



A Multi-attention Triple Decoder Deep Convolution Network for Breast Cancer Segmentation Using Ultrasound Images

Muhammad Junaid Umer¹ · Muhammad Sharif¹ · Mudassar Raza²

Received: 5 July 2023 / Accepted: 14 October 2023 / Published online: 13 November 2023
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

Breast cancer (BC) is a widely diagnosed deadly disease commonly present in middle-aged women around the globe. Ultrasound (U/S) imaging is widely used for the early prediction and segmentation of BC due to low radiation and cheapness. Manual BC segmentation from ultrasound imaging is a complex and laborious task due to inherited noise. Many deep learning-based breast cancer diagnostic methods are presented that can further be enhanced to improve the segmentation performance. This work proposed a U-shaped auto encoder-based multi-attention triple decoder convolution neural network for BC segmentation from U/S images. To capture multi-scale diverse spatial image features this work introduced a multi-scale convolution operation-based encoder network. To process the multi-scale learned diverse spatial features in the encoder path multi-scale triple decoder network is designed that was not found in earlier studies. To highlight the tumor region at different scales multi-attention mechanism is introduced in each decoder network. The multi-attention mechanism is designed to suppress the other region information and to highlight the tumor region features at different scales. The proposed deep network produced the segmentation dice of 90.45% on the UDIAT dataset and the segmentation dice of 89.13% on the BUSI dataset. The testing Jaccard index of 83.40% is recorded on the UDIAT dataset and a Jaccard index of 82.31% is recorded on the BUSI dataset. The result comparison with existing methods shows that our method achieved the highest results. The segmentation performance of the triple decoder-based BC segmentation model suggested that it can effectively be used to automate the manual breast cancer segmentation task from ultrasound images.

Keywords Breast cancer · Deep learning · Neural network · Ultrasound · Autoencoder · Segmentation

Introduction

Breast cancer is a mostly spreading disease frequently diagnosed in middle-aged women around the globe with the second highest death rate after lung cancer. According to a recent BC survey, about 287,850 new cases will be diagnosed in 2022 in America. Approximately 43,250 women will die due to BC this year. About 82% of BC-diagnosed women are above the age of 50 and the highest death ratio of 91% is observed in this age group [1, 2]. The early detection of BC with proper treatment can help to overcome the mortality rate. Currently, different imaging technologies

including mammography, ultrasonography, and biopsy are utilized to detect BC manually by pathologists. The manual BC detection method is a complex and laborious task due to inherited noise which is also subjective and may be diagnosed differently by different pathologists [3]. Computer-aided diagnosis of BC can play a vital role in normalizing the detection differences of this deadly disease. Automating the manual diagnosis of BC deep learning is playing a novel role [4, 5]. Many deep learning-based methods are being proposed at large scale for BC segmentation as studies show that ninety percent of BC positively diagnosed individuals may be cured by proper treatment [6].

Breast ultrasound (U/S) imaging is a cheap and low-radiation technology that can effectively be used for BC detection at its early stage. The manual diagnosis of BC using breast U/S imaging is a difficult task and needs expert clinicians for the final diagnosis. Due to advancements in technology, deep learning is being used to aid the diagnosis process of different diseases [7]. Deep learning showed exceptional performance in medical image segmentation

✉ Muhammad Sharif
sharif@ciitwah.edu.pk

¹ Department of Computer Science, COMSATS University Islamabad, Wah Campus, Wah Cantt, Pakistan

² Department of Computer Science, HITEC University Taxila, Taxila, Pakistan

specifically BC detection and segmentation [8, 9]. Recently, Vigil et al. [10] proposed a BC segmentation method by applying a dual-intended mechanism on a U/S imaging dataset. Their method achieved a 78.5% accuracy to accurately classify the BC. Wang et al. automated the segmentation of BC from U/S images using different flavors of existing ResNet models. A unique combination of pre-trained models was utilized to produce the automatic segmentation method [11]. Ayana et al. [12] introduced another transfer learning-based method for BC segmentation from U/S images and reported a classification accuracy of 99% on the U/S dataset. Recently, U-net based models have gained increased attention in the segmentation of BC due to their high performance as Umer et al. [13] proposed BC detection using the feature selection method. A connected U-net-based automated method for BC segmentation is presented in [14]. In the connected BC segmentation method, the two U-net models are used in concatenated fission with the help of skip connections.

This work proposed a U-shaped auto encoder-based multi-attention triple decoder (MATD) convolution network for BC segmentation from U/S imaging. The proposed method introduced a multi-scale convolution-based encoder network for diverse spatial image feature extraction. The learned input image features at different scales are then processed through a multi-scale triple decoder network for BC segmentation. In each decoder network to highlight the tumor region attention mechanism is also utilized for best segmentation performance. The outputs of the multi-scale decoder blocks were concatenated for the accurate segmentation of BC. The main contributions of the proposed multi-attention triple decoder convolution network are given below.

- A single multi-scale encoder block is introduced in the contracting path to capture diverse image features at different scales.
- To transform the multi-scale learned image features, a triple decoder network is introduced to process the learned image feature separately.
- To highlight the tumor region at different scales, a multi-attention mechanism is implemented in each decoder network for accurate segmentation of BC.
- The output of the multi-attention triple decoder network is concatenated to predict the segmentation mask of input breast U/S images.
- The comparison of the proposed multi-attention triple decoder convolution network with the existing method is also carried out.

Related Work

BC detection and segmentation are an active area of research due to the very high death ratio of this disease. BC detection and segmentation are being automated by using deep

learning-based automated models. Deep learning is gaining increased attention in the medical field due to its outstanding detection performance. Many deep learning-based BC segmentation models are presented for the detection and segmentation of BC from the U/S modality. Luo et al. [15] implemented two parallel networks for the segmentation-guided classification of BC using the U/S imaging dataset and attention mechanism. Their proposed BC segmentation-guided classification achieved an accuracy of 90.78%. Yan et al. [16] proposed a method for BC segmentation by implementing an attention mechanism to reduce spatial information loss during feature learning. For feature learning a hybrid method with a dilation factor was introduced for efficient feature learning and their method achieved a segmentation IOU of 81.8%. In another recent work, a double attention-based global and local guided U-net-based method with cascaded convolution and residual connections is introduced for accurate BC segmentation [17]. The attention mechanisms in the segmentation methods have gained increased attention due to higher performance. A lot of BC segmentation methods in literature utilized attention mechanisms. A dual attention mechanism with a multi-scale convolution process for BC segmentation from U/S images is proposed in [18]. Another attention-based method for BC segmentation is presented by Farooq et al. [19].

Due to high performance in segmentation tasks, autoencoder-based U-shaped methods with different implementations are being utilized to automate the BC segmentation task from U/S images. Tong et al. [20] introduced a novel loss function with an embedded attention-based deep learning model for BC segmentation from the U/S imaging dataset. In their method to replace the conventional convolution method, a residual convolution method was implemented for performance enhancement. In another recent work, a quantization-assisted BC segmentation method is presented with a fusion mechanism [21]. Vianna et al. [22] presented the comparative analysis of U-net and SegNet deep CNN models for BC segmentation from U/S images and reported a dice of 86.3% with U-net and dice of 81.1% with SegNet. A mixed attention loss-based U-net model with four attention loss functions and a selective kernel method was proposed in [23]. Their modified version for BC segmentation produced a dice of 92.2%. A fusion method for BC segmentation from U/S images by implementing the residual convolutional attention method was introduced in [24]. In their method, a fusion attention mechanism was introduced with the channel attention method, and dice of 92.1% was reported for the BC segmentation task from U/S images. Li et al. [25] introduced a multi-scale fusion and focal loss function in U-net for BC segmentation and reported a dice of 95.35. In their method, a dilated convolution operation with different scales was introduced to enhance the segmentation performance.

BC segmentation from U/S images is being automated by using deep learning-based semantic image segmentation methods. A spatial attention mechanism for BC classification from U/S images using deep CNN is presented by Lu et al. [26]. In their method, a pre-trained ResNet model was utilized for feature learning. Punn et al. [27] proposed an inception-based model using cross spatial attention method for BC segmentation from the U/S imaging dataset. In their BC segmentation model, residual connections were utilized in the attention mechanism. Wang et al. recently introduced a breast lesion localization method using a U/S dataset [28]. For breast lesion localization, an enhancement method was implemented and then the segmented data was processed through the classification model for the classification task. A contour optimization method in a deformed U-net with an adversarial mechanism is presented in [29] for BC segmentation from the U/S imaging dataset. A dice of 89.7% was achieved by their modified segmentation model. Cho et al. [30] proposed a new method for BC localization and classification consisting of multi-stage segmentation with a residual fusion mechanism. In their method in the first step classification is performed to check the malignancy of input images and the classified images are further passed to the segmentation network of BC localization. Gong et al. [31] recently presented a U/S imaging-based BC classification method. BC segmentation by using a breast mammography dataset was proposed by Peng et al. [32] in which a hybrid loss function was introduced. Lou et al. [33] introduced a U-shaped deep CNN model for BC segmentation from U/S images. In their work, a context-aware-based fusion mechanism was introduced with the residual pyramid. The proposed context-level fusion method tried to overcome the contextual gap during the feature learning process and their method achieved higher performance.

From the above discussion of previous deep learning-based methods for BC segmentation, it can be concluded that most of the previous studies utilized a fixed receptive field-based model. The encoder paths in existing methods are developed with the limitation of single or fixed-sized convolution operations. Furthermore, the multi-scale encoder with multi-scale decoding mechanisms is rarely discussed. In this regard to enhance the segmentation performance, this work proposed a single multi-scale encoder-based multi-attention triple decoder convolution network for BC segmentation from the U/S imaging dataset.

Proposed Method

This work proposed a single multi-scale encoder-based multi-attention triple decoder convolution network for BC segmentation from the U/S imaging dataset. The proposed MATD convolution network is presented in Fig. 1. The proposed BC segmentation model is composed of a single multi-scale

encoder and a multi-decoder network. The multi-decoder network is designed to tackle the problem of fixed receptive fields. The proposed multi-decoder network is developed by using three decoders with three different receptive fields including 3×3 , 5×5 , and 7×7 decoders. The contracting path of the proposed multi-attention triple decoder convolution network utilized multi-scale convolution operations. The encoder of the proposed convolution network is composed of a 3×3 convolution block containing a convolution operation of a 3×3 kernel followed by 30% dropout, batch normalization, and ReLU activation. The convolution result of the 3×3 convolution block is passed to the 5×5 convolution block, and finally, the convolution result of the 5×5 convolution block is passed to the 7×7 convolution block. After the multi-scale convolution operation, a max pooling layer is implemented. The convolution operation, activation operation, and max pooling method are mathematically defined in Eqs. 1, 2, and 3, respectively, where W is the weight, φ , and bias factor.

$$Conv_{img} = \sum_{-k}^k kernel_{stride}^{size} * (Input_{img} \cdot W) + \varphi \tag{1}$$

$$Activation_{ReLU} = f(Conv_{img}) = \begin{cases} 0, & Conv_{img} < 0 \\ 1, & Conv_{img} \geq 0 \end{cases} \tag{2}$$

$$Max_pooling_{img} = MAX_{stride}^{win} * Input_{img} \tag{3}$$

The contracting path of our convolution network is implemented with four multi-scale encoder blocks. Each encoder block utilized a batch normalization operation to reduce the overfitting problem. The batch normalization procedure is mathematically described in Eqs. 4–6. For batch normalization of the input U/S image dataset, the mean of the inputted images was calculated by using Eq. 4, the variance of the inputted images was computed utilizing Eq. 5, and finally, to normalize the input, a batch-level normalization operation was implemented using Eq. 6 where ω and τ are learning parameters.

$$\overline{Mean}_{m_batch} = \frac{1}{N} \sum_1^N Input_{img} \tag{4}$$

$$V_{\sigma^2 m_batch} = \frac{1}{N} \sum_1^N (Input_{img} - \overline{Mean}_{m_batch})^2 \tag{5}$$

$$\widehat{Batch}_{normal} = \frac{Input_{img} - \overline{Mean}_{m_batch}}{\sqrt{V_{\sigma^2 m_batch} + \epsilon}} \xrightarrow{\text{produced}} \tau \widehat{Batch}_{normal} + \omega \tag{6}$$

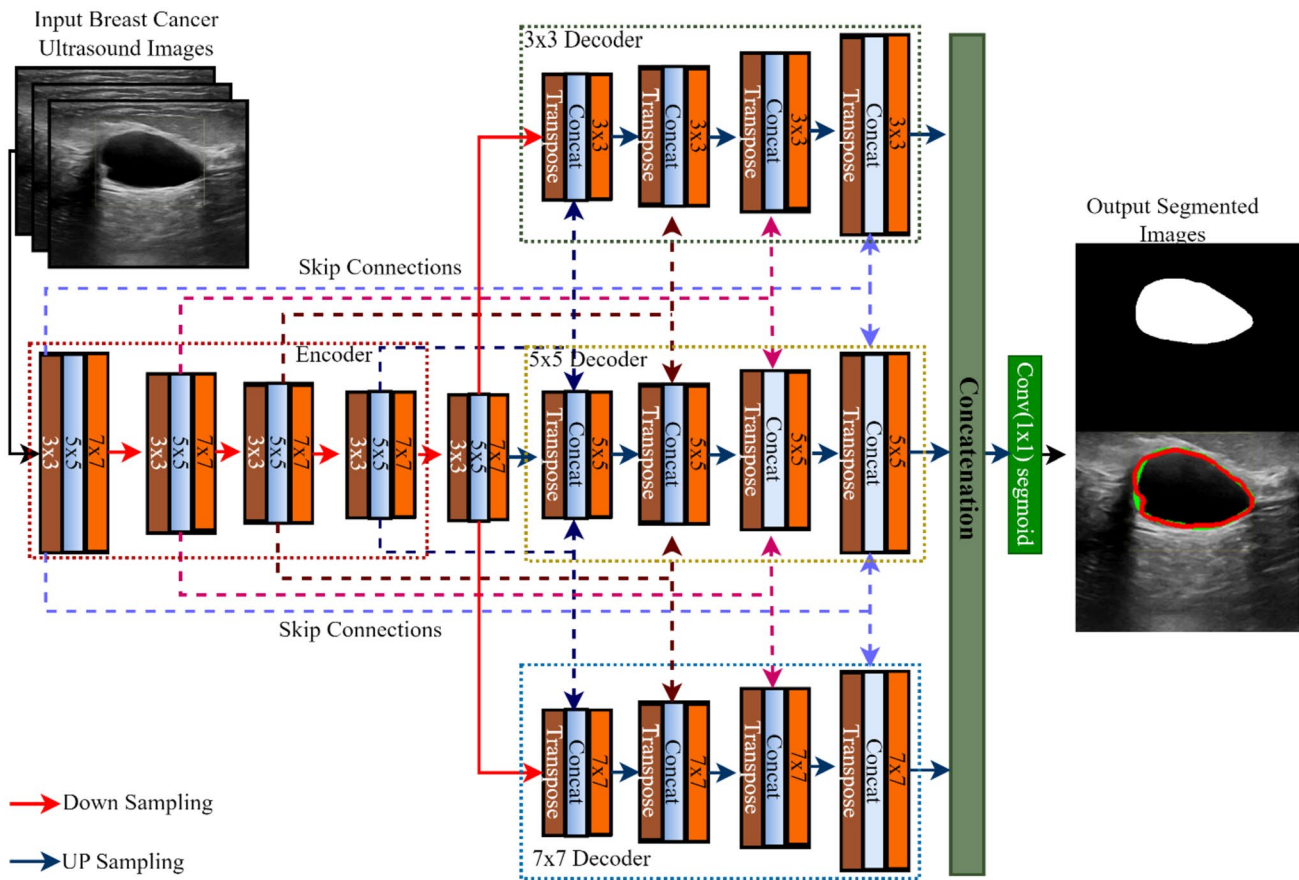


Fig. 1 The general workflow of the proposed U-shaped multi-attention triple decoder convolution network for BC segmentation from U/S images

After the encoder path, a multi-scale bridge block was implemented to transform the learned image features. The learned highly discriminative image features at different scales are transformed into the multi-decoder network by using skip connections. In this work, three different receptive fields including three, five, and 7 sized kernels were utilized to capture the diverse spatial features, and three decoders were utilized to regenerate the BC segmented images from these features. For each convolution scale, a separate decoder was implemented to enhance the segmentation performance. In each decoder network, an attention mechanism was used to highlight the tumor region at a different scale. Finally, the outputs of the multi-scale decoders were concatenated and a 1×1 convolution operation was carried out to get the segmented output of breast U/S images. The expansion path of the proposed convolution network is implemented by using a 3×3 decoder, a 5×5 decoder, and a 7×7 decoder. In each multi-scale decoder network, the up-convolution operation of the respective receptive field was carried out. The learned images feature of the 3×3 convolution block of the encoder was transformed into the 3×3 decoder network by using skip

connections. The learned features from the 5×5 encoder block were transformed to the 5×5 decoder network by using skip connections, and similarly, the learned image features from the 7×7 convolution block of the encoder were transformed to the 7×7 decoder network with the help of skip connections. Each decoder network contained four decoder blocks of respective receptive fields and in each decoder block a transpose convolution operation, attention mechanism, and concatenation operation followed by respective convolution operation are implemented. The BC U/S dataset is passed to the encoder of the MATD convolution model and the segmented BC images are outputted from the 1×1 output layer. The input image size of the proposed multi-attention triple decoder convolution network is 128×128 .

Single Multi-scale Encoder Network

This work introduced a MATD convolution network for BC segmentation from U/S images. For this purpose, a single multi-scale encoder network is designed for diverse spatial feature learning. In most of the previous BC segmentation

methods, fixed sized convolution operation mechanism is utilized that is unable to accurately segment the different-sized tumors. Moreover, the fixed receptive field may miss the spatial information that can be handled using multi-scale convolution operation. In this regard, we proposed a multi-scale single-encoder network for feature learning. The proposed multi-scale single encoder network includes four encoders and in each encoder, a 3×3 convolution, 5×5 convolution, and 7×7 convolution is implemented to accurately segment the diversely shaped breast tumors. Each multi-scale convolution block is implemented using respective block-sized convolution operation followed by activation that was carried out using ReLU function, batch normalization that was implemented at mini-batch size, thirty percent dropout layer, and max pooling operation. The learned spatial multi-scale features are transformed into a decoder network using skip connections. For each scale of encoder convolution operation, a separate decoder of the respective scale is designed to enhance the BC segmentation performance.

Multi-scale Triple Decoder Network

This work introduced a single multi-scale encoder and multi-scale triple decoder network for BC segmentation from the U/S imaging dataset. The proposed multi-scale triple decoder network is introduced to process the learned multi-scale features separately. For this purpose, three decoder networks with different receptive fields including a 3×3 decoder, 5×5 decoder, and 7×7 decoder are developed. The learned input features in the encoder at scale 3×3 are transformed to the 3×3 decoder network using skip connections. The learned input features in the encoder at scale 5×5 are transformed to the 5×5 decoder network using skip connections. Similarly, the learned input features in the encoder at scale 7×7 are transformed to the 7×7 decoder network using skip connections. To get the final predicted segmentation mask the outcomes of the multi-scale decoder network are combined using a concatenation operation. Each multi-scale decoder network was developed using four decoder blocks. In each decoder block of the multi-scale decoders network a transpose convolution operation of the respective scale, a concatenation operation to concatenate the transpose operation output with the learned image features transformed through skip connections, and an up convolution operation of the respective scale was carried out. The experimental outcomes validated that the proposed single multi-scale encoder and multi-attention triple decoder network for BC segmentation from U/S images achieved the best segmentation results due to separate processing of multi-scale learned features in the multi-scale triple decoder network.

Multi-attention in Triple Decoder Network

In this work, a multi-attention triple decoder network is proposed in which three multi-scale decoder networks are introduced to retain the spatial information of learned features at different scales. The proposed multi-attention triple decoder mechanism is presented in Fig. 2. The multi-scale learned spatial image features in encoder blocks are passed to the triple decoder network through skip connections. To accurately segment BC using U/S images, the learned multi-scale feature maps are processed separately by three decoders. For multi-scale attention of breast tumor region, an attention mechanism in each decoder block was introduced for high segmentation performance. The input of the 3×3 decoder attention mechanism is the outcome of the transpose operation of the decoder network and learned spatial image features of the encoder block with 3×3 convolution operation which are transformed using skip connections. In the 3×3 decoder attention block to highlight the tumor region, 1×1 convolution operations are performed at both inputs of the 3×3 decoder block. In the next step, an element-wise summation operation was implemented followed by ReLU activation, one more 1×1 convolution, and sigmoid activation. The output of this activation is further multiplied with 3×3 learned encoder features to suppress the other irrelevant information and to attain more information from the tumor region. Similarly, the attention mechanism was implemented in 5×5 decoder blocks and 7×7 decoder blocks to capture tumor region spatial features at different scales.

The multi-attention mechanism of the proposed segmentation model is implemented in multi-scale decoder networks separately and then the output of the multi-decoder networks is concatenated to retain the multi-scale learned spatial features information.

Datasets

This work utilized two publicly available breast U/S datasets for the evaluation of the proposed MATD convolution network for BC segmentation from U/S imaging datasets. All the results were computed by using 30:70 ratios for testing and training the multi-attention triple decoder convolution network. Detail of each utilized dataset is presented below in bullets.

BUSI dataset: The first dataset used to evaluate the proposed MATD convolution network for BC segmentation from U/S images is the BUSI dataset. The BUSI dataset is available freely for research purposes and is being utilized extensively for BC segmentation and classification tasks. The contributor to this dataset is Dhabyani et al. [34]. The collection process of this dataset is carried out at Baheya Hospital. A total of 780 breast U/S images were collected

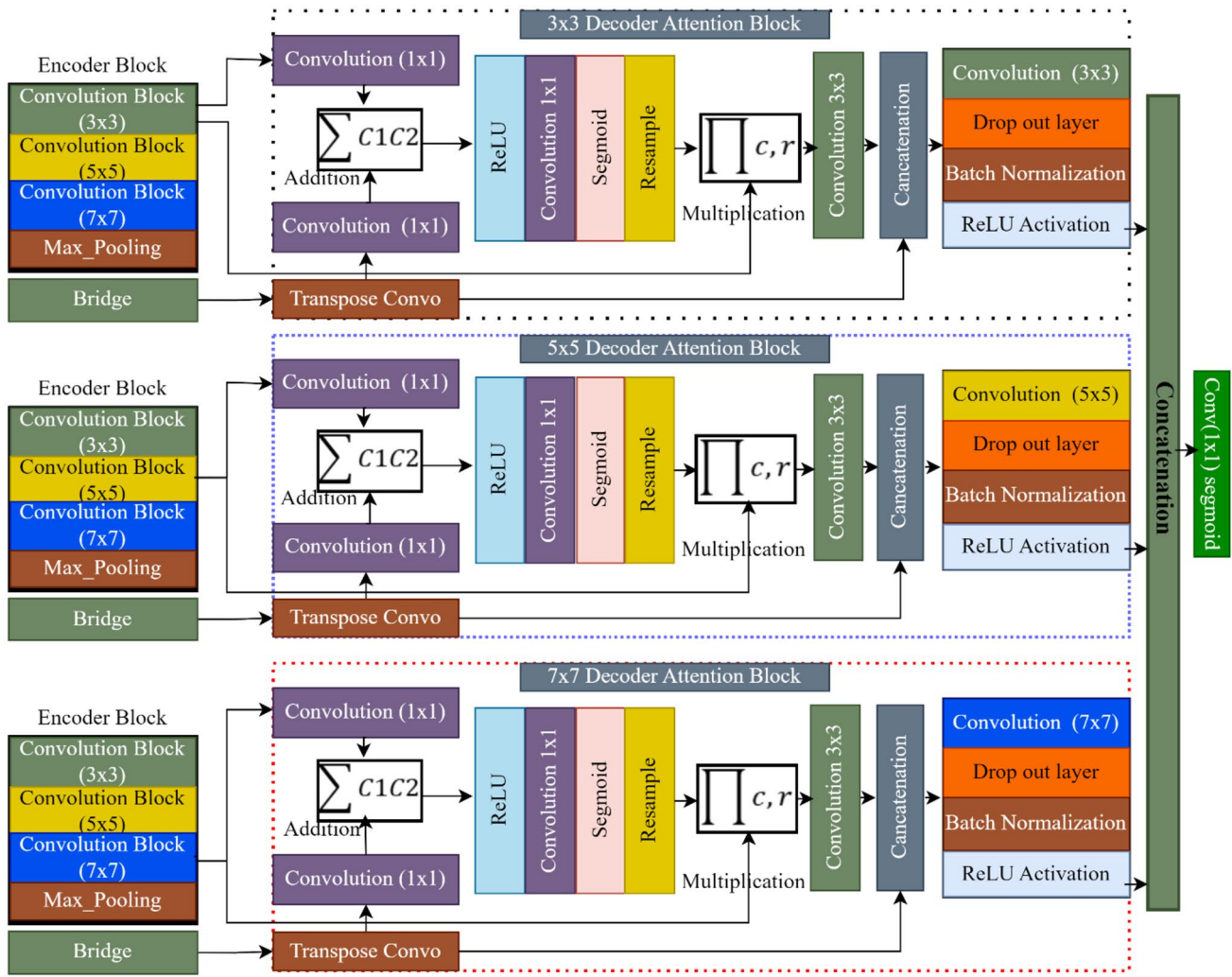


Fig. 2 Multi-attention mechanism of the proposed U-shaped multi-attention triple decoder convolution network for BC segmentation from U/S images

and annotated by radiologists. The collected dataset is available for research purposes with ground truth images and has an average image size of 500×500 .

UDIAT dataset: The second UDIAT open-source dataset was utilized to test the proposed MATD convolution network for BC segmentation from U/S images. The UDIAT U/S dataset is contributed by Yap et al. [35] and is a collection of only 163 BC U/S images with ground truths. The data of this repository is available in PNG format with an average size of 500×500 . The UDIAT breast U/S imaging dataset was prepared and annotated in UDIAT Diagnostic Centre.

Evaluation Metrics

Performance evaluation of the proposed U-shaped multi-attention triple decoder convolution network for BC segmentation from the U/S imaging dataset was carried out by using different metrics such as dice similarity coefficient

(DC), recall (Re), precision (Pr), Jaccard coefficient (JC), and accuracy (Ac). The formulation of each used metric is given in Eqs. 7, 8, 9, 10, 11. Where $ground_{truth}$ is showing the ground truth and $predicted_{seg}$ are representing predicted masks, $True_{Positive}$, $True_{Negative}$, $False_{positive}$, and $False_{negative}$ are showing the true positive, true negative, false positive, and false negative respectively.

$$\begin{aligned}
 & \text{Dice Coefficient} (ground_{truth}, predicted_{seg}) \\
 &= \frac{2 | ground_{truth(i)} \cap predicted_{seg} |}{| ground_{truth(i)} + predicted_{seg(i)} |} \tag{7}
 \end{aligned}$$

$$\text{Jaccard}(ground_{truth}, predicted_{seg}) = \frac{| ground_{truth(i)} \cap predicted_{seg(i)} |}{| ground_{truth(i)} \cup predicted_{seg(i)} |} \tag{8}$$

$$Recall = \frac{True_{Positive}}{True_{Positive} + False_{negative}} \quad (9)$$

$$Precision = \frac{True_{Positive}}{True_{Positive} + Fl_{pos}} \quad (10)$$

$$Accuracy = \frac{True_{Positive} + True_{negative}}{True_{Positive} + True_{negative} + False_{positive} + False_{negative}} \quad (11)$$

Results

The experimental outcomes of the proposed U-shaped multi-attention triple decoder convolution network for BC segmentation from the U/S imaging dataset are presented in this section. The implementation of our segmentation methodology was carried out by using the TensorFlow library in Python version 3.6. For the implementation of the proposed MATD convolution network, Dell precision corei7 m4800 workstation with 20 GB of RAM, and 2 GB of NVidia graphic card was utilized. For the evaluation of our MATD convolution network method, two publicly available breast U/S image datasets were utilized. The proposed BC segmentation method was trained from scratch using both U/S datasets for 90 epochs using an initial learning rate of 0.0001, Adam optimizer, and mini-batch size of 8.

Experiment 1 on UDIAT Dataset

In the first experiment, the BC segmentation results are computed by using the proposed MATD convolution network on the breast U/S imaging UDIAT dataset. The segmentation results of the proposed segmentation network are tabulated in Table 1. In this experiment, the BC segmentation results of the U-shaped multi-attention triple decoder convolution network were computed by using different configurations. A DC of 66.34% was recorded with a single decoder network, a DC of 84.48% was recorded with a double decoder network, a dice of 88.83% was recorded with a triple decoder network, and a high dice of 90.45% was recorded by using proposed MATD convolution network. The multi-attention

triple decoder network significantly improves the segmentation dice scores. The Jaccard index for BC segmentation using a single decoder was 50.65% which was improved by using a double decoder network and recorded as 73.85%. The proposed triple decoder network improves the Jaccard index at 80.84% and finally, the proposed multi-scale attention-based triple decoder network achieved the highest Jaccard of 83.40%. This experiment validated that the proposed MATD convolution network performed better with multi-attention triple decoder implementation on the UDIAT BC U/S dataset.

The visual segmentation performance of the proposed U-shaped multi-attention triple decoder convolution network on the UDIAT dataset is presented in Fig. 3. To show the image level BC segmentation from U/S images using the proposed MATD convolution network, these results were computed by randomly choosing 24 breast U/S images from the testing set of the UDIAT repository. The results are presented in the form of predicted segmentation of BC using a MATD convolution network and respective ground truth images. For more understanding of the visual outcomes of the U-shaped multi-attention triple decoder convolution network on the UDIAT repository, a DC score of each tested image is also shown in the results.

Experiment 2 on BUSI Dataset

In the second experiment, the BC segmentation results are computed by using the proposed MATD convolution network on the breast U/S imaging BUSI dataset. The segmentation outcomes using the BUSI repository are given in Table 2. In this experiment, the segmentation outcomes of the proposed U-shaped multi-attention triple decoder convolution network were also computed by using different configurations. A DC of 64.88% was recorded with a single decoder network, a DC of 83.01% was recorded with a double decoder network, a dice of 86.12% was recorded with a triple decoder network, and a high dice of 89.13% was recorded by using proposed MATD convolution network. The multi-attention triple decoder network significantly improved the segmentation dice scores on the BUSI dataset. The Jaccard index for BC segmentation using a single decoder was 49.58% which was improved by using a double decoder network and recorded as 71.55%. The proposed triple decoder network improves

Table 1 Segmentation outcomes using MATD convolution network for BC segmentation from U/S images on the UDIAT dataset

Method	Ac (%)	DC (%)	JC (%)	Pr (%)	Re (%)
Proposed model (single decoder)	80.11	66.34	50.65	77.74	75.45
Proposed model (double decoder)	90.15	84.48	73.85	88.60	82.75
Proposed model (triple decoder)	95.12	88.83	80.84	94.97	85.93
Proposed model (triple decoder + multi-attention)	97.69	90.45	83.40	95.50	87.14

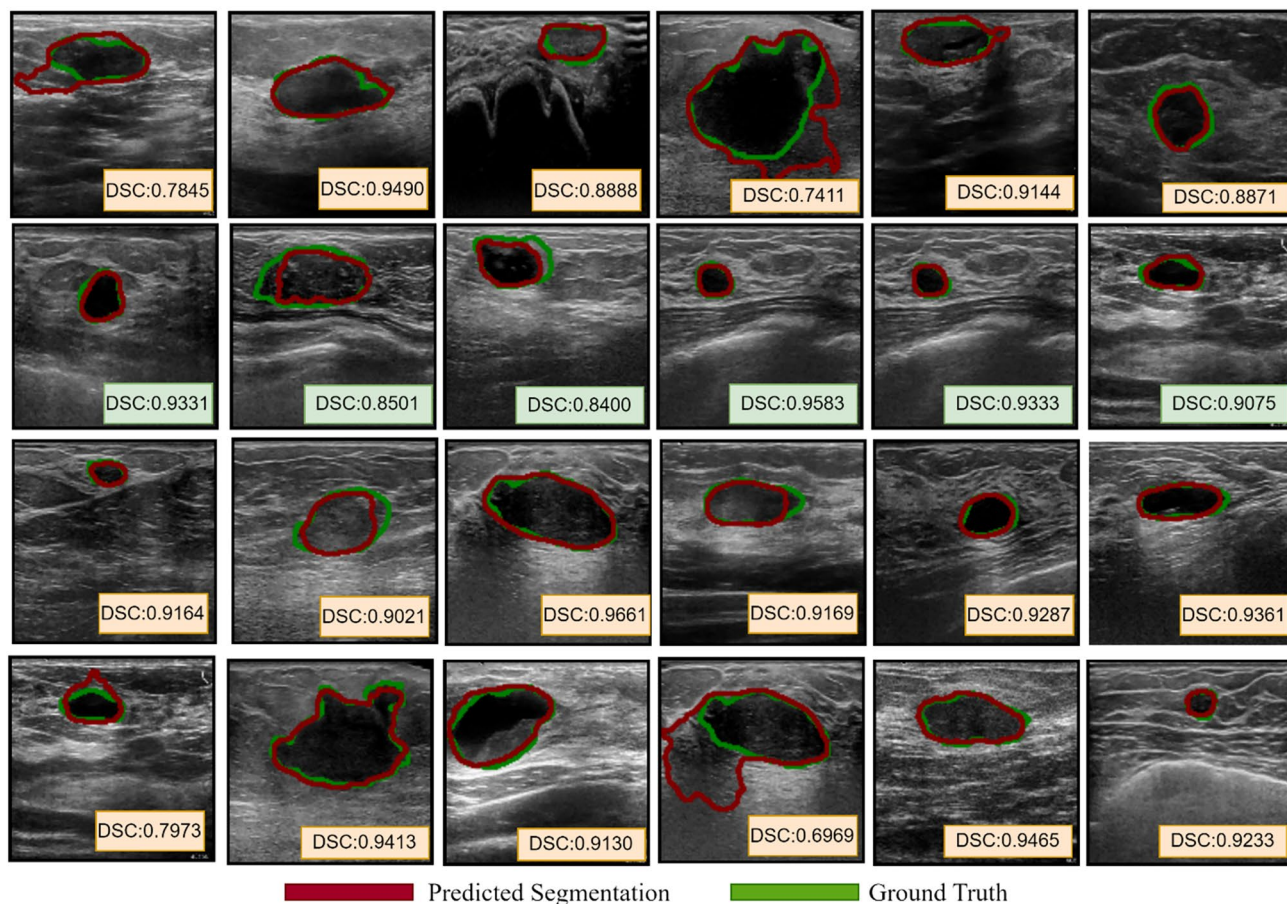


Fig. 3 The Segmentation visual results of the MATD convolution network for BC segmentation from U/S images on UDIAT dataset

the Jaccard index at 79.01%, and finally, the proposed multi-scale attention-based triple decoder network achieved the highest Jaccard of 82.31% on the BUSI dataset. This experiment validated that the proposed MATD convolution network performed better with multi-attention triple decoder implementation on the BUSI BC U/S dataset.

The visual image level breast tumors localized segmentation results of the proposed U-shaped multi-attention triple decoder convolution network on the BUSI dataset are given in Fig. 4. The visual BC segmentation results as localized tumors from U/S images using the proposed U-shaped multi-attention triple decoder convolution network were computed by randomly choosing 24 breast U/S

images from the testing set of the BUSI repository. The results are presented in the form of predicted segmentation of BC using a MATD convolution network and respective ground truth images. For more understanding of the visual outcomes of the U-shaped multi-attention triple decoder convolution network on the BUSI repository a DC score of each tested image is also shown in the results.

Comparison with Existing Methods Using BUSI Dataset

The segmentation result comparison of BC segmentation using the U-shaped multi-attention triple decoder convolution network with existing techniques on the BUSI

Table 2 Segmentation outcomes of the proposed U-shaped multi-attention triple decoder convolution network for BC segmentation from U/S images using the BUSI dataset

Method	Ac (%)	DC (%)	JC (%)	Pr (%)	Re (%)
Proposed model (single decoder)	78.55	64.88	49.58	75.45	73.15
Proposed model (double decoder)	88.44	83.01	71.55	86.12	80.11
Proposed model (triple decoder)	94.11	86.12	79.01	92.11	83.22
Proposed model (triple decoder + multi-attention)	96.99	89.13	82.31	94.55	89.03

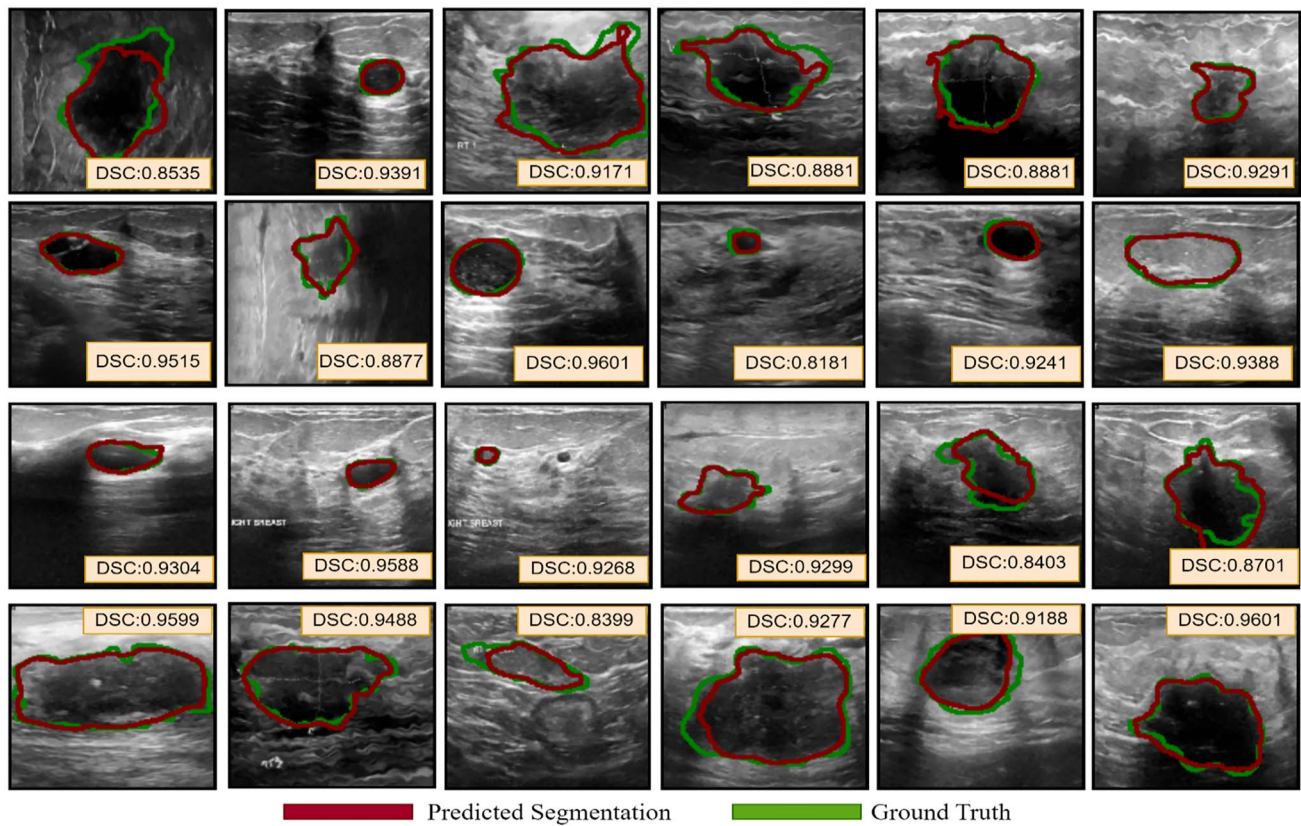


Fig. 4 Segmentation visual results of the proposed U-shaped multi-attention triple decoder convolution network for BC segmentation from U/S images using the BUSI dataset

repository is given in Table 3. For the fair comparison of the U-shaped multi-attention triple decoder convolution network, different well-known image segmentation methods are selected. The comparison was carried out by implementing five existing methods on BUSI dataset including U-Net by Ronneberger et al. [36], U-Net++ by Zhou et al. [37], DeepLabv3+ by Chen et al. [38], PSP-Net by Zhao et al. [39], and MSU-Net by the Su et al. [40]. To further highlight the contributions of this work, three state-of-the-art methods including [41–43] are used to compare the performance of our model with their reported results on BUSI dataset. This comparison concluded that the proposed MATD convolution network method achieved the best DC of 89.13% using the BUSI repository. The bar graphs of the proposed MATD convolution network are presented in Fig. 5 for a better understanding of the comparison. The visual predictions in terms of localized tumors on output images using the proposed U-shaped multi-attention triple decoder convolution network, U-Net, U-Net++, PSP-Net, MSU-Net, and DeepLabv3+ are given in Fig. 6. The predicted localized breast tumor segmentation-based comparison showed that the proposed MATD convolution network achieved the highest DC score. This comparison was

conducted by using four random U/S images from BUSI repository. From the comparison of the proposed method with existing methods, it can be concluded that proposed multi-scale encoder performed the vital role in extracting

Table 3 Comparison of the proposed U-shaped multi-attention triple decoder convolution network for BC segmentation from U/S images with existing methods on the BUSI repository

Method	Year	DC (%)	JC (%)	Pr (%)	Re (%)
U-Net [36]	2015	73.8	66.1	86.2	74.5
PSP-Net [39]	2017	74.7	62.3	81.4	73.5
U-Net++ [37]	2018	72.2	62.2	83.6	71.4
MSU-Net [40]	2021	82.1	70.4	82.9	74.2
DeepLabv3+ [38]	2018	80.2	69.6	84.2	82.2
ESTAN [41]	2020	81.1	72.1	–	89.1
Yang et al. [42]	2022	79.1	70.1	–	82.1
Zhang et al. [43]	2022	74.4	–	78.8	74.7
Chen et al. TransUnet [44]	2021	79.00	71.47	81.66	80.78
Zhao et al. [45]	2023	79.21	71.52	84.00	79.97
Proposed method		89.1	82.3	94.5	89.0

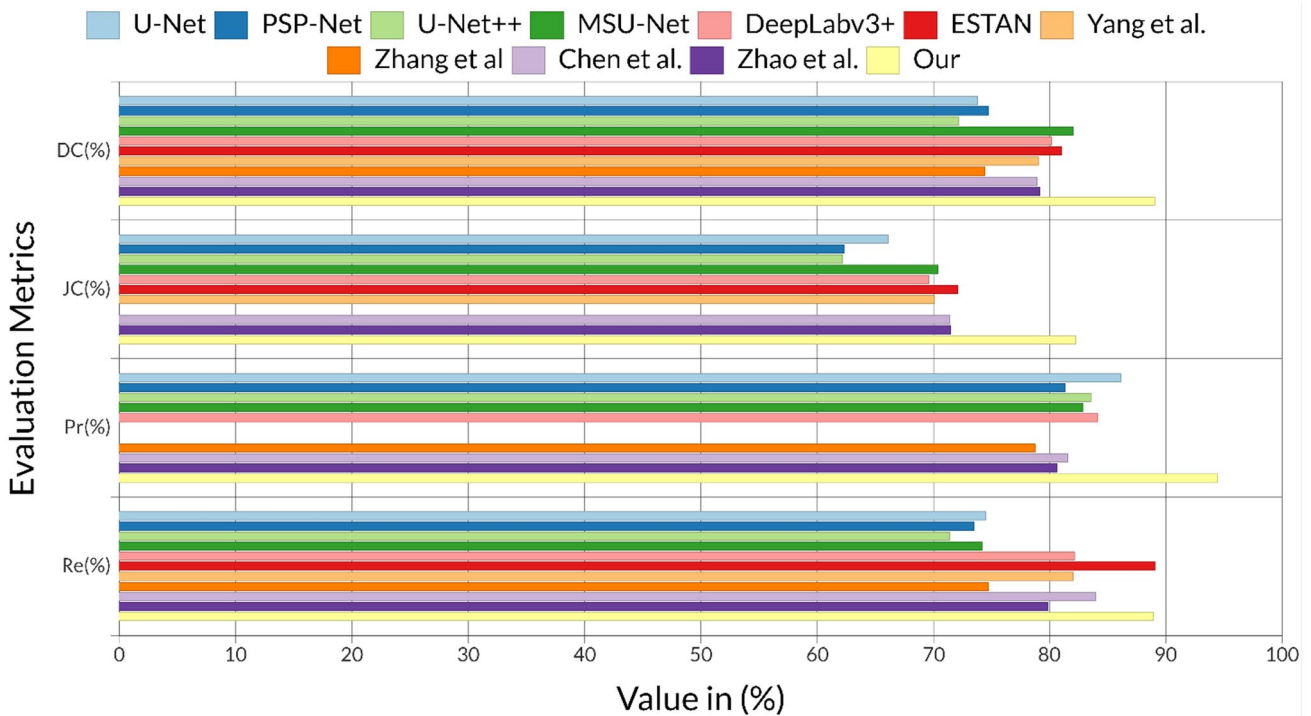


Fig. 5 Bar graphs comparison of proposed MATD convolution network for BC segmentation from U/S images using the UDIAT dataset

the diverse features that was not introduced in earlier studies. The newly introduced multi-scale decoder network with multi-attention mechanism significantly improved the segmentation performance due multi-scale attention mechanism. The more importance of different components of

proposed multi-decoder network on BUSI dataset is shown in ablation study in Table 2. The addition of each decoder in the MATD network significantly improved the segmentation performance which is showing the uniqueness of the proposed model.

Fig. 6 Visual segmentation including dice scores comparison of the U-shaped multi-attention triple decoder convolution network for BC segmentation from U/S images with existing methods on the BUSI repository

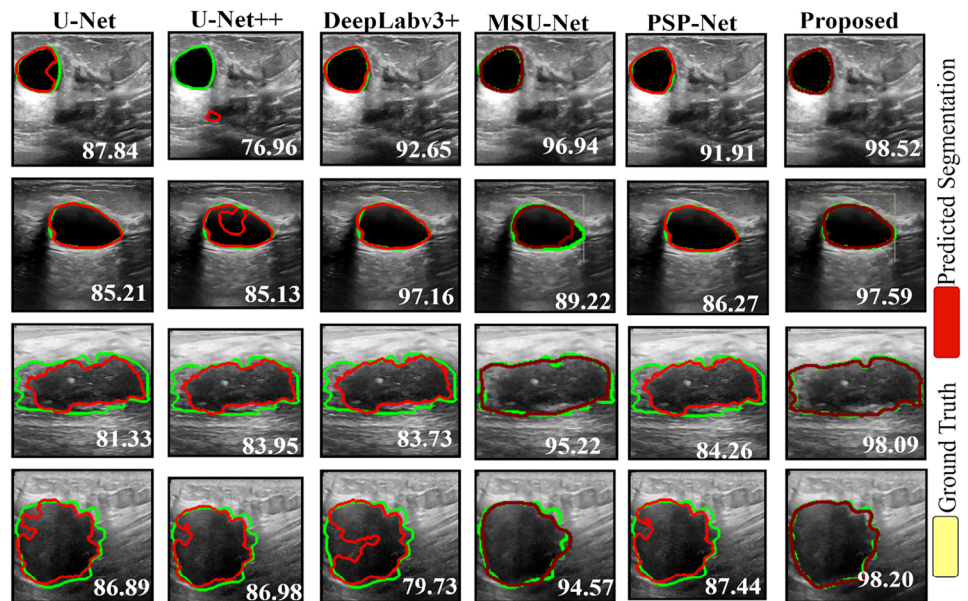


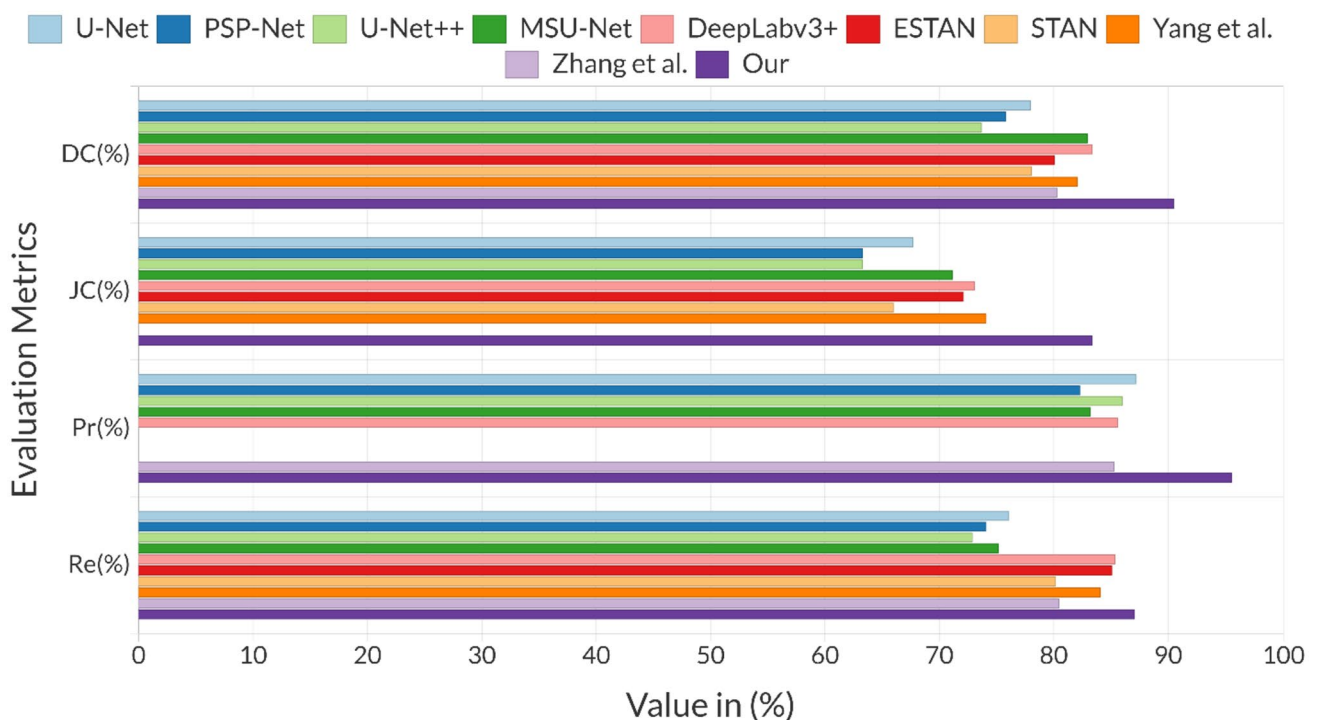
Table 4 Comparison of the proposed U-shaped multi-attention triple decoder convolution network for BC segmentation from U/S images with existing methods using the UDIAT repository

Method	Year	DC (%)	JC (%)	Pr (%)	Re (%)
U-Net [36]	2015	78.0	67.7	87.2	76.1
PSP-Net [39]	2017	75.8	63.3	82.3	74.1
U-Net++ [37]	2018	73.7	63.3	86.0	72.9
MSU-Net [40]	2021	83.0	71.2	83.2	75.2
DeepLabv3+ [38]	2018	83.4	73.1	85.6	85.4
ESTAN [41]	2020	80.1	72.1	-	85.1
STAN [46]	2020	78.1	66.0	-	80.2
Yang et al. [42]	2022	82.1	74.1	-	84.1
Zhang et al. [43]	2022	80.3	-	85.3	80.5
Proposed method	-	90.5	83.4	95.5	87.1

Comparison with Existing Methods Using UDIAT Dataset

The segmentation result comparison of BC segmentation using the proposed U-shaped multi-attention triple decoder convolution network with existing techniques on the UDIAT repository is tabulated in Table 4. For the fair comparison of the U-shaped multi-attention triple decoder convolution network, different well-known image segmentation methods are selected. The comparison was carried out by implementing five existing methods on UDIAT dataset including U-Net by Ronneberger et al. [36], U-Net++ by

Zhou et al. [37], DeepLabv3+ by Chen et al. [38], PSP-Net by Zhao et al. [39], and MSU-Net by the Su et al. [40]. To further highlight the contributions of this work, four state-of-the-art methods including [41–43, 46] are used to compare the performance of our model with their reported results on UDIAT dataset. This comparison concluded that the proposed MATD convolution network method achieved the best DC of 90.45% on the UDIAT dataset. The line graphs of the proposed MATD convolution network are presented in Fig. 7 for a better understanding of comparison. The visual predictions in terms of localized tumors of the proposed U-shaped multi-attention triple decoder convolution network method with five implemented methods including U-Net, U-Net++, PSP-Net, MSU-Net, and DeepLabv3+ are presented in Fig. 8. The predicted localized breast tumors segmentation-based comparison showed that the proposed MATD convolution network achieved the highest DC score. This comparison was conducted by using four random U/S images from the UDIAT repository. From the comparison of the proposed method with existing methods, it can be concluded that proposed multi-scale encoder performed the vital role in extracting the diverse features that was not introduced in earlier studies. The newly introduced multi-scale decoder network with multi-attention mechanism significantly improved the segmentation performance due multi-scale attention mechanism. The more importance of different components of proposed multi-decoder network on UDIAT dataset is shown in ablation study in Table 1. The

**Fig. 7** Line graphs comparison of proposed MATD convolution network for BC segmentation from U/S images using the UDIAT dataset

spatial features at different scales are then transformed into multi-scale triple decoder networks. Each decoder network utilized the attention mechanism to highlight the tumor region at different scales. The results of the proposed MATD convolution network model show that the multi-attention triple decoder network significantly improves the segmentation outcomes. The proposed MATD convolution network produced the highest DC of 90.45% and DC of 89.13% on the UDIAT and BUSI datasets, respectively. The real-life application of the U-shaped multi-attention triple decoder convolution network is the implementation of this method in hospitals for BC segmentation from U/S images. The computation efficiency of the proposed model with other competitors' methods is presented in Table 5. The computation comparison is conducted based on the number of training parameters, inference time, trained model size, and mean frame per second. The computation comparison showed that the MATD method is best suited for BC segmentation with low computation cost.

Conclusion

This work proposed a MATD convolution network for BC segmentation from U/S images. The proposed segmentation method introduced a multi-scale convolution method in the encoder path of the model. The multi-scale learned spatial image features are transformed into a triple decoder network for predicated segmentation mask regeneration. The triple decoder network was implemented at different scales to handle each learned scaled spatial image feature in the encoder path. A multi-attention mechanism is also introduced in each decoder block to highlight the tumor region. The proposed MATD convolution network for BC segmentation from U/S images is composed of four encoder and four decoder blocks. For the transformation of the learned multi-scale spatial image features in the contraction path, the skip connections were utilized. The segmentation results showed that the proposed multi-attention triple decoder network produced the highest segmentation DC. Two publicly available BC U/S image datasets were used to test the performance of the proposed MATD convolution network for BC segmentation from U/S images including BUSI and UDIAT datasets. The proposed method produced a DC score of 90.45% on the UDIAT U/S image dataset and a DC score of 89.13% on the BUSI U/S dataset. In the future, this work will further be enhanced to implement the multi-encoder framework for BC segmentation by using more U/S imaging datasets.

Data Availability This study utilized two publicly available datasets including the BUSI dataset by Dhabyani et al. [34] available at <https://scholar.cu.edu.eg/?q=afahmy/pages/dataset> and the UDIAT dataset B which can be accessed by mailing the principal author of the study Yap et al. [35].

Declarations

Ethics Approval This article does not contain any studies with human participants performed by any of the authors.

Competing Interest The authors declare no competing interests.

References

- Giaquinto AN, Sung H, Miller KD, Kramer JL, Newman LA, Minihan A, Jemal A, Siegel RL. Breast cancer statistics. *CA Cancer J Clin.* 2022. <https://doi.org/10.3322/caac.21754>.
- Sung H, Ferlay J, Siegel RL, 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 Cancer J Clin.* 2021;71:209–49.
- Alanazi SA, Kamruzzaman MM, Islam Sarker MN, Alruwaili M, Alhwaiti Y, Alshammari N, Siddiqi MH. Boosting breast cancer detection using convolutional neural network. *J Healthc Eng.* 2021;2021:1–11. <https://doi.org/10.1155/2021/5528622>.
- Murtaza G, Shuib L, Wahab AWA, Mujtaba G, Nweke HF, Al-garadi MA, Zulfiqar F, Raza G, Azmi NA. Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges. *Artif Intell Rev.* 2019;1–66.
- Debelee TG, Schwenker F, Ibenhal A, Yohannes D. Survey of deep learning in breast cancer image analysis. *Evol Syst.* 2020;11:143–63.
- Sabtu SN, Abdul Sani SF, Bradley DA, Looi LM, Osman Z. A review of the applications of Raman spectroscopy for breast cancer tissue diagnostic and their histopathological classification of epithelial to mesenchymal transition. *J Raman Spectrosc.* 2020;51:380–9.
- Umer MJ, Sharif M, Kadry S, Alharbi A. Multi-class classification of breast cancer using 6B-Net with deep feature fusion and selection method. *J Pers Med.* 2022;12:683. <https://doi.org/10.3390/jpm12050683>.
- Jahwar AF, Abdulazeez AM. Segmentation and classification for breast cancer ultrasound images using deep learning techniques: a review. In: 2022 IEEE 18th International Colloquium on Signal Processing & Applications (CSPA). IEEE. 2022;225–230
- Umer MJ, Sharif M, Wang S-H. Breast cancer classification and segmentation framework using multiscale CNN and U-shaped dual decoded attention network. *Expert Syst n/a:e13192.* <https://doi.org/10.1111/exsy.13192>.
- Vigil N, Barry M, Amini A, Akhloufi M, Maldague XPV, Ma L, Ren L, Yousefi B. Dual-intended deep learning model for breast cancer diagnosis in ultrasound imaging. *Cancers.* 2022;14:2663. <https://doi.org/10.3390/cancers14112663>.
- Wang Q, Chen H, Luo G, Li B, Shang H, Shao H, Sun S, Wang Z, Wang K, Cheng W. Performance of novel deep learning network with the incorporation of the automatic segmentation network for diagnosis of breast cancer in automated breast ultrasound. *Eur Radiol.* 2022;32:7163–72. <https://doi.org/10.1007/s00330-022-08836-x>.
- Ayana G, Park J, Jeong J-W, Choe S. A novel multistage transfer learning for ultrasound breast cancer image classification. *Diagnostics.* 2022;12:135. <https://doi.org/10.3390/diagnostics12010135>.
- Umer M, Sharif M, Alhaisoni M, Tariq U, Kim Y, Chang B. A framework of deep learning and selection-based breast cancer detection from histopathology images. *Comput Syst Sci Eng.* 2022;45:1001–1016. <https://doi.org/10.32604/csse.2023.030463>.
- Baccouche A, Garcia-Zapirain B, Castillo Olea C, Elmaghraby AS. Connected-UNets: a deep learning architecture for breast

- mass segmentation. *Npj Breast Cancer*. 2021;7:1–12. <https://doi.org/10.1038/s41523-021-00358-x>.
15. Luo Y, Huang Q, Li X. Segmentation information with attention integration for classification of breast tumor in ultrasound image. *Pattern Recognit*. 2022;124: 108427. <https://doi.org/10.1016/j.patcog.2021.108427>.
 16. Yan Y, Liu Y, Wu Y, Zhang H, Zhang Y, Meng L. Accurate segmentation of breast tumors using AE U-net with HDC model in ultrasound images. *Biomed Signal Process Control*. 2022;72: 103299. <https://doi.org/10.1016/j.bspc.2021.103299>.
 17. Chen G, Dai Y, Zhang J. C-Net: Cascaded convolutional neural network with global guidance and refinement residuals for breast ultrasound images segmentation. *Comput Methods Programs Biomed*. 2022;225: 107086. <https://doi.org/10.1016/j.cmpb.2022.107086>.
 18. Iqbal A, Sharif M. MDA-Net: multiscale dual attention-based network for breast lesion segmentation using ultrasound images. *J King Saud Univ - Comput Inf Sci*. 2021;S1319157821002895. <https://doi.org/10.1016/j.jksuci.2021.10.002>.
 19. Farooq MA, Gong ZX, Liu Y, Zubair M, Manzoor A, Zhang G. Breast cancer detection from ultrasound images using attention U-nets model. In: Fourteenth International Conference on Digital Image Processing (ICDIP 2022). SPIE. 2022;161–174.
 20. Tong Y, Liu Y, Zhao M, Meng L, Zhang J. Improved U-net MALF model for lesion segmentation in breast ultrasound images. *Biomed Signal Process Control*. 2021;68: 102721. <https://doi.org/10.1016/j.bspc.2021.102721>.
 21. Meraj T, Alosaimi W, Alouffi B, Rauf HT, Kumar SA, Damaševičius R, Alyami H. A quantization assisted U-Net study with ICA and deep features fusion for breast cancer identification using ultrasonic data. *PeerJ Comput Sci*. 2021;7: e805.
 22. Vianna P, Farias R, de Albuquerque Pereira WC. U-Net and SegNet performances on lesion segmentation of breast ultrasonography images. *Res Biomed Eng*. 2021;37:171–9. <https://doi.org/10.1007/s42600-021-00137-4>.
 23. Sannasi Chakravarthy SR, Rajaguru H. SKMAT-U-Net architecture for breast mass segmentation. *Int J Imaging Syst Technol*. 2022;32:1880–8. <https://doi.org/10.1002/ima.22781>.
 24. Zhao T, Dai H. Breast tumor ultrasound image segmentation method based on improved residual U-Net network. *Comput Intell Neurosci*. 2022;2022: e3905998. <https://doi.org/10.1155/2022/3905998>.
 25. Li J, Cheng L, Xia T, Ni H, Li J. Multi-scale fusion U-Net for the segmentation of breast lesions. *IEEE Access*. 2021;9:137125–39. <https://doi.org/10.1109/ACCESS.2021.3117578>.
 26. Lu S-Y, Wang S-H, Zhang Y-D. SAFNet: a deep spatial attention network with classifier fusion for breast cancer detection. *Comput Biol Med*. 2022;148: 105812. <https://doi.org/10.1016/j.compbiomed.2022.105812>.
 27. Punni NS, Agarwal S. RCA-IUNet: a residual cross-spatial attention-guided inception U-Net model for tumor segmentation in breast ultrasound imaging. *Mach Vis Appl*. 2022;33:27. <https://doi.org/10.1007/s00138-022-01280-3>.
 28. Wang Y, Yao Y. Breast lesion detection using an anchor-free network from ultrasound images with segmentation-based enhancement. *Sci Rep*. 2022;12:14720. <https://doi.org/10.1038/s41598-022-18747-y>.
 29. Wang J, Chen G, Chen S, Joseph Raj AN, Zhuang Z, Xie L, Ma S. Ultrasonic breast tumor extraction based on adversarial mechanism and active contour. *Comput Methods Programs Biomed*. 2022;225: 107052. <https://doi.org/10.1016/j.cmpb.2022.107052>.
 30. Woon Cho S, Rae Baek N, Ryoung Park K. Deep learning-based multi-stage segmentation method using ultrasound images for breast cancer diagnosis. *J King Saud Univ - Comput Inf Sci*. 2022. <https://doi.org/10.1016/j.jksuci.2022.10.020>.
 31. Gong X, Zhao X, Fan L, Li T, Guo Y, Luo J. BUS-net: a bimodal ultrasound network for breast cancer diagnosis. *Int J Mach Learn Cybern*. 2022;13:3311–28. <https://doi.org/10.1007/s13042-022-01596-6>.
 32. Peng C, Zhang Y, Meng Y, Yang Y, Qiu B, Cao Y, Zheng J. LMA-Net: A lesion morphology aware network for medical image segmentation towards breast tumors. *Comput Biol Med*. 2022;147: 105685. <https://doi.org/10.1016/j.compbiomed.2022.105685>.
 33. Lou M, Meng J, Qi Y, Li X, Ma Y. MCRNet: multi-level context refinement network for semantic segmentation in breast ultrasound imaging. *Neurocomputing*. 2022;470:154–69. <https://doi.org/10.1016/j.neucom.2021.10.102>.
 34. Al-Dhabyani W, Gomaa M, Khaled H, Fahmy A. Dataset of breast ultrasound images. *Data Brief*. 2020;28: 104863. <https://doi.org/10.1016/j.dib.2019.104863>.
 35. Yap MH, Pons G, Martí J, Ganau S, Sentís M, Zwigelaar R, Davison AK, Martí R. Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE J Biomed Health Inform*. 2018;22:1218–26. <https://doi.org/10.1109/JBHI.2017.2731873>.
 36. Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation. In: Navab N, Hornegger J, Wells WM, Frangi AF, editors. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*. Cham: Springer International Publishing; 2015. p. 234–41.
 37. Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, Liang J. UNet++: a nested U-Net architecture for medical image segmentation. In: Stoyanov D, Taylor Z, Carneiro G, Syeda-Mahmood T, Martel A, Maier-Hein L, Tavares JMRS, Bradley A, Papa JP, Belagiannis V, Nascimento JC, Lu Z, Conjeti S, Moradi M, Greenspan H, Madabhushi A, editors. *Deep Learning in medical image analysis and multimodal learning for clinical decision support*. Cham: Springer International Publishing; 2018. p. 3–11.
 38. Chen L-C, Zhu Y, Papandreou G, Schroff F, Adam H. Encoder-decoder with atrous separable convolution for semantic image segmentation. In: *Proceedings of the European conference on computer vision (ECCV)*. 2018;801–818.
 39. Zhao H, Shi J, Qi X, Wang X, Jia J. Pyramid scene parsing network. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017;2881–2890.
 40. Su R, Zhang D, Liu J, Cheng C. MSU-net: multi-scale U-net for 2D medical image segmentation. *Front Genet*. 2021;12: 639930.
 41. Shareef B, Vakanski A, Xian M, Freer PE. ESTAN: enhanced small tumor-aware network for breast ultrasound image segmentation. *ArXiv Prepr*. 2020;ArXiv200912894.
 42. Yang K, Suzuki A, Ye J, Nosato H, Izumori A, Sakanashi H. CTG-Net: Cross-task guided network for breast ultrasound diagnosis. *PLoS ONE*. 2022;17: e0271106.
 43. Zhang M, Huang A, Yang D, Xu R, Wu Y. Boundary-oriented network for automatic breast tumor segmentation in ultrasound images. Available SSRN 4098691.
 44. Chen J, Lu Y, Yu Q, Luo X, Adeli E, Wang Y, Lu L, Yuille AL, Zhou Y. TransUNet: transformers make strong encoders for medical image segmentation. 2021.
 45. Zhao X, Jia H, Pang Y, Lv L, Tian F, Zhang L, Sun W, Lu H. M²SSNet: multi-scale in multi-scale subtraction network for medical image segmentation. 2023.
 46. Shareef B, Xian M, Vakanski A. STAN: small tumor-aware network for breast ultrasound image segmentation. In: *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*. IEEE. 2020;1–5.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.