

Research Article

Image Segmentation of Triple-Negative Breast Cancer by Incorporating Multiscale and Parallel Attention Mechanisms

Qian Zhang ¹, Junbiao Xiao ¹ and Bingjie Zheng ²

¹School of Computer Science, Zhongyuan University of Technology, Zhengzhou 450007, China

²Department of Radiology, Affiliated Cancer Hospital of Zhengzhou University, Zhengzhou 450008, China

Correspondence should be addressed to Qian Zhang; qzhang@ha.edu.cn and Bingjie Zheng; zlyzjb2100@zzu.edu.cn

Received 25 May 2023; Revised 6 November 2023; Accepted 21 November 2023; Published 16 December 2023

Academic Editor: Sadiq Hussain

Copyright © 2023 Qian Zhang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Breast cancer is a highly prevalent cancer. Triple-negative breast cancer (TNBC) is more likely to recur and metastasize than other subtypes of breast cancer. Research on the treatment of TNBC is of great importance, and accurate segmentation of the breast lesion area is an important step in the treatment of TNBC. Currently, the gold standard for tumor segmentation is still sketched manually by doctors, which requires expertise in the field of medical imaging and consumes a great deal of doctors' time and energy. Automatic segmentation of breast cancer not only reduces the burden of doctors but also improves work efficiency. Therefore, it is of great significance to study the automatic segmentation technique for breast cancer lesion regions. In this paper, a deep-learning-based automatic segmentation algorithm for TNBC images is proposed. The experimental data were dynamic contrast-enhanced magnetic resonance imaging TNBC dataset provided by the Cancer Hospital of Zhengzhou University. The experiments were analyzed by comparing several models with UNet, Attention-UNet, ResUNet, and SegNet and using evaluation indexes such as Dice score and Iou. Compared to UNet, Attention-UNet, ResUNet, and SegNet, the proposed method improved the Dice score by 2.1%, 1.54%, 0.88%, and 9.65%, respectively. The experimental results show that the proposed deep-learning-based TNBC image segmentation model can effectively improve the segmentation performance of TNBC tumors.

1. Introduction

In recent years, breast cancer has been one of the diseases that seriously affect women's lives. Data published in GLOBOCAN 2020 [1], prepared by the National Cancer Institute, shows that breast cancer is already one of the most frequently diagnosed types of cancer.

The incidence of breast cancer is also rising [2], and it can be classified into four molecular subtypes: Luminal A, Luminal B, triple-negative, and HER-2 overexpression.

For the treatment of breast cancer, early and accurate determination of a patient's molecular subtype is crucial for selecting the most appropriate treatment. To achieve this goal, many studies have applied convolutional neural network algorithms [3] for the prediction and classification of breast cancer molecular subtypes using imaging data such as dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) data [4], breast MRI data [3], and mammography images [5, 6]. These methods have significant potential contributions

to disease prognosis and treatment outcomes. Additionally, introducing a deep-learning model for channel dimensional feature reconstruction can further improve the ability to predict breast cancer molecular subtypes [7].

Among these four molecular typologies, triple-negative breast cancer (TNBC) is the most malignant, accounting for approximately 10%–20% of all breast cancers [8]. TNBC is defined as a cancerous tissue that is negative for three characteristic receptors on immunohistochemistry, i.e., negative for estrogen receptor, progesterone receptor, and proto-oncogene HER-2. It is highly malignant, and patients have faster tumor proliferation, higher chances of metastasis, and poorer patient prognosis [9]. TNBC has higher chances of metastasis, faster tumor proliferation in patients, high recurrence, postoperative recurrence, and poor tumor prognosis compared to several other molecular subtypes of breast cancer [9]. Therefore, compared with several other molecular typing, the treatment of TNBC requires more active and precise intervention, and when treatment is carried out, the

earlier the treatment, the better the effect will be, and the detection and diagnosis of early TNBC lesion area is very important [10]. There are many researchers conducting research on breast cancer diagnosis, such as Iqbal et al. [11], to conduct a review of data, methods, and other aspects.

For the treatment of TNBC, it is an effective measure to predict the condition and formulate the treatment plan based on the shape, size, and other characteristics of the lesion area through timely examination and discovery of the mass in the lesion area. However, the current gold standard for tumor segmentation outlining is still manual outlining by doctors. When sketching, the sketcher should have professional knowledge in the field of medical images, and it will consume a lot of time and energy of the doctor, and the situation of sketching will change subjectively with the doctor's situation [12]. Due to the small area of the lesion region and the complex background, it is easy to cause missed segmentation, mis-segmentation, and other conditions that affect the segmentation accuracy. Therefore, the study of automatic segmentation technology for TNBC is of great significance.

The difficulties and challenges of breast cancer segmentation. In this paper, an image segmentation model for TNBC that incorporates multiscale and parallel attention mechanisms (PAMs) is proposed in this paper.

This paper makes three main contributions:

- (1) For the characteristics of breast tumor region boundary mode, the lesion size is not fixed, and there are voids; this paper is designed to use multiscale convolution to obtain more comprehensive feature information according to different sizes of convolution and to increase the extraction of small target lesion region features by using void convolution.
- (2) There are complex high-density glandular tissues in the lesion region, and there is no clear boundary between the glandular tissues and the lesion region, thus causing interference to the segmentation. In this paper, we use the PAM to improve the segmentation of the lesion region by enhancing the degree of attention to the feature information of the lesion region.
- (3) To address the loss of image information caused by the sampling pooling layer under the UNet network structure, this paper uses 3×3 convolution with step size 2 instead of a pooling operation.

2. Literature Review

With the development of computers, the application of computer technology has been of great help in the field of medical segmentation, and the application of computer technology in the detection and segmentation of TNBC is becoming increasingly widespread. Traditional image segmentation algorithms are based on machine-learning algorithms, which are mainly divided into threshold-based segmentation, region-based segmentation, and edge-based segmentation. Kirthika et al. [13] achieved extractive segmentation of tumors by selecting a target value, searching for the optimal threshold value by using

the cuckoo-search (CS) algorithm until this target threshold is found, and using active-contour at the boundary based on convergent pixel enhancement. Arjmand et al. [14] achieved the differentiation of background, healthy tissue, and lesion areas by using a clustering method to segment breast tumors, using k-means clustering based on the characteristics of the lesion areas that differ from the surrounding background areas, and optimizing the initialized center of mass by the CS algorithm. Militello et al. [15] proposed a semiautomated interactive method based on the spatial fuzzy C-mean (FCM) algorithm for segmenting masses on DCE-MRI of the breast. Zebari et al. [16] proposed based on median, mean, and entropy. The segmentation task is performed by computing adaptive thresholds for binarization of breast images. Shen et al. [17] proposed to first extract the initial edges of breast tumors based on the grayscale distribution of ultrasound images using the grayscale threshold segmentation localization method and then correct the edges based on the gradient information of the image grayscale using the dynamic planning method, to accurately extract the edges of breast tumors from ultrasound images. Chakraborty et al. [18] proposed a multilevel threshold method controlled by gradient and intensity for the detection of mass focal regions, using gradient and intensity information to detect potential mass loci for the segment of breast tumors. Feng et al. [19] achieved segmentation of breast tumors by constructing the regional terms of the posterior probability-based activity contour model in the wavelet domain, while using the fuzzy velocity function to construct the activity contour. Jha et al. [20] used a hybrid segmentation FCMs and convolutional neural network (CNN) model for breast cancer risk prediction. Liu et al. [21] proposed an ultrasound image feature extraction algorithm combining edge features and morphological feature information, which has a significant effect on the extraction of edge features.

With the development of computers, classification detection methods can effectively identify different regions or structures in images and provide important prior information for segmentation algorithms. For example, Haq et al. [22] conducted classification research on breast tissue through a deep-learning model. Haq et al. [23] used supervised and semi-supervised feature selection techniques to detect breast cancer. Agbley et al. [24] classified breast tumors by magnifying histopathological images combined with fusion.

Although these methods provide an important aid to physicians in manual segmentation, traditional machine-learning algorithms also require high quality for medical image data. As the data size increases, these methods are not well adapted to medical image data with larger data sizes. When processing is performed, it still consumes a lot of labor and is costly, and the model generalization of this algorithm is insufficient. With the development of deep learning, deep-learning neural networks are gradually used for breast tumor segmentation. The segmentation of breast tumors by deep-learning network models reduces errors due to the subjectivity of medical personnel, reduces the workload of physicians, and does not rely on physicians' recognition of features [25, 26], while improving the speed and accuracy of physicians' diagnoses. There has been a great progress in semantic segmentation starting from

fully CNNs [27]. The UNet network model proposed by Ronneberger et al. [28] effectively mitigates the problem of deep feature loss in the upsampling phase by fusing deep and shallow features using cascade operations. Isensee et al. [29] proposed a deep learn-based segmentation method that can automatically configure itself. Gong et al. [30] designed an image edge reply assist task to design a DF-UNet model to enhance the characterization of lesion edges with a single CNN for image information acquisition, which is not comprehensive enough. Many researchers have focused on using multiscale neural networks instead of single CNNs, and Qin et al. [31] proposed integrating multiscale information fusion in the encoding path and constructing attention residual modules in the decoding path for problems such as low contrast at the boundary of the breast lesion region.

Early attention mechanisms were used in processing [32], and researchers introduced them to semantic segmentation tasks with good results. The attention mechanism enhances the extraction of features by the neural network by focusing the network attention on the key feature information through the attention module. Vaswani et al. [33] proposed a network architecture, transformer, which is entirely based on the attention mechanism and completely eliminates recursion and convolution. Oktay et al. [34] proposed the Attention-UNet model, the first time an attention mechanism was used in medical segmentation, which proposes attention gates added to jump connections to improve segmentation by suppressing irrelevant regions in the input image while highlighting salient features of specific local regions. Fan et al. [35] proposed a parallel reverse attention network for polyp segmentation. Luo et al. [36] used multiscale residual units to replace two adjacent convolutional blocks of UNet during downsampling to enhance the focus on morphological size differences, followed using cross-layer attention-guided networks to focus on focal regions during the upsampling phase and the introduction of void space pyramidal pooling as a bridging module for segmentation networks to enhance the characterization of lesions. Fu et al. [37], in CVPR2019, proposed a dual attention model by introducing spatial and channel latitude to weight all location features and selectively aggregating the features at each location. It makes capturing feature information more effective. Jha et al. [38] proposed a transformer-based residual network for segmentation. Liu et al. [39] used multiscale convolution as the basic module to extract features more comprehensively, using two-domain attention serials, thus setting weights on the features to improve the edge recognition and boundary-keeping ability of the network, thus improving the segmentation performance. Segmentation of TNBC is difficult due to the blurred boundaries between the tumor area and the surrounding normal tissue, and differences in the shape and size of the tumor [40]. Moreover, it is challenging to detect segmented masses, and low contrast, blurred borders, and varying sizes make segmentation difficult [41].

3. Materials and Methods

3.1. UNet Module. In 2015, the UNet network model was proposed in the paper, which was originally designed for

medical segmentation [42]. Initially, this aspect was solved in the cellular-level segmentation task, and it is widely used in the semantic segmentation direction with its superior segmentation effect. UNet consists of two parts: the encoding structure on the left and the symmetric decoding structure on the right. The encoding part of UNet is the extraction of features, which is achieved by alternating convolutional and downsampling pooling layers to gradually reduce the dimensionality of the feature map. The decoding part recovers the small feature map to the original graphic resolution large by convolutional layer and upsampling layer. And the shallow features and deep features are fused by cascading to obtain better target details. However, when UNet performs feature extraction, a single convolution cannot extract feature information at more scales, and there is an impact on fine segmentation.

3.2. Network Model Design. To address the problems in breast DCE-MRI tumor region segmentation, this paper proposes a network model for TNBC segmentation incorporating multiple scales of parallel attention mechanisms (MSPAMUNet) by improving the UNet network. The model uses an encoder–decoder backbone structure, and the model consists of three parts, the encoding layer, the attention module, and the decoding layer. The multiscale convolution module is introduced in the coding stage to realize the extraction of feature information at different scales by convolution of different sizes and to realize the effective extraction and fusion of semantic information in low-level feature maps and high-level feature maps so that the network can extract richer feature information. A parallel attention module is added between encoding and decoding to suppress irrelevant interference information, and attention operations are performed after the encoding stage to enhance the capture of location information and increase the extraction of feature information of tumor edge contours. The pooling operation in the UNet network structure causes a loss of image information, which has an impact on the segmentation. In this method, convolution is used in the pooling operation to compensate for this deficiency. So, in this paper, convolution is taken instead of the pooling operation, and the feature map is reduced to the segmented map in the final decoding stage. Because the convolution operation uses the features of the local neighborhood, it preserves more image information. The structure of the designed model is shown in Figure 1.

3.3. Multiscale Modules. The purpose of the encoding phase is to generate feature mappings by extracting the features and classifying the feature information. According to the characteristics of TNBC lesions, the lesion size is not fixed; there are cavities, sometimes there are large differences, the morphology is also diverse, and the intensity of the lesion area is also different [21]. A larger receptive field will lead to a decrease in localization accuracy, while a smaller receptive field will lead to a decrease in classification accuracy and a loss in feature information extraction. Therefore, in this paper, the feature information is extracted by using different scales of convolutional blocks to extract different levels of features through different sizes of perceptual fields and fuse

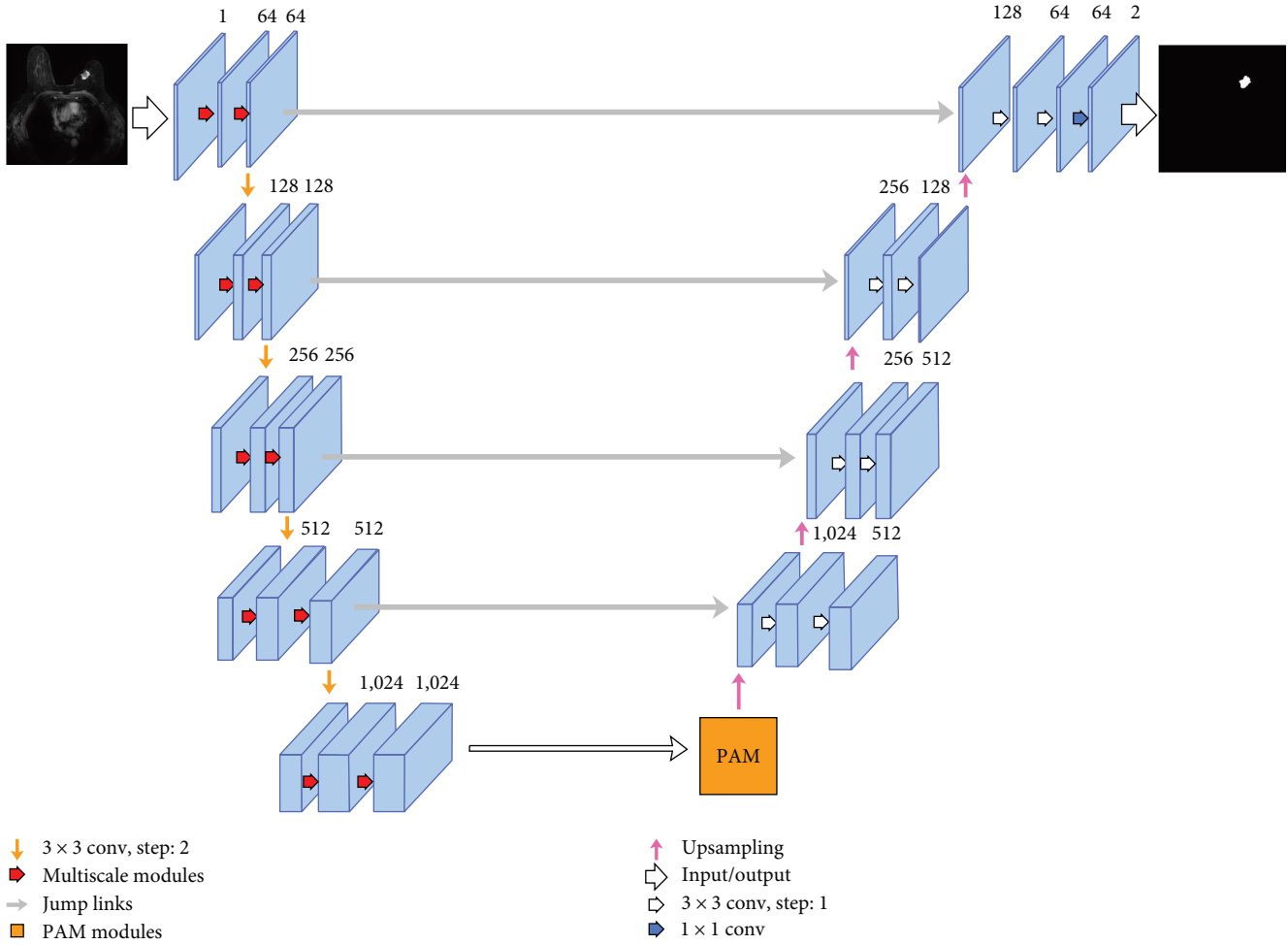


FIGURE 1: MSPAMUNet model.

them according to several different sizes of convolutional fusion operations to speed up the recognition speed while ensuring that the recognition accuracy and the number of parameters of the network are similar [43]. The feature of null convolution is used to increase the perceptual field without losing feature information. Eventually, more useful pixel information is extracted from the image. In this paper, we borrow the idea of GoogleNet to design the multibranch structure and borrow the idea of ASPP [44] to add the hole convolution in the convolutional branch and use multiple small convolutional kernels instead of large convolutional kernels. By replacing large convolutional kernels with multiple small convolutional kernels, the same receptive field can be obtained while the number of parameters can be greatly reduced. The same effect is obtained as in this paper by two 3×3 convolutional kernels instead of one 5×5 convolutional kernel with reduced parameters. The multiscale module in this paper consists of four branches, where the 1×1 convolution kernel is designed to adaptively change the number of channels. The first branch consists of a 1×1 convolution, 3×3 convolution; the second branch consists of a 1×1 convolution, a 3×3 convolution, and a 3×3 with a void rate of 3 and a 3×3 with a void rate of 5 in series, And

after each branch, a normalization (batch normalization, BN) operation after each branch, borrowing the idea of inceptionv2 [45], by adding BN layer can increase the perceptual field while speeding up the convergence speed, and finally fuse the three branches feature information.

The fourth branch is a channel directly connected to the output, which prevents the problem of gradient disappearance and gradient explosion as the network level deepens through this residual structure. Finally, the four branches are fused and activated by the rectified linear unit (ReLU) operation to obtain more comprehensive feature information. The structural model is shown in Figure 2.

3.4. Parallel Attention Module. In the encoding stage, the feature information obtained after fusing multiscale convolutional operations will be more comprehensive. However, different channels represent distinct feature information, and various categories of feature information have varying importance. Therefore, in order to highlight the lesion area and enhance attention towards it, we need to assign weights to the rich feature information. In this paper, by introducing the PAM module, the feature information weights are redistributed to make the network focus more on the region of the

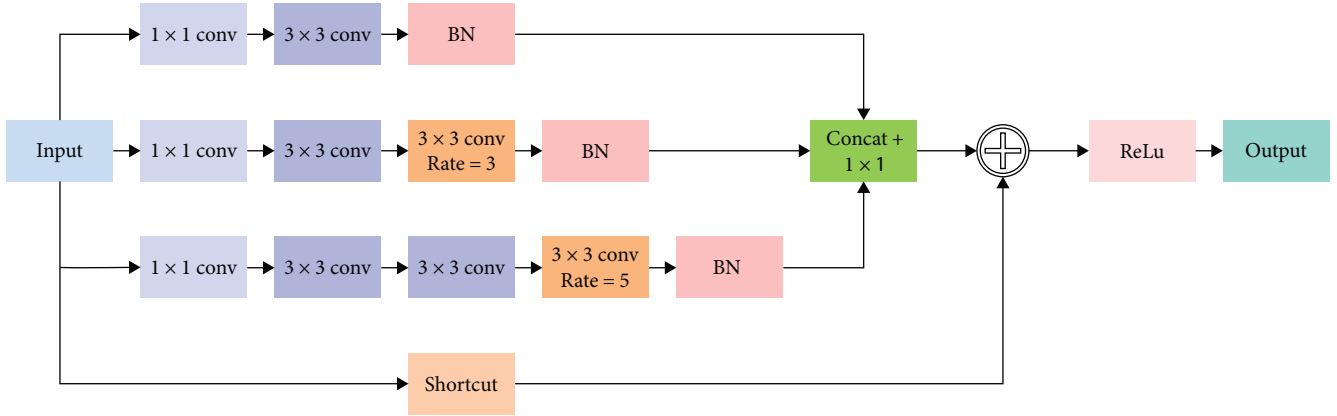


FIGURE 2: Multiple scale module.

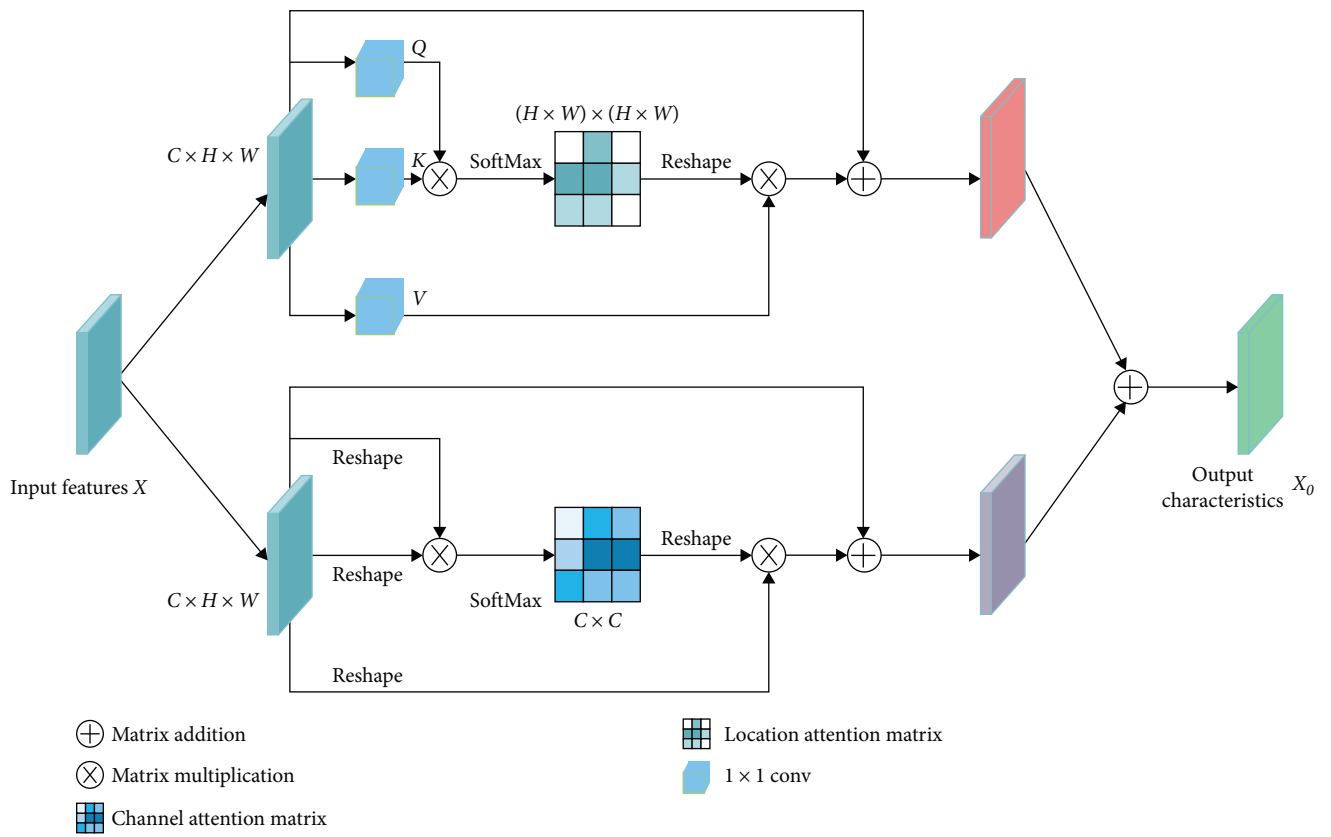


FIGURE 3: Parallel attention module.

breast tumor. The model structure is shown in Figure 3. This parallel attention module consists of two branches where input features X are $C \times H \times W$ with H representing height, W representing width, and C being the number of channels. The input features X in the branch of the location attention module are obtained after convolution by three 1×1 convolutions and latitude transformations to obtain $Q \in \mathbb{R}^{C \times H \times W}$, $K \in \mathbb{R}^{C \times H \times W}$, $V \in \mathbb{R}^{C \times H \times W}$, and normalized by the SoftMax layer to obtain the location attention feature map, as shown in Equation (1).

$$S_{ji} = \frac{\exp(Q_i \cdot K_j)}{\sum_{i=1}^N \exp(Q_i \cdot K_j)}, \quad (1)$$

where S_{ji} denotes the i th element, Q_i denotes the i th element of the matrix Q , and K_j , similarly, N is the number of elements in the channel, while V is the matrix multiplied with the obtained feature map S and subjected to the latitude transformation operation. Finally, the result is multiplied

by the scale factor g , where g is initialized to 0 and is a learnable parameter that is constantly updated during training and learning. The input feature X is summed, such as in Equation (2), to get the final output X_0 .

$$X_j = g \sum_{i=1}^N (S_{ji} V_i) + X_j, \quad (2)$$

where X_j denotes the j th element of the j th input X and V_i is the same.

The channel attention module branches, which differs from position attention in that for the input feature X , the input feature map is $H \times W \times C$. Then the channel attention feature map is computed directly, the reconstruction operation is performed, and then the matrix multiplication operation is performed afterward. Finally, the channel attention feature map is obtained by SoftMax and then weighted with the original channel features to obtain the final output feature information. Finally, the dual-attention features are fused in parallel to obtain more complementary features and to solve the problem of insufficient feature representation in a single feature vector [46]. The channel attention feature map $X \in R^{C \times C}$, as shown in Equation (3).

$$C_{ji} = \frac{\exp[X_i \cdot (X_j)^T]}{\sum_{i=1}^C \exp[X_i \cdot (X_j)^T]}. \quad (3)$$

C_{ji} indicates the impact of the i th channel on the j th channel, X_i indicates the i th element of matrix X , $(X_j)^T$ indicates the j th element of the transpose matrix, the formula and then matrix multiplication with the input feature X_i , in multiplying with the scale factor g , and summation operation with the input feature X_j , and finally obtain the output feature X_0 , as shown in Equation (4), where g is also initialized to 0 and is a learnable parameter that is constantly updated during training and learning.

$$X_j = g \sum_{i=1}^C (C_{ji} X_j) + X_j. \quad (4)$$

4. Results

4.1. Dataset. The data used in this experiment were the DCE-MRI TNBC image dataset provided by the Cancer Hospital of Zhengzhou University. The study was approved by the institutional ethics committee, and the need for written informed consent was waived by the institutional ethics committee due to the retrospective nature of the study.

We recruited a total of 31 women with a clinical diagnosis of breast cancer. These patients must be clinically diagnosed to make sure they have breast cancer. In the patients recruited, we selected the appropriate data for the experiment. Subsequently, we checked the quality of MRI scan

images and selected qualified images to ensure the accuracy and reliability of the analysis.

The dataset used shows the morphology of the lesion and the surrounding tissue. The patients are all female patients, and the data set also contains information about the images as well as patient information. Therefore, the patient breast cancer data were first desensitized before being used in the experimental study. The format of the dataset slices is Dicom format, with a size of 512×512 pixels, and each slice corresponds to a label bit depth of 8b. The experimental datasets were divided into training, validation, and test datasets of 1,621, 91, and 120, respectively, and the final datasets were approximately divided into training (90%), validation (5%), and test (5%), respectively. Figure 4 shows some triple-negative breast tumor samples, and the red marked parts are the location of the lesion area, and it can be seen from the figure that the triple-negative breast lesion area exists in shape and size.

4.2. Data Preprocessing. When training a model, it is not only the design of the model itself that is important but also the preprocessing of the data. The good or bad data preprocessing often affects the segmentation effect of the model. In this experiment, the following preprocessing operations were performed for TNBC DEC-MRI images.

- (1) In patient MRI breast tumor data, there are many frames that do not contain the target lesion region, so it is necessary to select appropriate frames from the MRI sequence to prevent these irrelevant frame sequences from interfering with the feature learning of the lesion region during training. In this method, the frame sequence suitable for training is selected by threshold filtering. Through cleaning, the final available data set is 458 pieces, which can reduce the training time during segmentation and reduce the interference of irrelevant feature information.
- (2) Medical image tumor segmentation is a typical classification of unbalanced segmentation. While in breast cancer MRI images, the percentage of irrelevant tissue in the total area of MRI slices of each patient is large, as shown in Figure 5, the focal area of normal patient MRI images has a small percentage of the whole image, and the background area has a large percentage, which seriously affects the segmentation of breast cancer focal area. To reduce the interference of background regions to the training, the network is able to learn feature information more fully. Its original image size is 512×512 pixels of DCE-MRI image, which is cropped from the original 512×512 pixels to a 128×128 pixel image containing the lesion area, and then the cropped image is normalized. Figure 5(a) shows the original data, Figure 5(b) shows the cropped processed graph, and Figure 5(c) shows the segmentation labels.
- (3) The amount of data also has a deep impact on how well deep learning is trained. Therefore, in this paper, the breast cancer image data were augmented, and

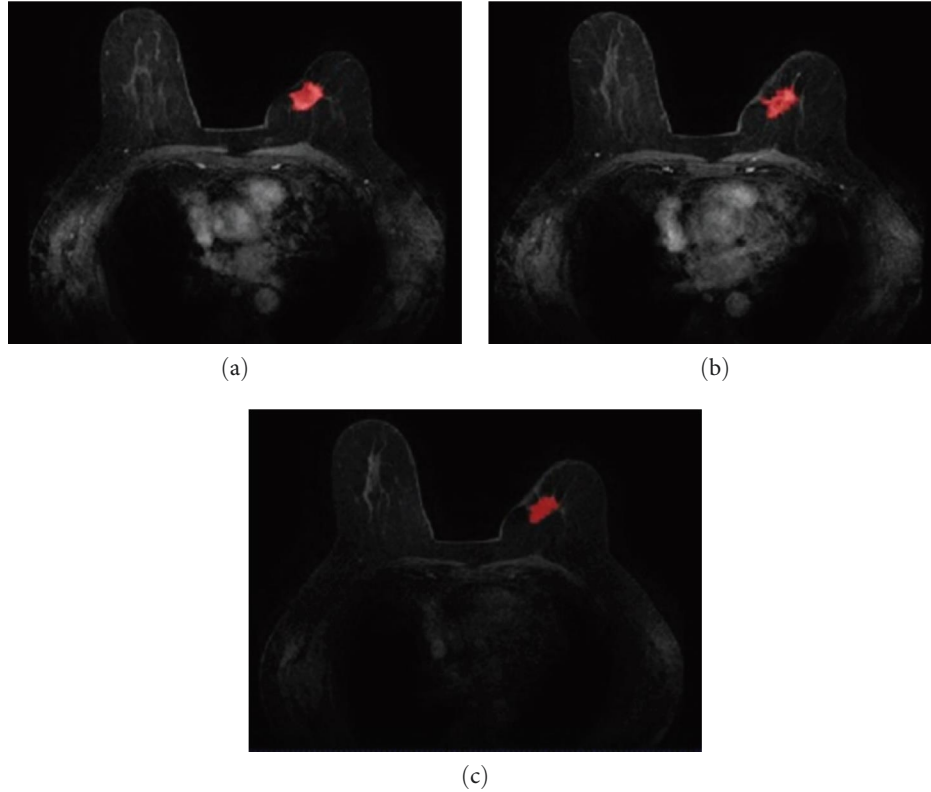


FIGURE 4: Partial MRI images of triple-negative breast cancer: (a) example Figure 1; (b) example Figure 2; (c) example Figure 3.

the original data were quadrupled by the affine transformation with 90° , 180° and left–right transformations, and the dataset was extended to 1,832 sheets, thus preventing overfitting due to too little data.

4.3. Experimental Setting and Evaluation Indexes. The experimental model uses the Pytorch framework with an AMD Ryzen 7 5,800H CPU and an NVIDIA GeForce RTX 3,050 GPU. In the experiment, the input image specification of the network model is $128 \times 128 \times 1$, the batch size is set to 16, and the number of training epochs is set to 100. The learning rate adopts a dynamic adjustment strategy, and its initial value is set to 0.01, and the learning rate is decayed according to the number of iterations during the training process, and the learning rate is set to be reduced to half of the original one each time, to the learning rate is reduced to half of the original rate each time to enhance the stability of the model, and the optimizer uses Adam to change the learning rate. In this paper, we adopt the evaluation offender method commonly used in medical segmentation and take the evaluation indexes of similarity coefficient (Dice), intersection ratio (Iou), accuracy (Acc), sensitivity (Sens), and specificity (Spec) to analyze the experimental results.

$$\text{Dice} = \frac{2TP}{FP + 2TP + FN}, \quad (5)$$

$$\text{Iou} = \frac{V_{\text{seg}} \cap V_{\text{gt}}}{V_{\text{seg}} \cup V_{\text{gt}}}, \quad (6)$$

$$\text{Acc} = \frac{TP + TN}{TP + FP + TN + FN}, \quad (7)$$

$$\text{Sens} = \frac{TP}{TP + FN}, \quad (8)$$

$$\text{Spec} = \frac{TN}{TN + FP}, \quad (9)$$

where TP indicates the overlap between the number of pixels in the region where the true target appears, and the number of pixels in the region where the predicted target appears, i.e., the number of pixels where the prediction is a positive case, and the true case is a positive case. FN denotes the number of pixels predicted to be negative but real positive cases, FP denotes the number of pixels predicted to be positive but real negative cases, TN denotes the number of pixels predicted to be positive and real negative cases, and V_{seg} and V_{gt} denote the result of model segmentation and the original label, respectively.

4.4. Loss Function. The choice of loss function also performs segmentation, and the target loss function plays a crucial role in determining whether the model can converge quickly. For medical image data, the segmentation task is only to tumor region and nontumor regions, and the number of pixels occupied by the tumor region is much smaller than the number of nontumor regions and the commonly used loss function is cross-entropy loss function. The cross-entropy loss function is usually used for classification, but when

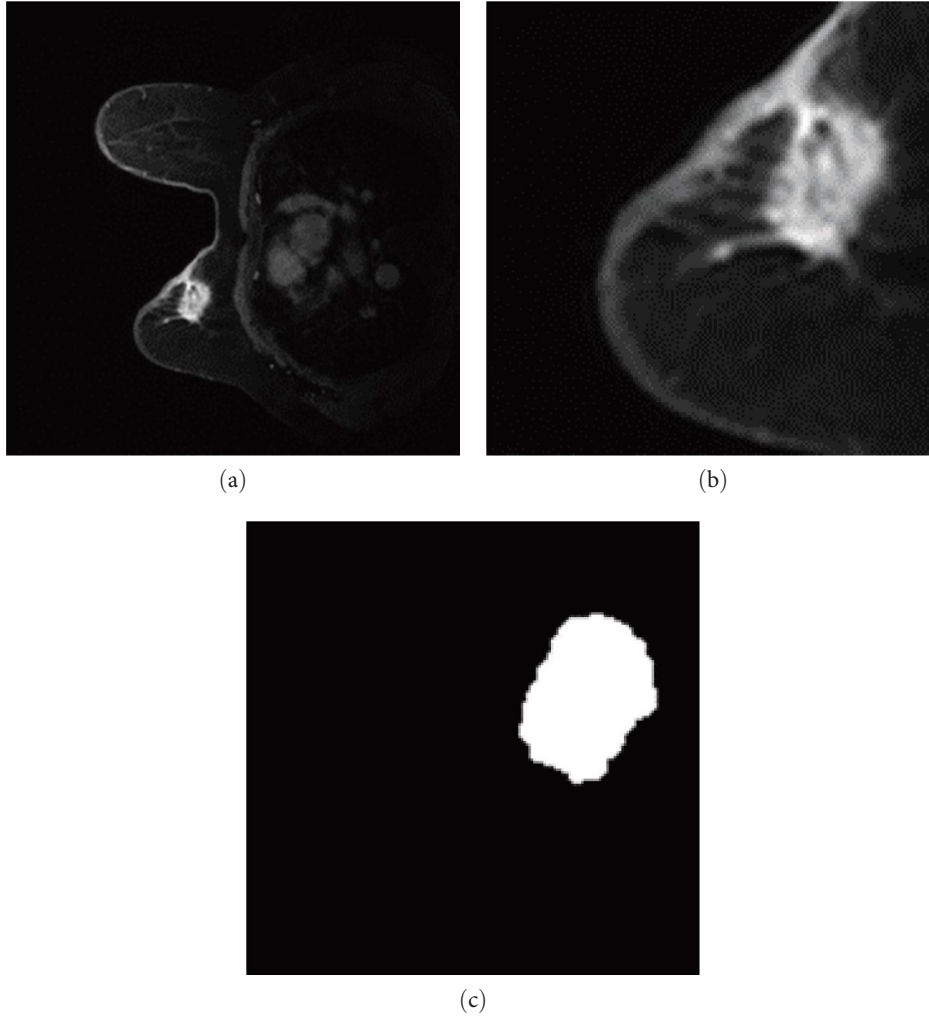


FIGURE 5: Raw MRI image preprocessing: (a) original image; (b) cropped image; (c) label.

segmenting breast tumors, it essentially also does a binary classification of background and foreground pixels [47], but due to the characteristics of medical images, this loss function has the obvious disadvantage that it makes the model biased toward the background, resulting in poor results. Therefore, the experiments in this paper use the weighted cross-entropy loss function, which is used to make the judgment of segmentation results by calculating entropy pixel-by-pixel and category-by-category. In this paper, we weight the positive samples by adding a weight parameter to each category based on the cross-loss to get better results for the data in the case of imbalance. The weighted cross-entropy formula is shown as Equation (10).

$$L_{CE} = - \sum_{c=1}^C \sum_{i=1}^R \sum_{j=1}^O W_c y_{i,j,c} \ln l_{i,j,c}, \quad (10)$$

where $l_{i,j,c}$ denotes the probability that the pixel in row i column j belongs to class c , where $c = \{0, 1\}$ denotes the two categories of background and tumor, and W_c denotes the weighting coefficient of each category; in this paper, the

weights focus on the foreground tumor parameter, and the weighting coefficient set for the tumor region is 0.7 and for the background region is 0.3. C denotes the category and R denotes the number of image rows. O denotes the number of image columns. $y_{i,j,c} \in \{0, 1\}$ is whether the pixel in row i and column j belongs to category c , and is 1 if it is, and 0 otherwise.

4.5. Ablation Experiments. In order to prove that several improvements in this paper can effectively improve the segmentation performance of the network model, we do ablation experiments to show that the multiscale module and PAM can improve the segmentation performance in different degrees. The fine-tuned UNet network (replacing the downsampling pooling operation with convolution) is used as the segmentation base reference baseline, and then the multiscale module and the PAM module are added separately for analysis and comparison, while ensuring that other influencing factors are the same.

Table 1 shows that the model segmentation performance is improved after adding the multiscale module adopted in this paper, which indicates that the extraction of feature

TABLE 1: Segmentation model on the data set ablation analysis results.

Model	Dice	Iou	Acc	Spec	Sens
BaseL	0.8338	0.7150	0.9913	0.9921	0.9615
BaseL + MS	0.8439	0.7300	0.9912	0.9922	0.9654
BaseL + PAM	0.8345	0.7160	0.9893	0.9825	0.9798
MSPAMUNet	0.8544	0.7457	0.9960	0.9927	0.9607

Note: Bolded font is the optimal value for each line.

TABLE 2: Results of the comparative analysis of the different models on the dataset.

Model	Dice	Iou	Acc	Spec	Sens
UNet	0.8334	0.7144	0.9888	0.9897	0.9780
Att-UNet	0.8390	0.7227	0.9898	0.9902	0.9848
ResUNet	0.8423	0.7276	0.9912	0.9926	0.9531
SegNet	0.7579	0.6102	0.9852	0.9854	0.9842
MSPAMUNet	0.8544	0.7457	0.9960	0.9927	0.9607

Note: Bolded font is the optimal value for each line.

information is stronger, and the representation of features is more accurate after adding this module. The analysis of the results showed that the evaluation of the model was also improved by adding the PAM alone, and the improvement was not too high by adding this module alone. Through the analysis, some of the labels in the labeled data had similar marginal breast tissue and lesion areas, and since some of the labels are of poor quality, it leads to some segmentation errors when adding the position attention module to the model, resulting in some of the evaluation coefficients being very low and leading to a decrease in the final average evaluation coefficient, but the evaluation indexes were still improved, proving the advantage of adding this module. Then, the two modules are fused and fine-tuned in the downsampling stage, and the obtained feature information of the lesion region is richer and more biased to the lesion region, which can finally improve the model performance overall. The Dice coefficients, Iou, and Acc of the method in this paper are improved, which verifies the superiority of the method in this paper.

4.6. Comparison Experiments. To verify the segmentation effect of the model in this paper, it is compared with the following classical network models: the UNet, Attention-UNet, ResUNet [48], and SegNet [49], respectively. Among them, the network structures of UNet, Attention-UNet (abbreviated here as Att-UNet), and ResUNet are related to the method proposed in this paper and are based on the UNet image segmentation method for segmentation. Among them, SegNet and this paper are concerned with feature extraction and processing by the method.

From Table 2, it can be seen that, by comparison, the metrics of this paper perform well and are higher than the segmentation effect of other networks in Dice coefficient and Acc metrics, and compared with UNet, Dice coefficient, Iou, Acc, and Spec are improved by 2, 3.13, 0.72, and 0.3 percentage points, respectively. Compared with the other three models, the Dice coefficient, Iou value, and Acc also improved,

where the Spec value ResUNet was the best. Through comparison experiments, it is verified that the performance of segmentation is largely enhanced by the method in this paper. Figure 6 shows the segmentation effect of several models.

5. Discussion

Through comparison, it can be seen that the model can both achieve the requirement of segmenting the background and the focal area, but due to the dense breast tissue and focal boundary-blurring, the UNet boundary segmentation is not as detailed as the model in this paper, and ResUNet and SegNet are mainly processed by upsampling the features, but they are not strong in capturing the smaller focal areas, which can easily cause missed segmentation or mis-segmentation. The Att-UNet model does not extract enough feature information for the focal area, resulting in missed segmentation, while in this paper, the fusion of spatial and positional attention is used in some edge parts to effectively suppress the interference information and achieve a greater degree of detail segmentation. It can be seen that each model is unable to segment the lesion region perfectly, because one of the reasons is that there are errors in the labels themselves, leading to mislearning during model training and learning, and the other is that the network model does not have enough learning ability for detail information and appears to have poor edge detail processing.

The box plots of the dice coefficients of several models are shown in Figure 7, and it also shows that the evaluation metrics of the method in this paper are more concentrated in dice coefficients compared to several other models.

It is proved that the model in this paper is more stable compared with other models, and the minimum dice coefficient value is higher than the minimum value of other models. Due to the quality of the test data, there are some image labels with poor quality, thus generating outliers, but after

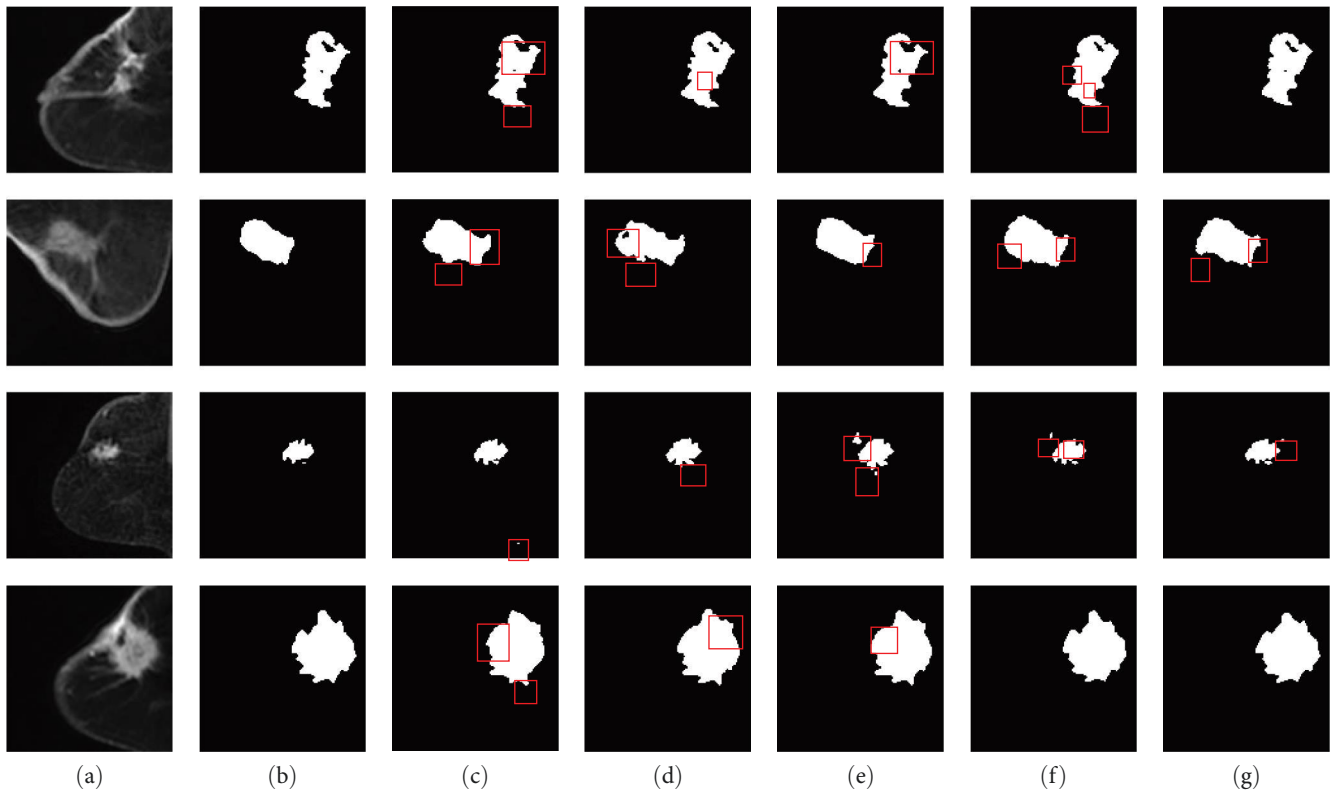


FIGURE 6: Five model segmentation results: (a) cropped image; (b) GT; (c) UNet; (d) Att-UNet; (e) ResUNet; (f) SegNet; (g) MSPAMUNet.

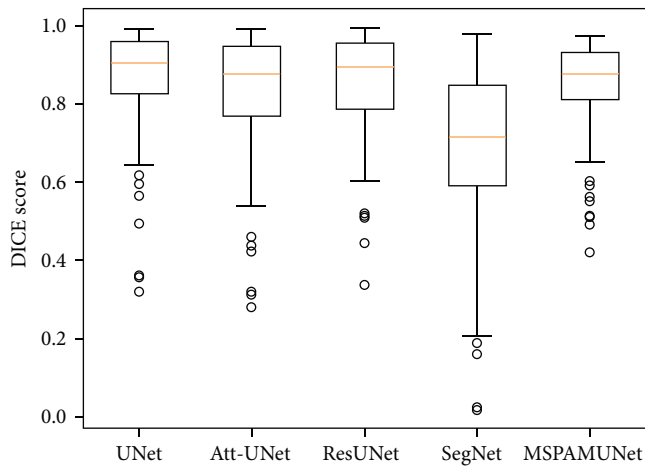


FIGURE 7: Dice box diagram of different models.

comparing the results of this method for poor quality data segmentation is also better than other models, and the overall dice values of the outliers are also on the upper side.

To further analyze the segmentation effect of the model on different data sets, in our experiments, we also use two databases to evaluate the performance of the proposed method. The public datasets INbreast [50] and CBIS-DDSM [51] datasets were combined into one dataset for split testing to obtain the corresponding evaluation metrics. The method in this paper still has good segmentation results on the public dataset. As shown in Figure 8, this paper adopts

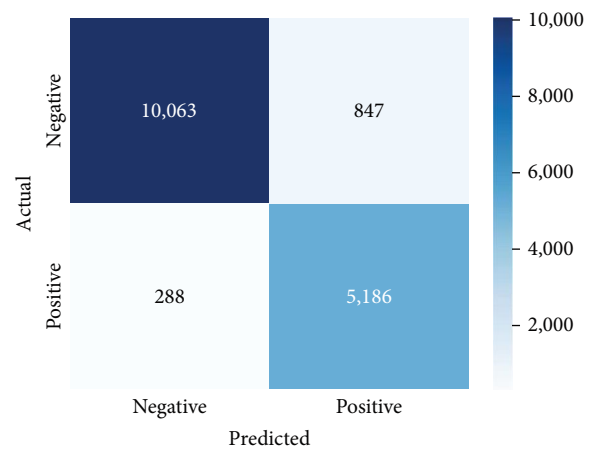


FIGURE 8: Random confusion matrix.

the method of random image extraction for image confusion matrix analysis, and the results show that the model performs well in the classification of positive and negative samples. Specifically, in Table 3, we can see that the average score of the model is better than other models in dice metrics, Iou, Acc, and Spec.

6. Conclusion and Future Works

In this paper, we propose a segmentation algorithm based on the fusion of multiscale PAMs for the MRI lesion segmentation problem of TNBC. Regarding multiscale information

TABLE 3: Results of comparative analysis of different models on public datasets.

Model	Dice	Iou	Acc	Spec	Sens
UNet	0.6535	0.485	0.8701	0.9255	0.8505
Att-UNet	0.6464	0.4775	0.8846	0.9278	0.8703
ResUNet	0.6408	0.4714	0.8631	0.9132	0.8677
SegNet	0.6395	0.4700	0.8883	0.9281	0.8618
MSPAMUNet	0.6852	0.5211	0.8914	0.9409	0.8249

Note: Bolded font is the optimal value for each line.

acquisition, this paper utilizes a combination of convolutions with different sizes, convolutions with varying null rates, and residuals to ensure the acquisition of receptive fields in different sizes and prevent the disappearance of network gradients. Additionally, we introduce the mechanism (fusion of position attention and channel attention) to enhance feature expression. Experimental validation was performed on the DCE-MRI breast cancer dataset provided by Henan Cancer Hospital, in which the Dice coefficient, Iou, and Acc reached 85.44%, 74.57%, and 99.60%, respectively, and better segmentation results were obtained.

In this paper, for the TNBC MRI lesion segmentation problem, we designed a fusion segmentation model different from others according to the advantages of the two modules and carried out experimental verification on the dataset. The experimental results demonstrate the effectiveness of the method proposed in this paper. The model in this paper also has some limitations, which may increase the complexity of the model after adding the attention module and is prone to overfitting when the dataset is small. Therefore, in the future work, we will continue to study and learn better network models and segmentation models with data generalization capabilities, so that they can be applied to more scenarios. By improving the applicability of the model and applying it to higher dimensional data, the segmentation performance of the model will be improved, and the model will be robust.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This work was supported by the Key Technologies R&D Program of Henan Province (no. 222102210281), the National Natural Science Foundation of China under Grant (nos. 62206252 and 82202270), the Key Scientific Research Projects of Universities in Henan Province (no. 24B520048), and the Zhongyuan Institute of Technology Superior Discipline Strength Enhancement Program (no. SD202230).

References

- [1] H. Sung, J. Ferlay, R. L. Siegel et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021.
- [2] J. Ferlay, M. Colombet, I. Soerjomataram et al., "Estimating the global cancer incidence and mortality in 2018: GLOBOCAN sources and methods," *International Journal of Cancer*, vol. 144, no. 8, pp. 1941–1953, 2019.
- [3] R. Ha, S. Mutasa, J. Karcich et al., "Predicting breast cancer molecular subtype with MRI dataset utilizing convolutional neural network algorithm," *Journal of Digital Imaging*, vol. 32, no. 2, pp. 276–282, 2019.
- [4] Y. Zhang, J.-H. Chen, Y. Lin et al., "Prediction of breast cancer molecular subtypes on DCE-MRI using convolutional neural network with transfer learning between two centers," *European Radiology*, vol. 31, no. 4, pp. 2559–2567, 2021.
- [5] V. K. Singh, H. A. Rashwan, S. Romani et al., "Breast tumor segmentation and shape classification in mammograms using generative adversarial and convolutional neural network," *Expert Systems with Applications*, vol. 139, Article ID 112855, 2020.
- [6] V. K. Singh, S. Romani, J. T. Barrena et al., "Classification of breast cancer molecular subtypes from their micro-texture in mammograms using a VGGNet-based convolutional neural network," in *International Conference of the Catalan Association for Artificial Intelligence*, pp. 76–85, 2017.
- [7] X.-X. Zhou, L. Zhang, Q.-X. Cui et al., "A channel-dimensional feature-reconstructed deep learning model for predicting breast cancer molecular subtypes on overall b-value diffusion-weighted MRI on-weighted MRI," *Journal of Magnetic Resonance*, 2023.
- [8] Z. Z. Hao and H. E. Shen, "Current situation of neoadjuvant therapy for triple negative breast cancer," *Lingnan Modern Clinics in Surgery*, vol. 22, no. 2, pp. 205–209, 2022.
- [9] L. Liu and J. Zhang, "Research progress on targeted therapy of triple negative breast cancer," *Tianjin Pharmaceutical*, vol. 46, no. 12, pp. 1363–1368, 2018.
- [10] M. Schmidt and A.-S. Heimes, "Immunomodulating therapies in breast cancer—from prognosis to clinical practice," *Cancers*, vol. 13, no. 19, Article ID 4883, 2021.
- [11] M. S. Iqbal, W. Ahmad, R. Alizadehsani, S. Hussain, and R. Rehman, "Breast cancer dataset, classification and detection using deep learning," *Healthcare*, vol. 10, no. 12, Article ID 2395, 2022.
- [12] M. Bretthauer and M. Kalager, "Principles, effectiveness and caveats in screening for cancer," *British Journal of Surgery*, vol. 100, no. 1, pp. 55–65, 2013.
- [13] A. Kirthika, N. S. M. Raja, R. Sivakumar, and S. Arunmozhi, "Assessment of tumor in breast MRI using Kapur's thresholding and active contour segmentation," in *International Conference on System, Computation, Automation and Networking (ICSCAN)*, pp. 1–4, 2020.
- [14] A. Arjmand, S. Meshginiet, R. Afrouzian, and A. Farzamnia, "Breast tumor segmentation using k-means clustering and cuckoo search optimization," in *9th International Conference on Computer and Knowledge Engineering (ICCCKE)*, pp. 305–308, 2019.
- [15] C. Militello, L. Rundo, M. Dimarco et al., "Semi-automated and interactive segmentation of contrast-enhancing masses on breast DCE-MRI using spatial fuzzy clustering," *Biomedical*

- Signal Processing and Control*, vol. 71, Article ID 103113, 2022.
- [16] D. A. Zebari, D. Q. Zeebaree, A. M. Abdulazeez, H. Haron, and H. N. A. Hamed, "Improved threshold based and trainable fully automated segmentation for breast cancer boundary and pectoral muscle in mammogram images," in *IEEE Access*, vol. 8, pp. 203097–203116, IEEE, 2020.
 - [17] J. L. Shen, Y. Y. Wang, and Y. Wang, "Boundary extraction of ultrasonic breast tumor image based on gray-level threshold segmentation and dynamic programming," *Space Medicine & Medical Engineering*, vol. 18, no. 4, pp. 281–286, 2005.
 - [18] J. Chakraborty, S. Mukhopadhyay, V. Singla, N. Khandelwal, and R. M. Rangayyan, "Detection of masses in mammograms using region growing controlled by multilevel thresholding," in *25th IEEE International Symposium on Computer-Based Medical Systems (CBMS)*, pp. 1–6, 2012.
 - [19] B. Feng, C. L. Li, Z. Li, and Z. S. Liu, "Breast cancer related LVI status prediction based on active contour model and radiomics," *Journal of Northeastern University (Natural Science)*, vol. 41, no. 2, pp. 193–199, 2020.
 - [20] S. Jha, S. Ahmad, A. Arya et al., "Ensemble learning-based hybrid segmentation of mammographic images for breast cancer risk prediction using fuzzy c-means and CNN model," *Journal of Healthcare Engineering*, vol. 2023, Article ID 1491955, 18 pages, 2023.
 - [21] Y. Liu, L. Ren, X. Cao, and Y. Tong, "Breast tumour recognition based on edge feature extraction using support vector machine," *Biomedical Signal Processing and Control*, vol. 58, Article ID 101825, 2020.
 - [22] A. U. Haq, J. P. Li, I. Khan et al., "DEBCM: deep learning-based enhanced breast invasive ductal carcinoma classification model in IoMT healthcare systems," *IEEE Journal of Biomedical and Health Informatics*, 2022.
 - [23] A. U. Haq, J. P. Li, A. Saboor et al., "Detection of breast cancer through clinical data using supervised and unsupervised feature selection techniques," *IEEE Access*, vol. 9, pp. 22090–22105, 2021.
 - [24] B. L. Y. Agbley, J. P. Li, A. U. Haq et al., "Federated fusion of magnified histopathological images for breast tumor classification in the internet of medical things," *IEEE Journal of Biomedical and Health*, 2023.
 - [25] A. Esteva, A. Robicquet, B. Ramsundar et al., "A guide to deep learning in healthcare," *Nature Medicine*, vol. 25, no. 1, pp. 24–29, 2019.
 - [26] L. Zhou, Z. Zhang, Y.-C. Chen, Z.-Y. Zhao, X.-D. Yin, and H.-B. Jiang, "A deep learning-based radiomics model for differentiating benign and malignant renal tumors," *Translational Oncology*, vol. 12, no. 2, pp. 292–300, 2019.
 - [27] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," pp. 3431–3430, 2015.
 - [28] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: convolutional networks for biomedical image segmentation," pp. 234–241, 2015.
 - [29] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, pp. 203–211, 2021.
 - [30] R. L. Gong, J. Shi, and J. Wang, "Hybrid supervised dual-channel feedback U-Net for segmentation of breast ultrasound images," *Journal of Image and Graphics*, vol. 25, no. 10, pp. 2206–2217, 2020.
 - [31] C. B. Qin, Z. Y. Song, J. Y. Zeng L. F. Tian, and F. Li, "Deeply supervised breast cancer segmentation combined with multi-scale and attention-residuals," *Optics and Precision Engineering*, vol. 29, no. 4, pp. 877–895, 2021.
 - [32] L. Shi, Y. Wang, Y. Cheng, and R. B. Wei, "Review of attention mechanism in natural language processing," *Data Analysis and Knowledge Discovery*, vol. 4, no. 5, pp. 1–14, 2020.
 - [33] A. Vaswani, N. Shazeer, N. Parmar et al., "Attention is all you need, advances in neural information processing systems," 2017.
 - [34] O. Oktay, J. Schlemper, L. L. Folgoc et al., "Attention U-net: learning where to look for the pancreas," 2018.
 - [35] D. P. Fan, G. P. Ji, T. Zhou et al., "PraNet: parallel reverse attention network for polyp segmentation," in *International Conference on Medical Image Computing and Computer-assisted Intervention*, pp. 263–273, Springer International Publishing, Cham, 2020.
 - [36] S. Q. Luo, J. Y. Chen, and H. J. Li, "Multiscale residual UNet based on attention mechanism to realize breast cancer lesion segmentation," *Journal of Computer Applications*, vol. 42, no. 3, pp. 818–824, 2022.
 - [37] J. Fu, J. Liu, H. Tian et al., "Dual attention network for scene segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (IEEE)*, pp. 3146–3154, 2019.
 - [38] D. Jha, N. K. Tomar, V. Sharma, and U. Bagci, "TransNetR: transformer-based residual network for polyp segmentation with multi-center out-of-distribution testing," 2023.
 - [39] X. Liu, Z. Lu, B. Li, B. Wang, and D. Wang, "Segmentation algorithm of breast tumor in dynamic contrast-enhanced magnetic resonance imaging based on network with multi-scale residuals and dual-domain attention," *Journal of Electronics & Information Technology*, vol. 45, no. 5, pp. 1–12, 2022.
 - [40] H. Y. Zhou, B. Feng, F. F. Qi, Z. S. Liu, and W. S. Long, "Combining MRF energy and DCE-MRI time-domain features for breast tumors segmentation algorithm," *Acta Photonica Sinica*, vol. 50, no. 6, Article ID 0610002, 2021.
 - [41] A. Oliver, J. Freixenet, J. Martí et al., "A review of automatic mass detection and segmentation in mammographic images," *Medical Image Analysis*, vol. 14, no. 2, pp. 87–110, 2010.
 - [42] T. Zhou, Y. Dong, B. Huo, S. Liu, and Z. J. Ma, "U-Net and its applications in medical image segmentation: a review," *Journal of Image and Graphics*, vol. 26, no. 9, pp. 2058–2077, 2021.
 - [43] D. F. Yuan, C. F. Chen, and F. M. Dong, "Research on residual network of image recognition based on multiscale split," *Computer Engineering*, vol. 48, no. 5, pp. 258–262, 2022.
 - [44] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, "DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 4, pp. 834–848, 2018.
 - [45] S. Ioffe and C. Szegedy, "Batch normalization: accelerating deep network training by reducing internal covariate shift," in *Proceedings of the 32nd International Conference on Machine Learning*, vol. 37, pp. 448–456, JMLR, 2015.
 - [46] A. J. Ma, P. C. Yuen, and J. H. Lai, "Linear dependency modeling for classifier fusion and feature combination, Institute of Electrical and Electronics Engineers," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 5, pp. 1135–1148, 2013.
 - [47] S. Aydin, N. Arica, E. Ergul, and O. Tan, "Classification of obsessive compulsive disorder by EEG complexity and

- hemispheric dependency measurements,” *International Journal of Neural Systems*, vol. 25, no. 3, Article ID 1550010, 2015.
- [48] X. Xiao, S. Lian, Z. Luo, and S. Li, “Weighted res-unet for high-quality retina vessel segmentation,” in *2018 9th International Conference on Information Technology in Medicine and Education (ITME)*, pp. 327–331, 2018.
- [49] V. Badrinarayanan, A. Kendall, and R. Cipolla, “SegNet: a deep convolutional encoder–decoder architecture for image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 12, pp. 2481–2495, 2017.
- [50] I. C. Moreira, I. Amaral, I. Domingues, A. Cardoso, M. J. Cardoso, and J. S. Cardoso, “INbreast: toward a full-field digital mammographic database,” *Academic Radiology*, vol. 19, no. 2, pp. 236–248, 2012.
- [51] R. S. Lee, F. Gimenez, A. Hoogi, K. K. Miyake, M. Gorovoy, and D. L. Rubin, “A curated mammography data set for use in computer-aided detection and diagnosis research,” *Scientific Data*, vol. 4, Article ID 170177, 2017.