Biomedical Signal Processing and Control 91 (2024) 105969

Contents lists available at ScienceDirect

Biomedical Signal Processing and Control

journal homepage: www.elsevier.com/locate/bspc

AAPFC-BUSnet: Hierarchical encoder–decoder based CNN with attention aggregation pyramid feature clustering for breast ultrasound image lesion segmentation

Sushma B. [a,*], Aparna Pulikala [b]

[a] Department of Electronics and Communication Engineering, CMR Institute of Technology, Bengaluru 560037, Karnataka, India
[b] Image Processing and Analysis Lab (iPAL), Department of Electronics and Communication Engineering, National Institute of Technology Karnataka-Surathkal, Mangalore 575025, Karnataka, India

ARTICLE INFO

Keywords:
Breast tumor
Convolutional neural network
Deep learning
Pyramid features
Semantic segmentation
Self attention mechanism
Ultrasound images

ABSTRACT

Breast cancer causes a serious menace to women's health and lives, underscoring the urgency of accurate tumor detection. Detecting both cancerous and non-cancerous breast tumors has become increasingly crucial, with ultrasound imaging emerging as a widely adopted modality for this purpose. However, identifying breast lesions in ultrasound images is a challenging task due to various tumor morphologies, geometry, similar color intensity distributions, and fuzzy boundaries, particularly irregularly shaped malignant tumors. This work proposes an encoder–decoder based U-shaped convolutional neural network (CNN) variant with an attention aggregation-based pyramid feature clustering module (AAPFC) to detect breast lesion regions. The network consists of the U-Net variant as a base network and AAPFC to fuse features extracted at the various levels of the base U-Net using a suitable feature fusion technique. Furthermore, the deformable convolution with adaptive self-attention mechanism is introduced to decode the pyramid features parallel to capture the various geometric features at multi-stages. Two public breast lesion ultrasound datasets consisting 263 malignant, 547 benign and 133 normal images are considered to evaluate the performance of the proposed model and state-of-the-art deep CNN-based segmentation models. The proposed model provides 96% accuracy, 68% Mean-IoU, 97% specificity, 82% sensitivity and 0.747 kappa score respectively. The conducted qualitative and quantitative performance analysis experiments show that the proposed model performs better in breast lesion segmentation on ultrasound images.

1. Introduction

Breast cancer is a common disease and a primary cause of early death in women [1]. In order to increase the survival rate, regular breast scanning and diagnosis are essential. The standard tests and procedures to investigate breast cancer involve a manual breast assessment, mammogram and breast ultrasound [2]. Manual breast examination involves doctors examining the breasts and lymph nodes beneath the armpit for lumps and other abnormalities. However, this method is not proven effective in detecting cancer or improving survival rates for women with breast cancer [3]. A mammogram is an X-ray of the breast, and it is tough to detect breast cancers due to breast tissue density variation in the X-ray image [4]. Ultrasound uses sound waves to image the breast, which produces more accurate results than mammography in individuals with dense breasts. Ultrasound imaging has become a standard modality and important for breast tumor screening because of its adaptability and easy usage [5]. Identifying breast lesions in Breast Ultrasound (BUS) images can be an arduous task, even for an experienced radiologist. This difficulty arises from the diverse geometries of tumors, the presence of indistinct boundaries, and variations in intensity, especially in the case of malignant tumors with irregular shapes [6]. Furthermore, the scarcity of highly skilled radiologists poses a significant challenge, particularly in economically disadvantaged and resource-constrained nations. Therefore, developing an accurate and entirely automatic computer aided breast lesion segmentation technique is vital and necessary. Segmenting BUS images is essential in diagnosing and locating the breast tumor using computer-aided diagnosis [7].

Many deep-learning techniques have been presented recently to segment breast lesions from ultrasound images. However, the segmentation accuracy of the lesions is deeply affected by intricate ultrasound

* Corresponding author.
E-mail address: sushma.b@cmrit.ac.in (Sushma B.).
URL: https://www.cmrit.ac.in (Sushma B.).

https://doi.org/10.1016/j.bspc.2024.105969
Received 30 October 2023; Received in revised form 3 January 2024; Accepted 28 January 2024
Available online 2 February 2024
1746-8094/© 2024 Elsevier Ltd. All rights reserved.

structures, reminiscent distribution of intensity, irregular tumor shapes and unrecognized breast tumors. Deep CNN-based methods have made significant advancements, simplifying the accurate identification of anomalous regions in medical images [8]. U-Net, with an encoder–decoder structure, was initially proposed for biomedical image segmentation [9]. The ultrasound images were preprocessed using contrast enhancement and speckle noise removal techniques to improve the network performance [10]. However, the simple frameworks fail to produce excellent segmentation performance on BUS images due to the complex patterns and influence of the tissue regions around the lesion [11]. Although these preprocessing methods can enhance network performance, they obliterate the objects' native spatial feature composition. To improve the performance, U-Net skip connections are redesigned to minimize the semantic gap between the encoder–decoder feature maps [12,13]. The network also made use of the multi-scale feature fusion capabilities of U-Net. It used an ensemble of U-Nets with various depths to solve the problem of undetermined network depth. Regular convolutions are replaced with dilated convolutions to expand the receptive fields of the breast anomalies [14]. However, extracting global features using dilated convolution in deeper networks is difficult as they tend to extract local features along with. Global feature extraction is required for segmenting breast lesions of variable sizes. Composite dilated convolutions help with the problems caused by various lesion sizes and shapes. Dilated convolutions, by themselves, are insufficient in addressing interference from nearby tissues and the presence of fuzzy borders when aiming to create an expanded receptive field [15].

An attention mechanism was introduced in skip connection to improve the segmentation performance. However, the mechanism is restricted to the finite receptive field as U-Net exploits the convolution operations, restricting the segmentation performance. U-Net Models included a dynamic kernel size selection process for breast lesion segmentation that enables them to adaptively change the size of their receptive fields based on various input information scales. The feature maps from different branches with varying kernel sizes are combined using softmax attention. The receptive fields in the resultant fusion layers have different dimensions depending on the attention applied to these branches [15]. A combination of spatial attention, channel attention and boundary detection modules are included in U-Net to create a global guiding network for the segmentation of breast lesions [16]. To underscore the importance of boundary accuracy in the automatic segmentation of tumors in Breast Ultrasound (BUS) images, a boundary selection module for automatic prioritization of unclear boundary regions with graph convolution-based boundary rendering module that harnesses global contour information was proposed in [17]. The above methods reduce the impact of various factors such as intricate structure, similar intensity spread, irregular tumor shapes and unrecognized tumors in ultrasound images on breast lesion segmentation performance.

Further, to achieve a segmentation performance close to ground truth, dilated convolutions and attention mechanisms are integrated into the models [18,19]. Although the segmentation performance was enhanced using dilated convolutions and attention strategies, it continued to have issues with fixed receptive fields and a single attention method. Adaptive spatial and channel attention mechanism is introduced into the U-Net model to learn more generic features. The adaptable attention mechanism (AAM) can assist the network in choosing more reliable representations from various perspectives compared to standard convolution operations with pre-set receptive fields [20]. However, fusing multilevel features to learn variable shapes and locations is crucial.

This work proposes an attention aggregation-based pyramid feature clustering module (AAPFC) to learn the multistage features effectively to provide a robust semantic representation with an edge constraint scheme. AAPFC extracts distinctive features by aggregating overall context features across the entire feature pyramid and disseminating them to every level. It aggregates the supporting information from neighborhood features to provide location-wise reconfiguration of kernels for content-aware feature up-sampling and down-sampling. To select more reliable interpretations from different perspectives, hybrid adaptable attention approaches are used in U-Net instead of the primary convolutional layers with predetermined kernel size. Various studies show that the proposed adaptable attention U-net significantly improves breast lesions segmentation. The main contribution of the proposed work are:

- A U-Net model consisting adaptable spatial and channel attention mechanism to select variable receptive fields with different shapes and size.
- Integrating the AAPFC model to the proposed network to effectively aggregate and fuse multilevel features. The model extracts distinct characteristics by acquiring broad contextual features across the entire feature hierarchy and distributing them to each level.
- The proposed model is assessed on several public BUS datasets. Several experiments on these datasets show how the proposed model outperforms the most recent segmentation models.

The remaining part of the paper is structured as follows: Section 2 introduces the network, the key components and the loss function for BUS image segmentation. The public dataset details, experimental setup and evaluation metrics are given in Section 3. The results and discussion is presented in Sections 4 and 5. Finally, the conclusion is given in Section 6.

2. Methodology

The details of the proposed network structure, main network building components and loss functions used in training are provided in this section.

2.1. Proposed network

The proposed encoder–decoder based network structure for BUS image segmentation of breast lesions is shown in Fig. 1. The main components of the structure are the adaptable attention module (AAM) and AAPFC module. The structure consists of 4-stages on the encoder and 4-stages on the decoder side, which are connected through a bridge. It includes eight AAMs, four on the encoder and four on the decoder side. The output features of each AAM are up-sampled on the encoder side and down-sampled on the decoder side. AAM is developed to learn more robust representations of the BUS images using channel and spatial geometrical features. Each encoder stage's output features are fed into the AAPFC model to produce the feature pyramid by fusing with next-level features from low to higher levels. A detailed discussion of the two main components is provided in the following sections.

2.2. Adaptable attention module (AAM)

The convolution layers of fixed kernel size of 3×3 in conventional U-Net are replaced by AAM, which is shown in Fig. 2(a). AAM is mainly designed to extract the feature maps of receptive fields of different sizes using regular convolutional layers of kernel sizes 5×5, 3×3 and 1×1 and dilation convolution layers of kernel size 3×3 with dilation rate-3. The feature maps of various receptive fields are later fed to the channel self-attention block (ChSAB) and spatial self-attention block (SpSAB) [21]. It enhances the network's capability to adapt to various inputs to characterize the BUS images more accurately. AAM consists of two stages and each stage consists ChSAB, CBNR 3×3, CBNR 1×1 and SpSAB.

ChSAB shown in Fig. 2(b) is developed to capture beneficial objective characteristics from various receptive fields to guide the network to learn more accurate feature descriptions. Its primary objective is to focus on the feature category. ChSAB helps the BUS segmentation model

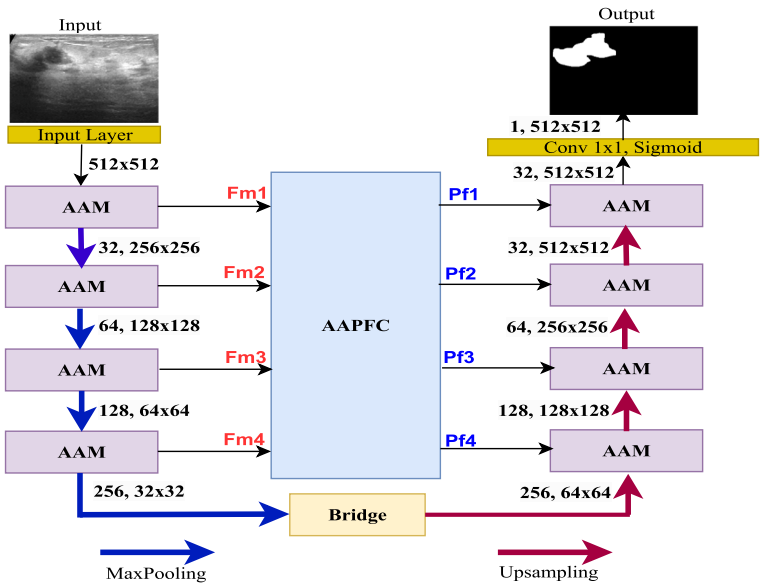


Fig. 1. Proposed U-Net based Network model for BUS image segmentation. The model consists four down-sampling AAM stages and four up-sampling AAM stages. Features from each AAM stage on the encoder side are clustered using AAPFC module and concatenated with decoder side AAM block.

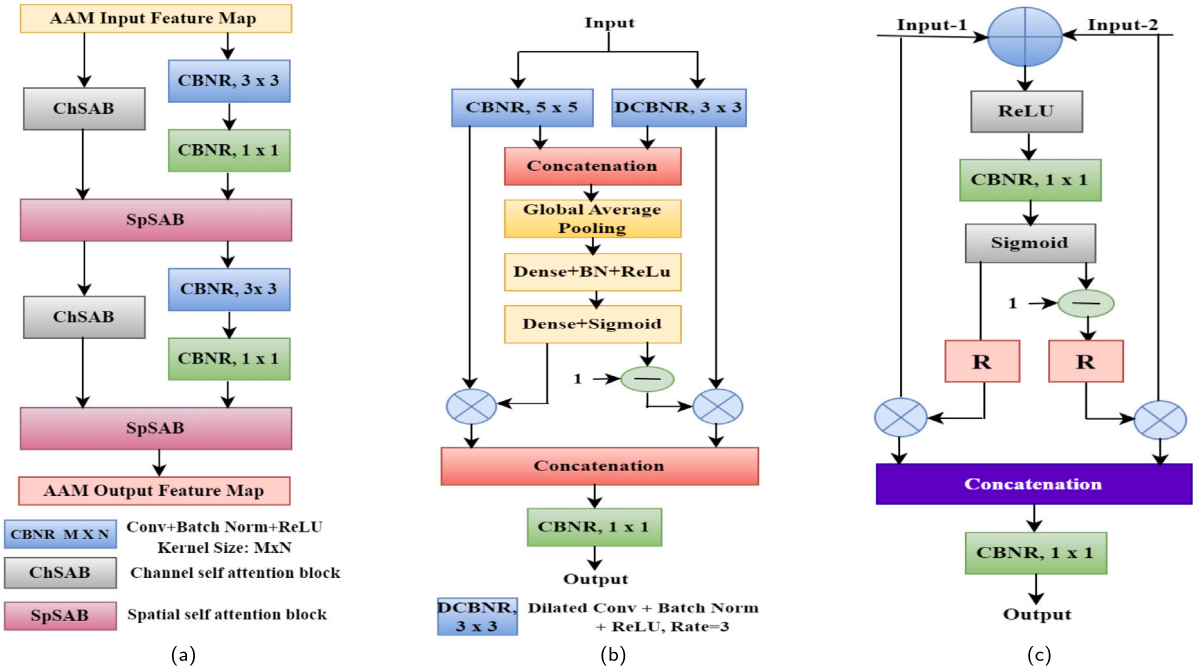


Fig. 2. (a) Adaptive attention model (AAM) (b) Channel self attention block (CSAB) (c) Spatial self attention block (SpSAB).

to select the most significant features from the channel dimension. The feature maps extracted from CBNR and DCBNR layers of size $c \times h \times w$ with kernel size 5×5 and 3×3 are concatenated, and global average is pooled to produce the new feature map of size $2c \times 1 \times 1$. The new feature map is input to a dense layer followed by batch normalization [22] and ReLU [23], and the final channel attention map (C_{map}) is obtained by passing it through dense layer and sigmoid activation function. C_{map} and $1 - C_{map}$ channel attention feature maps can assist the network in adaptively extracting more representative feature maps which are multiplied by CBNR and DCBNR outputs respectively. The resultant feature maps are combined and used as the input to next stage.

The SpSAB shown in Fig. 2(c) further improves the stability of the network, and the main objective is to locate the features accurately. The output feature maps of ChSAB and CBNR with kernel size 1×1 are inputs to SpSAB. The input feature maps are element wise added and passed through CBNR layer which will locate the object more precisely. Two feature maps s_{map} and $1 - s_{map}$ obtained by passing the output of CBNR through sigmoid activation are re-sampled to obtain maps with identical number of channels and matrix-multiplied by the inputs feature maps of SpSAB. The matrix multiplier outputs are fused and passed through convolution operation to give the final output.

2.3. Attention aggregation based pyramid feature clustering (AAPFC) model

Pyramid features enable segmenting of lesions of different scales by encoding the multi-scale feature maps [24]. The AAPFC module shown in Fig. 3 is designed to combine the pyramid features retrieved from multiple encoder stages with an efficient fusion strategy. AAPFC accepts

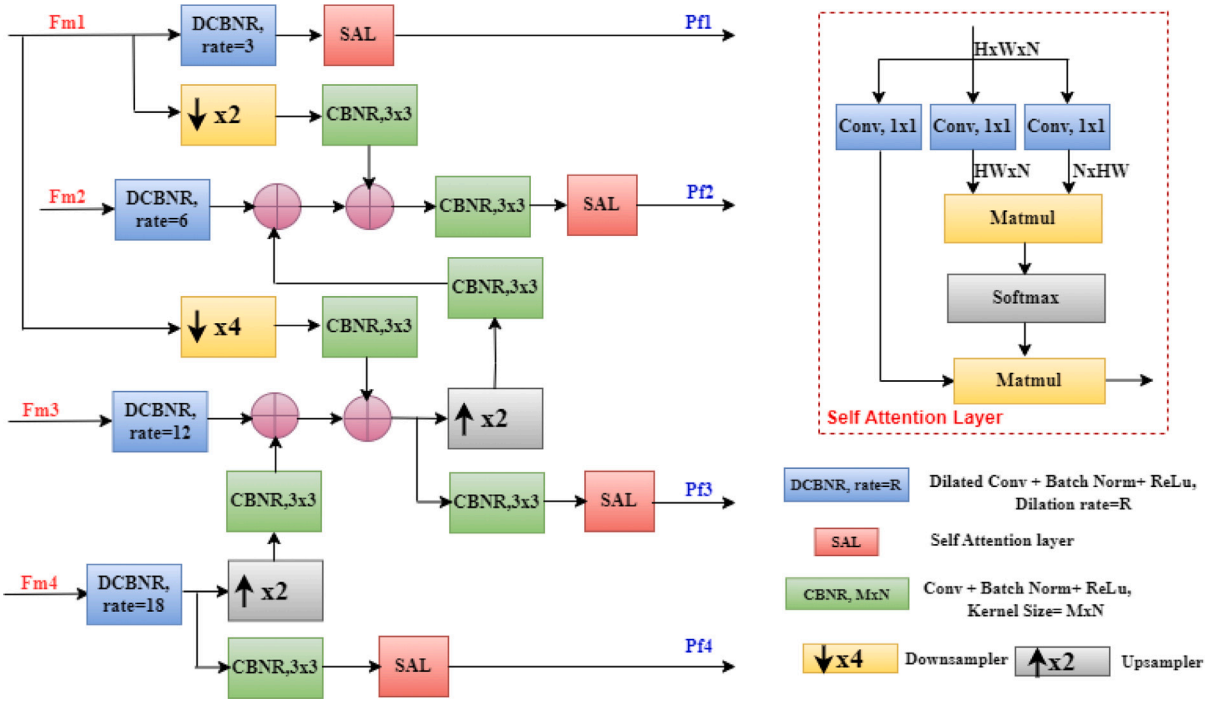


Fig. 3. Attention aggregation based pyramid feature clustering (AAPFC) module. The AAPFC module provides multiscale context features by accepting features from four encoder side AAM blocks. Upsampled multiscale context features fused with low level features are processed with SAL and fed to AAMs on decoder side.

Table 1
BUS datasets details.

Dataset	Malignant	Benign	Normal	Total	Image-Resolution
Dataset-A	53	110	Nil	163	760 × 570
BUSID	210	437	133	780	500 × 500

features from encoding stages 1–4, and to provide multi-scale context-relevant information, dilation convolutions with dilation rates 3,6,12 and 18 are used. In the AAPFC module, the upsampled higher-level features are processed by the convolution layer and fused with low-level features by addition. Finally, the fused features processed by self attention layer (SAL) [25] are fed to AAMs onto the decoder side. The proposed model extracts discriminating features by pooling multilevel context features and minimizes the loss. Without AAPFC, the features of different levels of the network are affected by information loss. The proposed approach will adaptively aggregate the pyramid structures' discriminative features.

2.4. Training loss function

The loss function is crucial in developing the segmentation model and impacts how the model learns. For developing the segmentation models, the loss function is crucial which influences the model learning. The performance of segmentation has been improved through many evaluations utilizing various models and loss functions. The most popular segmentation loss function is Binary Cross Entropy (BCE) [26], which results in a biased model that outputs incorrect breast tumor boundary maps. To overcome this the model is trained with a combination of BCE and boundary loss functions [27] given in (1).

$$Loss_{comb} = Loss_{BCE} + Loss_{BL} \tag{1}$$

Where $Loss_{BCE}$ is defined in (2).

$$\text{Loss}_{\text{BCE}} = -\sum_{m,n} y_g \log(y_p) + (1 - y_g)\log(1 - y_p) \tag{2}$$

Where, y_g and y_p are ground-truth and predicted segmentation masks respectively.

(3) to (8) defines the functions related to boundary-loss $Loss_{BL}$

$$y_g^b = \text{pool}\left(1 - y_g, \alpha\right) - \left(1 - y_g\right) \tag{3}$$

$$y_p^b = \text{pool}\left(1 - y_p, \alpha\right) - \left(1 - y_p\right) \tag{4}$$

Here, y_g is the ground truth segmentation map and y_p is the predicted segmentation map. *pool(.)* is a pixel wise max-pooling operator. α is an hyper-parameter set to 3.

$$y_g^{b,e} = \text{pool}\left(y_g^b, \beta\right), \quad y_p^{b,e} = \text{pool}\left(y_p^b, \beta\right) \tag{5}$$

$$P^c = \frac{\text{sum}\left(y_p^b \circ y_g^{b,e}\right)}{\text{sum}\left(y_p^b\right)}, \quad R^c = \frac{\text{sum}\left(y_g^b \circ y_p^{b,e}\right)}{\text{sum}\left(y_g^b\right)} \tag{6}$$

Here, $\circ$ and sum represent pixel-wise multiplication and summation of the two maps respectively.

$$BL^c = \frac{2P^c R^c}{P^c + R^c} \tag{7}$$

$$Loss_{BL} = 1 - BL^c \tag{8}$$

3. Dataset, experimental details and evaluation metrics

3.1. Dataset description

In this work, two public BUS datasets described in Table 1 with various resolutions are used to assess the performance of the proposed segmentation network. Dataset-A consists of 163 images of resolution 760 × 570 acquired by Siemens ACUSON Sequoia Ultrasound System [28]. The second dataset is the breast ultrasound images dataset

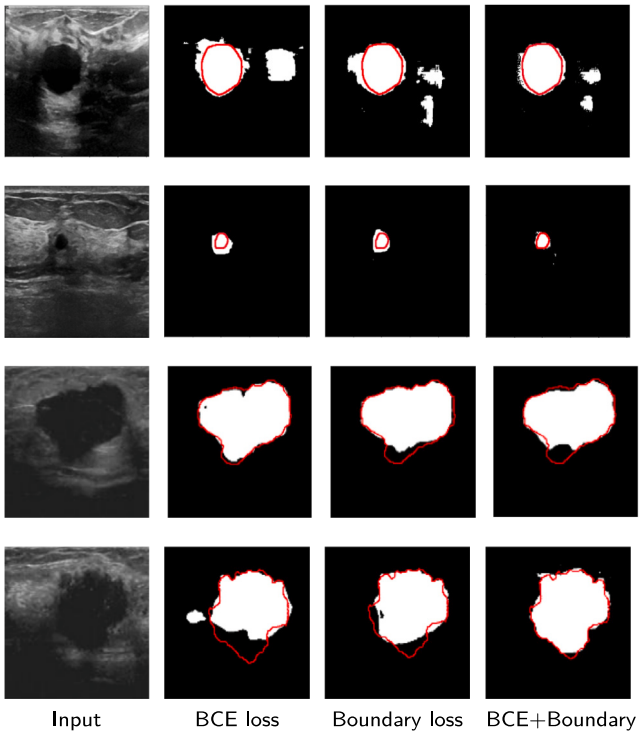


Fig. 4. Visual comparison of segmentation results for different loss functions. The ground-truth breast lesion boundary is given by red curve.

(BUSID) consisting of 780 images collected from 600 female patients with an image resolution of 500 × 500 pixels [29]. BUSID is collected from LOGIQ E9 and LOGIQ E9 AGILE ultrasound systems at Baheya Hospital, which is the first hospital in Egypt focused on diagnosing and treating breast cancer in its earliest stages [29]. For training the network, images and the corresponding masks are resized into 512 × 512. The final dataset consists of 943 images, including 755 train images and 188 test images.

3.2. Experimental settings

Two datasets are considered to demonstrate the stability and efficiency of the proposed method. A substantial number of experiments are conducted for ablation studies to understand the effectiveness of various components and parameters used in the model, robustness analysis and comparison with the state-of-the-art segmentation models. A four-fold cross-validation approach is used for the experiments, in which the dataset is split into four folds, with three folds used for model training and one fold left out for validation in each iteration. The training and test data for each fold do not overlap during the training procedure. The average of three iterations on the Dice index, Mean-IoU, Jaccard similarity index, sensitivity, and specificity are used to assess the model's performance. Robustness analysis is done for the segmentation of benign and malignant lesions, different attention mechanisms and various fusion strategies.

Network implementation and training of the proposed network for 50 epochs is done using Keras, a deep learning API using Tensorflow Python library as backend on an NVIDIA Tesla P100 GPU with 16 GB RAM and developed using Windows 10-OS. Adam optimizer is used for the supervised training with the learning rate for the model is initially fixed at 0.0001 and decays by a factor of 2 for every 10 epochs. The network parameters are updated by loss function computed using (1) during back propagation.

3.3. Evaluation metrics

Image segmentation metrics such as Mean Intersection over Union (Mean-IoU) defined in (9) and Dice similarity coefficient (DSC) defined in (10) are used to evaluate the segmentation performance of the proposed model. Mean-IoU and DSC are the most common image segmentation metrics which measure the overlap between predicted and ground truth masks [30]. To evaluate the segmentation model's classification performance, specificity given in (11) and sensitivity given in (12) are used [31]. Pixel accuracy (Acc) given in (13) is another popular evaluation metric in semantic segmentation defined as the ratio of accurate positive and negative predictions to total correct predictions.

An additional method for calculating the degree of agreement between the predicted masks and ground truth is the Cohen's kappa (κ) given in (14). Kappa offers relative accuracy with reference to the other classes which contributes to the sensation of originality and dependability in identification of an instance of a medical diagnostic [32]. The range of κ is between −1 and +1 considered as the worst best score respectively. An instance where the model correctly predicted the positive class is known as a true positive (TP). A true negative (TN) is a similar outcome where the model accurately predicts the negative class. An outcome when the model predicts the positive class wrongly is known as a false positive (FP). False negative (FN) refers to the classification of negative classes as positive.

$$Mean - IoU = \frac{TP}{TP + FP + FN} \tag{9}$$

$$DSC = \frac{2TP}{2TP + FP + FN} \tag{10}$$

$$specificity = \frac{TN}{TN + FP} \tag{11}$$

$$sensitivity = \frac{TP}{TP + FN} \tag{12}$$

$$Acc = \frac{TP + TN}{TP + TN + FN + FP} \tag{13}$$

$$\kappa = \frac{(TP + TN) - f_c}{(TP + TN + FN + FP) - f_c} \tag{14}$$

$$f_c = \frac{(TN + FN)\,(TB + FP) + (FP + TP)\,(FN + TP)}{TP + TN + FN + FP} \tag{15}$$

4. Experimental results

In this section, the ablation studies are conducted on the parameters and components of the proposed network to justify the design of the network. Next, the stability of the network is analyzed for various factors. Finally, the performance of the proposed model for BUS image segmentation is compared with the popular benchmarking segmentation models.

4.1. Ablation studies for loss function

The first ablation study is conducted to analyze the significance of combination of loss functions. The loss function plays a critical role on segmentation performance. From the visual comparisons shown in Fig. 4, it is observed that although all the losses perform well, segmentation results are more precise with BCE + Boundary loss. This is further strengthened from the quantitative summary of the results is given in Table 2.

Table 2
Ablation study of combination and individual loss functions of the proposed network on the segmentation performance (Mean ± STD).

Loss function	Mean-IoU	DSC	Specificity	Sensitivity	Accuracy	kappa
$Loss_{BCE}$	62.28 ± 0.33	71.74 ± 1.31	97.16 ± 0.04	78.18 ± 1.62	95.36 ± 0.49	0.725
$Loss_{BL}$	64.94 ± 0.59	74.73 ± 1.22	97.23 ± 0.09	80.92 ± 1.54	95.92 ± 0.64	0.738
$Loss_{BCE}$+$Loss_{BL}$	67.27 ± 0.61	78.72 ± 0.54	97.32 ± 0.19	82.22 ± 1.43	96.12 ± 0.28	0.747

Table 3
Impact of various kernel-sizes and dilation rates on BUS image segmentation (Mean ± STD).

Kernel-Size	Dilation rate	Mean-IoU	DSC	Specificity	Sensitivity	Accuracy	Kappa
All kernels are of Size 3 × 3	2	64.81 ± 0.51	77.89 ± 0.49	96.54 ± 0.26	81.03 ± 1.85	94.67 ± 0.26	0.718
All kernels are of Size 5 × 5	3	65.94 ± 0.49	78.24 ± 0.42	96.84 ± 0.34	81.62 ± 1.78	94.86 ± 0.48	0.725
Both 3 × 3 and 5 × 5 (Proposed)	3	67.27 ± 0.61	78.72 ± 0.54	97.32 ± 0.19	82.22 ± 1.43	96.12 ± 0.78	0.747

Table 4
Overall segmentation result (Mean ± STD) obtained by conducting ablation studies on various network components.

Network	Mean-IoU	DSC	Specificity	Sensitivity	Accuracy	Kappa
UNet-Basic	57.27 ± 0.50	70.20 ± 1.90	96.18 ± 0.67	78.32 ± 2.48	94.31 ± 0.42	0.716
UNet-CAM-Conv	59.76 ± 0.71	73.34 ± 2.13	96.29 ± 0.37	79.46 ± 1.94	94.57 ± 0.76	0.724
UNet-SAM-Conv	60.56 ± 0.42	73.91 ± 1.83	96.46 ± 0.41	79.98 ± 1.89	94.79 ± 0.21	0.729
UNet-AAM-Conv	65.91 ± 0.57	76.69 ± 0.89	96.54 ± 0.26	81.12 ± 1.85	94.98 ± 0.48	0.731
UNet-SAG-Skip	59.85 ± 0.84	73.79 ± 1.46	96.35 ± 0.34	79.91 ± 1.83	94.21 ± 0.32	0.714
UNet-FPN	63.92 ± 0.83	75.16 ± 0.82	95.76 ± 0.28	80.03 ± 1.49	93.89 ± 0.41	0.715
UNet-FAM	64.39 ± 0.73	75.92 ± 0.87	96.04 ± 0.31	80.79 ± 1.81	94.98 ± 0.29	0.718
UNet-AAPFC	65.45 ± 0.79	76.34 ± 0.47	96.84 ± 0.34	80.92 ± 1.78	95.02 ± 0.64	0.725
UNet-AAM-AAPFC	**67.27 ± 0.61**	**78.72 ± 0.54**	**97.32 ± 0.19**	**82.22 ± 1.43**	**96.12 ± 0.78**	**0.747**

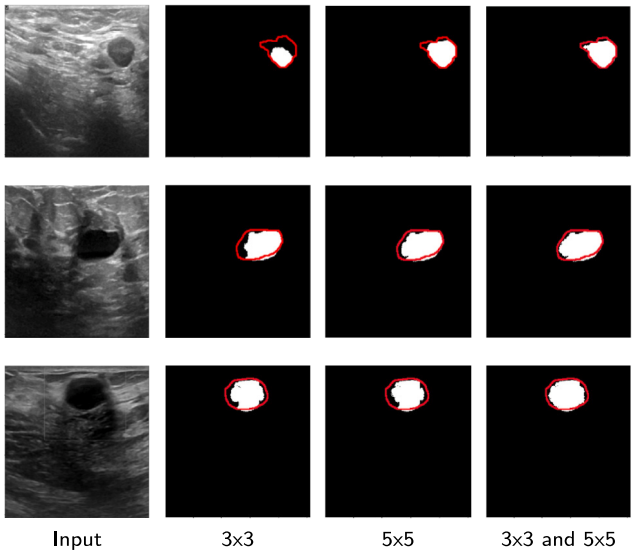


Fig. 5. Visual comparison of segmentation results for different receptive field sizes. The ground-truth breast lesion boundary is given by red curve.

4.2. Ablation studies for receptive field size choice

To analyze the effect of various receptive field sizes, all the 5 × 5 kernel sizes are first modified to 3 × 3 and dilation rates in dilated convolutions are reduced to 2. Next, all 3 × 3 kernel size is changed to 5 × 5 kernel to analyze the effect of larger receptive fields and dilation rate is set to 3. To assess the performance of different components in the network, ablation studies are conducted by considering basic U-Net as a benchmark model. A visual comparison of segmentation results for various receptive field sizes is shown in Fig. 5. The comparison results of various receptive field sizes are given in Table 3. According to the results given in Table 3, the combination of 3 × 3 and 5 × 5 kernels which capture both smaller and larger receptive fields, provide better BUS segmentation than using only 3 × 3 kernels or 5 × 5 kernels alone.

4.3. Ablation studies to prove the importance of AAM and AAPFC modules

The segmentation results of the proposed model is compared with basic U-Net, convolution layers integrated with channel attention mechanism (UNet-CAM-Conv), spatial attention mechanism (UNet-SAM-Conv) and proposed Adaptive attention module (UNet-AAM-Conv). Segmentation performance is analyzed by conducting ablation studies on using spatial attention gate in skip connections (UNet-SAG-Skip), different fusion strategies such as feature pyramid network (UNet-FPN) [24], Feature aligned module (UNet-FAM) [33] and AAPFC (UNet-AAPFC). The segmentation performance of the proposed network consisting AAM in feature extraction blocks of the UNet and AAPFC as a feature fusion techniques (UNet-AAM-AAPFC) is compared with UNet-Basic, UNet-CAM-Conv, UNet-SAM-Conv, UNet-AAM-Conv, UNet-SAG-Skip, UNet-FPN, UNet-FAM, UNet-AAPFC. The ablation studies findings of various methods are listed in Table 4, which dictate that integrating AAM and AAPFC components to the benchmark U-Net enables the network to learn accurate predictions from BUS images.

4.4. Comparative analysis with the state-of-the-art segmentation methods

The proposed network performance on BUS image segmentation is compared with the state-of-the-art deep CNN models to assess the robustness and efficiency. The benchmarking models considered for comparison are U-Net [9], U-Net++ [13], DeepLabV3+ [34], SegNet [35], CE-Net [36], SK-UNet [15] and SMU-Net [11]. Table 5 presents the quantitative assessment results of the benchmark and existing segmentation techniques performed on four-fold cross-validation. Table 5 indicates that the proposed method provides better results across four evaluation metrics. The receiver operating characteristic (ROC) curves of various segmentation methods is shown in Fig. 6. The ROC curves are used most often in medical image analysis as a method of validating the results [37]. The degree of certainty of accurate predictions corresponding to a method is indicated by ROC curve and the area under the ROC curve. The proposed approach achieves the highest AUC value when compared to other approaches. The visual comparison of the segmentation results of all the methods is shown in Fig. 7. When compared to benchmark methods, the proposed method

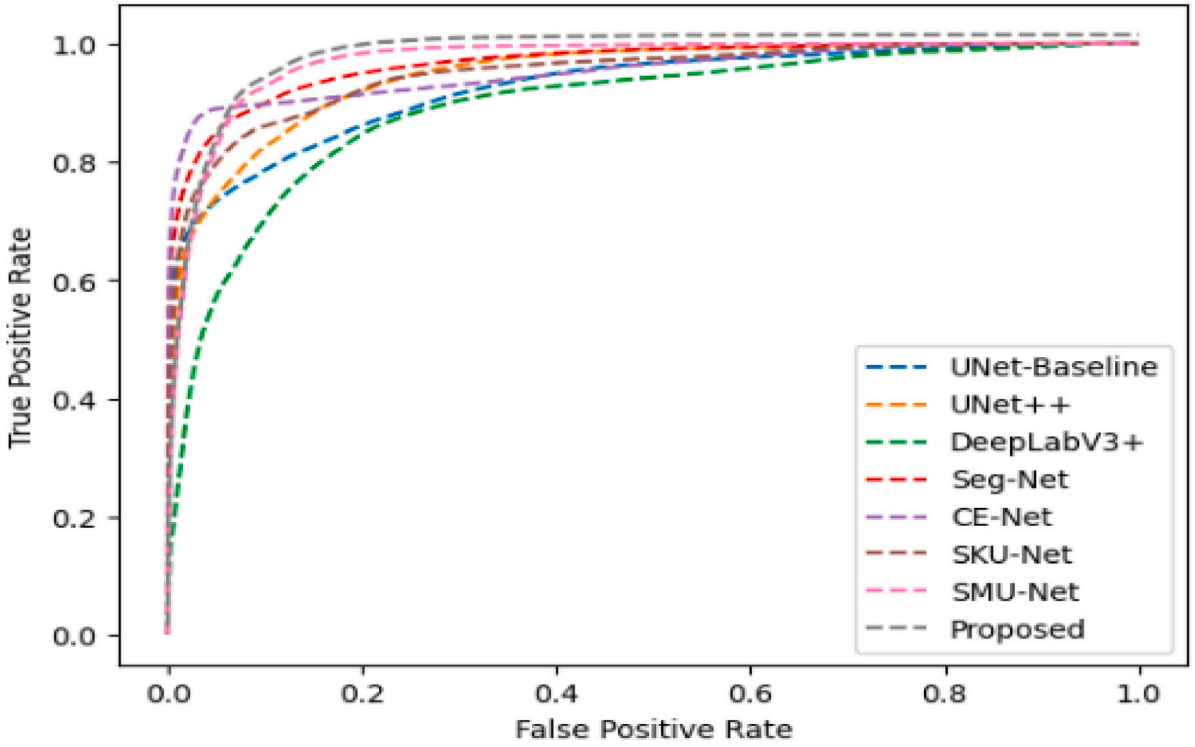


Fig. 6. ROC curves of proposed and benchmark segmentation methods on overall dataset.

Table 5
Breast lesion segmentation performance (Mean ± STD) comparison between benchmark models and proposed model.

		UNet-Baseline	UNet++	DeepLabV3+	Seg-Net	CE-Net	SK-UNet	SMU-Net	Proposed
Overall	Mean-IoU	57.27 ± 0.50	58.18 ± 0.61	60.31 ± 0.42	63.69 ± 0.41	59.64 ± 0.74	65.11 ± 0.76	65.94 ± 0.67	**67.27 ± 0.61**
	DSC	70.20 ± 1.90	71.58 ± 2.09	73.31 ± 1.94	75.45 ± 0.31	73.37 ± 1.84	75.29 ± 1.80	76.12 ± 1.24	**78.72 ± 0.54**
	Specificity	96.18 ± 0.67	97.14 ± 0.62	96.79 ± 0.59	96.99 ± 0.86	97.23 ± 0.71	97.35 ± 0.64	97.41 ± 0.72	**97.32 ± 0.19**
	Sensitivity	78.32 ± 2.48	72.55 ± 2.66	76.49 ± 2.15	82.83 ± 2.34	81.55 ± 1.83	81.99 ± 2.16	81.62 ± 1.63	**82.22 ± 1.43**
	Accuracy	94.36 ± 0.79	95.49 ± 0.18	94.66 ± 0.62	95.18 ± 0.59	95.45 ± 0.86	95.79 ± 0.71	95.75 ± 0.64	**96.06 ± 0.38**
	Kappa	0.717	0.743	0.724	0.726	0.728	0.731	0.729	**0.747**

generates better visual results and the predicted masks of the proposed are nearer to the ground-truth masks of the breast lesion region indicated by red curve. Also, the proposed model can segment the lesions better in unclear and low-contrast images under the supervision of channel and spatial feature maps at low-level and high-level semantic maps. Overall assessment and visual results indicate that the proposed method produces the better results in segmentation of lesions in BUS images.

4.5. Stability analysis

The segmentation performance of normal, malignant and benign lesions in BUS images by the presented is compared with the other methods to assess the stability of the proposed method. Malignant lesions exhibit irregular shapes with larger sizes and faded surroundings compared to benign lesions. The gray level difference in malignant lesions is blurrier compared to benign lesions. The ROC curves of various segmentation methods on malignant and benign breast lesions shown in Fig. 8 confirms the presented method's stability. The ROC-Curves show that the presented method produces better performance on both benign and malignant lesion segmentation. Table 6 evaluates the effect of normal, benign and malignant lesions on the segmentation performance of BUS images of the proposed and various methods. The performance is tested separately on normal, benign and malignant lesions. From the results, it has been observed that the segmentation of malignant lesions is more challenging than benign lesions. The model provides better benign lesion segmentation performance and focusses more on segmenting benign lesions due to irregular and blurred malignant regions and more images with benign lesions than malignant lesions. The difference can be reduced by increasing the number of malignant images. It signifies that when a sufficiently large and well-balanced dataset is utilized for network training, model can also perform satisfactorily on malignant lesion segmentation.

The lesion size on BUS images varies significantly as a diagnostic feature, which is quite difficult for performance enhancement. The suggested model segments lesions of varying sizes with a reasonable level of performance. The study also suggests that, size is not a primary feature for the proposed model in segmenting the lesions. Irregular shaped lesions segmentation is quite difficult and affects the performance of the model. To study the models adaptability for various shapes, the dataset is divided into images consisting regular and irregular shaped lesions. Separate test are conducted on images of both the shapes and the proposed model provides promising results as seen in Table 7 irrespective of the shape which confirms the stability of the model on lesion shape.

4.6. Validation analysis

To validate the generalized performance of the model, 52 BUS images acquired by GE Voluson E10 ultrasonic medical device are

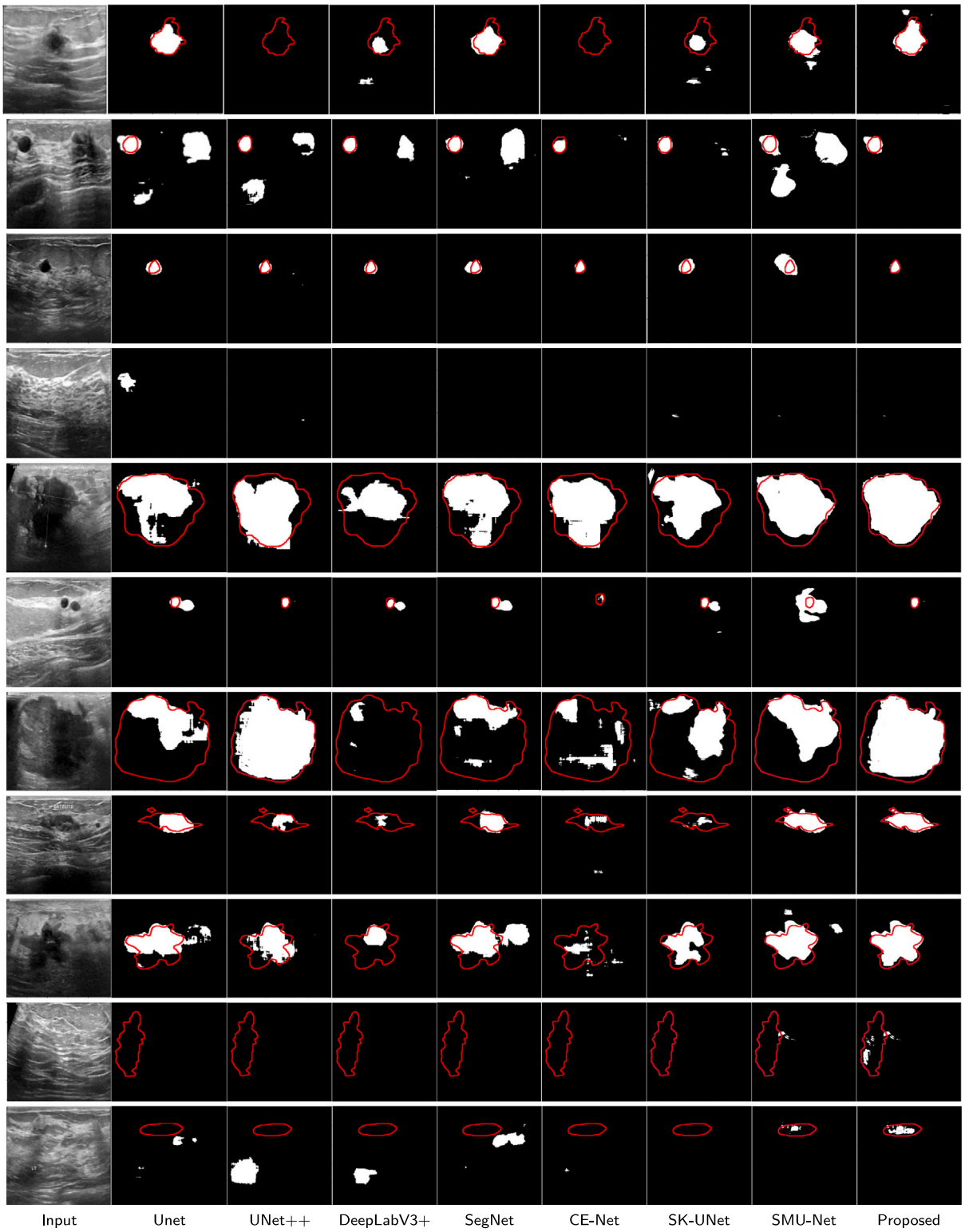
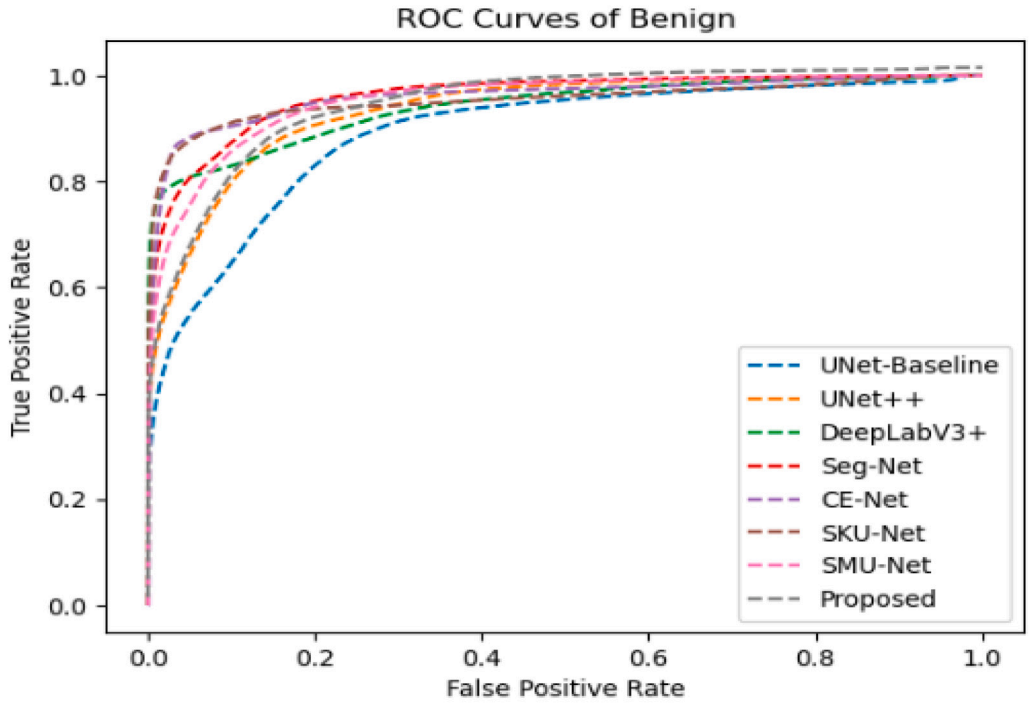


Fig. 7. Visual comparison of segmentation results of the proposed and benchmark methods. The ground-truth breast lesion boundary is given by red curve.

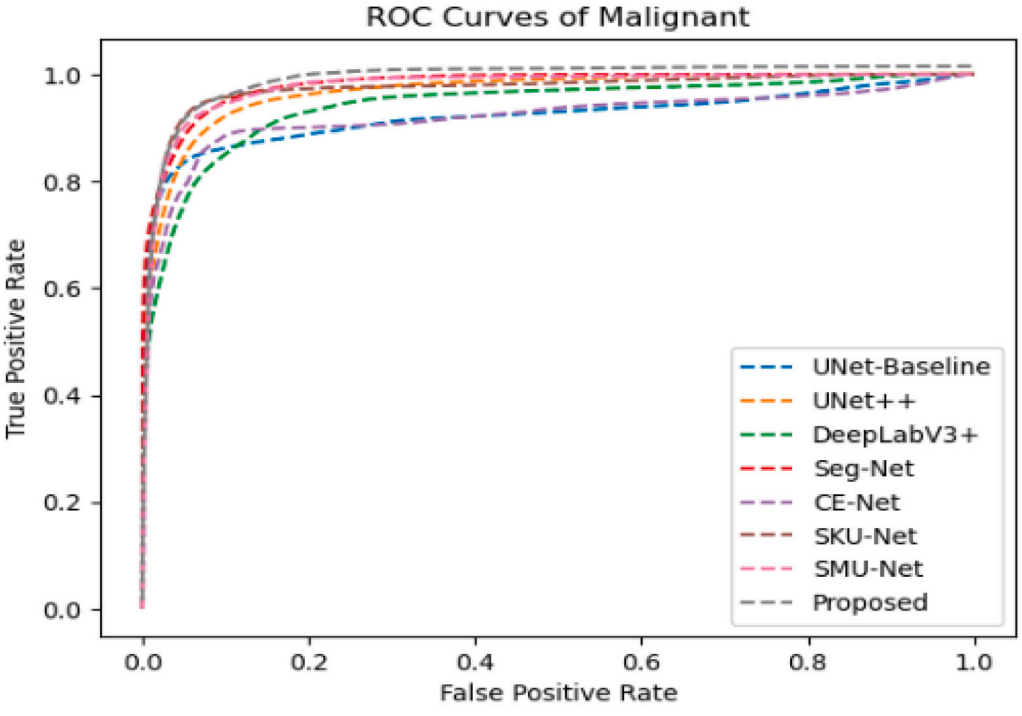
utilized in this work [38]. These images are not used in training the model and there are significant differences compared to images used for training. The model is trained using Dataset-A and BUSID and to test the robustness of the proposed model, 52 BUS images are used as external data. The model still achieves better results of 69.88% DSC, 63.98% Mean-IoU, 95.94% Specificity, 80.92% sensitivity, 95% Accuracy and 0.714 Kappa. These results show that the model has better stability and more appropriate for breast lesion segmentation.

5. Discussion

In this current study, in-depth examination and performance analysis of several U-Net models is done for segmenting breast lesions in BUS images. Furthermore, the performances of the proposed model is compared with state-of-the-art U-Net based deep learning models such as Baseline U-Net, U-Net++, Seg-Net, CE-Net, SK-UNet, SMU-Net and DeepLabV3+ to test the competitiveness. Though some studies have



(a)



(b)

Fig. 8. ROC curves of different segmentation methods on benign and malignant breast lesions.

used transfer learning techniques by modifying few layers of the popular networks for classification of medical images with abnormalities and achieved the better performance [39,40], in this study only U-Net based models are considered as they are proven to be efficient for segmenting the area of interest in the images. Accurate identification of breast lesion position, shape and volume is essential for precise prediction of cancers in BUS images [41]. In this study, a pyramid feature clustering module for attention aggregation-based breast lesion segmentation is presented along with an adaptive attention module featuring channel and spatial attention method. The relative performance of the proposed

Table 6
Segmentation performance (Mean ± STD) of BUS normal images without lesions (normal), images consisting malignant and benign lesions by different networks.

		UNet-Baseline	UNet++	DeepLabV3+	Seg-Net	CE-Net	SK-UNet	SMU-Net	Proposed
Malignant	Mean-IoU	51.68 ± 1.12	52.83 ± 0.52	53.27 ± 1.41	63.69 ± 0.41	51.52 ± 0.59	60.07 ± 0.64	60.15 ± 0.58	**60.65 ± 0.86**
	DSC	65.57 ± 1.76	66.92 ± 1.85	66.98 ± 1.84	72.26 ± 0.58	65.87 ± 1.64	67.56 ± 1.86	72.16 ± 1.53	**72.75 ± 1.24**
	Specificity	93.45 ± 0.97	94.89 ± 0.58	94.33 ± 0.68	98.09 ± 0.66	94.07 ± 0.62	94.67 ± 0.95	94.24 ± 0.77	**94.86 ± 2.06**
	Sensitivity	70.88 ± 2.69	74.96 ± 2.36	71.54 ± 1.80	81.95 ± 1.74	74.55 ± 2.58	76.04 ± 0.67	77.04 ± 1.72	**77.75 ± 1.97**
	Accuracy	91.57 ± 0.97	93.02 ± 0.58	92.47 ± 0.68	96.21 ± 0.76	92.38 ± 0.82	92.92 ± 0.55	92.53 ± 0.81	**92.98 ± 0.74**
	Kappa	0.695	0.707	0.705	0.732	0.705	0.708	0.711	**0.712**
Benign	Mean-IoU	58.81 ± 1.27	65.38 ± 0.97	61.36 ± 1.09	64.25 ± 1.36	62.96 ± 1.22	66.92 ± 0.96	69.34 ± 0.86	**71.78 ± 0.92**
	DSC	70.79 ± 1.82	75.93 ± 2.14	74.55 ± 2.26	75.62 ± 2.65	75.67 ± 1.82	76.25 ± 1.86	78.72 ± 2.17	**82.09 ± 2.04**
	Specificity	97.72 ± 0.62	97.48 ± 0.74	97.95 ± 0.60	95.98 ± 0.88	98.38 ± 0.60	98.08 ± 0.73	98.07 ± 0.59	**98.14 ± 0.82**
	Sensitivity	74.35 ± 2.10	81.55 ± 2.21	76.95 ± 2.10	83.94 ± 2.21	83.02 ± 1.25	82.97 ± 2.01	80.87 ± 2.98	**84.26 ± 2.15**
	Accuracy	95.97 ± 0.39	95.82 ± 0.48	96.17 ± 0.32	94.29 ± 0.89	96.61 ± 0.09	96.32 ± 0.38	96.37 ± 094	**96.53 ± 0.27**
	Kappa	0.733	0.731	0.734	0.720	0.740	0.735	0.736	**0.740**
Normal	Mean-IoU	44.18 ± 0.85	47.86 ± 0.98	46.42 ± 0.91	47.45 ± 0.85	46.35 ± 0.88	49.72 ± 0.77	51.04 ± 0.84	**56.98 ± 0.97**
	DSC	56.12 ± 1.89	59.64 ± 1.66	56.87 ± 1.67	59.66 ± 2.15	61.42 ± 1.78	59.89 ± 1.71	60.92 ± 1.82	**67.04 ± 2.58**
	Specificity	96.09 ± 0.71	96.19 ± 0.83	96.23 ± 0.84	96.12 ± 0.68	97.16 ± 0.81	96.75 ± 0.89	96.81 ± 0.92	**96.91 ± 1.03**
	Sensitivity	65.32 ± 2.54	68.03 ± 2.14	65.23 ± 2.22	52.56 ± 2.08	70.67 ± 1.98	70.58 ± 2.28	70.21 ± 2.95	**70.99 ± 2.63**
	Accuracy	94.42 ± 0.82	94.32 ± 0.74	94.62 ± 0.49	94.49 ± 073	95.49 ± 0.62	94.92 ± 0.72	94.99 ± 0.67	**95.12 ± 0.72**
	Kappa	0.721	0.722	0.725	0.724	0.732	0.728	0.731	**0.733**

Table 7
Segmenting performance on regular and irregular shaped lesions.

	Regular	Irregular
Mean-IoU	71.36 ± 0.58	69.96 ± 0.82
DSC	78.16 ± 0.87	77.61 ± 2.10
Specificity	97.60 ± 1.97	96.37 ± 0.94
Sensitivity	82.75 ± 1.25	81.69 ± 0.76
Accuracy	96.76 ± 0.98	96.03 ± 0.71
Kappa	0.749	0.748

model is tested through segmentation comparisons utilizing statistical analysis and standard comparison criteria. In order to evaluate the efficiency, ablation studies were first carried out on the parameters of the proposed framework and its individual components. The results of the experiments show that the network performs exceptionally well.

By the findings of experimental outcomes, these are the conclusions made with respect to the benchmark and proposed models. Generally, Modified skip connections in a variant model of baseline U-Net such as U-Net++, gives better segmentation results compared to original U-Net. U-Net++ model's performance shows that the skip connection with convolution layers will better fuse low-level encoder features with high-level decoder features to provide better segmentation. The U-Net based models with attention mechanisms such as SK-Net show that introducing an attention system can also increase the network's segmentation performance. However, according to visual comparison results, SK-Net has poor segmentation results, proving the network variant sensitivity to the attention mechanism. In SegNet, feature location information is used in the U-shaped network and provides better segmentation compared to most of the segmentation approaches. According to quantitative comparison, SK-Net achieves better performance. Better segmentation performance is achieved quantitatively and visually when spatial attention with large receptive field is introduced as shown in the proposed model.

According to the visual segmentation findings shown in Fig. 7, most of the models miss detecting certain details for the small-sized breast lesions. As shown in the last two rows, a type of breast lesion which does not show a significant difference compared to the surrounding tissue is undetected by all the models. Many models fail to detect breast lesions with high unevenness and they cannot acquire accurate tumor boundaries. The visual results show that some normal tissues are segmented as lesions that are not anticipated in breast lesion detection in addition to the segmentation of lesions. It is also observed that certain lesions lack segmentation. This might have happened as a result of the models being overfitted during training, which is due to intensity homogeneities found in BUS images. In addition to having few missed detections and failing to identify certain kinds of lesions, the suggested model outperforms current techniques by a large margin.

The hyper-parameters such as the learning rate is chosen directly. The model parameters are optimized using the gradient-based optimization method known as Adam. Gradient-based approaches usually settle in the local minima of an error surface. The BUSI segmentation problem is an imbalanced class problem in which lesions have very few representations in the training process compared to healthy tissues and the number of pixels corresponding to lesions is very less compared to that of healthy tissues. Model hyper-parameter tuning can be done using random search technique to improve the training performance [42]. To get around the issues of local minima and class imbalance, meta-heuristics can be utilized in place of gradient-based learning algorithms [43]. Training and testing the model on diversified dataset is necessary, since the nature of data acquired from different devices varies. Appropriate preprocessing steps can be included to correct the homogeneous intensity distributions and fuzzy boundaries to improve the segmentation performance.

The stability analysis of the network demonstrates the better generalization ability and emphasizes the benefits of channel and spatial attention mechanisms and aggregated feature clustering methods. Experimental results obtained by conducting segmentation on benign and malignant lesions demonstrate the stability of various methods. The proposed model still performs better than other models under various breast lesion intensities, sizes and shapes. Based on the results of the studies conducted, it can be concluded that the proposed model performed better in BUSI segmentation compared to existing methods. Still, it can be seen from our results that the method has some shortcomings. Our approach still needs to be optimized for more complex structured BUSI lesion segmentation to increase the accurate detection and to reduce the miss rate. Obtaining accurate lesion boundaries still remains a challenging task. The low-intensity variations between surrounding tissues and breast lesions significantly impact on the accuracy of breast lesion segmentation. To address the above issues the combination of novel boundary, breast lesion texture and shape aware loss functions [44] can be introduced to train the network and performance can be assessed. It is also required to come up with a large and balanced dataset with all types of lesion characteristics.

6. Conclusion

To address the existing issues in breast lesions segmentation, a hierarchical encoder–decoder based CNN with attention mechanism and pyramid feature clustering is proposed in this research. The spatial and channel attention module with different kernel sizes to select variable receptive fields can direct the model to segment the lesions of

different shapes and sizes and enables the complex breast lesion region segmentation. The AAPFC module is integrated to aggregate and fuse features from different stages effectively. The model extracts distinct characteristics across the entire feature hierarchy and distributes them to each level. Experimental findings demonstrate that the proposed model achieves better efficiency and stability than state-of-the-art BUSI segmentation models.

CRediT authorship contribution statement

Sushma B.: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Writing – original draft. **Aparna Pulikala:** Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

[1] M. Arnold, E. Morgan, H. Rumgay, A. Mafra, D. Singh, M. Laversanne, J. Vignat, J.R. Gralow, F. Cardoso, S. Siesling, et al., Current and future burden of breast cancer: Global statistics for 2020 and 2040, Breast 66 (2022) 15–23.

[2] W. Ren, M. Chen, Y. Qiao, F. Zhao, Global guidelines for breast cancer screening: a systematic review, Breast 64 (2022) 85–99.

[3] I. Jatoi, Screening clinical breast examination, Surg. Clinics 83 (4) (2003) 789–801.

[4] D. Schopper, C. de Wolf, How effective are breast cancer screening programmes by mammography? review of the current evidence, Eur. J. Cancer 45 (11) (2009) 1916–1923.

[5] C.H. Lee, D.D. Dershaw, D. Kopans, P. Evans, B. Monsees, D. Monticciolo, R.J. Brenner, L. Bassett, W. Berg, S. Feig, et al., Breast cancer screening with imaging: recommendations from the society of breast imaging and the ACR on the use of mammography, breast MRI, breast ultrasound, and other technologies for the detection of clinically occult breast cancer, J. Am. College Radiol. 7 (1) (2010) 18–27.

[6] D.-M. Koh, N. Papanikolaou, U. Bick, R. Illing, C.E. Kahn Jr., J. Kalpathi-Cramer, C. Matos, L. Martí-Bonmatí, A. Miles, S.K. Mun, et al., Artificial intelligence and machine learning in cancer imaging, Commun. Med. 2 (1) (2022) 133.

[7] Y. Zhang, M. Xian, H.-D. Cheng, B. Shareef, J. Ding, F. Xu, K. Huang, B. Zhang, C. Ning, Y. Wang, BUSIS: a benchmark for breast ultrasound image segmentation, in: Healthcare, vol. 10, (4) MDPI, 2022, p. 729.

[8] G. Litjens, T. Kooi, B.E. Bejnordi, A.A.A. Setio, F. Ciompi, M. Ghafoorian, J.A. Van Der Laak, B. Van Ginneken, C.I. Sánchez, A survey on deep learning in medical image analysis, Med. Image Anal. 42 (2017) 60–88.

[9] O. Ronneberger, P. Fischer, T. Brox, U-net: Convolutional networks for biomedical image segmentation, in: Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18, Springer, 2015, pp. 234–241.

[10] R. Almajalid, J. Shan, Y. Du, M. Zhang, Development of a deep-learning-based method for breast ultrasound image segmentation, in: 2018 17th IEEE International Conference on Machine Learning and Applications, ICMLA, IEEE, 2018, pp. 1103–1108.

[11] Z. Ning, S. Zhong, Q. Feng, W. Chen, Y. Zhang, SMU-net: saliency-guided morphology-aware U-net for breast lesion segmentation in ultrasound image, IEEE Trans. Med. Imaging 41 (2) (2021) 476–490.

[12] Z. Zhou, M.M. Rahman Siddiquee, N. Tajbakhsh, 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: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, Springer, 2018, pp. 3–11.

[13] Z. Zhou, M.M.R. Siddiquee, N. Tajbakhsh, J. Liang, Unet++: Redesigning skip connections to exploit multiscale features in image segmentation, IEEE Trans. Med. Imaging 39 (6) (2019) 1856–1867.

[14] Y. Hu, Y. Guo, Y. Wang, J. Yu, J. Li, S. Zhou, C. Chang, Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model, Med. Phys. 46 (1) (2019) 215–228.

[15] M. Byra, P. Jarosik, A. Szubert, M. Galperin, H. Ojeda-Fournier, L. Olson, M. O'Boyle, C. Comstock, M. Andre, Breast mass segmentation in ultrasound with selective kernel U-net convolutional neural network, Biomed. Signal Process. Control 61 (2020) 102027.

[16] C. Xue, L. Zhu, H. Fu, X. Hu, X. Li, H. Zhang, P.-A. Heng, Global guidance network for breast lesion segmentation in ultrasound images, Med. Image Anal. 70 (2021) 101989.

[17] R. Huang, M. Lin, H. Dou, Z. Lin, Q. Ying, X. Jia, W. Xu, Z. Mei, X. Yang, Y. Dong, et al., Boundary-rendering network for breast lesion segmentation in ultrasound images, Med. Image Anal. 80 (2022) 102478.

[18] Y. Yan, Y. Liu, Y. Wu, H. Zhang, Y. Zhang, L. Meng, Accurate segmentation of breast tumors using AE U-net with HDC model in ultrasound images, Biomed. Signal Process. Control 72 (2022) 103299.

[19] H. Lee, J. Park, J.Y. Hwang, Channel attention module with multiscale grid average pooling for breast cancer segmentation in an ultrasound image, IEEE Trans. Ultrason. Ferroelectr. Freq. Control 67 (7) (2020) 1344–1353.

[20] G. Chen, L. Li, Y. Dai, J. Zhang, M.H. Yap, AAU-net: An adaptive attention U-net for breast lesions segmentation in ultrasound images, IEEE Trans. Med. Imaging (2022).

[21] J. Zheng, L. Yang, Y. Li, K. Yang, Z. Wang, J. Zhou, Lightweight vision transformer with spatial and channel enhanced self-attention, in: Proceedings of the IEEE/CVF International Conference on Computer Vision, 2023, pp. 1492–1496.

[22] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International Conference on Machine Learning, pmlr, 2015, pp. 448–456.

[23] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in: Proceedings of the 27th International Conference on Machine Learning, ICML-10, 2010, pp. 807–814.

[24] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, S. Belongie, Feature pyramid networks for object detection, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 2117–2125.

[25] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A.N. Gomez, Ł. Kaiser, I. Polosukhin, Attention is all you need, Adv. Neural Inf. Process. Syst. 30 (2017).

[26] M. Yi-de, L. Qing, Q. Zhi-Bai, Automated image segmentation using improved PCNN model based on cross-entropy, in: Proceedings of 2004 International Symposium on Intelligent Multimedia, Video and Speech Processing, 2004, IEEE, 2004, pp. 743–746.

[27] H. Kervadec, J. Bouchtiba, C. Desrosiers, E. Granger, J. Dolz, I.B. Ayed, Boundary loss for highly unbalanced segmentation, Med. Image Anal. 67 (2021) 101851.

[28] M.H. Yap, M. Goyal, F. Osman, R. Martí, E. Denton, A. Juette, R. Zwiggelaar, Breast ultrasound region of interest detection and lesion localisation, Artif. Intell. Med. 107 (2020) 101880.

[29] W. Al-Dhabyani, M. Gomaa, H. Khaled, A. Fahmy, Dataset of breast ultrasound images, Data Brief 28 (2019) 104863, (2020).

[30] A.A. Taha, A. Hanbury, Metrics for evaluating 3D medical image segmentation: analysis, selection, and tool, BMC Med. Imaging 15 (1) (2015) 1–28.

[31] Z. Wang, E. Wang, Y. Zhu, Image segmentation evaluation: a survey of methods, Artif. Intell. Rev. 53 (2020) 5637–5674.

[32] A.J. Viera, J.M. Garrett, et al., Understanding interobserver agreement: the kappa statistic, Fam. Med. 37 (5) (2005) 360–363.

[33] Z. Huang, Y. Wei, X. Wang, W. Liu, T.S. Huang, H. Shi, Alignseg: Feature-aligned segmentation networks, IEEE Trans. Pattern Anal. Mach. Intell. 44 (1) (2021) 550–557.

[34] L.-C. Chen, Y. Zhu, G. Papandreou, F. Schroff, H. Adam, Encoder-decoder with atrous separable convolution for semantic image segmentation, in: Proceedings of the European Conference on Computer Vision, ECCV, 2018, pp. 801–818.

[35] V. Badrinarayanan, A. Kendall, R. Cipolla, Segnet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 39 (12) (2017) 2481–2495.

[36] Z. Gu, J. Cheng, H. Fu, K. Zhou, H. Hao, Y. Zhao, T. Zhang, S. Gao, J. Liu, Ce-net: Context encoder network for 2d medical image segmentation, IEEE Trans. Med. Imaging 38 (10) (2019) 2281–2292.

[37] J.V. Carter, J. Pan, S.N. Rai, S. Galandiuk, ROC-ing along: Evaluation and interpretation of receiver operating characteristic curves, Surgery 159 (6) (2016) 1638–1645.

[38] Z. Zhuang, N. Li, A.N. Joseph Raj, V.G. Mahesh, S. Qiu, An RDAU-NET model for lesion segmentation in breast ultrasound images, PLoS One 14 (8) (2019) e0221535.

[39] S.-Y. Lu, D.R. Nayak, S.-H. Wang, Y.-D. Zhang, A cerebral microbleed diagnosis method via featurenet and ensembled randomized neural networks, Appl. Soft Comput. 109 (2021) 107567.

[40] S. Lu, S.-H. Wang, Y.-D. Zhang, Detection of abnormal brain in MRI via improved AlexNet and elm optimized by chaotic bat algorithm, Neural Comput. Appl. 33 (2021) 10799–10811.

[41] Y. Wang, Y. Yao, Breast lesion detection using an anchor-free network from ultrasound images with segmentation-based enhancement, Sci. Rep. 12 (1) (2022) 14720.

[42] J. Bergstra, Y. Bengio, Random search for hyper-parameter optimization., J. Mach. Learn. Res. 13 (2) (2012).

[43] M. Kaveh, M.S. Mesgari, Application of meta-heuristic algorithms for training neural networks and deep learning architectures: A comprehensive review, Neural Process. Lett. 55 (4) (2023) 4519–4622.

[44] S. Li, C. Zhang, X. He, Shape-aware semi-supervised 3D semantic segmentation for medical images, in: Medical Image Computing and Computer Assisted Intervention–MICCAI 2020: 23rd International Conference, Lima, Peru, October 4–8, 2020, Proceedings, Part I 23, Springer, 2020, pp. 552–561.