IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 22 May 2024, accepted 8 July 2024, date of publication 16 July 2024, date of current version 14 August 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3429386

RESEARCH ARTICLE

Breast Cancer Segmentation From Ultrasound Images Using Multiscale Cascaded Convolution With Residual Attention-Based Double Decoder Network

MUHAMMAD JUNAID UMER[1], **MUHAMMAD IRFAN SHARIF**[2], **AND JUNGEUN KIM**[3,4]
[1]Department of Computer Science, COMSATS University Islamabad, Wah Campus, Wah Cantt 47040, Pakistan
[2]Department of Information Sciences, University of Education Lahore, Jauharabad Campus, Jauharabad, Pakistan
[3]Department of Software, Kongju National University, Cheonan 31080, South Korea
[4]Department of Computer Science and Engineering, Kongju National University, Cheonan 31080, South Korea

Corresponding authors: Muhammad Irfan Sharif (irfan.sharif@ue.edu.pk) and Jungeun Kim (jekim@kongju.ac.kr)

This work was supported in part by the Technology Development Program of MSS under Grant S3033853 and in part by the Research Grant of Kongju National University in 2024.

ABSTRACT Accurate segmentation of breast cancer (BC) in ultrasound images is a complicated task due to the variable nature of ultrasound images. Recently, many techniques are suggested to accurately segment BC using ultrasound imaging and deep learning. A multiscale cascaded convolution with residual attention-based double decoder network for BC segmentation is presented in this study. A multiscale cascaded convolution operation-based encoder path is designed to overcome the problem of a single scale feature learning process. The proposed multiscale convolution operation helps to extract the diverse semantic spatial features. In the segmentation process, a residual attention-based double decoder network is proposed. The proposed attention mechanism is implemented to take out the more prominent features from the tumor region and to suppress the other information that can mislead the segmentation model during training. The double decoding mechanisms are introduced to capture the highly diverse spatial features that are learned in the encoder path. For experimental purposes two publically available ultrasound image datasets namely BUSI and UDIAT are utilized. The proposed U-shaped multiscale cascaded convolution with residual attention-based double decoder network achieved the segmentation dice of 91.38% with 81.67% of the Jaccard index, 94.43% precision, and 87.76% recall score on the UDIAT dataset. A dice score of 90.55% is recorded with a Jaccard of 80.87%, 93.53 % precision score, and 88.46% recall score on the BUSI dataset. The results of the multiscale cascaded convolution with residual attention-based double decoder network validated that it can effectively be used for breast cancer segmentation tasks.

INDEX TERMS Breast cancer, u-net, residual, encoder, segmentation.

I. INTRODUCTION

Breast cancer disease is being diagnosed on a large scale and is mostly present in middle-aged women worldwide, which has a very high mortality rate. BC was declared as one of the most frequently diagnosed cancers in females in 2020, statistics of breast cancer show that newly diagnosed breast cancer in females reaches the highest number of 2261419 cases with many new deaths of 684996 [1]. Breast cancer disease is declared as the highest death ratio disease after lung cancer according to a cancer survey [2]. Therefore, early and accurate detection of BC can help to reduce the death ratio

The associate editor coordinating the review of this manuscript and approving it for publication was N. Ramesh Babu.

of this disease through early and proper treatment. Initial diagnosis of BC is accomplished manually by using medical imaging such as mammography and ultrasonography. The final diagnosis of extracted specimens after surgery is then carried out using microscopy by the pathologists [3]. At this stage of the final diagnostic process, accurate classification of breast cancer is highly required from the pathologists which may lead to another surgery if not diagnosed properly. Due to a lack of objective standards analyzing the visual results of microscopic images is subjective and time taking [4]. The high number of breast cancer patients worldwide necessitates the computer-aided diagnosis of this deadly disease for early detection. The death ratio of this deadly disease can be minimized with early and accurate detection before going to the critical stage [5]. According to a survey, about 90 percent of BC patients may be recovered if diagnosed early and treated properly [6].

For initial diagnosis of this disease, breast ultrasound (BUS) imaging is considered as the cheapest and safest method. The BUS imaging-based manual detection of BC is a complicated task, due to the complex nature of BUS images [7]. Computer-aided diagnosis (CAD) of this disease can be an alternative and effective solution to replace the manual diagnosis process. Recently, with the advancements in technology, many CADs are presented for the early examination of different diseases [8], [9]. A CAD solution based on multi-scale dual attention mechanism is presented in [10] using BUS images. Meraj et al [11] recently introduced a study to accurately categorizing the BC using ultrasound modality. A CNN with a quantization technique was introduced as a U-shaped model for the segmentation task, containing fusion mechanism. Yap et al. [12] introduced a multi-model solution for BC segmentation from BUS images. Three different methods were utilized for BC segmentation including transfer learning (TL) with alexnet and FCN, a U-net-based method, and a multi-level patched-based method using LeNet. This work also ensured the availability of breast ultrasound dataset B for research purposes. Singh et al. [13] recently developed a CAD system for BC segmentation and shape classification. For BC localization task generator and discriminator-based CNN model using an adversarial network with encoding and decoding mechanism was proposed. After the successful segmentation of the breast tumors in the next step, the shape classification of the segmented tumors was carried out with the help of the CNN model.

BC segmentation from BUS imaging requires an expert radiologist for diagnosis due to the variable nature of BUS images. Many DL-based methods to replace manual segmentation of BC are presented in the literature. This study proposed a multiscale cascaded convolution with a residual attention-based double decoder network to accurately segment the BC from BUS images. The proposed method introduced a cascaded convolution mechanism with different receptive fields to handle the missed spatial information problem that occurred due to fixed receptive fields in the feature learning process. A residual attention mechanism with double decoding implementation is developed to get the highly discriminative image features from the tumor region. In the first step, the BUS images are passed to the proposed U-shaped multiscale cascaded convolution with residual attention-based double decoder network which is convolved through the multiscale cascaded convolution block for feature learning in the encoder path. The learned highly discriminative image features in the encoder network are then transformed into the residual attention-based double decoder network using skip connections. Finally, the segmented BUS images with predicted segmentation are taken from the output layer of 1×1 convolution. The experiments validated that the proposed multiscale cascaded convolution with residual attention-based double decoder network achieved the best segmentation dice similarity coefficient (DSC) scores. The main contributions of the multiscale cascaded convolution with residual attention-based double decoder network are presented below in bullets.

- This work designed a multiscale cascaded convolution operation block for highly discriminative feature learning.
- This work introduced a residual attention mechanism to enhance the segmentation performance by using an attention mechanism in the double decoding fission.
- This work introduced a double decoding method with residual skip connections to increase the segmentation DSC.

II. RELATED WORK

Due to technological advancements, BC segmentation is being carried out with the assistance of DL and machine learning (ML) [14], [15]. Recently different methods that utilized DL for early detection and classification of BC [16], [17], [18], [19] are presented. A et al. [20] presented a meta-heuristic procedure to fine-tune the CNN model to accurately segment the breast tumor before classification. A classification accuracy of 98% was achieved by their classification scheme. Vigil et al. [21] developed a dual-stage convolution network for combined BC localization and classification with 85.3% of accuracy. Yan et al. [22] recently introduced a DL solution for accurate segmentation of BC from BUS images by using attention enhance u-net with hybrid dilated convolution and reported a segmentation accuracy of 95.81%. A hybrid solution for BC segmentation by implementing MobileNetv2 and VGG16 as encoder paths of the u-shaped CNN was presented by [23]. Zhang et al. [24] developed an attention-based network for BC segmentation. In their work, two attention mechanisms including soft and hard attention with the multitask learning convolution method were proposed. The contraction path of the segmentation model was developed with dense layers and attention gates. The binary classification of the segmented tumors was also performed. Recently Inan et al. [25] introduced a technique for BC classification guided by the segmentation method. Different pre-trained CNN was utilized to perform the both segmentation and classification of BC. A DSC of 63.4%

was achieved by their method with a classification accuracy of 78.92% in the classification stage. Lei et al. [26] introduced a CAD system for BC diagnosis by implementing a region-based convolutional network. Their method was composed of five different parts or subnets and a dice of 85% was recorded in the segmentation of breast tumors.

Recently, many DL-based solutions are developed to replace the manual segmentation process of BC localization tasks. Ilesanmi et al. [27] introduced a BC localization model by first preprocessing the BUS images from the histogram equalization technique followed by the variant-enhanced mechanism before the concatenated convolution operations. Their method achieved a DSC of 89.73% on the benign class of breast tumors and a DSC of 89.62% on the malignant class. Huang et al. [28] recently suggested a BC segmentation solution using fuzzy logic. The breast segmentation process was completed in two steps, in the initial step wavelet features and contrast enhancement method were implemented as a preprocessing step then the augmented BUS dataset was converted into the fuzzy logic domain. Finally, the features from the fuzzy domain model were post-processed using the method of conditional random fields. Their method achieved an intersection over union score of 81.29% in the breast segmentation task. In another work, Kim et al. [29] introduced a DL model for BC diagnosis by using multiple pre-trained CNN models and reported the AUC value from 0.92 to 0.96. Tong et al. [30] presented a study to automate the manual BC localization process by introducing four integrated attention loss functions in an attention-based network. In the encoder path of the network, residual blocks were utilized for rich feature learning and a segmentation accuracy of more than 80% was achieved by their proposed setup of BC segmentation. Luo et al. [31] developed a BC classification method using a segmentation attention mechanism. In their work first, the segmentation task was carried out, and then in the second step, the segmented BUS images were fed to two deep CNN models for the classification task, and an accuracy of 90.78% was attained.

In another recent work, Zhang et al. [32] proposed an ANN for BC localization and classification using the back-propagation method. Their work reported a segmentation accuracy of 97.3% inferring that their model achieved the highest segmentation accuracy with hidden layers. Xue et al. [33] introduced a method to implement the global guidance mechanism for BC segmentation by developing a CNN model containing global guidance blocks to extract the spatial and channel-level features. To enhance the segmentation results a shallow boundary detection method was implemented and a dice coefficient of 87.1% was achieved. Ragab et al. [34] introduced a segmentation-assisted classification framework for BC diagnosis by using an ensemble of CNN models. Yu at el. [35] developed another method for BC segmentation by using the residual connection and dense block in the encoder path. Punn and Agarwal [36] presented a cross-spatial filter mechanism for BC segmentation with attention and inception model by implementing cross residual attention in the encoder path for better feature learning and reported a dice of 0.93. Wang et al. [37] developed a BC segmentation model from BUS images by using residual blocks in the segmentation model and reported a dice coefficient of 87.00%. Chen et al. [38] proposed a residuals cascaded CNN model. A global guidance mechanism with attention was also introduced in their work. Zhai et al. [39] developed a BC localization solution by using two generators for feature generation and one discriminator-based network. Pan et al. [40] presented an LSTM method for BC segmentation from BUS images by using an attention mechanism and reported a dice score of 0.81. Karunanayake et al. [41] suggested an intelligent fusion model for BC segmentation.

From the literature survey, it can be concluded that most of the earlier work in BC segmentation was performed with the limitation of fixed receptive [11], [12], [42] that can further be improved by using a multiscale convolution mechanism. In the previous BC segmentation methods, only a single decoding mechanism was presented that can be improved by using a dual decoding mechanism. In the previous BC segmentation methods, residual connections-based models were rarely discussed which can also be used to enhance BC segmentation performance.

III. PROPOSED METHODOLOGY

This work proposed a BC segmentation model by introducing multi-scale cascaded convolution with a residual attention-based decoder network. The proposed u-shaped segmentation model is presented in Fig. 1, which shows the encoder and decoder paths of the proposed multi-scale cascaded convolution with a residual attention-based decoder network in detail. The BUS images are inputted to the proposed multi-scale cascaded convolution with a residual attention-based double decoder network and the output BUS images in the form of segmented images are taken from the output layer. The proposed multi-scale cascaded convolution with a residual attention-based decoder network is composed of four encoder blocks, one multiscale cascaded convolution-based bridge block, and four decoder blocks. After the decoder blocks, to get the segmented output a 1×1 convolution followed by sigmoid activation is utilized. The encoder path of the proposed method is designed by introducing multi-scale cascaded convolution operations to learn the diverse image features. Each encoder structure is composed of a 3×3 convolution block, 5×5 convolution block, 7×7 convolution block, and max pooling operation block. These multiscale convolution blocks in each encoder are implemented in a cascaded order such that the output of the one convolution operation block is the input of the next multi-scale cascaded convolution operation. The decoder blocks of the proposed multi-scale cascaded convolution with a residual attention-based decoder network are comprised of a transpose convolution operation, attention block, first concatenation operation, multiscale cascaded convolution block, second concatenation block, and a 3×3 convolution block.

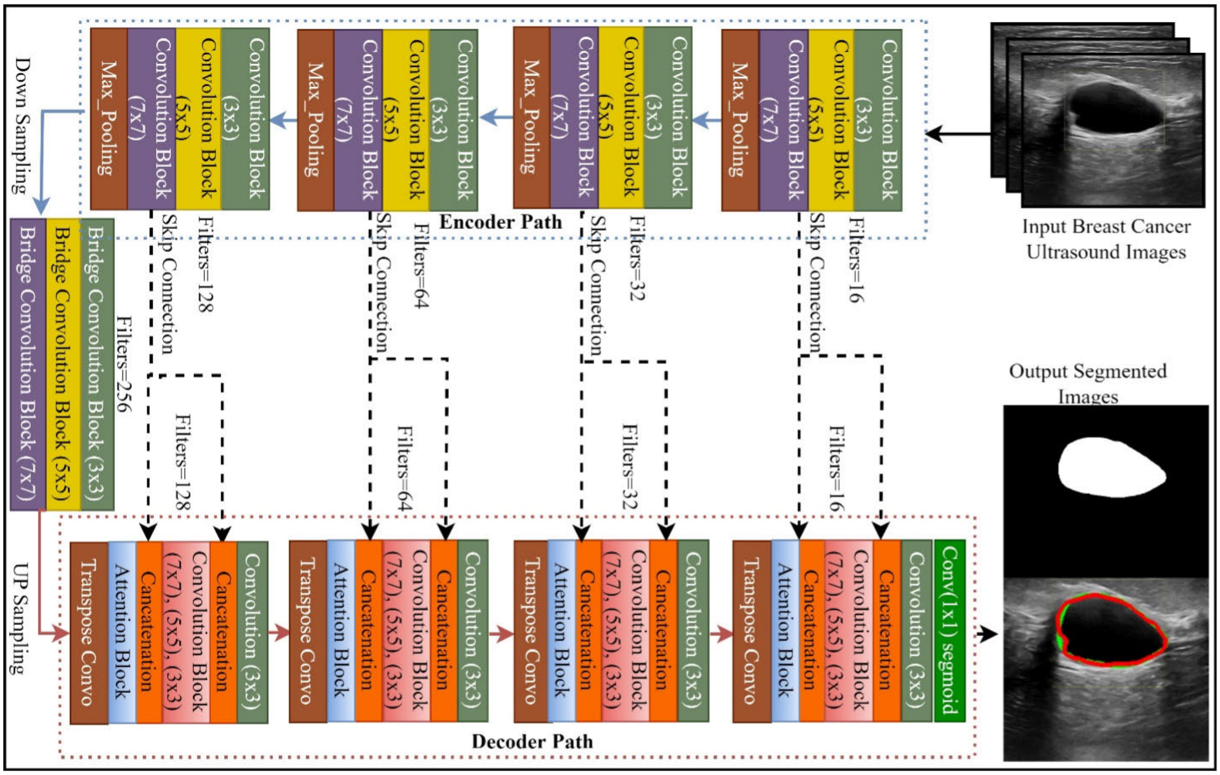

FIGURE 1. **The proposed multi-scale cascaded convolution with residual attention-based double decoder network internal structure diagram.**

The decoder path of our multi-scale cascaded convolution with a residual attention-based double decoder network is implemented with a residual attention mechanism by using skip connections. The skip connections are utilized to transfer the features that are learned in the encoder path to the decoder path for the reconstruction of segmented images. The proposed segmentation model utilized the residual skip connections in which the learned features of the encoder blocks are transformed to each corresponding decoder block in two stages. The number of kernels in each decoder and encoder block is increased gradually from 16 to 128. The encoder path performed the down-sampling of the input BUS image dataset to learn the discriminative spatial features for the segmentation task while the decoder path performed the up-sampling task to regenerate the BUS input image into its original form to show the segmented tumor region. The input image dimensions of our multi-scale cascaded convolution with residual attention decoder network is 128×128. In the first step of multi-scale cascaded convolution with a residual attention-based decoder network, the BUS images are inputted into the proposed segmentation model. The input BUS images are convolved through the different encoder blocks for feature learning. The input BUS images are down-sampled in each encoder block to capture the discriminative spatial image features. In the second step, the down-sampled BUS images are sent to the decoder path by using the bridge and skip connection mechanism for up-sampling. The expansion or decoder path of the multi-scale cascaded convolution with a residual attention-based decoder network performed the up-sampling task to regenerate the input BUS images. The output final segmented BUS images are taken out from the output layer that is implemented with 1×1 convolution and sigmoid activation function.

A. MULTI SCALE CASCADED CONVOLUTION BLOCK

The proposed multi-scale cascaded convolution with a residual attention-based decoder network for BC segmentation introduced a multiscale cascaded convolution mechanism to get over the issue of a fixed receptive field. The proposed multi-scale cascaded convolution process is implemented with a three-scale convolution mechanism to capture the diverse image features. The proposed multiscale cascaded convolution block is presented in Fig. 2. In the first step of the proposed cascaded convolution block, a 3×3 convolution operation is implemented with a stride of two and the output of this 3×3 operation is passed to the batch normalization layer after applying a 50 % dropout operation. After the batch normalization layer, a ReLU function is implemented.

The output of the activation layer is then convolved again with a 5×5-sized kernel followed by the 50% dropout, batch normalization, and ReLU activation. In the last step of the cascaded convolution block, a convolution operation

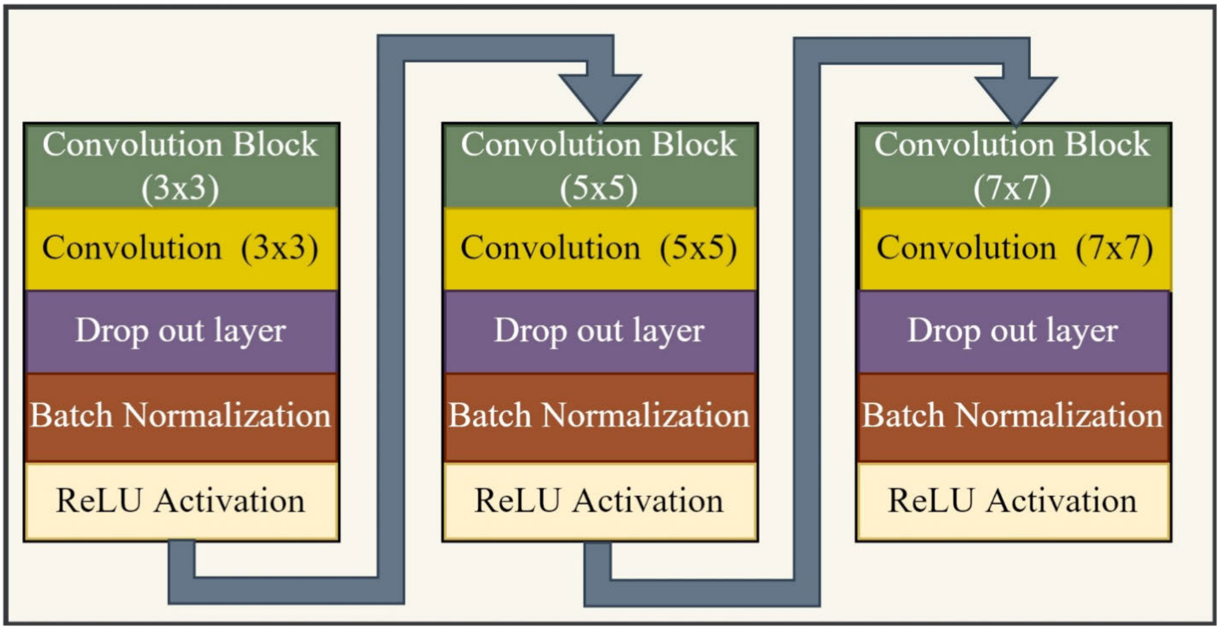

FIGURE 2. The internal structural view of the proposed multi-scale cascaded convolution block.

of 7 × 7 is implemented with 50% dropout, batch normalization, and ReLU activation. Different dropout values such as 30%, 40%, 50%, and 60% are used and the best results are generated by applying the 50% dropout. The proposed multiscale cascaded convolution blocks significantly improve the model performance in terms of the dice coefficient. Most of the previous solutions for automatic BC localization utilized fixed-sized convolution operations which are unable to extract large spatial information. The variable size and nature of breast tumors make it difficult to accurately segment the tumor region in this regard this work implemented a multiscale cascaded convolution method to enhance the segmentation performance. The multiscale convolution operation is mathematically defined in equations 1 to 4. Where Cas_{Conv1} is the cascaded convolution operation with $\varphi_{2,2}^{3\times3}$ receptive field 3 × 3 and stride 2, Cas_{Conv2} is the cascaded convolution operation with $\varphi_{2,2}^{5\times5}$ receptive field 5 × 5 and stride 2, Cas_{Conv3} is the cascaded convolution operation with $\varphi_{2,2}^{7\times7}$ receptive field 7 × 7 and stride 2, and BC_{inImg} is the input BUS image. W and δ represent the weight of the model and bias and shows the cascaded concatenation operation.

$$Cas_{\mathrm{Conv1}} = \left(\varphi_{2,2}^{3\times3} X \left(BC_{inImg}.W\right) + \delta\right) \quad (1)$$

$$Cas_{\mathrm{Conv2}} = \left(\varphi_{2,2}^{5\times5} X \left(Cas_{\mathrm{Conv1}}.W\right) + \delta\right) \quad (2)$$

$$Cas_{\mathrm{Conv3}} = \left(\varphi_{2,2}^{7\times7} X \left(Cas_{\mathrm{Conv2}}.W\right) + \delta\right) \quad (3)$$

$$Msacle_{\mathrm{Conv}} = (((Cas_{\mathrm{Conv1}}).Cas_{\mathrm{Conv2}}).Cas_{\mathrm{Conv3}})) \quad (4)$$

B. RESIDUAL ATTENTION BASED DECODER

This work introduced a residual attention-based two-stage decoder network for accurate localization and segmentation of BC from BUS imaging datasets. The residual attention-based decoder is presented in Fig. 3. The proposed residual attention-based two-stage decoder network achieved the best performance due to its residual attention mechanism which gives more attention to the tumor region. The multi-scale diverse features learned in the encoder path from the encoder block of the multiscale convolution operation block are inputted to the attention mechanism which takes two inputs. The second input of the residual attention operation is the output of the transpose convolution operation. Both inputs of the residual attention mechanism are summed up with an element-wise summation after a 1 × 1 convoluting operation. After this addition, a ReLU activation is applied followed by the second 1 × 1 convolution operation. After the second convolution operation, a sigmoid activation is implemented and then resampling is carried out to multiply this result with the encoder block learned features. After the element-wise multiplication operation, a 3 × 3 convolution operation is performed.

The result of this attention block is further sent as input to the concatenation operation to concatenate it with transpose convolution output. The concatenated output was then passed to a multiscale cascaded convolution block for further processing. The proposed multiscale cascaded convolution block performed the three convolution operations with the different receptive fields. Each convolution operation of the multiscale block is followed by the dropout operation, a batch normalization that works based on batch size, and a ReLU activation operation. The computed outcomes of the multiscale cascaded convolution block are then processed by using the concatenation operation on the encoder's learned features as a second stage. Finally, the

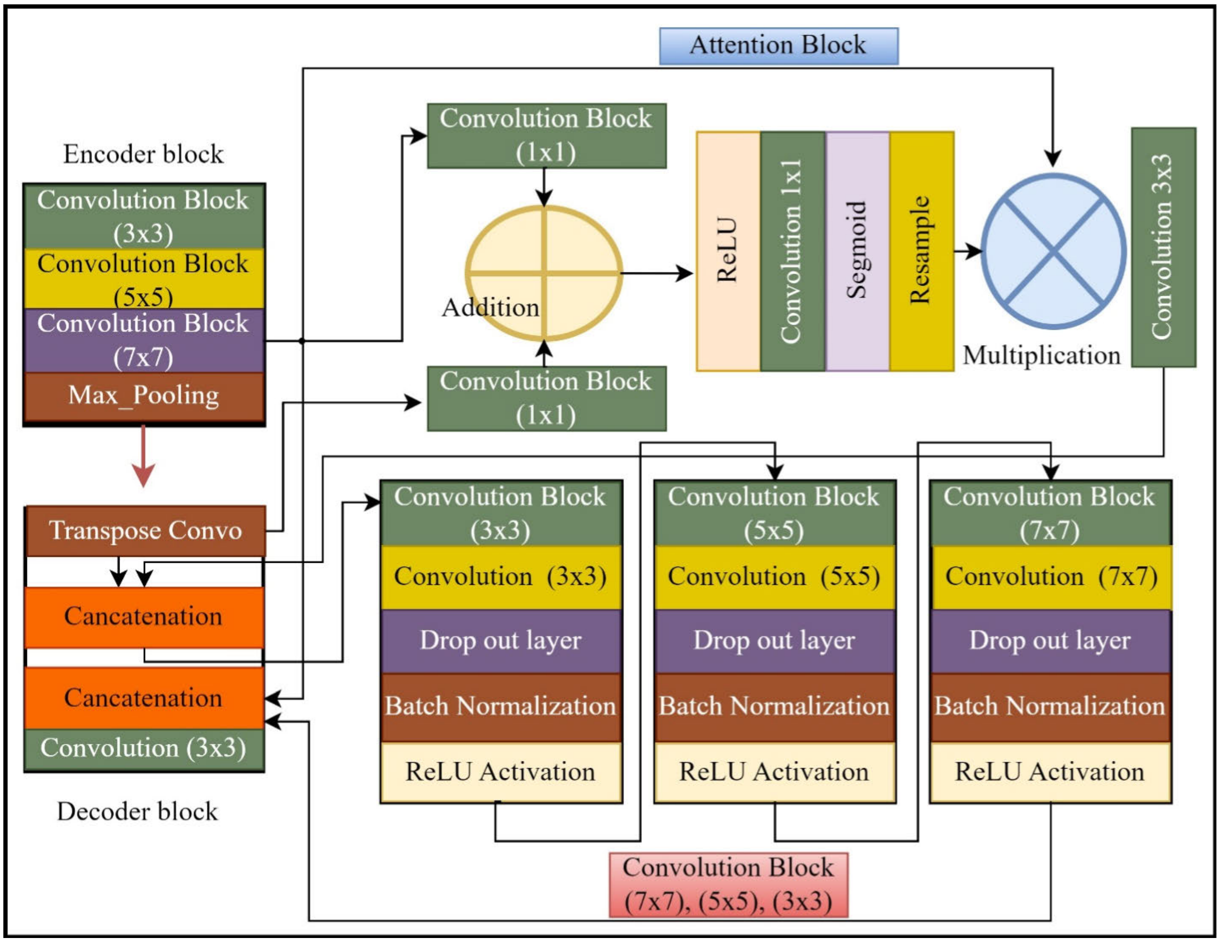

FIGURE 3. The internal structural view of residual attention-based double decoder network of proposed multi-scale cascaded convolution.

output of the concatenation step is processed through the final 3×3 convolution operation. The novelty of this proposed method is the residual attention-based double decoder network that is achieved by introducing the attention mechanism in the residual connections-based decoder. The learned features in the multiscale encoder network are transferred to the attention mechanism as well as the concatenation layer to achieve the residual function of the decoder for better segmentation performance. In the concatenation step the residual connections of multiscale cascaded convolution operation and attention mechanism are concatenated to reconstruct the BC segmented and localized output images. The results show that the proposed residual attention-based double decoder model achieved the best results on BUS imaging datasets.

C. DATASET

For the training and testing of multiscale cascaded convolution with a residual attention-based double decoder network, two imaging repositories are used. Further details about the utilized BUS imaging datasets are given below.

- **BUSI Dataset:** The first dataset that is utilized in this work is the BUSI dataset. This dataset contained BUS images with respective ground truth images and is openly accessible for research analysis. This dataset was contributed byAl-Dhabyani et al. [43]. The BUSI repository was initially composed at Baheya Hospital for research purposes and was annotated by an expert radiologist. This data set is formed up of a total of 780 BUS images from different six hundred female patients having age ranges from 25 years to 75 years. The average image size of this repository is 500×500.
- **UDIAT Dataset:** UDIAT dataset that was provided by Yap et al. [12] is openly accessible for research analysis and is the second BUS image dataset that is utilized in this work. This dataset is a small BC segmentation dataset that only contains 163 BUS PNG images and ground truths. The dataset was prepared in UDIAT diagnostic Centre and annotated by expert radiologists. The average image resolution of this repository is about 560×500. The sample dataset images from both utilized BUS image datasets with respective ground truth are presented in Fig. 4.

IV. RESULTS

The experimental outcomes of the proposed multiscale cascaded convolution with residual attention-based double

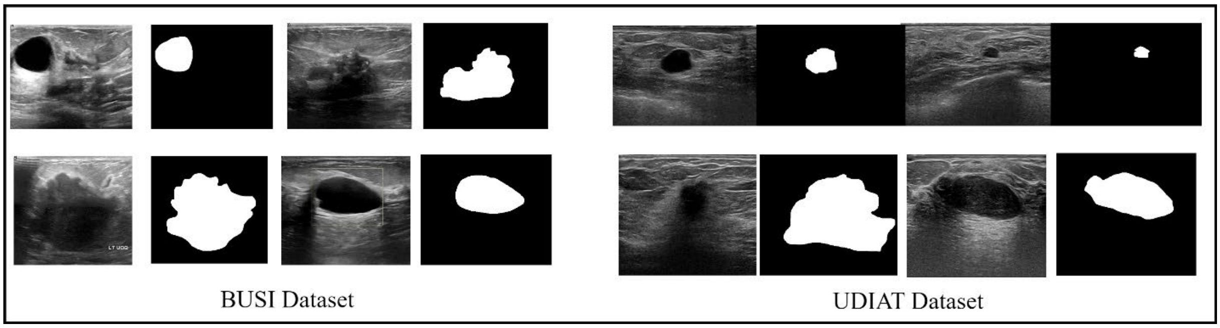

FIGURE 4. The sample BUS datasets images that are utilized to test the proposed BC segmentation model.

decoder network are given in this section. The results are calculated by using two BUS imaging datasets including BUSI and UDIAT. The datasets were partitioned into three parts such as training testing and validation sets by using 60, 20, and 20 percent of the data respectively. The implementation of the proposed multiscale cascaded convolution with residual attention-based double decoder network is accomplished by using the Dell Precision M4800 system which has 20 GB of RAM, and 2GB of dedicated NVidia graphic card Quadro K100M. The proposed segmentation model was implemented by using Python 3.6 and tensor flow. The hyperparameters for the training of the proposed model are set as, the initial learning rate in model training is 0.001 with Adam optimizer and momentum rate of 0.9 the learning rate is reduced with the patience of 6 epochs dynamically, the mini_batch size is set to 8, and the model was trained for 100 epochs with the early stopping condition.

A. EXPERIMENTAL OUTCOMES USING UDIAT DATASET

The experimental outcomes by utilizing the proposed multiscale cascaded convolution with residual attention double decoder network on the UDIAT dataset are presented in Table 1. These results are generated and presented by using four different configurations of the proposed multiscale cascaded convolution with residual attention double decoder network. In the first order of the experiment, to show the importance of each component of the proposed model the segmentation results are generated by implementing the fixed-sized receptive field convolution operation. A DSC of 82.35% was achieved on the UDIAT dataset by using the fixed-sized convolution method. A DSC of 85.38% was achieved by implementing the multiscale cascaded convolution operation in proposed methodology. Thirdly, the results were collected by using the residual attention mechanism and a DSC of 87.88% was recorded. Finally, in the last configuration, the results were generated by using the proposed multiscale cascaded convolution operation with residual attention-based double decoder, and the highest segmentation DSC of 91.38 % was achieved. The complete results including AC, DSC, JSC, Pre, and Rec on the first used BUS imaging repository are presented in Table 1.

To show the higher performance of the proposed multiscale cascaded convolution with residual attention-based double decoder network the visual outcomes on the UDIAT BUS image dataset are given in Figure. 5. The visual outcomes of the proposed multiscale cascaded convolution with residual attention-based double decoder network method are presented with the DSC score of each output image. The visual comparison of the proposed segmentation model is presented to show the difference between predicted segmented and ground truths.

B. EXPERIMENTAL OUTCOMES USING BUSI DATASET

The breast cancer localization outcomes of the proposed multiscale cascaded convolution with residual attention-based double decoder network using BUSI dataset are given in detail in Table 2. All BC segmentation results on this dataset were computed by using AC, DSC, JSC, Pre, and Rec evaluation metrics in percent. In this experiment, the highest DSC of 90.55 was recorded with the proposed multiscale cascaded convolution plus residual attention-based dual decoder. The lowest segmentation DSC was achieved with a fixed receptive field mechanism.

The visual segmentation results by using the multiscale cascaded convolution with residual attention-based dual decoder model on the BUSI ultrasounds are shown in Figure. 6. These comparisons are computed by randomly selecting the eighteen BUS images from the testing set and are presented to show the difference between predicted results from multiscale cascaded convolution with residual attention-based dual decoder network and ground truths. The DSC score of each tested BUS sample is also given with each resulting BUS image.

C. COMPERISON WITH EXISTING METHODS ON THE UDIAT DATASET

The segmentation results computed by using the multiscale cascaded convolution with residual attention-based dual decoder network method and existing BC segmentation approaches on the UDIAT repository are given in Table 3. The segmentation results of the proposed multiscale cascaded convolution with residual attention-based dual decoder

TABLE 1. The segmentation outcomes of multi-scale cascaded convolution with residual attention-based double decoder network using the UDIAT dataset.

Methods	Ac(%)	DSC(%)	JSC(%)	Pre(%)	Rec(%)
Proposed model (Fixed receptive field)	91.11	82.35	70.80	91.69	81.15
Proposed model (Multiscale cascaded convolution)	93.11	85.38	75.48	93.89	82.06
Proposed model (residual attention)	95.11	87.88	79.11	95.62	84.14
Proposed model (multiscale cascaded convolution + residual attention)	98.67	**91.38**	81.67	94.43	87.76

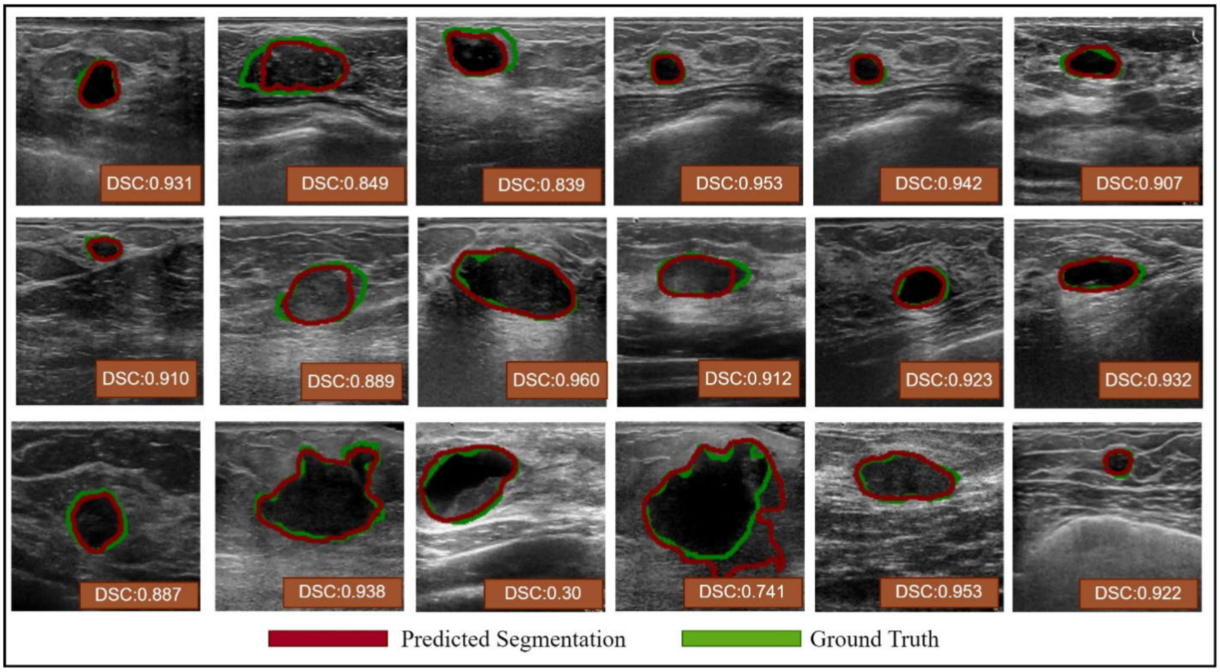

FIGURE 5. The visual segmentation outcome comparison of randomly selected BUS images with ground truths using the proposed multiscale cascaded convolution with residual attention-based dual decoder network on the UDIAT dataset, DSC score of each tested image is present with each image.

network are compared by implementing the well-known existing deep learning-based segmentation methods. The U-Net by Ronneberger et al. [42] produced a dice coefficient of 77.98% while U-Net++ which is an extended version of U-Net and was presented by Zhou et al. [44] produced a dice of 73.64%. A segmentation Dice of 83.36% is achieved with DeepLabv3+ which was introduced by Chen et al. [45], and a dice score of 75.76% is produced by PSP-Net [46]. The MSU-Net which was developed in [47] produced a dice score of 83.01%. The comparative analysis of the multiscale cascaded convolution and residual attention-based double decoder network with existing studies on reported results without implementation including [48], [49], [50], [51], [52], [53] are also given in Table 3.

The comparative analysis represented that the proposed multiscale cascaded convolution with residual attention-based dual decoder produced the best DSC of 91.38 % on the first utilized BUS imaging dataset. The graphical representation of the comparison of the multiscale cascaded convolution with residual attention-based double decoder network is given in Fig. 7. The output images-based comparison of the proposed segmentation method with five existing implemented models is given in Fig. 8. This comparison was accomplished by using four breast samples taken from the UDIAT dataset which showed that the proposed multiscale cascaded convolution with residual attention-based dual decoding strategy achieved the best DSC score.

TABLE 2. The proposed multi-scale cascaded convolution with residual attention-based double decoder network segmentation results using the BUSI dataset.

Methods	Ac(%)	DSC(%)	JSC(%)	Pre(%)	Rec(%)
Proposed model (Fixed receptive field)	91.13	81.11	77.12	90.55	80.44
Proposed model (Multiscale cascaded convolution)	92.45	84.12	74.13	92.15	81.78
Proposed model (residual attention)	94.34	86.44	78.13	94.13	83.18
Proposed model (multiscale cascaded convolution + residual attention)	97.69	**90.55**	80.87	93.53	88.46

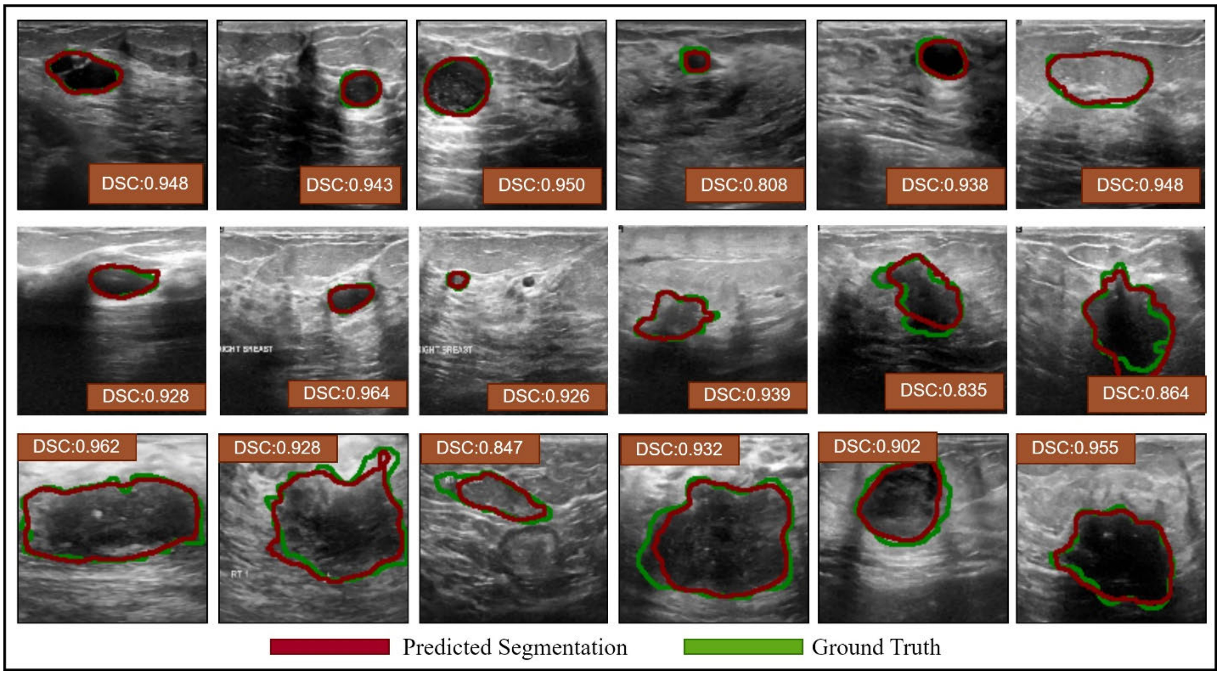

FIGURE 6. The visual segmentation outcome comparison of randomly selected BUS images with ground truths using the proposed multiscale cascaded convolution with residual attention-based dual decoder network on the BUSI dataset.

D. COMPERISON WITH EXISTING METHODS ON THE BUSI DATASET

The segmentation results generated by using the proposed multiscale cascaded convolution with residual attention-based dual decoder network method and existing methods on the BUSI repository are given in Table 4. The segmentation results of the proposed multiscale cascaded convolution with residual attention-based dual decoder network are compared by implementing the well-known existing deep learning-based segmentation methods. The U-Net which was introduced by Ronneberger et al. [42] produced a dice coefficient of 73.90% with a precision score of 86.12% while U-Net++ which is an extended version of U-Net and was presented by Zhou et al. [44] produced a dice of 72.15%. A segmentation Dice of 80.12% is achieved with DeepLabv3+ which was introduced by Chen et al. [45], and a dice score of 74.60% and JSC of62.25% is produced by PSP-Net [46]. The MSU-Net which was developed in study [47] produced a dice score of 82.09%. The comparative analysis of the multiscale cascaded convolution and residual attention-based double decoder network with existing studies on reported results without implementation including [48], [49], [51], [52], [53] also given in Table 4. From these segmentation results, it can be concluded that the highest DSC of 90.55% was achieved by the proposed multiscale cascaded convolution with residual attention-based double decoder network. The graphical comparison in the form of line graphs of the multiscale cascaded convolution with residual attention-based double decoder network is given in Fig. 9.

TABLE 3. Comparative analysis of the multi-scale cascaded convolution and residual attention-based double decoder network for BC segmentation with existing models on the UDIAT dataset.

Methods	Year	DSC(%)	JSC(%)	Pre(%)	Rec(%)
U-Net++ [44]	2018	73.64	63.23	85.90	72.86
DeepLabv3+ [45]	2018	83.36	73.06	85.53	85.37
ESTAN [49]	2020	80.00	72.00	-	85.00
Yang et al. [51]	2022	82.00	74.00	-	84.00
STAN [50]	2020	78.20	65.90	-	80.10
Zhang et al.[53]	2023	86.85	78.46	86.04	82.51
Zhang et al. [52]	2023	88.73	81.22	88.68	-
U-Net [42]	2015	77.98	67.66	87.17	76.10
Ma et al. [48]	2023	86.73	75.74	-	88.34
MSU-Net [47]	2021	83.01	71.13	83.11	75.15
PSP-Net [46]	2017	75.76	63.24	82.24	74.11
Proposed method		**91.38**	81.67	94.43	87.76

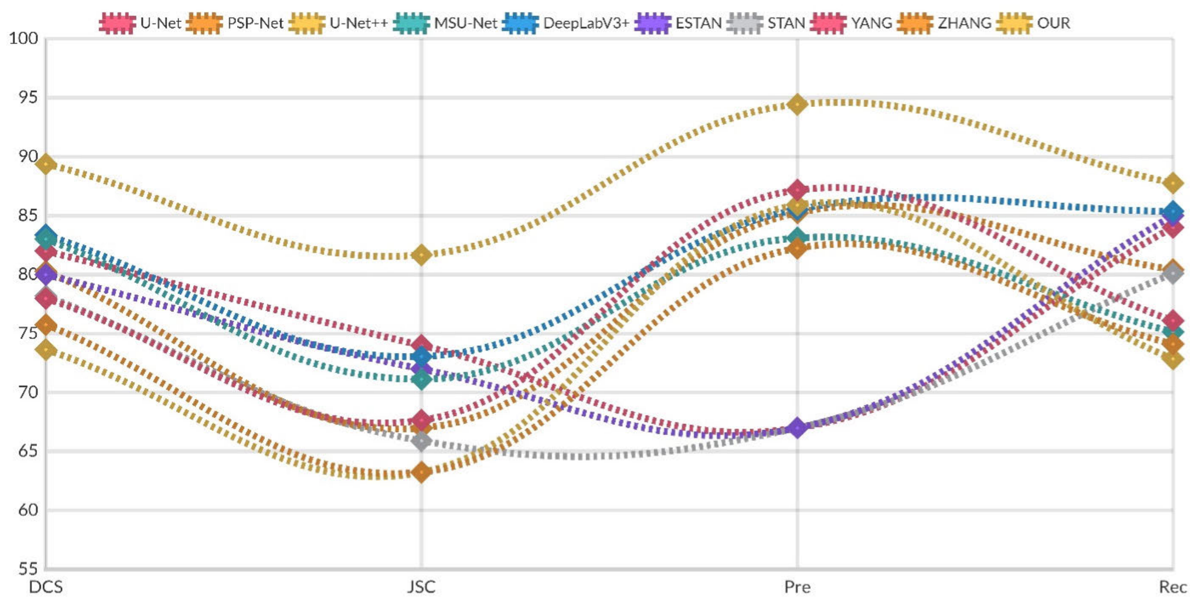

FIGURE 7. The graphical representation of the multiscale cascaded convolution with residual attention-based double decoder network using the UDIAT dataset.

The output images-based comparison of the proposed multiscale cascaded convolution and residual attention-based dual decoder network with five implemented models using the BUSI repository is given in Fig. 10. This comparison

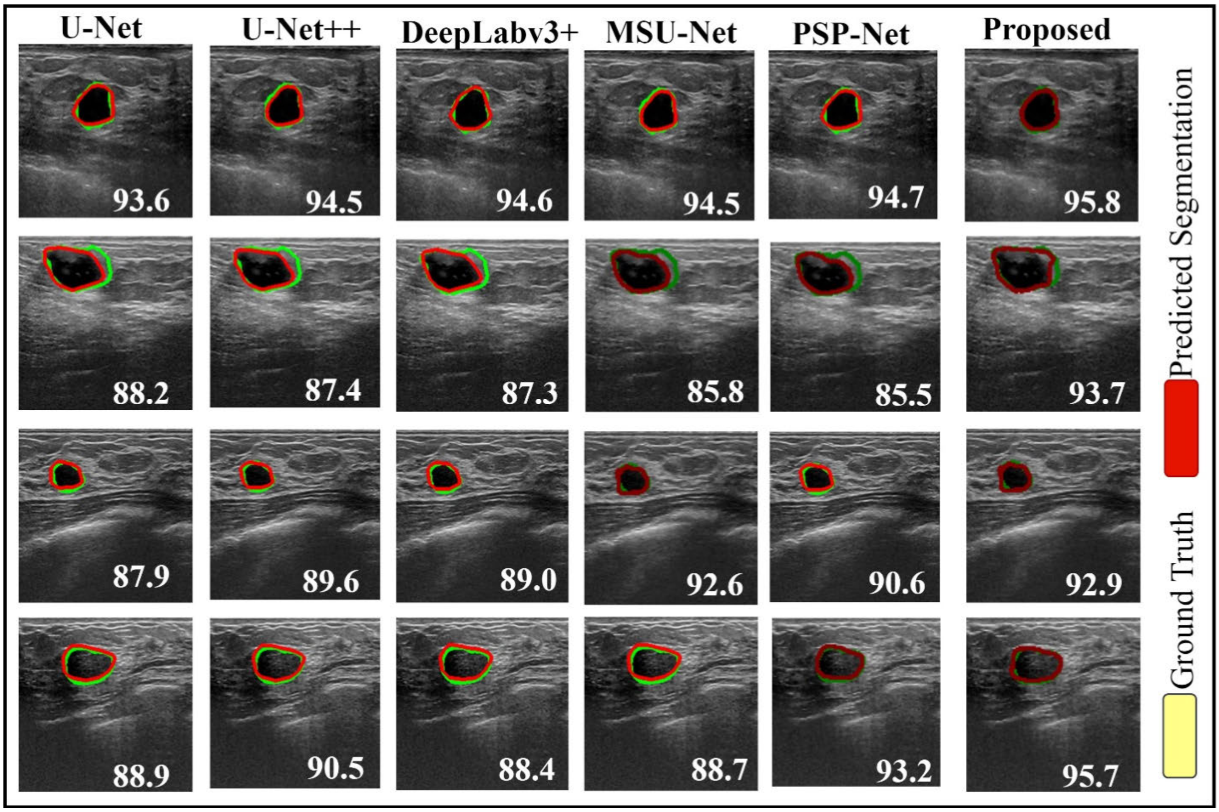

FIGURE 8. The visual segmentation outcome comparison of randomly selected BUS with existing studies using the proposed multiscale cascaded convolution with residual attention-based dual decoder network on the UDIAT dataset, DSC score of each tested image is present with each image.

TABLE 4. The segmentation results comparison of the proposed multi-scale cascaded convolution with residual attention-based double decoder network for BC segmentation with existing studies using the BUSI dataset.

Methods	Year	DSC(%)	JSC(%)	Pre(%)	Rec(%)
MSU-Net [47]	2021	82.09	70.32	82.80	74.13
Yang et al. [51]	2022	79.00	70.00	-	82.00
DeepLabv3+ [45]	2018	80.12	69.55	84.13	82.30
ESTAN [49]	2020	81.00	72.00	-	89.00
U-Net++ [44]	2018	72.15	62.13	83.55	71.32
Zhang et al.[53]	2023	79.54	70.33	82.75	82.51
Zhang et al. [52]	2023	83.11	75.26	-	86.08
U-Net [42]	2015	73.90	66.01	86.12	74.40
Ma et al. [48]	2023	82.64	69.73	-	82.78
PSP-Net [46]	2017	74.66	62.25	81.31	73.44
Proposed method		**90.55**	80.87	93.53	88.46

was performed by using the four breast samples from the BUSI dataset which showed that the proposed multiscale cascaded convolution with residual attention-based double decoder achieved the best DSC scores. The DSC scores of

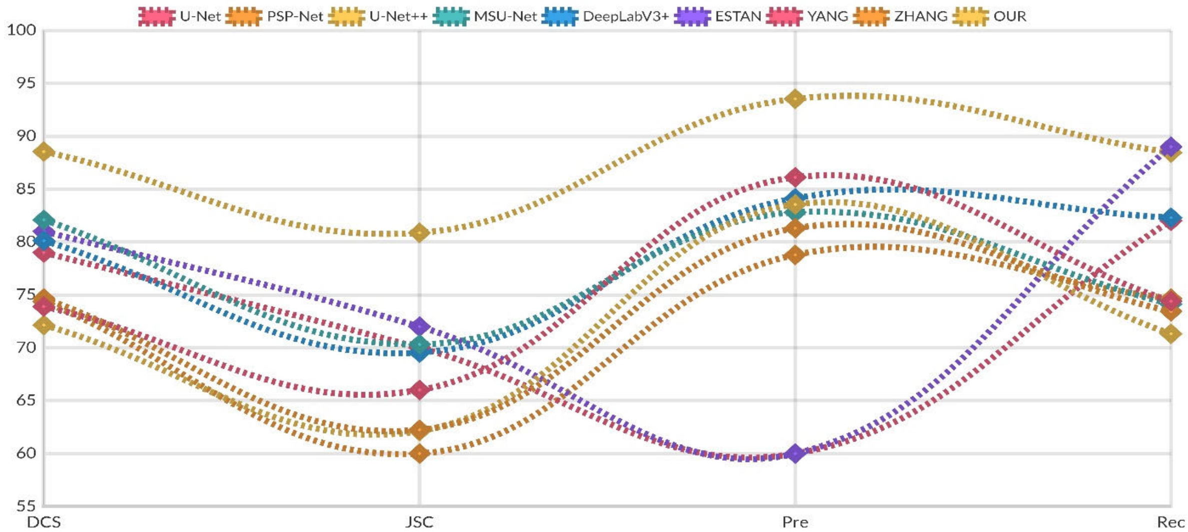

FIGURE 9. The comparison line graphs of proposed multiscale cascaded convolution with residual attention-based double decoder network using BUSI dataset.

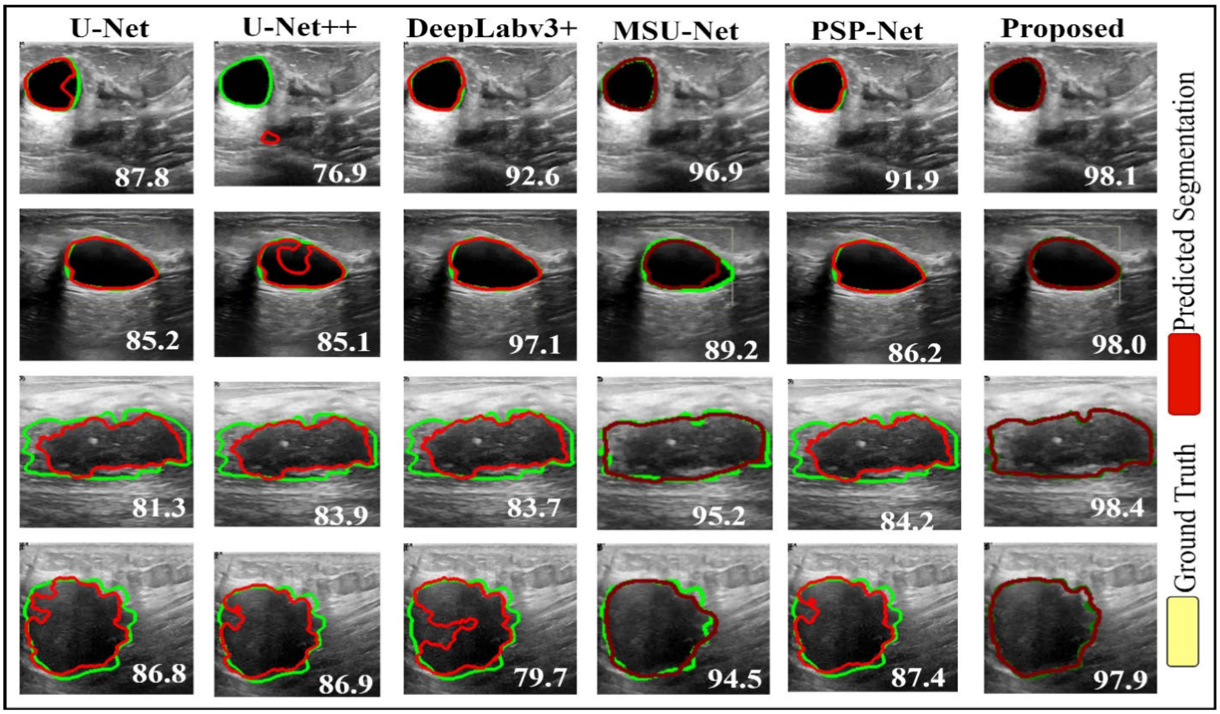

FIGURE 10. The visual segmentation outcome comparison of randomly selected BUS with previous studies using the proposed multiscale cascaded convolution with residual attention-based dual decoder on the BUSI dataset, DSC score of each tested image is present with each image.

each tested BUS image with implemented existing models are also shown with each tested image. The comparison shows that the proposed multiscale feature learning mechanism with residual attention method significantly enhanced the segmentation performance that was not used in previous studies.

Furthermore, the comparison highlighted that multi-scale feature learning method with dual decoding strategy outperformed the existing methods due to diverse feature learning.

V. DISCUSSION

Segmentation of the BC by using BUS images is a complicated task due to the thermal noise of the ultrasound instrument. Different studies have been proposed that utilize deep learning for the automation of the manual segmentation process. Most of the earlier segmentation methods tried to automate the segmentation process by implementing the fixed receptive field-based convolution operations. The fixed receptive field convolution methods are unable to capture the high-level spatial image features. Some of the previous methods applied the attention mechanism to enhance breast cancer localization performance by giving more weightage to the tumor region. This work proposed a multiscale cascaded convolution with a residual attention-based double decoding mechanism for breast cancer segmentation from the BUS dataset. To overcome the problem of a fixed receptive field a multiscale cascaded convolution block is introduced in the proposed method. The results of multiscale cascaded convolution with residual attention-based dual decoder network method show that multiscale cascaded convolution significantly increases the segmentation performance. Secondly, a residual attention-based double decoder network was implemented in this work. The multiscale cascaded convolution with residual attention-based double decoder network produced the best DSC of 91.38% using UDIAT dataset and DSC of 90.55% using BUSI repository which shows that our method can effectively be used for the BC segmentation task.

VI. CONCLUSION

This work proposed a U-shaped multiscale cascaded convolution with residual attention-based double decoder network to accurately segment the input BUS image datasets. The proposed method comprised four encoder blocks, one bridge block, and four decoder blocks. The encoder blocks were implemented with the multiscale cascaded convolution procedure. The output of the encoder block in the form of learning image features is transformed into decoder blocks by using the bridge and skip connection. For the segmentation of BUS images, a residual attention process is developed in decoder blocks. The residual attention with a double decoding mechanism is implemented to capture the high-level diverse spatial features from the tumor regions. For experiments, two publicly available BUS image datasets were utilized. The proposed U-shaped multiscale cascaded convolution with residual attention-based double decoder network for BC segmentation achieved the best segmentation results. A DSC of 91.38% was attained using UDIAT dataset while a DSC of 90.55% was reported on the BUSI dataset. This work will further be enhanced in the future to apply the GAN network in contraction and expansion paths to improve the segmentation accuracy.

REFERENCES

[1] H. Sung, J. Ferlay, R. L. Siegel, M. Laversanne, I. Soerjomataram, A. Jemal, and F. Bray, "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clinicians*, vol. 71, no. 3, pp. 209–249, May 2021.

[2] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA, Cancer J. Clinicians*, vol. 68, no. 6, pp. 394–424, Nov. 2018, doi: 10.3322/caac.21492.

[3] T. G. Debelee, F. Schwenker, A. Ibenthal, and D. Yohannes, "Survey of deep learning in breast cancer image analysis," *Evolving Syst.*, vol. 11, no. 1, pp. 143–163, Mar. 2020.

[4] S. A. Alanazi, M. M. Kamruzzaman, M. N. Islam Sarker, M. Alruwaili, Y. Alhwaiti, N. Alshammari, and M. H. Siddiqi, "Boosting breast cancer detection using convolutional neural network," *J. Healthcare Eng.*, vol. 2021, pp. 1–11, Apr. 2021, doi: 10.1155/2021/5528622.

[5] J. M. Yeh, K. P. Lowry, C. B. Schechter, L. R. Diller, G. O'Brien, O. Alagoz, G. T. Armstrong, J. M. Hampton, M. M. Hudson, W. Leisenring, Q. Liu, J. S. Mandelblatt, D. L. Miglioretti, C. S. Moskowitz, P. C. Nathan, J. P. Neglia, K. C. Oeffinger, A. Trentham-Dietz, and N. K. Stout, "Breast cancer screening among childhood cancer survivors treated without chest radiation: Clinical benefits and cost-effectiveness," *JNCI, J. Nat. Cancer Inst.*, vol. 114, no. 2, pp. 235–244, Feb. 2022, doi: 10.1093/jnci/djab149.

[6] S. N. Sabtu, S. F. A. Sani, D. A. Bradley, L. M. Looi, and Z. Osman, "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.*, vol. 51, no. 3, pp. 380–389, Mar. 2020.

[7] J. Geisel, M. Raghu, and R. Hooley, "The role of ultrasound in breast cancer screening: The case for and against ultrasound," *Seminars Ultrasound, CT MRI*, vol. 39, no. 1, pp. 25–34, Feb. 2018, doi: 10.1053/j.sult.2017.09.006.

[8] M. J. Umer, J. Amin, M. Sharif, M. A. Anjum, F. Azam, and J. H. Shah, "An integrated framework for COVID-19 classification based on classical and quantum transfer learning from a chest radiograph," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 20, p. e6434, Sep. 2022, doi: 10.1002/cpe.6434.

[9] M. J. Umer, M. Sharif, S. Kadry, and A. Alharbi, "Multi-class classification of breast cancer using 6B-net with deep feature fusion and selection method," *J. Personalized Med.*, vol. 12, no. 5, p. 683, Apr. 2022, doi: 10.3390/jpm12050683.

[10] A. Iqbal and M. Sharif, "MDA-Net: Multiscale dual attention-based network for breast lesion segmentation using ultrasound images," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 34, no. 9, pp. 7283–7299, Oct. 2022, doi: 10.1016/j.jksuci.2021.10.002.

[11] T. Meraj, W. Alosaimi, B. Alouffi, H. T. Rauf, S. A. Kumar, R. Damaševičius, and H. Alyami, "A quantization assisted U-Net study with ICA and deep features fusion for breast cancer identification using ultrasonic data," *PeerJ Comput. Sci.*, vol. 7, p. e805, Dec. 2021.

[12] M. H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwiggelaar, A. K. Davison, and R. Martí, "Automated breast ultrasound lesions detection using convolutional neural networks," *IEEE J. Biomed. Health Informat.*, vol. 22, no. 4, pp. 1218–1226, Jul. 2018, doi: 10.1109/JBHI.2017.2731873.

[13] V. K. Singh, H. A. Rashwan, S. Romani, F. Akram, N. Pandey, M. M. K. Sarker, A. Saleh, M. Arenas, M. Arquez, D. Puig, and J. Torrents-Barrena, "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network," *Expert Syst. Appl.*, vol. 139, Jan. 2020, Art. no. 112855, doi: 10.1016/j.eswa.2019.112855.

[14] M. J. Umer and M. I. Sharif, "A comprehensive survey on quantum machine learning and possible applications," *Int. J. E-Health Med. Commun.*, vol. 13, no. 5, pp. 1–17, Oct. 2022, doi: 10.4018/ijehmc.315730.

[15] M. J. Umer, M. Sharif, M. Raza, and S. Kadry, "A deep feature fusion and selection-based retinal eye disease detection from OCT images," *Expert Syst.*, vol. 40, no. 6, Jul. 2023, doi: 10.1111/exsy.13232.

[16] E. Michael, H. Ma, H. Li, and S. Qi, "An optimized framework for breast cancer classification using machine learning," *BioMed Res. Int.*, vol. 2022, pp. 1–18, Feb. 2022, doi: 10.1155/2022/8482022.

[17] K. Jabeen, M. A. Khan, M. Alhaisoni, U. Tariq, Y.-D. Zhang, A. Hamza, A. Mickus, and R. Damaševičius, "Breast cancer classification from ultrasound images using probability-based optimal deep learning feature fusion," *Sensors*, vol. 22, no. 3, p. 807, Jan. 2022.

[18] M. J. Umer, M. Sharif, M. Alhaisoni, U. Tariq, Y. J. Kim, and B. Chang, "A framework of deep learning and selection-based breast cancer detection from histopathology images," *Comput. Syst. Sci. Eng.*, vol. 45, no. 2, pp. 1001–1016, 2023, doi: 10.32604/csse.2023.030463.

[19] M. J. Umer, M. Sharif, and S. Wang, "Breast cancer classification and segmentation framework using multiscale CNN and U-shaped dual decoded attention network," *Expert Syst.*, Nov. 2022, doi: 10.1111/exsy.13192.

[20] A. Ahila, M. Poongodi, S. Bourouis, S. S. Band, A. Mosavi, S. Agrawal, and M. Hamdi, "Meta-heuristic algorithm-tuned neural network for breast cancer diagnosis using ultrasound images," *Frontiers Oncol.*, vol. 12, Jun. 2022, Art. no. 834028, doi: 10.3389/fonc.2022.834028.

[21] N. Vigil, M. Barry, A. Amini, M. Akhloufi, X. P. V. Maldague, L. Ma, L. Ren, and B. Yousefi, "Dual-intended deep learning model for breast cancer diagnosis in ultrasound imaging," *Cancers*, vol. 14, no. 11, p. 2663, May 2022, doi: 10.3390/cancers14112663.

[22] Y. Yan, Y. Liu, Y. Wu, H. Zhang, Y. Zhang, and L. Meng, "Accurate segmentation of breast tumors using AE U-Net with HDC model in ultrasound images," *Biomed. Signal Process. Control*, vol. 72, Feb. 2022, Art. no. 103299, doi: 10.1016/j.bspc.2021.103299.

[23] I. B. Ahmed, W. Ouarda, and C. Ben Amar, "Hybrid UNET model segmentation for an early breast cancer detection using ultrasound images," in *Computational Collective Intelligence* (Lecture Notes in Computer Science), N. T. Nguyen, Y. Manolopoulos, R. Chbeir, A. Kozierkiewicz, and B. Trawiński, Eds., Cham, Switzerland: Springer, 2022, pp. 464–476, doi: 10.1007/978-3-031-16014-1_37.

[24] G. Zhang, K. Zhao, Y. Hong, X. Qiu, K. Zhang, and B. Wei, "SHA-MTL: Soft and hard attention multi-task learning for automated breast cancer ultrasound image segmentation and classification," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 16, no. 10, pp. 1719–1725, Oct. 2021, doi: 10.1007/s11548-021-02445-7.

[25] M. S. K. Inan, F. I. Alam, and R. Hasan, "Deep integrated pipeline of segmentation guided classification of breast cancer from ultrasound images," *Biomed. Signal Process. Control*, vol. 75, May 2022, Art. no. 103553, doi: 10.1016/j.bspc.2022.103553.

[26] Y. Lei, X. He, J. Yao, T. Wang, L. Wang, W. Li, W. J. Curran, T. Liu, D. Xu, and X. Yang, "Breast tumor segmentation in 3D automatic breast ultrasound using mask scoring R-CNN," *Med. Phys.*, vol. 48, no. 1, pp. 204–214, Jan. 2021, doi: 10.1002/mp.14569.

[27] A. E. Ilesanmi, U. Chaumrattanakul, and S. S. Makhanov, "A method for segmentation of tumors in breast ultrasound images using the variant enhanced deep learning," *Biocybernetics Biomed. Eng.*, vol. 41, no. 2, pp. 802–818, Apr. 2021, doi: 10.1016/j.bbe.2021.05.007.

[28] K. Huang, Y. Zhang, H. D. Cheng, P. Xing, and B. Zhang, "Semantic segmentation of breast ultrasound image with fuzzy deep learning network and breast anatomy constraints," *Neurocomputing*, vol. 450, pp. 319–335, Aug. 2021, doi: 10.1016/j.neucom.2021.04.012.

[29] J. Kim, H. J. Kim, C. Kim, J. H. Lee, K. W. Kim, Y. M. Park, H. W. Kim, S. Y. Ki, Y. M. Kim, and W. H. Kim, "Weakly-supervised deep learning for ultrasound diagnosis of breast cancer," *Sci. Rep.*, vol. 11, no. 1, Dec. 2021, Art. no. 1, doi: 10.1038/s41598-021-03806-7.

[30] Y. Tong, Y. Liu, M. Zhao, L. Meng, and J. Zhang, "Improved U-Net MALF model for lesion segmentation in breast ultrasound images," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102721, doi: 10.1016/j.bspc.2021.102721.

[31] Y. Luo, Q. Huang, and X. Li, "Segmentation information with attention integration for classification of breast tumor in ultrasound image," *Pattern Recognit.*, vol. 124, Apr. 2022, Art. no. 108427, doi: 10.1016/j.patcog.2021.108427.

[32] L. Zhang, Z. Jia, X. Leng, and F. Ma, "Artificial intelligence algorithm-based ultrasound image segmentation technology in the diagnosis of breast cancer axillary lymph node metastasis," *J. Healthcare Eng.*, vol. 2021, Jul. 2021, Art. no. e8830260, doi: 10.1155/2021/8830260.

[33] C. Xue, L. Zhu, H. Fu, X. Hu, X. Li, H. Zhang, and P.-A. Heng, "Global guidance network for breast lesion segmentation in ultrasound images," *Med. Image Anal.*, vol. 70, May 2021, Art. no. 101989, doi: 10.1016/j.media.2021.101989.

[34] M. Ragab, A. Albukhari, J. Alyami, and R. F. Mansour, "Ensemble deep-learning-enabled clinical decision support system for breast cancer diagnosis and classification on ultrasound images," *Biology*, vol. 11, no. 3, p. 439, Mar. 2022, doi: 10.3390/biology11030439.

[35] K. Yu, S. Chen, and Y. Chen, "Tumor segmentation in breast ultrasound image by means of res path combined with dense connection neural network," *Diagnostics*, vol. 11, no. 9, p. 1565, Aug. 2021, doi: 10.3390/diagnostics11091565.

[36] N. S. Punn and S. Agarwal, "RCA-IUnet: A residual cross-spatial attention-guided inception U-Net model for tumor segmentation in breast ultrasound imaging," *Mach. Vis. Appl.*, vol. 33, no. 2, p. 27, Feb. 2022, doi: 10.1007/s00138-022-01280-3.

[37] K. Wang, S. Liang, and Y. Zhang, "Residual feedback network for breast lesion segmentation in ultrasound image," in *Medical Image Computing and Computer Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), M. de Bruijne, P. C. Cattin, S. Cotin, N. Padoy, S. Speidel, Y. Zheng, and C. Essert, Eds., Cham, Switzerland: Springer, 2021, pp. 471–481, doi: 10.1007/978-3-030-87193-2_45.

[38] G. Chen, Y. Dai, and J. Zhang, "C-Net: Cascaded convolutional neural network with global guidance and refinement residuals for breast ultrasound images segmentation," *Comput. Methods Programs Biomed.*, vol. 225, Oct. 2022, Art. no. 107086, doi: 10.1016/j.cmpb.2022.107086.

[39] D. Zhai, B. Hu, X. Gong, H. Zou, and J. Luo, "ASS-GAN: Asymmetric semi-supervised GAN for breast ultrasound image segmentation," *Neurocomputing*, vol. 493, pp. 204–216, Jul. 2022, doi: 10.1016/j.neucom.2022.04.021.

[40] P. Pan, H. Chen, Y. Li, N. Cai, L. Cheng, and S. Wang, "Tumor segmentation in automated whole breast ultrasound using bidirectional LSTM neural network and attention mechanism," *Ultrasonics*, vol. 110, Feb. 2021, Art. no. 106271, doi: 10.1016/j.ultras.2020.106271.

[41] N. Karunanayake, W. Lohitvisate, and S. S. Makhanov, "Artificial life for segmentation of fusion ultrasound images of breast abnormalities," *Pattern Recognit.*, vol. 131, Nov. 2022, Art. no. 108838, doi: 10.1016/j.patcog.2022.108838.

[42] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention—MICCAI* (Lecture Notes in Computer Science), N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, Eds., Cham, Switzerland: Springer, 2015, pp. 234–241, doi: 10.1007/978-3-319-24574-4_28.

[43] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset of breast ultrasound images," *Data Brief*, vol. 28, Feb. 2020, Art. no. 104863, doi: 10.1016/j.dib.2019.104863.

[44] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, "UNet++: A nested U-Net architecture for medical image segmentation," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support* (Lecture Notes in Computer Science), D. Stoyanov, Z. Taylor, G. Carneiro, T. Syeda-Mahmood, A. Martel, L. Maier-Hein, J. M. R. S. Tavares, A. Bradley, J. P. Papa, V. Belagiannis, J. C. Nascimento, Z. Lu, S. Conjeti, M. Moradi, H. Greenspan, and A. Madabhushi, Eds., Cham, Switzerland: Springer, 2018, pp. 3–11, doi: 10.1007/978-3-030-00889-5_1.

[45] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, and H. Adam, "Encoder–decoder with atrous separable convolution for semantic image segmentation," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2018, pp. 801–818.

[46] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid scene parsing network," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6230–6239.

[47] R. Su, D. Zhang, J. Liu, and C. Cheng, "MSU-Net: Multi-scale U-Net for 2D medical image segmentation," *Frontiers Genet.*, vol. 12, Feb. 2021, Art. no. 639930.

[48] Z. Ma, Y. Qi, C. Xu, W. Zhao, M. Lou, Y. Wang, and Y. Ma, "ATFE-Net: Axial transformer and feature enhancement-based CNN for ultrasound breast mass segmentation," *Comput. Biol. Med.*, vol. 153, Feb. 2023, Art. no. 106533, doi: 10.1016/j.compbiomed.2022.106533.

[49] B. Shareef, A. Vakanski, M. Xian, and P. E. Freer, "ESTAN: Enhanced small tumor-aware network for breast ultrasound image segmentation," 2020, *arXiv:2009.12894*.

[50] B. Shareef, M. Xian, and A. Vakanski, "Stan: Small tumor-aware network for breast ultrasound image segmentation," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1–5.

[51] K. Yang, A. Suzuki, J. Ye, H. Nosato, A. Izumori, and H. Sakanashi, ''CTG-Net: Cross-task guided network for breast ultrasound diagnosis,'' *PLoS ONE*, vol. 17, no. 8, Aug. 2022, Art. no. e0271106.

[52] H. Zhang, J. Lian, Z. Yi, R. Wu, X. Lu, P. Ma, and Y. Ma, ''HAU-Net: Hybrid CNN-transformer for breast ultrasound image segmentation,'' *Biomed. Signal Process. Control*, vol. 87, Jan. 2024, Art. no. 105427, doi: 10.1016/j.bspc.2023.105427.

[53] M. Zhang, A. Huang, D. Yang, and R. Xu, ''Boundary-oriented network for automatic breast tumor segmentation in ultrasound images,'' *Ultrason. Imag.*, vol. 45, no. 2, pp. 62–73, Mar. 2023, doi: 10.1177/01617346231162925.

MUHAMMAD JUNAID UMER received the B.S. degree in CS from the Federal Urdu University of Art, Science, and Technology, Islamabad, Pakistan, in 2016, the M.S. degree in CS from the National University of Science and Technology (NUST), SEECS Campus, Islamabad, in 2019, and the Ph.D. degree from COMSATS University Islamabad, Wah Campus, Pakistan, in 2023. His research interests include image processing, machine learning, and deep learning, medical image processing, breast cancer detection, pattern recognition, and computer vision.

MUHAMMAD IRFAN SHARIF received the Ph.D. degree from the University of Electronic Science and Technology of China, Chengdu, China. He is currently working as an Assistant Professor with the Department of Information Sciences, University of Education Lahore, Jauharabad Campus, Jauharabad, Pakistan. His research interests include medical imaging, machine learning, computer vision, artificial intelligence, and pattern recognition.

JUNGEUN KIM received the Ph.D. degree in knowledge service engineering from Korea Advanced Institute of Science and Technology (KAIST). He is currently an Associate Professor with Kongju National University (KNU). Before joining KNU, he was a Senior Researcher at the Artificial Intelligence Research Laboratory, Electronics and Telecommunications Research Institute (ETRI). His research interests include data mining, artificial intelligence, big data analysis with distributed processing platforms, and open data platforms.