms*, vol. 17, no. 6, p. 257, 2024.
- [8] Cai, F., Wen, J., He, F., Xia, Y., Xu, W., Zhang, Y., ... & Li, J., “SC-Unext: A Lightweight Image Segmentation Model With Cellular Mechanism for Breast Ultrasound Tumor Diagnosis,” *Journal of Imaging Informatics in Medicine*, vol. 37, no. 4, pp. 1505-1515, 2024.
- [9] Yu, K., Chen, S., & Chen, Y., “Tumor Segmentation in Breast Ultrasound Image by Means of Res Path Combined With Dense Connection Neural Network,” *Diagnostics*, vol. 11, no. 9, p. 1565, 2021.
- [10] Horsch, K., Giger, M. L., Venta, L. A., & Vyborny, C. J., “Automatic Segmentation of Breast Lesions on Ultrasound,” *Medical physics*, vol. 28, no. 8, pp. 1652-1659, 2001.
- [11] Chen, C. M., Chou, Y. H., Han, K. C., Hung, G. S., Tiu, C. M., Chiou, H. J., & Chiou, S. Y., “Breast Lesions on Sonograms: Computer-Aided Diagnosis With Nearly Setting-Independent Features and Artificial Neural Networks,” *Radiology*, vol. 226, no. 2, pp. 504-514, 2003.
- [12] Chang, R. F., Wu, W. J., Moon, W. K., & Chen, D. R., “Improvement in Breast Tumor Discrimination by Support Vector Machines and Speckle-Emphasis Texture Analysis,” *Ultrasound in medicine & biology*, vol. 29, no. 5, pp. 679-686, 2003.
- [13] Ilesanmi, A. E., Chaumrattanakul, U., & Makhanov, S. S., “Methods for the Segmentation and Classification of Breast Ultrasound Images: A Review,” *Journal of ultrasound*, vol. 24, no. 4, pp. 367-382, 2021.
- [14] Avci, H., & Karakaya, J., “Segmentation Techniques in Mammography: Their Value in Breast Cancer Diagnosis,” *Journal of Clinical Practice & Research*, vol. 46, no. 6, p. 548, 2024.
- [15] Al-Fahaidy, F. A., Al-Fuhaidi, B., AL-Darouby, I., AL-Abady, F., AL-Qadry, M., & AL-Gamal, A., “A Diagnostic Model of Breast Cancer Based on Digital Mammogram Images Using Machine Learning Techniques,” *Applied Computational Intelligence and Soft Computing*, 2022, no. 1, p. 3895976, 2022.
- [16] Kamil, M. Y., & Jassam, A. L. A., “Analysis of Tissue Abnormality in Mammography Images Using Gray Level Co-Occurrence Matrix Method,” In *Journal of Physics: Conference Series*, vol. 1530, no. 1, p. 012101, 2020.
- [17] Xian, M., Zhang, Y., Cheng, H. D., Xu, F., Zhang, B., & Ding, J., “Automatic Breast Ultrasound Image Segmentation: A Survey,” *Pattern Recognition*, vol. 79, pp. 340-355, 2018.
- [18] Badawy, S. M., Mohamed, A. E. N. A., Hefnawy, A. A., Zidan, H. E., GadAllah, M. T., & El-Banby, G. M., “Automatic Semantic Segmentation of Breast Tumors in Ultrasound Images Based on Combining Fuzzy Logic and Deep Learning—A Feasibility Study,” *PloS one*, vol. 16, no. 5, p. e0251899, 2021.
- [19] Hesamian, M. H., Jia, W., He, X., & Kennedy, P., “Deep Learning Techniques for Medical Image Segmentation: Achievements and Challenges,” *Journal of digital imaging*, vol. 32, no. 4, pp. 582-596, 2019.
- [20] O. Ronneberger, P. Fischer, and T. Brox, “U-Net: Convolutional Networks for Biomedical Image Segmentation,” in *Proceedings of MICCAI*, vol. 9351, pp. 234-241, 2015.
- [21] Chen, G., Li, L., Dai, Y., Zhang, J., & Yap, M. H., “AAU-net: An Adaptive Attention U-net for Breast Lesions Segmentation in Ultrasound Images,” *IEEE Transactions on Medical Imaging*, vol. 42, no. 5, pp. 1289-1300, 2022.

- [22] Paszke, A., Chaurasia, A., Kim, S., & Culurciello, E., "ENet: A Deep Neural Network Architecture for Real-Time Semantic Segmentation," arXiv preprint arXiv:1606.02147, 2016.
- [23] Zhou, Q., Wang, Q., Bao, Y., Kong, L., Jin, X., & Ou, W., "LAEDNet: A Lightweight Attention Encoder–Decoder Network for Ultrasound Medical Image Segmentation," *Computers and Electrical Engineering*, vol. 99, p. 107777, 2022.
- [24] Guo, H., Wang, S., Dang, H., Xiao, K., Yang, Y., Liu, W., ... & Wan, Y., "LightBTseg: A lightweight Breast Tumor Segmentation Model Using Ultrasound Images via Dual-Path Joint Knowledge Distillation," In *2023 China Automation Congress (CAC)*, pp. 3841-3847, IEEE, 2023.
- [25] Anari, S., Sadeghi, S., Sheikhi, G., Ranjbarzadeh, R., & Bendeckache, M., "Explainable Attention Based Breast Tumor Segmentation Using a Combination of UNet, ResNet, DenseNet, and EfficientNet Models," *Scientific Reports*, vol. 15, no. 1, pp. 1027, 2025.
- [26] Ben Ahmed, I., Ouarda, W., & Ben Amar, C., "Hybrid UNET Model Segmentation for an Early Breast Cancer Detection Using Ultrasound Images," In *International Conference on Computational Collective Intelligence*, pp. 464-476, 2022.
- [27] Wang, T., Liu, J., & Tang, J., "A Cross-Scale Attention-Based U-net for Breast Ultrasound Image Segmentation," *Journal of Imaging Informatics in Medicine*, pp. 1-14, 2025.
- [28] Chen, L. C., Papandreou, G., Schroff, F., & Adam, H., "Rethinking Atrous Convolution for Semantic Image Segmentation," arXiv preprint arXiv:1706.05587, 2017.
- [29] Shia, W. C., Hsu, F. R., Dai, S. T., Guo, S. L., & Chen, D. R., "Semantic Segmentation of the Malignant Breast Imaging Reporting and Data System Lexicon on Breast Ultrasound Images by Using DeepLab v3+," *Sensors*, vol. 22, no. 14, p. 5352, 2022.
- [30] Zhang, S., Liao, M., Wang, J., Zhu, Y., Zhang, Y., Zhang, J., ... & Wang, W., "Fully Automatic Tumor Segmentation of Breast Ultrasound Images With Deep Learning," *Journal of Applied Clinical Medical Physics*, vol. 24, no. 1, p. e13863, 2023.
- [31] Yeung, M., Sala, E., Schönlieb, C. B., & Rundo, L., "Unified Focal Loss: Generalising Dice and Cross Entropy-Based Losses to Handle Class Imbalanced Medical Image Segmentation," *Computerized Medical Imaging and Graphics*, vol. 95, p. 102026, 2022.
- [32] Pramanik, P., Pramanik, R., Schwenker, F., & Sarkar, R., "DBU-Net: Dual Branch U-Net for Tumor Segmentation in Breast Ultrasound Images," *Plos one*, vol. 18, no. 11, p. e0293615, 2023.
- [33] Pramanik, P., Roy, A., Cuevas, E., Perez-Cisneros, M., & Sarkar, R., "DAU-Net: Dual Attention-Aided U-Net for Segmenting Tumor in Breast Ultrasound Images," *Plos one*, vol. 19, no. 5, p. e0303670, 2024.
- [34] Rahmani, A. A., Shirazi, A. A. B., & Behnam, H., "Automatic Breast Mass Segmentation in Ultrasound Images With U-Net and Resolution Enhancement Blocks," *Biomedical Signal Processing and Control*, vol. 94, p. 106270, 2024.
- [35] Pan, L., Tang, M., Chen, X., Du, Z., Huang, D., Yang, M., & Chen, Y., "M2UNet: Multi-Scale Feature Acquisition and Multi-Input Edge Supplement Based on UNet for Efficient Segmentation of Breast Tumor in Ultrasound Images," *Diagnostics*, vol. 15, no. 8, p. 944, 2025.