Pattern Recognition 124 (2022) 108427

Contents lists available at ScienceDirect

Pattern Recognition

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

Segmentation information with attention integration for classification of breast tumor in ultrasound image

Yaozhong Luo[a], Qinghua Huang[b,*], Xuelong Li[b]

[a] School of Electronic and Information Engineering, South China University of Technology, 510641, China
[b] School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an, Shaanxi 710072, China

ARTICLE INFO

Article history:
Received 16 March 2021
Revised 25 October 2021
Accepted 8 November 2021
Available online 11 November 2021

Keywords:
Computer-aided diagnosis
Breast ultrasound
Deep convolution neural network
Feature combination

ABSTRACT

Breast cancer is one of the most common forms of cancer among women worldwide. The development of computer-aided diagnosis (CAD) technology based on ultrasound imaging to promote the diagnosis of breast lesions has attracted the attention of researchers and deep learning is a popular and effective method. However, most of the deep learning based CAD methods neglect the relationship between two vision tasks tumor region segmentation and classification. In this paper, taking into account some prior knowledges of medicine, we propose a novel segmentation-to-classification scheme by adding the segmentation-based attention (SBA) information to the deep convolution network (DCNN) for breast tumors classification. A segmentation network is trained to generate tumor segmentation enhancement images. Then two parallel networks extract features for the original images and segmentation enhanced images and one channel attention based feature aggregation network is to automatically integrate the features extracted from two feature networks to improve the performance of recognizing malignant tumors in the breast ultrasound images. To validate our method, experiments have been conducted on breast ultrasound datasets. The classification results of our method have been compared with those obtained by eleven existing approaches. The experimental results show that the proposed method achieves the highest Accuracy (90.78%), Sensitivity (91.18%), Specificity (90.44%), F1-score (91.46%), and AUC (0.9549).

1. Introduction

Breast cancer is one of the most common forms of cancer among women worldwide and has brought a great threat to women's health. Because of the complex etiology that makes it difficult for the medical community to provide effective preventive measures, early detection and diagnosis are the keys to reducing the death rate (40% or more) [1]. Mammography, magnetic resonance imaging (MRI), Mammography, and ultrasound are the most commonly used techniques for the discovery and diagnosis of breast cancer. In recent years, more and more attention has been paid to ultrasound imaging because it has the following merits: no radiation, faster imaging, higher sensitivity and accuracy, and lower cost [2]. However, ultrasound is an operator-dependent method. Moreover, reading ultrasound images requires well-trained and experienced radiologists and the diagnostic results of physicians with different medical attainments are very different. Because of the individual difference among the patients and the poor image quality, radiologists cannot always make correct judgments in the diagnosis of breast ultrasound images. To reduce the inter-observer variation among different radiologists and help generate more reliable and accurate diagnostic conclusions, the computer-aided diagnosis (CAD) system has been proposed [3,4]. With the rapid development of computer technology and artificial intelligence, the CAD has become one of the hotspots in modern medical image research, and its clinical value has been demonstrated in real diagnoses to assist radiologists in breast tumor classification and recognition.

Generally, the key of traditional breast tumor classification method includes two aspects, which are feature extraction and machine learning classifier [5]. In the traditional method, extracted features are explicitly designed or handcrafted, mainly texture features and morphological features [6]. Commonly used texture feature extraction methods include gray level co-occurrence matrix (GLCM) [7], auto-covariance coefficients [9], spatial gray-level dependence (SGLD) [10], correlation of the co-occurrence matrix [11], global texture feature [12], wavelet transform [13], shearlet-based features [14] and so on. Among them, there are some texture extraction methods, such as GLCM and wavelet transform, which calculate the matrix or transform domain representation firstly, and then calculate multiple statistic values as texture features. However, some methods, such as the global texture feature, will be di-

* Corresponding author.
E-mail address: enicarwhw@qq.com (Q. Huang).

https://doi.org/10.1016/j.patcog.2021.108427
0031-3203/

rectly represented as texture features. Common morphological features include margin sharpness [8], depth to width ratio [11,15], branch pattern [16], circularity [11,16], and so on. Although the number of handcrafted features has reached tens of thousands, these features are shallow and of low order, which may not fully characterize the heterogeneous patterns within the tumor. Moreover, the handcrafted features are highly dependent on the experience of experts and depend greatly on a good understanding of tumors at the radiological level.

In recent years, popular deep learning methods are capable of automatic rather than manual feature extraction. The deep convolution network (DCNN) has widely been recognized as a reliable approach to learning predictive features for image classification, object detection, semantic segmentation, and other computer vision tasks. At present, DCNN has also become popular in the community of medical imaging analysis [17,18]. For breast ultrasound image classification, several studies have been proposed and obtained better classification results than that of the classical method. In the research of breast ultrasound classification, there are two kinds of deep learning methods. One is the training of CNNs from scratch. However, training a CNN from scratch requires a large dataset because the CNN is often deep to obtain better performance. The amount of ultrasound datasets is usually small, which is an important characteristic of medical data analysis. Training a deep CNN from scratch on a small amount of data set will lead to over-fitting, resulting in poor performance. Some studies [19–21] put attention into training shallow networks for breast ultrasound classification. Masud et al. [19], Wang et al. [20], and Zeimarani et al. [21] proposed custom convolutional neural networks for breast ultrasound classification, respectively. Among them, the methods in [19–21] have mainly one, two, and four convolution layers, respectively, combined with fully connected, pooling, dropout, and batch normalization layers. Another kind of method is based on transfer learning. The network pre-trained on a large labeled dataset is applied for the breast ultrasound dataset. The pre-trained network can be directly used for feature extraction, and can also be fine-tuned on the available datasets. In 2019, Daoud et al. [22], Hijab et al. [23] and Cao et al. [24] applied pre-trained AlexNet, VGG16 network, and DenseNet121 and fine-tuned on the ultrasound images to discriminate between benign and malignant breast masses, respectively. Fujioka et al. [25] applied GoogleNet based DCNN architectures in discriminating between benign and malignant breast masses from breast ultrasound images and had equal or better diagnostic performance compared to radiologists. Byraa et al. [26] designed a matching layer to convert grayscale ultrasound images to RGB images and fine-tuned pre-trained VGG19 network for breast tumor classification. Tanaka et al. [27] constructed an ensemble network by combining VGG19 and ResNet152 fine-tuned on breast ultrasound images for breast mass classification. Zhuang et al. [28] employed the image decomposition method to obtain fuzzy enhanced and bilateral filtered images input to a pre-trained VGG model for breast mass classification and got good improved results. At present, the more efficient deep learning method is to apply different pre-trained models then fine-tuning on breast ultrasound dataset.

However, the current ultrasound images classification methods based on deep learning still have some problems. One significant difficulty is that the amount of ultrasound datasets is usually small, which is an important characteristic of medical data analysis. Compared with natural images, ultrasound data are more difficult to collect and the use of data must be agreed upon by the patients. Labeling ultrasound images is extremely strict and requires experienced radiologists or biopsies. At the same time, the types of breast tumors, as well as corresponding imaging results, may be different and diverse. A small amount of dataset may be unable to cover all types. How to extract the most effective and robust feature of breast ultrasound is the most important content to improve the diagnosis performance.

In addition, the development of medicine and imaging provides us with a lot of diagnostic knowledge. In the analysis of a small amount dataset, it is an effective method to make up for the insufficient number of samples to introduce prior knowledge into the design of a deep learning network for breast tumor classification. Considering actual clinical diagnosis, radiologists judge the type of tumors according to the clinical information of tumors, such as location, shape, size, depth to width ratio, circularity, edge integrity, lobulation, margin sharpness, and other factors [29]. In 2003, the Breast Imaging-Reporting and Data System (BI-RADS) [30] was developed by the American College of Radiology and the formed scoring data concluded experts' practical experience. The scoring metrics are summarized from massive data and have been widely applied in the classification of breast cancer as an international standard.

However, because the deep learning model is an automatic feature extraction method, the extracted features are not completely consistent with the commonly accepted BI-RADS, and more rely on the characteristics of the training data. When analyzing a small amount of dataset, it is difficult to learn the most representative features in the deep learning model. In BI-RADS, ten of the eighteen scoring metrics, including shape, orientation, margin integrality, margin ambiguity, angular, microlobulated, spiculated, posterior feature, calcification in mass, and architectural distortion are strongly associated with tumor segmentation. These metrics are an important basis for the clinical diagnosis of breast ultrasound [31]. If we can guide the deep learning model to focus more on these features, it would be very helpful for the classification of breast cancers. Therefore, the segmentation of tumors can further guide the feature extraction related to tumor classification and improve the performance [32]. However, in the previous deep learning classification methods, only the identified region of interest (ROI) image containing breast tumor is taken as the input of the network, and no tumor contour prior information is added. These methods can extract but not enhance tumor segmentation related clinical criteria, which makes the classification performance difficult to improve. In clinical diagnosis, radiologists can focus on the tumor region and automatically eliminate the interference of normal tissues, which is impossible for the existing deep learning CAD methods. Therefore, it is necessary to design the network and enhance the features containing the segmentation information from the standards of clinical diagnosis.

Hence, we propose a novel architecture to enhance the tumor segmentation and the related clinical diagnostic information in feature extraction. To obtain the segmentation priori information, we propose a segmentation-to-classification framework to realize these goals. Unlike the traditional DCNN framework, the system adds a parallel feature extraction branch that contains segmentation prior information. The segmentation result is obtained by the segmentation network and applied as the input of the convolution network to extract features. Then a feature aggregation network based on an attention mechanism is used to synthesize the features of two parallel feature extraction branches. Further details for considering the framework are described as follows.

First, adding a parallel branch based on segmentation for breast tumor classification is desirable. As shown in [19–28], the commonly applied breast ultrasound classification deep network did not consider the role of segmentation prior information in breast classification. Due to the lack of training samples, the most robust and relevant features may not be extracted. Fortunately, there are a series of solutions in the classical breast tumor classification method. In classical methods, feature extraction based on segmentation results plays an important role in breast tumor classification [33], which is closely related to a series of clinical features and

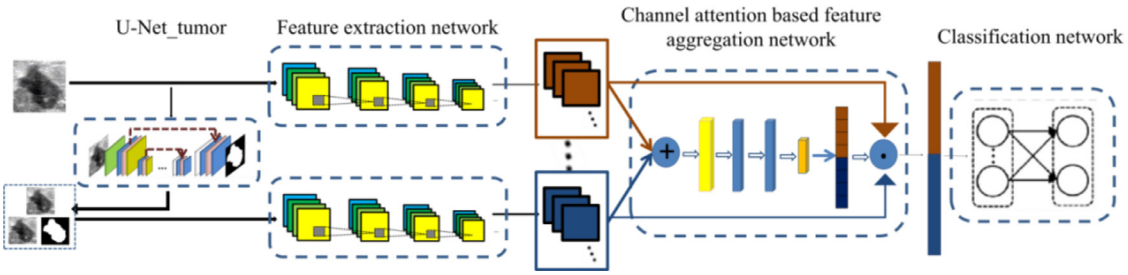

Fig. 1. The flowchart of the proposed algorithm. The framework consists of a segmentation model, two parallel feature extraction networks, a channel attention based feature aggregation network, and a classification network.

breast discrimination. Segmentation results for benign and malignant breast tumor classification have been proved to be an effective feature to improve performance. Therefore, it is necessary to add a feature extraction branch including segmentation prior. Two parallel branches are used for feature extraction of ultrasound tumors, which enriches the feature extraction and can get more effective features for classification.

Secondly, the channel attention based feature aggregation network is necessary. There is redundant information in feature extraction of two parallel networks, so it is helpful for breast image classification to use an aggregation network to aggregate the features extracted by two parallel networks, enhance useful features, remove redundant features and get more effective feature representation. The channel attention network can analyze the importance of different channels of two feature networks, and weigh the different channel features to get more effective features.

In summary, the main contributions of this paper are summarized as follows:

(1) A new classification scheme is proposed in which the segmentation information is put into the network as a priori information for breast ultrasound diagnosis. The segmentation network is trained to obtain the tumor regions and form segmentation enhanced images, and the features of the segmentation enhanced images and original images are extracted in parallel for classification.
(2) An channel attention method is proposed for feature aggregation of two parallel networks to enhance features useful for classification and get more effective feature representation. Because the segmentation enhanced information is added to the network, and the channel attention method is carried out to get the most informative features, our approach is called a segmentation-based attention network (SBANet).
(3) The experimental results demonstrate that the segmentation-to-classification framework is able to improve diagnostic performance.

This paper is organized as follows. Section 2 introduces the proposed method in detail. Next, the experimental results and comparisons among different methods are presented in Section 3. Finally, we provide some discussion and draw the conclusion in Section 4.

2. Methods

In this section, the overall architecture of the proposed method is illustrated, including the architecture of the segmentation model, the two parallel feature extraction networks, the channel attention based feature aggregation network, and the final classification network.

2.1. Overall architecture

The proposed architecture is shown in Fig. 1. Unlike the traditional framework, our system is a segmentation-to-classification framework that consists of four components:

(1) Segmentation model

To obtain the shape and morphology priori for the subsequent classification tasks and avoid the huge workload of professional radiologists labeling in practical use, a segmentation model is firstly trained to estimate the edges of the tumors. The segmentation network receives the sample of ultrasound images as input and outputs the segmented binary images. After processing by the morphological opening and closing algorithms, we obtain segmented images containing the tumors and the tissues.

(2) Two parallel feature extraction networks

After obtaining the segmentation results, we combine the original image and the segmented image to form a segmentation enhanced image. To obtain a more robust image representation, two parallel feature extraction networks are designed corresponding to the original images and segmentation enhancement images, respectively. The features are obtained from the output of two parallel feature extraction networks.

(3) Feature aggregation of two networks

After feature extraction, a feature aggregation network is designed to focus on the relationship between different channel feature maps obtained by two networks and automatically learn the importance of different channel features to generate a more effective feature representation. The feature maps are multiplied by the adaptive importance weights to obtain the more effective feature representation.

(4) Classification network

After obtaining the feature representation of the images, a classification network is trained to obtain the final classification of breast ultrasound images.

Our breast tumor classification scheme involves three problems: (1) How to automatically obtain the tumor segmentation and form an effective segmentation enhanced image for feature extraction; (2) How to extract features on the ultrasound images and avoid over-fitting; (3) How to automatically learn the importance of different channel features extracted by two networks and generate more effective feature representation. Sections 2.2, 2.3, and 2.4 will introduce the solutions to the problems, respectively.

2.2. Segmentation model

Tumor segmentation plays an important role in ultrasound image classification and is taken as the pre-procedure for feature extraction and image representation in the classical diagnosis framework [34,35]. Owing to its precise segmentation ability on very

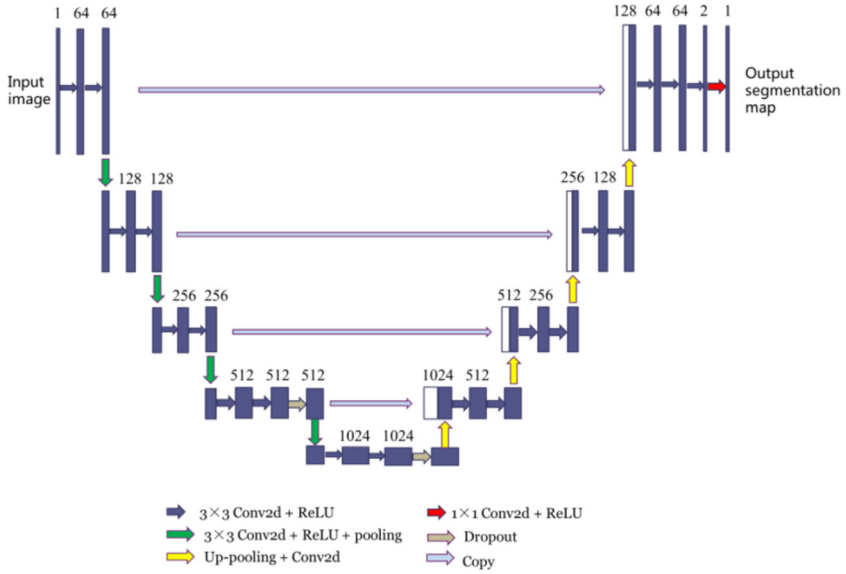

Fig. 2. The architecture of U-Net_tumor. The left part is the encoder and the right is the decoder. The encoder part consists of 5 blocks and the decoder part consists of 4 blocks.

few training images, U-shape convolutional neural network (U-Net) [36] has been widely used in biomedical image segmentation due to its encoder-decoder architecture and unique skip connections. To obtain the shape and morphology information of the breast tumor for the subsequent classification tasks, U-Net is employed as the segmentation network in the proposed model. The specific network structure and parameters in our work are shown in Fig. 2 and we refer to it as U-Net_tumor. U-Net_tumor is similar to the work of [36], whose encoder part consists of 5 blocks and the decoder part consists of 4 blocks. Blocks 1–3 have repeated application of two convolutions (unpadded convolutions), each followed by a rectified linear unit (ReLU) and a max-pooling operation with stride 2 for down-sampling. Blocks 6–9 are the same, with an up-sampling of the feature map followed by a convolution, a concatenation with the correspondingly down-sampled feature map from the encoder part, and two convolutions, each followed by a ReLU. In addition to the above content, to reduce the influence of over-fitting, our work adds the dropout method in block 4 and block 5. The value of the dropout rate in our work is 0.5.

The original image is applied as the input and the binary image in which tumor and normal tissue are 1 and 0, respectively as the output to train U-Net_tumor. After the breast tumor image is segmented by the U-Net_tumor, a binary image containing the object (tumor) and the background (tissue) is obtained. Next, morphological opening and closing are conducted to refine the tumor contour, with an opening to reduce the spicules and closing to fill the holes. An elliptical kernel is used for both opening and closing operations. After processing by the morphological opening and closing algorithms, the segmented images containing the tumors and the tissues are obtained.

2.3. Feature extraction network

After segmentation of the tumor, the ultrasound images and their corresponding segmentation results are obtained. In order to make the final classification, we need to extract features of the ultrasound images. In the currently proposed method, ResNet50 is chosen as the feature extraction network because of its powerful feature representation ability. In order to avoid over-fitting, transfer learning is applied and the pre-trained parameters are applied to initialize the parameters of the network and fine-tuned for adapting the breast ultrasound dataset. Due to the lack of public large-scale labeled medical ultrasound datasets, we choose non-medical ImageNet as the dataset for pre-training. Although the ImageNet is different from ultrasound datasets, the low-level features are universal to most of the image analysis tasks [37]. Hence, transferred parameters can help provide a powerful set of features, which reduce the need for a large dataset as well as the training time and memory cost. On this basis, fine-tuning the model can extract the most suitable features for specific ultrasound image classification. As shown in Fig. 3, the ResNet50 model consists of a stack of similar residual blocks and there is an identity mapping path to connect the input and output of the residual block. To combine the information of original images and segmentation enhanced images, two ResNets are designed as feature extraction networks for original images and segmentation enhancement images, which are called ResNet_ORI and ResNet_SEI, respectively.

2.3.1. ResNet_ORI

ResNet_ORI is applied to extract the features of the original images. After assigning pre-train parameters on the ImageNet dataset, ResNet_ORI is fine-tuned on the breast ultrasound dataset, to adapt the target domain dataset. The input of ResNet_ORI is original breast ultrasound images and the output of the last convolution layer of Conv5_x is the feature representation of the original image. More technically, let $\mu_{\text{ori}} = (\mu_{ori}^{(0)}, \mu_{ori}^{(1)}, \ldots, \mu_{ori}^{(K-1)})$ be a tensor that outputs from the last convolution layer of Conv5_x and K represents the number of channels in the layer. Therefore, μ_{ori} is the feature representation of ResNet_ORI.

2.3.2. ResNet_SEI

As introduced above, the corresponding segmentation results of ultrasound images are obtained and applied to form a segmentation enhanced image. Because a binary image cannot represent the features associated with texture and shape, the segmented binary image cannot be directly applied as the input of the deep convolution network to obtain image features. Therefore, we propose a new 3-channel artificial RGB image method by concatenating the original image with its segmented binary images. In the current study, three forms of composition are applied and compared in an experiment. Form 1 is that two channels are original ultrasound images and the other channel is the segmented binary

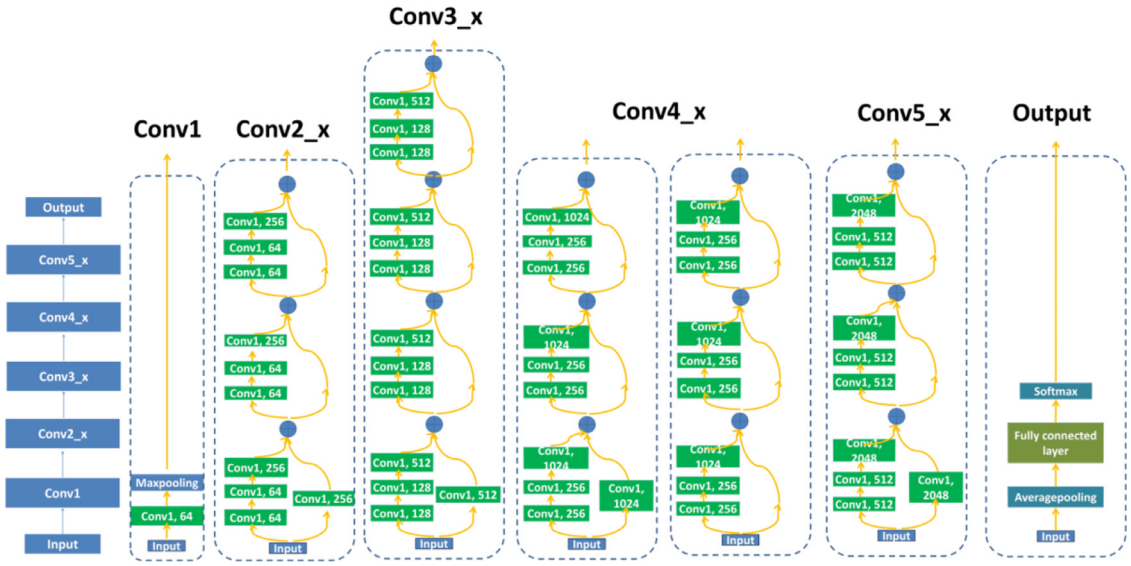

Fig. 3. The specific network structure of ResNet50. It contains six series of convolution blocks, namely Conv1, Conv2, Conv3, Conv4, Conv5, and output block, respectively. The design of the residual blocks reduces the problem of vanishing gradient in the network so that the deeper network can be trained to obtain a deeper level semantic representation.

image. Form 2 is that one channel is the original ultrasound image and the other two channels are the segmented binary images. Form 3 is that one channel is the original ultrasound image and the others are the segmented binary images and segmented binary images with a black and white reversal. In the current study, several experiments are carried out with the SBANet in three forms of the segmentation enhanced methods, and the Form 1 method is called the Proposed method, and others are applied as comparison algorithms.

Let I_{ori} denote the original gray-scale image, I_{seg} denote the segmented binary image, $I_{seg\_resversal}$ denote the segmented binary image with black and white reversal, and I_{mul_1}, I_{mul_2}, I_{mul_3}, denote the three forms of segmentation attention images, respectively. Therefore, $I \in R^{M \times M \times C}$, and M represents the size of width and height, while C denotes the channel number. The three forms of composition process are as follow:

$$I_{mul\_1} = expand\_dims(I_{ori}, I_{ori}, I_{seg}) \tag{1}$$

$$I_{mul\_2} = expand\_dims(I_{ori}, I_{ori}, I_{seg}) \tag{2}$$

$$I_{mul\_3} = expand\_dims\left(I_{ori}, I_{ori}, I_{seg\_resversal}\right) \tag{3}$$

where operator *expand_dims* combines three one-channel gray-scale images into a three-channel image. As shown in Fig. 4, the superimposed image is equivalent to highlight the tumor region and segmentation enhanced information including tumor shape-related features and edge-related texture features. The subsequent ResNet_SEI can automatically learn these features which enhance the segmentation information. The specific network structure is consistent with them in ResNet_ORI. Similarly, let $\mu_{sei} = (\mu_{sei}^{(0)}, \mu_{sei}^{(1)}, \ldots, \mu_{sei}^{(K-1)})$ be a tensor that is the output of the last convolution layer of Conv5_x and K represents the number of nodes in the layer. Meanwhile, $\mu_{sei} \in R^{K\times 1}$ is the feature representation of Sub-network_SEI.

2.4. Channel attention based feature aggregation network

Because the features extracted by the two networks have both similar and different contents, in order to learn the importance of different channel features extracted from two networks and pay more attention to the most informative channel features, a feature aggregation network is proposed to aggregate the features of original and segmentation enhanced images automatically. Firstly, the feature maps extracted by the two networks are concatenated. Given the extracted feature maps of two networks μ_{ori}, μ_{sei}, the concatenated feature maps $\mu_{concat} = concatenate\ (\mu_{ori}, \mu_{sei}) = (\mu_{ori}^{(0)}, \mu_{ori}^{(1)}, \ldots, \mu_{ori}^{(K-1)}, \mu_{sei}^{(0)}, \mu_{sei}^{(1)}, \ldots, \mu_{sei}^{(K-1)})$. Here, *concatenate* is an operator which concatenates two tensors into a tensor, μ_{ori} and μ_{sei} are the feature maps of DCNN for original images and segmentation enhancement images. Then, the channel attention method, which mainly includes the squeezing step and excitation step, is used to weigh different channel features. The squeezing step is a global average pooling operation, which is performed on every concatenated feature map to obtain the channel-level global features. The whole spatial feature on a channel is encoded as a global feature. After that, according to the global features, the excitation step learns the relationship between the channels, and gets the adaptive weights of different channels, namely attention score. Two fully connected layers, with activation layer ReLu and sigmoid, respectively, are designed to get the attention score of each channel. Finally, the final feature representation is obtained by multiplying the attention score with the concatenated feature maps. The architecture of our feature aggregation and classification network is shown in Fig. 5. The classification network is composed of a fully connected layer and a softmax layer. The aggregated feature for a breast ultrasound image can be formulated as follows:

$$\mathbf{z} = \mathbf{F}_{sq}(\mu_{concat}) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W}\mu_{concat}(i,j)$$

$$\alpha^{(l)} = \mathbf{F}_{ex}(\mathbf{z}, \mathbf{W}) = \sigma(g(\mathbf{z}, \mathbf{W})) = \sigma(\mathbf{W}_2\delta(\mathbf{W}_1\mathbf{z}))$$

$$\alpha^{(l)} = \mathbf{F}_{ex}(\mathbf{z}, \mathbf{W}) = \sigma(g(\mathbf{z}, \mathbf{W})) = \sigma(\mathbf{W}_2\delta(\mathbf{W}_1\mathbf{z}))$$

$$\mu_{agg}^{(l)} = (\mu_{concat}^{(l)} \times \alpha^{(l)})$$

$$\mu_{agg} = \left(\mu_{agg}^{(0)}, \mu_{agg}^{(1)}, \cdots, \mu_{agg}^{(2K-1)}\right) \tag{4}$$

where δ refers to the ReLU function, σ refers to the softmax function, $\mathbf{W}_1 \in R^{\frac{C}{r}\times C}$, $\mathbf{W}_2 \in R^{C\times\frac{C}{r}}$, $\mathbf{F}_{sq}$ and $\mathbf{F}_{ex}$ refer to squeeze and

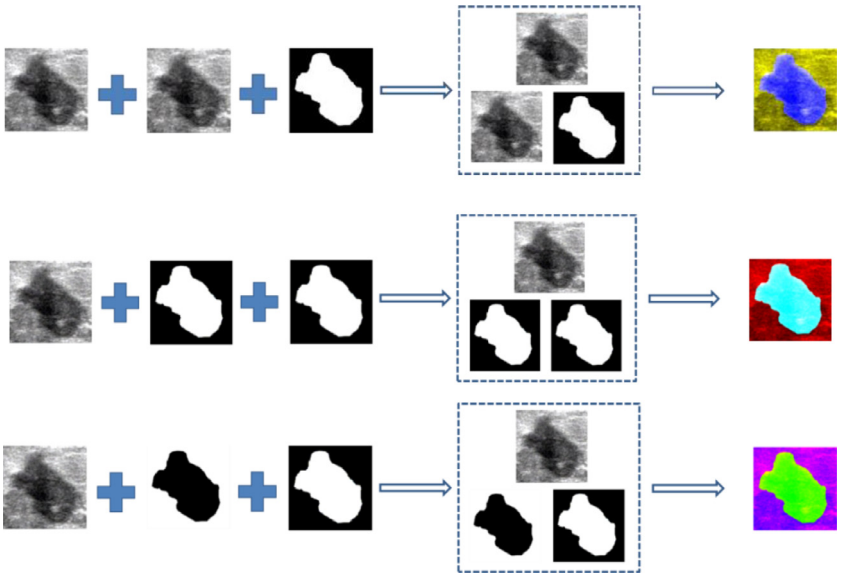

Fig. 4. Three forms of segmentation enhanced images. From the first line to the third line, they are Form 1, Form 2, and Form 3, respectively. The original gray-scale image and the segmented binary image are superimposed into a three-channel image, which is equivalent to highlight the tumor region.

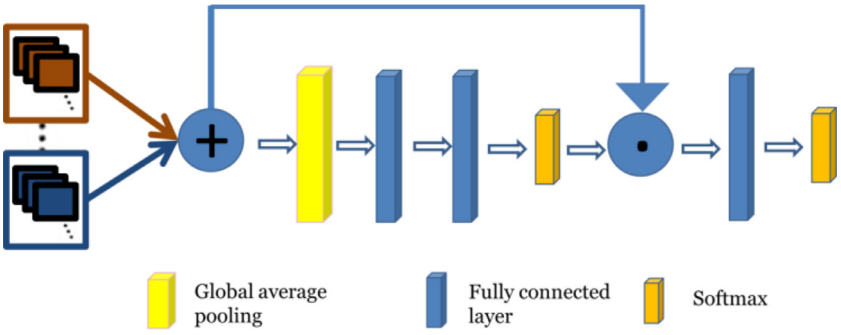

Fig. 5. The specific network structure of feature aggregation and classification network networks.

excitation operator, respectively, and μ_{agg} refers to the aggregated feature.

2.5. Training process

During the training, the segmentation network is firstly training on dataset A. After that, the classification networks including the two feature extraction networks on the original image and the segmentation enhanced image, attention based feature aggregation network, and classification network are trained on dataset B to obtain the final classification results. Adam algorithm is applied as the optimizer to update all parameters in the back-propagation algorithm for the training of the network and the initial learning rate is set to 0.0001. In order to train better and faster and avoid falling into local optimum, we first fine-tuned two feature extraction networks separately, and then train the general model together.

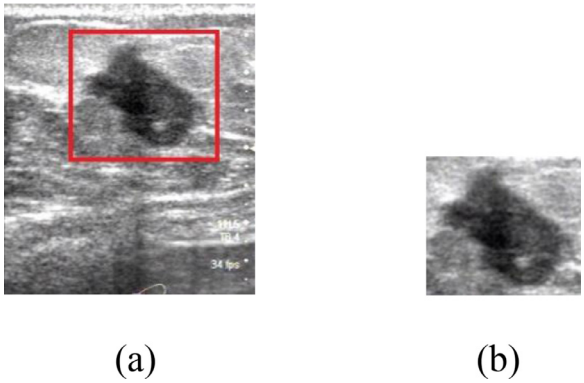

Fig. 6. Example of extracting a TCI, (a) the whole image, and (b) the TCI image.

2.6. Experimental methods

2.6.1. Data description

We developed the proposed method with the Python language using the Keras framework and ran it on a computer with an RTX 2080Ti for acceleration. To validate our method, experiments had been conducted on breast ultrasound datasets. All experimental protocols were approved by the Human Subject Ethics Committee of the South China University of Technology. Informed consent was obtained from all patients for use of their information in the research without violating their privacy. The 1994 clinical breast ultrasound images were provided by the Cancer Center of Sun Yat-sen University and were taken from an HDI 5000 SonoCT System (Philips Medical Systems) with an L12–5 50 mm Broadband Linear Array at the imaging frequency of 7.1 MHz. The proposed CAD system aimed to classify a previously identified ROI [38,39]. The ROIs, called tumor-centered images (TCIs) were obtained by manual segmentation as shown in Fig. 6. The size of the whole breast ultrasound image is 1280 × 1040, while the size of each TCI is different. To facilitate the subsequent analysis,

Table 1
The cardinality of Dataset A and Dataset B.

	Dataset A	Dataset B
Benign	160	786
Malignant	132	916
Total	292	1702

all TCIs are resized into a unified 224 × 224. These original ultrasound images and TCIs are one-channel gray images, and the format is JPG. In order to avoid over-fitting, we applied transfer learning, and the ResNet50 model pre-trained on ImageNet is used in our research. Because the model is trained on the RGB image dataset, to adapt to this situation, we change the one-channel gray image into a three-channel RGB image by repeating it three times. So in two parallel feature extraction networks, the input of ResNet_ORI is the three-channel image which is obtained by repeating the original TCI image three times. And in the ResNet_SEI network, the input is the segmentation enhanced superimposed image which is equivalent to replacing one of the channels with a segmented binary image, and the effect of three channels is shown in Fig. 4. Because not all the data have the label of segmentation and classification at the same time, our breast data was divided into two datasets, which we refer to as dataset A and dataset B. The breast images of dataset A have both diagnostic labels and tumor contours labels. The contours of the tumors were manually delineated by three experienced radiologists. The breast images of dataset B have only diagnostic labels. All cases had been verified by physician pathology. Dataset A comprises 292 images (160 benign and 132 malignant) in total. And the dataset B comprises 1702 images (786 benign and 916 malignant) in total. Table 1 summarizes the numbers of images in the two datasets. And the code is available at https://github.com/Jax-Luo/Segmentation-Information-with-Attention-Integration-for-Classification-of-Breast-tumor-in-Ultrasound/tree/master.

2.6.2. Comparative experiment

In the experiment, we firstly trained a segmentation network on dataset A. After that, we trained classification networks on dataset B combining the original image information and the segmentation results obtained through the trained segmentation network. During training, data augmentations such as rotation, flip, zoom, shearing, shift were applied. In order to validate our method, 10-fold cross-validation was utilized to evaluate the performance of the proposed SBANet for the breast ultrasound tumor classification method. 1702 samples of dataset B were randomly divided into 10 equal parts. In order to ensure that the ratio of benign and malignant samples of each subset was the same as the ratio of the whole data set, which can prevent the deviation caused by the inconsistent data distribution, we stratified sample the data set to divide it. In each training iteration, we selected one as the test set and the remaining sample sets as the training set. The average results of ten experiments were applied to evaluate the performance of each classification method.

The arrangement of the comparative experiment is divided into three steps. The first step is the comparison with other methods. To demonstrate the advantages of the proposed method, we also compared the proposed method with the other eleven breast image classification works based on deep learning in recent years. These eleven methods are Masud et al. [19], Wang et al. [20], Zeimarani et al. [21], Daoud et al. [22], Hijab et al. [23], Byra et al. [26], Cao et al. [24], Tanaka et al. [27], Fujioka et al. [25], Zhuang et al. [28] and the baseline ResNet50. Among them, the methods in [19–21] are custom networks, which have mainly one, two, and four convolution layers, respectively, combined with fully connected, pooling, dropout, and batch normalization layers. The rest of the works apply pre-trained models and fine-tuned them on breast ultrasound images. The works in [22,23,25], are based on per-trained AlexNet, VGG16, GoogleNet Inception V2 networks, respectively. There are many pre-trained methods applied in [24], and we take DenseNet based method in our experiment. And the method in [26] proposed a matching layer before pre-trained the VGG19 network. The work in [27] adopted an ensemble method, in which VGG19 and ResNet151 are combined, and the final result is obtained by averaging the out of two networks. And [28] applies an image decomposition method and forms a three-channel image including the original image, enhanced image, and de-noised image as the input of pre-trained VGG16. All of these eleven methods have been studied recently in breast classification, and have achieved good results. In particular, because ResNet50 is the backbone network of our algorithm, comparing the proposed method with ResNet50, can better validate the effectiveness of our algorithm. The second step is the ablation experiment of the proposed method. The third step is the comparison of three forms of segmentation enhanced methods in SBANet. In addition, the comparison algorithms were programmed in the same system and ran in the same hardware configuration. The TCIs were all the same for the classification methods, which did not affect the comparison.

2.6.3. Evaluation metrics

For evaluating the effectiveness of the classification methods, five metrics were taken to assess the performance, which include accuracy, sensitivity, specificity, F1-score, and AUC. Accuracy represents the ratio of the correctly predicted samples to the total samples. It reflects the overall classification accuracy and performance and does not consider whether the predicted samples are benign or malignant. Sensitivity represents the proportion of the sample with a malignant tumor that is correctly identified as a malignant case. The higher the sensitivity, the stronger the ability to detect malignant tumors. Specificity represents the proportion of the sample with a benign tumor that is correctly identified as a benign case. The higher the value of specificity, the greater the ability to detect benign tumors. F1-score is defined as the harmonic average of sensitivity rate and specificity rate, which is used to balance the two metrics to get a comprehensive evaluation. AUC is defined as the area under the ROC curve, which is a performance index to measure the overall quality of the system. Each of the five metrics has its significance to measure the performance of different work. The corresponding formulas of the metrics are as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{5}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{6}$$

$$\text{Specificity} = \frac{TN}{FP + TN} \tag{7}$$

$$\text{F1} - \text{score} = \frac{2 * \text{Sensitivity} * \text{Specificity}}{\text{Sensitivity} + \text{Specificity}} \tag{8}$$

where, TP and FP indicate the number of malignant tumors correctly classified to be malignant and misclassified to be benign tumors, respectively. TN and FN indicate the number of benign tumors correctly classified to be benign and misclassified to be malignant tumors, respectively.

3. Results

3.1. Comparison with other methods

Table 2 summarizes the performance of different methods for breast tumor classification. It shows the comparison with the

Table 2
Comparison result with other methods.

Classification method	Accuracy	Sensitivity	Specificity	F1-score	AUC
Masud et al. [19]	0.7844	0.7645	0.8281	0.8147	0.8701
Wang et al. [20]	0.7920	0.7580	0.8580	0.8243	0.8596
Zeimarani et al. [21]	0.8449	0.8350	0.8595	0.8602	0.9183
Daoud et al. [22]	0.8713	0.8692	0.8751	0.8823	0.9388
Hijab et al. [23]	0.8508	0.8533	0.8516	0.8629	0.9244
Byraa et al. [26]	0.8291	0.8252	0.8374	0.8454	0.9051
Cao et al. [24]	0.8719	0.8765	0.8765	0.8821	0.9455
Tanaka et al. [27]	0.8778	0.8702	0.8922	0.8895	0.9492
Fujioka et al. [25]	0.8748	0.8778	0.8748	0.8851	0.9426
Zhuang et al. [28]	0.8619	0.8552	0.8743	0.8745	0.9298
ResNet50	0.8802	0.8870	0.8825	0.8899	0.9448
Proposed method	**0.9078**	**0.9118**	**0.9044**	**0.9146**	**0.9549**

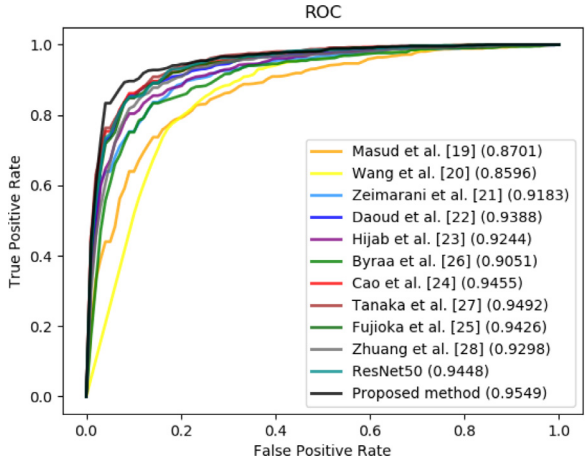

Fig. 7. ROC curves of different classification methods on breast ultrasound dataset.

recent breast image classification works based on deep learning. From the experimental results, the methods applying the pre-training model have better results than other custom networks, which shows that transfer learning is more effective for a small amount of dataset. In addition, our proposed methods have achieved the best results. The proposed method gets the highest Accuracy (90.78%), Sensitivity (91.18%), Specificity (90.44%), F1-score (91.46%), and AUC (0.9549). Moreover, from Table 2, we can see that compared with ResNet50, our methods have achieved better results in five metrics. This means that our methods have higher classification accuracy, better malignant detection rate, and lower misdiagnosis rate. This shows that our design mechanism of feature combination representation, adding segmentation information with attention integration into the classification network, is more excellent and effective, and achieves significant improvement.

Moreover, to evaluate the effectiveness of the proposed methods, we made a comparison based on the ROC metric, and the results are shown in Fig. 7. It can be observed that our proposed method performs better than other methods. The ROC of the proposed method was located at the outmost edge of the graph and achieved the highest AUC values than all the other methods. Fig. 8 illustrates the comparison of classification performance for different methods on benign and malignant tumors. In contrast, our method achieves both the highest classification rates in benign and malignant images.

3.2. Ablation experiment

In order to demonstrate the effectiveness of the novel segmentation enhancement feature extraction network and the attention based feature aggregation network of the proposed framework, we made an ablation experiment, and the results are shown in Table 3. In Table 3, the method denoted as Proposed_ORI represents the proposed method that removes the segmentation enhancement feature extraction network and the attention based feature aggregation network. Similarly, Proposed_SEI represents the proposed method that removes the feature extraction network for original images and the attention based feature aggregation network. Proposed_CAT represents the proposed method whose attention based feature aggregation network is replaced by the original feature combination module which directly concatenates the feature representation of the two networks to form the total feature. Proposed_ADD represents the proposed method whose attention based feature aggregation network is replaced by another feature combination module that directly adds the feature representation of the two networks to form the total feature. As seen in Table 3, the Proposed method, Proposed_CAT, and Proposed_ADD outperform Proposed_ORI and Proposed_SEI, which demonstrates that the DCNN combining segmentation enhanced information is effective for breast cancer classification. At the same time, the proposed method outperforms Proposed_CAT and Proposed_ADD, which means that channel attention based feature aggregation network is effective for breast ultrasound classification.

3.3. Experiment of three segmentation enhanced methods

In addition, we evaluated three forms of segmentation enhancement methods on the breast ultrasound classification. Ten-fold cross validation was performed and the averaged results were shown in Table 4. From the comparison results, whether SBANet_form1, SBANet_form2, or SBANet_form3, have achieved improved results than the baseline. It proves that the proposed SBANet for breast ultrasound classification methods is more effective and robust. Further analysis, SBANet_form1 has a slightly better classification effect than SBANet_form2 and SBANet_form3, which revealed that introducing the segmentation result into one channel is a little more effective than into two channels. The most important reason is that the proposed scheme is a weak-semi supervised method, in which segmentation error is introduced, even if it is not very large. The result of the segmentation is shown in Fig. 9. Most tumors can be well segmented, but there are still errors in the segmentation process.

3.4. Attention map analysis and visualization

In addition, to show the effectiveness of the feature aggregation network based on the attention mechanism more intuitively, we analyzed the attention maps of the two sub-networks and the proposed aggregated network. In order to visualize class-specific attention, we applied class activation mapping to generate attention

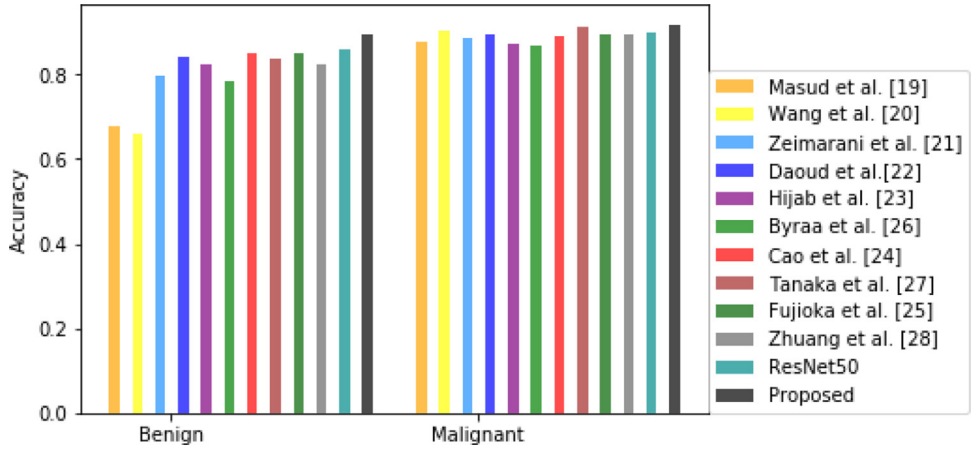

Fig. 8. Classification analysis for different breast tumors classification methods.

Table 3
Ablation experimental results.

Classification method	Accuracy	Sensitivity	Specificity	F1-score	AUC
Proposed_ORI	0.8802	0.8870	0.8825	0.8899	0.9448
Proposed_SEI	0.8837	0.8841	0.8868	0.8931	0.9507
Proposed_CAT	0.8989	0.9031	0.8951	0.9065	0.9539
Proposed_ADD	0.8948	0.8931	0.8989	0.9035	0.9533
Proposed method	**0.9078**	**0.9118**	**0.9044**	**0.9146**	**0.9549**

Table 4
Experiment results of three forms of SBA methods.

Classification method	Accuracy	Sensitivity	Specificity	F1-score	AUC
SBANet_form2	0.9013	0.9037	0.8990	0.9086	0.9509
SBANet_form3	0.8966	0.8994	0.8942	0.9042	0.9519
Proposed method (SBANet_form1)	**0.9078**	**0.9118**	**0.9044**	**0.9146**	**0.9549**

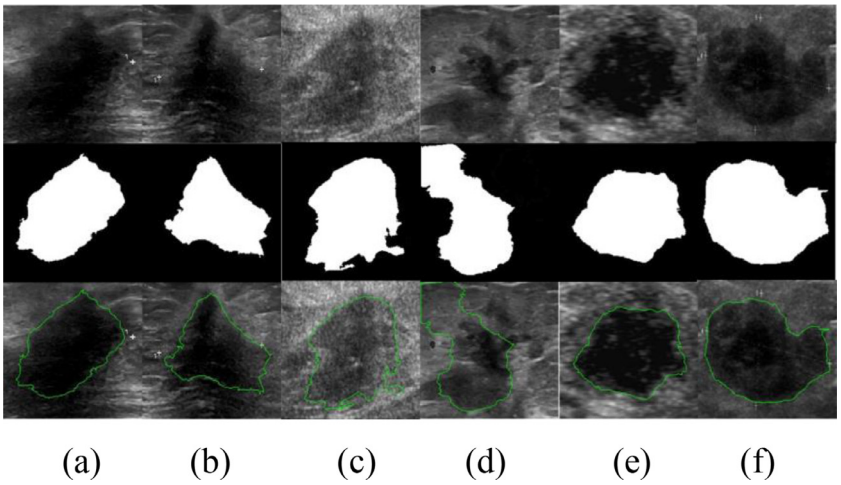

Fig. 9. Segmentation results for the breast tumors, the images of the first line are the original TCIs, the second line is the segmented binary results, the third line is the segmentation results delineated on the TCI.

maps of breast images, and the results were shown in Fig. 10. In this case, the aggregation layer relearns how to optimize the combination of features based on the two parallel feature networks. It can be seen that both ResNet_ORI and ResNet_SEI can notice the key areas of tumor classification, including boundary, calcification, and so on. Because the segmentation information is added to the ResNet_SEI, as shown on the attention map, it pays more attention to the edge-related areas. The proposed aggregation network, on the other hand, is selectively and dynamically integrated, rather than simply superimposing the attention area of ResNet_ORI and ResNet_SEI. Although the interpretability of deep learning is still a difficult research field, from the attention maps, we can see that the proposed feature aggregation network can effectively integrate the regions of interest of the two networks, and it is not a simple superposition, but a selective dynamic combination. On the other hand, it can effectively fuse the features extracted from the two networks for breast tumor classification.

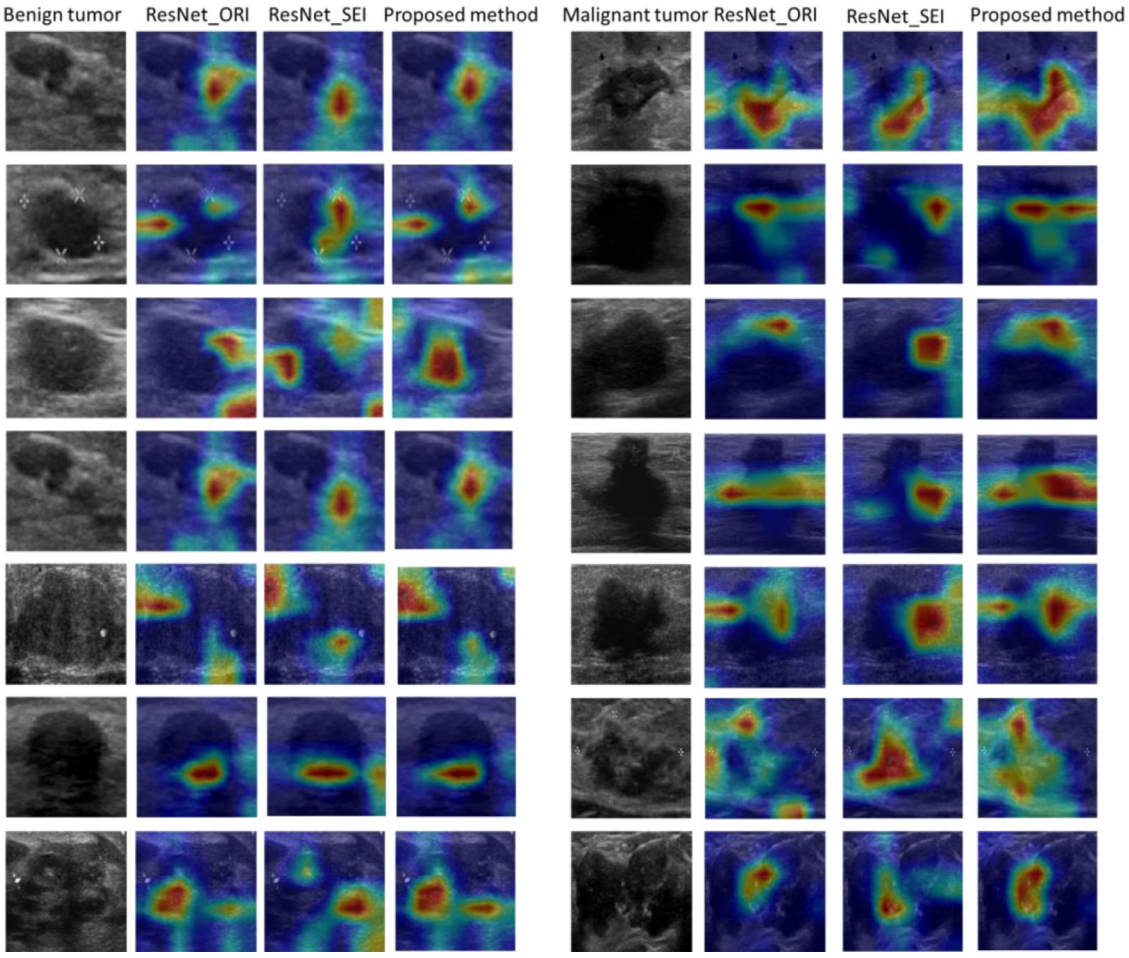

Fig. 10. Examples of obtained attention map from ResNet_ORI, ResNet_SEI and proposed aggregation method, respectively.

4. Discussion and conclusions

In this paper, we propose a novel breast tumor classification scheme for ultrasound images based on combining segmentation information and the attention integration method. Different from the previous deep learning based breast classification works which only extract the features on breast ultrasound images, our work aims at combining medical knowledge in the network, and design the segmentation network to obtain the tumor contour and region to extract more effective features for classification. We carry out ten-fold cross validation experiments to evaluate the performance of the proposed method and compare it with other works and the baseline. The results show that our work has a better effect on breast tumor classification than other methods and the baseline. It triumphs other works and the baseline on breast ultrasound datasets with up to 2.76% improvement in the metric Accuracy and 2.47% improvement in the metric F1-score compared to the baseline. As shown in Table 2, our method, compared with other breast classification work, including the transfer learning methods based on pre-trained models and shallow network methods, obtains better classification performance and the highest AUC and accuracy, which indicates that our method can classify breast more effectively. At the same time, as shown in Fig. 8, our method has higher classification accuracy in both benign and malignant images. It shows that our method has better feature extraction in both benign and malignant samples. The sensitivity and specificity are shown in Table 2 also show all this. As shown in Fig. 8, the ROC of the proposed method was located at the outmost edge of the graph and achieved the highest AUC values than all the other methods. In addition, as shown in Table 3, the performance of every aggregation method, including directly concatenating and directly adding feature representation of the two networks, is better than that of a single branch network, which indicates that adding a parallel branch based on segmentation prior is meaningful for breast classification. At the same time, ablation experiments also show that the feature aggregation network based on attention achieves higher performance. This proves that the feature aggregation network based on attention can effectively preserve the effective features and remove the redundant features. At the same time, compared with other feature aggregation methods, this method has higher performance. This proves that our works can extract more effective and classification-related features. In addition, the segmentation is automatic in our scheme, which can greatly enhance the operability of the system. If the segmentation needs manual drawn, it is not only a great burden for the operator but also difficult to carry out in many cases, because the ultrasound segmentation labeling needs professional radiologists, many novices or inexperienced people cannot do it well.

Breast cancer is a high incidence disease that brings great harm to women's health all over the world. Therefore, it is very important to detect and diagnose tumors as quickly and accurately as possible through imaging methods in the early stage. Ultrasound is one of the most common detection methods and has its unique advantages compared with mammography and biopsy. Our work can provide an effective auxiliary diagnostic tool for the classification of benign and malignant breast tumors, and help to effectively

identify malignant tumors in early screening. In addition, the study has also some enlightenment for breast ultrasound classification and even other types of medical image diagnosis, because more effective feature extraction is the basis for further image analysis.

The potential limitation of our work is that our classification scheme is not an end-to-end method and needs two stages, which increases training time and complexity. In addition, the step to obtain TCIs requires user's participation which may have a significant influence on the following classification. Consequently, how to automatically extract TCI from the breast ultrasound image and put the segmentation-based attention information into the network intermediate step instead of obtaining the segmentation results first are two of our future studies. In addition, there is more medical prior information, such as BIRADS information, which can be put into the network for the classification of benign and malignant breast tumors.

Declaration of Competing Interest

None.

Acknowledgments

This work was supported in part by the National Key Research and Development Program under Grants 2018AAA0102100 and 2017YFC0110602, Shaanxi Provincial Foundation for Distinguished Young Scholars under Grant 2019JC-13, the National Natural Science Foundation of China under Grants 62071382 and 82030047, and the Innovation Capability Support Program of Shaanxi under Program 021TD-57.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.patcog.2021.108427.

References

[1] Y. Li, J. Li, Y. Wang, Y. Zhang, J. Chu, C. Sun, Z. Fu, Y. Huang, H. Zhang, H. Yuan, Roles of cancer/testis antigens (CTAs) in breast cancer, Cancer Lett. 399 (2017) 64–73, doi:10.1016/j.canlet.2017.02.031.
[2] Q. Huang, Y. Huang, Y. Luo, F. Yuan, X. Li, Segmentation of breast ultrasound image with semantic classification of superpixels, Med. Image Anal. 61 (2020) 101657, doi:10.1016/j.media.2020.101657.
[3] C. Barata, M.E. Celebi, J.S. Marques, Explainable skin lesion diagnosis using taxonomies, Pattern Recognit. 110 (2020) 107413, doi:10.1016/j.patcog.2020.107413.
[4] X. Chen, L. Yao, T. Zhou, J. Dong, Y. Zhang, Momentum contrastive learning for few-shot COVID-19 diagnosis from chest CT images, Pattern Recognit. 113 (2020) 107826, doi:10.1016/j.patcog.2021.107826.
[5] H.D. Cheng, J. Shan, W. Ju, Y.H. Guo, L. Zhang, Automated breast cancer detection and classification using ultrasound images: a survey, Pattern Recognit. 43 (1) (2010) 299–317, doi:10.1016/j.patcog.2009.05.012.
[6] W.G. Flores, W.C.D. Pereira, A.F.C. Infantosi, Improving classification performance of breast lesions on ultrasonography, Pattern Recognit. 48 (4) (2015) 1125–1136, doi:10.1016/j.patcog.2014.06.006.
[7] W. Gomez, W.C.A. Pereira, A.F.C. Infantosi, Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound, IEEE Trans. Med. Imaging 31 (10) (2012) 1889–1899, doi:10.1109/TMI.2012.2206398.
[8] T. Tan, B. Platel, H. Huisman, C.I. Sánchez, R. Mus, N. Karssemeijer, Computer-aided lesion diagnosis in automated 3-D breast ultrasound using coronal spiculation, IEEE Trans. Med. Imaging 31 (5) (2012) 1034–1042, doi:10.1109/TMI.2012.2184549.
[9] K.M. Amin, A.I. Shahin, Y. Guo, A novel breast tumor classification algorithm using neutrosophic score features, Measurement 81 (2016) 210–220, doi:10.1016/j.measurement.2015.12.013.
[10] R. Liao, T. Wan, Z. Qin, Classification of benign and malignant breast tumors in ultrasound images based on multiple sonographic and textural features, in: Proceedings of the International Conference on Intelligent Human-Machine Systems and Cybernetics, 2011, doi:10.1109/IHMSC.2011.127.
[11] K.M. Prabusankarlal, P. Thirumoorthy, R. Manavalan, Assessment of combined textural and morphological features for diagnosis of breast masses in ultrasound, Hum. Centric Comput. Inf. Sci. 5 (2015) 12, doi:10.1186/s13673-015-0029-y.
[12] S.M. Hsu, W.H. Kuo, F.C. Kuo, Y.Y. Liao, Breast tumor classification using different features of quantitative ultrasound parametric images, Int. J. Comput. Assist. Radiol. Surg 14 (4) (2019) 623–633, doi:10.1007/s11548-018-01908-8.
[13] G.Y. Chen, W.F. Xie, Pattern recognition with SVM and dual-tree complex wavelets, Image Vis. Comput. 25 (6) (2007) 960–966, doi:10.1016/j.imavis.2006.07.009.
[14] S. Zhou, J. Shi, J. Zhu, Y. Cai, R. Wang, Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image, Biomed. Signal Process. Control 8 (6) (2013) 688–696, doi:10.1016/j.bspc.2013.06.011.
[15] A. Takemura, A. Shimizu, K. Hamamoto, Discrimination of breast tumors in ultrasonic images using an ensemble classifier based on the adaboost algorithm with feature selection, IEEE Trans. Med. Imaging 29 (3) (2010) 598–609, doi:10.1109/TMI.2009.2022630.
[16] Y.L. Huang, D.R. Chen, Y.R. Jiang, S.J. Kuo, H.K. Wu, W.K. Moon, Computer-aided diagnosis using morphological features for classifying breast lesions on ultrasound, Gynecol. Off. J. Int. Soc. Ultrasound Obstet. Gynecol. 32 (4) (2010) 565–572, doi:10.1002/uog.5205.
[17] J. Zhou, F. Pan, W. Li, H. Hu, W. Wang, Q. Huang, Feature fusion for diagnosis of atypical hepatocellular carcinoma in contrast-enhanced ultrasound, IEEE Trans. Ultrason. Ferroelectr. (2021), doi:10.1109/TUFFC.2021.311059.
[18] L. Bi, D.D. Feng, M. Fulham, J. Kim, Multi-label classification of multi-modality skin lesion via hyper-connected convolutional neural network, Pattern Recognit. 107 (2020) 107502, doi:10.1016/j.patcog.2020.107502.
[19] M. Masud, A.E.E. Rashed, M.S. Hossain, Convolutional neural network-based models for diagnosis of breast cancer, Neural Comput. Appl. 1 (2020) 1–12, doi:10.1007/s00521-020-05394-5.
[20] F. Wang, X. Liu, N. Yuan, B. Qian, L. Ruan, C. Yin, C. Jin, Study on automatic detection and classification of breast nodule using deep convolutional neural network system, J. Thorac. Dis. 12 (9) (2020) 4690–4701, doi:10.21037/jtd-19-3013.
[21] B. Zeimarani, M.G.F. Costa, N.Z. Nurani, S.R. Bianco, W.C.D. Pereira, C.F.F. Costa, Breast lesion classification in ultrasound images using deep convolutional neural network, IEEE Access 8 (2020) 133349–133359, doi:10.1109/ACCESS.2020.3010863.
[22] M.I. Daoud, S. Abdel-Rahman, R. Alazrai, Breast ultrasound image classification using a pre-trained convolutional neural network, in: Proceedings of the 15th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 2019, doi:10.1109/SITIS.2019.00037.
[23] A. Hijab, M.A. Rushdi, M.M. Gomaa, A. Eldeib, Breast cancer classification in ultrasound images using transfer learning, in: Proceedings of the 5th International Conference on Advances in Biomedical Engineering (ICABME), 2019, doi:10.1109/ICABME47164.2019.8940291.
[24] Z. Cao, L. Duan, G. Yang, T. Yue, Q. Chen, An experimental study on breast lesion detection and classification from ultrasound images using deep learning architectures, BMC Med. Imaging 19 (1) (2019) 51, doi:10.1186/s12880-019-0349-x.
[25] T. Fujioka, K. Kubota, M. Mori, Y. Kikuchi, L. Katsuta, M. Kasahara, G. Oda, T. Ishiba, T. Nakagawa, U. Tateishi, Distinction between benign and malignant breast masses at breast ultrasound using deep learning method with convolutional neural network, Jpn. J. Radiol. 37 (6) (2019) 466–472, doi:10.1007/s11604-019-00831-5.
[26] M. Byra, M. Galperin, H. Ojeda-Fournier, L. Olson, M. O'Boyle, C. Christopher, M. Andre, Breast mass classification in sonography with transfer learning using a deep convolutional neural network and color conversion, Med. Phys. 46 (2) (2019) 746–755, doi:10.1002/mp.13361.
[27] H. Tanaka, S. Chiu, T. Watanabe, S. Kaoku, T. Yamaguchi, Computer-aided diagnosis system for breast ultrasound images using deep learning, Phys. Med. Biol. 64 (23) (2019) 235013, doi:10.1088/1361-6560/ab5093.
[28] Z.M. Zhuang, Y.Q. Kang, A.N.J. Raj, Y. Yuan, W.L. Ding, S.M. Qiu, Breast ultrasound lesion classification based on image decomposition and transfer learning, Med. Phys. 47 (12) (2020) 6257–6269, doi:10.1002/mp.14510.
[29] C.M. Chen, Y.H. Chou, K.C. Han, G.S. Hung, C. Tiu, H.J. Chiou, S.Y. Chiou, Breast lesions on sonograms: computer-aided diagnosis with nearly setting-independent features and artificial neural networks, Radiology 226 (2) (2003) 504–514, doi:10.1148/radiol.2262011843.
[30] A.S. Hong, E.L. Rosen, M.S. Soo, J.A. Baker, BI-RADS for sonography: positive and negative predictive values of sonographic features, AJR Am. J. Roentgenol. 184 (4) (2005) 1260–1265, doi:10.2214/ajr.184.4.01841260.
[31] Q. Huang, Y. Chen, L. Liu, D. Tao, X. Li, On combining biclustering mining and AdaBoost for breast tumor classification, IEEE Trans. Knowl. Data Eng. 32 (4) (2020) 728–738, doi:10.1109/TKDE.2019.2891622.
[32] M. Xian, Y. Zhang, H.D. Cheng, Fully automatic segmentation of breast ultrasound images based on breast characteristics in space and frequency domains, Pattern Recognit. 48 (2) (2015) 485–497, doi:10.1016/j.patcog.2014.07.026.
[33] Y. Zhou, H.J. Chen, Y.F. Li, Q. Liu, X. Xu, S. Wang, P.T. Yap, D.G. Shen, Multi-task learning for segmentation and classification of tumors in 3D automated breast ultrasound images, Med. Image Anal. 70 (2021) 101918, doi:10.1016/j.media.2020.101918.
[34] M. Xian, Y.T. Zhang, H.D. Cheng, F. Xu, B.Y. Zhang, J.R. Ding, Automatic breast ultrasound image segmentation: a survey, Pattern Recognit. 79 (2018) 340–355, doi:10.1016/j.patcog.2018.02.012.
[35] Q. Huang, Z. Miao, S. Zhou, C. Chang, X. Li, Dense prediction and local fusion of superpixels: a framework for breast anatomy segmentation in ultrasound image with scarce data, IEEE Trans. Instrum. Meas. 70 (2021) 5011508, doi:10.1109/TIM.2021.3088421.

[36] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional Networks for biomedical image segmentation, in: Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), 2015, pp. 234–241, doi:10.1007/978-3-319-24574-4_28.
[37] M.A. Morid, A. Borjali, G.D. Fiol, A scoping review of transfer learning research on medical image analysis using ImageNet, Comput. Biol. Med. 128 (2021) 104115, doi:10.1016/j.compbiomed.2020.104115.
[38] W.X. Liao, P. He, J. Hao, X.Y. Wang, R.L. Yang, D. An, L.G. Cui, Automatic identification of breast ultrasound image based on supervised block-based region segmentation algorithm and features combination migration deep learning model, IEEE J. Biomed. Health 24 (4) (2020) 984–993, doi:10.1109/JBHI.2019.2960821.
[39] S. Han, H.K. Kang, J.Y. Jeong, M.H. Park, W. Kim, W.C. Bang, Y.K. Seong, A deep learning framework for supporting the classification of breast lesions in ultrasound images, Phys. Med. Biol. 62 (19) (2017) 7714–7728, doi:10.1088/1361-6560/aa82ec.

Yaozhong Luo is currently a Ph.D. student at the School of Electronic and Information Engineering, South China University of Technology. His research interests include ultrasonic image analysis, pattern recognition and machine learning.

Qinghua Huang received the Ph.D. degree in biomedical engineering from the Hong Kong Polytechnic University, Hong Kong, in 2007. In 2008, he joined the School of Electronic and Information Engineering, South China University of Technology, China. Now he is a full professor with School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an, China. His research interests include multi-dimensional ultrasonic imaging, pattern recognition, medical image analysis, machine learning for medical data, and intelligent computation.

Xuelong Li is a full professor with School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an, China.