Computer Methods and Programs in Biomedicine 208 (2021) 106221

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

Breast ultrasound tumor image classification using image decomposition and fusion based on adaptive multi-model spatial feature fusion

Zhemin Zhuang, Zengbiao Yang, Alex Noel Joseph Raj, Chuliang Wei, Pengcheng Jin, Shuxin Zhuang*

Department of Electronic Engineering, Shantou University, Shantou, Guangdong, China

ARTICLE INFO

Article history:
Received 23 March 2021
Accepted 26 May 2021

Keywords:
Breast ultrasound tumor image classification
Deep learning
Image decomposition
Image fusion
Adaptive spatial feature fusion

ABSTRACT

Background and Objective: Breast cancer is a fatal threat to the health of women. Ultrasonography is a common method for the detection of breast cancer. Computer-aided diagnosis of breast ultrasound images can help doctors in diagnosing benign and malignant lesions. In this paper, by combining image decomposition and fusion techniques with adaptive spatial feature fusion technology, a reliable classification method for breast ultrasound images of tumors is proposed.

Methods: First, fuzzy enhancement and bilateral filtering algorithms are used to process the original breast ultrasound image. Then, various decomposition images representing the clinical characteristics of breast tumors are obtained using the original and mask images. By considering the diversity of the benign and malignant characteristic information represented by each decomposition image, the decomposition images are fused through the RGB channel, and three types of fusion images are generated. Then, from a series of candidate deep learning models, transfer learning is used to select the best model as the base model to extract deep learning features. Finally, while training the classification network, adaptive spatial feature fusion technology is used to train the weight network to complete deep learning feature fusion and classification.

Results: In this study, 1328 breast ultrasound images were collected for training and testing. The experimental results show that the values of accuracy, precision, specificity, sensitivity/recall, F1 score, and area under the curve of the proposed method were 0.9548, 0.9811, 0.9833, 0.9392, 0.9571, and 0.9883, respectively.

Conclusion: Our research can automate breast cancer detection and has strong clinical utility. When compared to previous methods, our proposed method is expected to be more effective while assisting doctors in diagnosing breast ultrasound images.

1. Introduction

Breast cancer is one of the most common cancers among women, and it is also the leading cause of cancer-related deaths worldwide. According to the International Agency for Research on Cancer [1], in 2018, it was estimated that there were 2.1 million new cases of breast cancer and 627 thousand new deaths worldwide. In 2020, the number of new breast cancer cases exceeded that of lung cancer, making it the most prevalent form of cancer [2]. Early detection of breast cancer can effectively improve the survival time of patients, reduce mortality, and improve the quality of life of patients [3,4,5]. Ultrasonography is widely used in the early detection of breast diseases because of its painless and convenient operation and efficient real-time performance [6–8]. However, because of the high sensitivity of ultrasonic instruments, it is susceptible to the influence of different tissues of the human body and the environment, resulting in a large amount of speckle noise, which interferes with the diagnoses of doctors. Moreover, ultrasound surgeons with diverse levels of experience may obtain different diagnostic results based on breast ultrasound images, and the rate of missed diagnoses is 10-30% [9–11]. To cope with this problem, researchers have proposed various computer-aided diagnosis (CAD) methods [12–15].

In [12], the authors studied 1370 benign and 688 malignant breast ultrasound images from 1422 patients using Inception V3,

* Corresponding author.
E-mail address: sxzhuang@stu.edu.cn (S. Zhuang).

https://doi.org/10.1016/j.cmpb.2021.106221
0169-2607/

a residual network (ResNet) with 50 layers, Xception, and a convolutional neural network (CNN) composed of three convolution layers. The four fine-tuned deep learning models were used as feature extractors to extract the deep learning features of breast ultrasound images based on transfer learning. Finally, the extracted deep learning features were concatenated and classified using a simple artificial neural network (ANN).

In [13], the researchers first processed the original breast ultrasound images by filtering and segmenting the images to obtain the lesion area. Then, the size of the processed image was adjusted by clipping. Finally, the obtained dataset was augmented through left-right and up-down flips. For the expanded dataset, the fine-tuned Visual Geometry Group 19 (VGG19), Squeezenet, ResNet18, and GoogLeNet models were used for classification based on transfer learning, and GoogLeNet with the best classification result was selected as the target model for the extraction of deep learning features. As the extracted deep learning features contain redundant information, which affects the classification results, researchers used the feature selection method to reduce the redundant information and obtain the final optimal feature set for classification.

In [14], first, the breast ultrasound images from two different imaging modes, i.e., strain elastography and B-mode images, were segmented by a supervised block-based region segmentation algorithm. The fine-tuned VGG19 model was used as the deep learning feature extractor for 69 and 130 cases of malignant and benign lesions, respectively. The extracted deep learning features of the two types of ultrasound images were concatenated and sent to the ANN for classification.

In [15], the original breast ultrasound and mask images of the lesion area were used to obtain an ultrasound image of the lesion area. Then, the original breast ultrasound, lesion, and mask images were combined into RGB images. The above four types of images were trained and classified using different deep learning models. The model with the best classification results was selected as the base model for each type of image. Finally, the deep learning features extracted from each base model were fused and classified using ensemble learning.

It can be noted that classification methods based on deep learning are widely used in the automatic classification of breast ultrasound images, and researchers constantly propose strategies based on multi-image and multi-model fusion. Studies have shown that using different types of breast ultrasound images can help in distinguishing between benign and malignant breast lesions [14,15]. However, in the above studies, the categories of ultrasound images used for classification were limited. Further, the literature [12,14] shows that using deep learning features extracted from different deep learning models can improve the accuracy of breast ultrasound image classification. However, the concatenation method cannot effectively fuse the different deep learning features. In [15], the deep learning features were fused by artificial weighting, which partially improved the classification results when compared to the concatenation method. Moreover, as the malignant breast tumor infiltrates into the edge tissues, it forms a fuzzy infiltrating zone at the edge of the lesion, which helps in distinguishing between benign and malignant breast ultrasound images. Therefore, the following contributions can be observed in the present study:

(1) The original breast ultrasound and mask images of the lesion area were processed to generate a variety of decomposition images.
(2) Three types of RGB images were constructed by combining the decomposed, original ultrasound, and mask images.
(3) Adaptive spatial feature fusion technology was used to fuse the deep learning features extracted from RGB images.

2. Methods

Owing to the inherent speckle noise and the interference of the area outside the lesion, it is difficult to obtain good classification results using only the original breast ultrasound image [13,29]. Therefore, the classification of breast ultrasound images based on multi-image or multi-model fusion [12,14,15] has recently become a hot topic. In our experiment, various characteristics used in the clinical diagnosis of breast ultrasound images were studied and combined with multi-feature fusion to distinguish between benign and malignant breast ultrasound images, as shown in Fig. 1.

First, based on the original breast ultrasound and mask images of the lesion area, image processing techniques, such as fuzzy enhancement, bilateral filtering, and image morphology operation, are used to obtain the fuzzy enhancement, bilateral filtering, edge complexity, edge blur, and lesion images. The above characteristic images can effectively highlight the benign and malignant morphological characteristics of breast lesions in ultrasound images. The original image, mask image, and the above five characteristic images are called decomposition images in this paper. Then, considering the differences in the features expressed in the decomposed images, different fusion images are obtained through RGB channel fusion. The deep learning features of each type of fusion image are extracted based on different transfer learning-based base models. Finally, the adaptive spatial feature fusion technology is used to achieve an adaptive assignment of weights for different deep learning features. An ANN performs the final classification.

2.1. Breast ultrasound image decomposition

To enhance the characteristics of breast ultrasound images, such as edge complexity, edge blur, calcification, echo within the lesion, and aspect ratio, in our research, based on the original breast ultrasound and mask images of the lesion area, fuzzy enhancement, bilateral filtering, edge complexity, lesion, and edge blur images are obtained through image decomposition.

The fuzzy enhancement image can enhance the contrast between the lesion and other non-lesion areas of the original breast ultrasound images and effectively reduce the impact of non-lesion regions on the discrimination between benign and malignant lesions. The fuzzy enhancement algorithm adopted in this study primarily achieves image enhancement through the following three steps: (1) The membership function is designed to transform the image from the spatial domain to the fuzzy set domain; (2) The fuzzy enhancement arithmetic operator is designed to process the image in the fuzzy set domain; (3) According to the membership function in the first step, the image is transformed from the fuzzy set domain to the spatial domain. From the research in [16], the following can be noted: if we assume that the image threshold obtained by the Otsu algorithm [17] is T, then the best enhancement result of breast ultrasound image can be obtained when $T/5$ is taken as the initial value of transition point P_c in the fuzzy enhancement algorithm and the number of iterations of fuzzy enhancement is $r = 1$, as shown in Fig. 2(b).

Because the original ultrasound breast image contains more inherent speckle noise, it interferes with the identification of benign and malignant lesions. Bilateral filtering can effectively reduce speckle noise while preserving the original image information. The bilateral filter used in our experiment consists of two functions, as shown in equation (1): $G\sigma_S(||p - q||)$ is determined by the spatial distance between two pixels in the image; $G\sigma_r(|Ip - Iq|)$ is determined by the gray value difference between two pixels. BIp is the gray value of P in the filtered image BI. Wp is a normalizing parameter. Through these two functions, the bilateral filter can simultaneously consider the spatial information and gray information of the image to smooth the speckle noise of the original breast ultra-

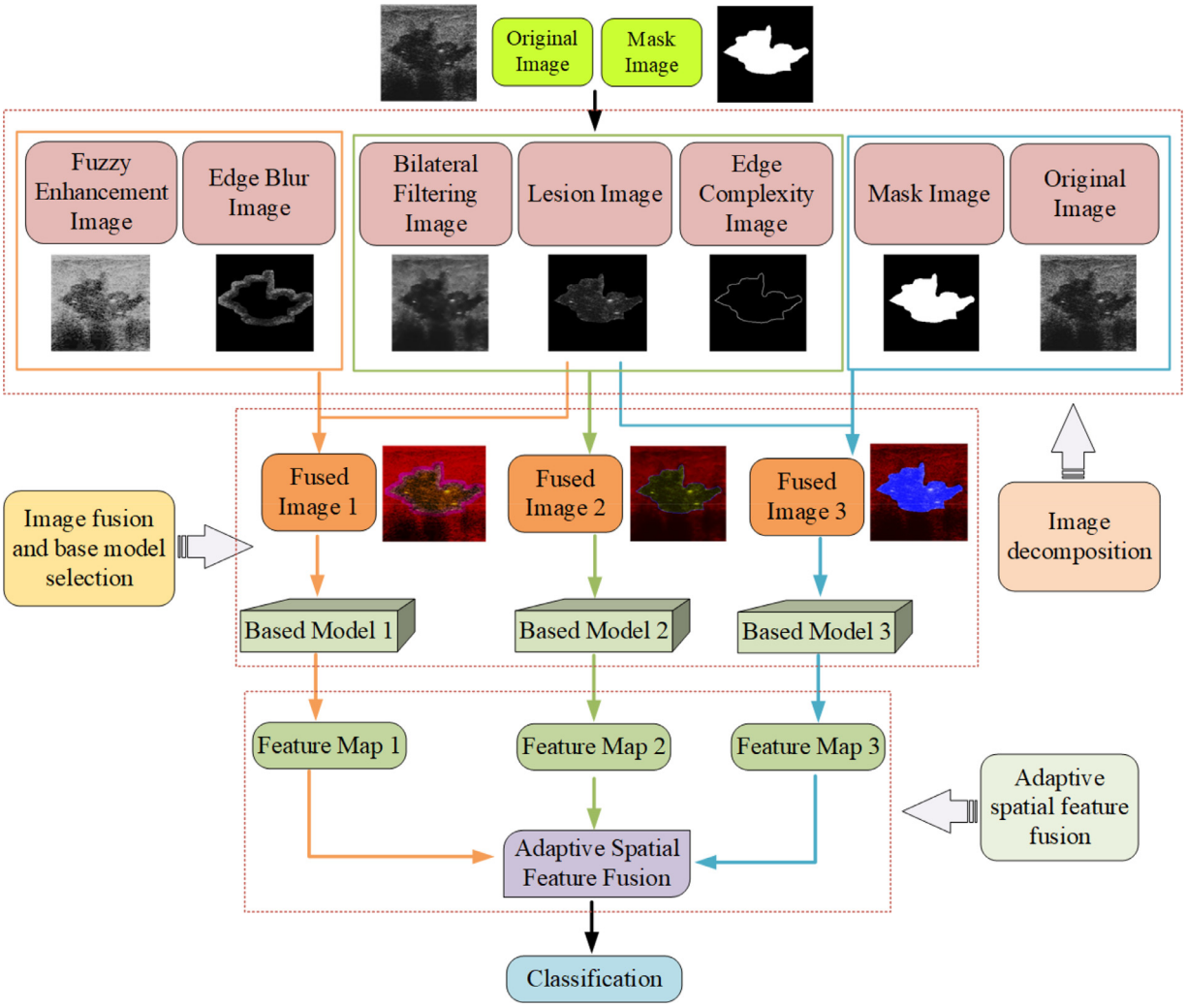

Fig. 1. Framework of the proposed method

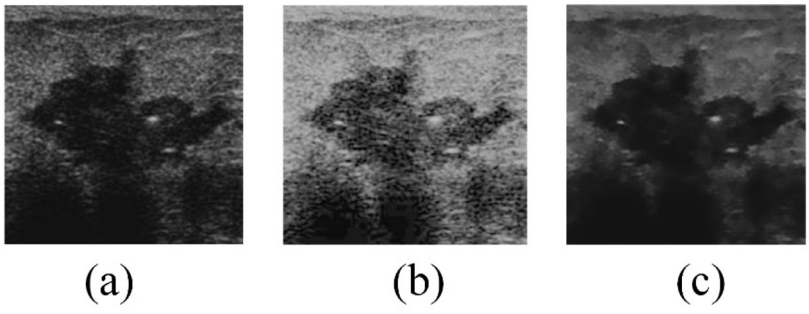

Fig. 2. Fuzzy enhancement and bilateral filtering: (a) Original breast ultrasound image; (b) Fuzzy enhancement image; (c) Bilateral filtering image

sound image and retain the edge information of the lesion area, as shown in Fig. 2(c).

$$BIp = \frac{1}{Wp}\sum_{q\in S} G\sigma_S(||p-q||)G\sigma_r(|Ip-Iq|) \quad (1)$$

Owing to the irregular infiltration of malignant breast lesions into surrounding tissues, the edges of the lesions are more complicated than in the case of benign lesions. To highlight the edge complexity of breast tumors, the Canny operator was used to extract the edge of the lesion in the lesion mask image, as shown in Fig. 3(c).

There is a wealth of information on breast ultrasound image lesions that can aid in distinguishing between benign and malignant lesions. First, malignant and benign tumors have different growth characteristics; consequently, they have different aspect ratio characteristics; moreover, calcification and echo in the lesion area are often used as important indicators to evaluate benign and malignant tumors [18]. Therefore, in this study, the original breast ultrasound image was multiplied by the mask image, and the lesion area was extracted as the lesion image, as shown in Fig. 3(d).

Edge blur is a critical indicator used by clinical doctors to distinguish between benign and malignant breast tumors. However, researchers often only focus on the edge complexity of the lesion area, thus ignoring edge ambiguity research. This study first dilated the edge complexity image and then multiplied the dilated binary image with the original breast ultrasound image to obtain the edge

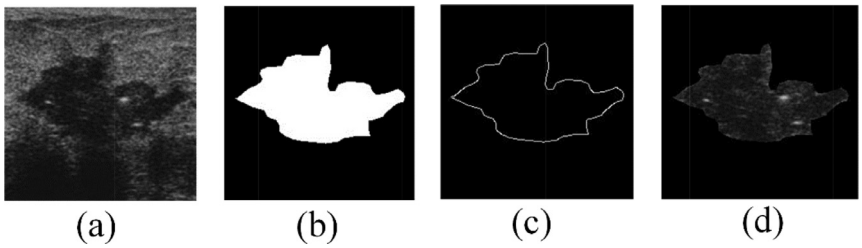

Fig. 3. Edge and lesion extraction: (a) Original breast ultrasound image; (b) Mask image; (c) Edge complexity image; (d) Lesion image

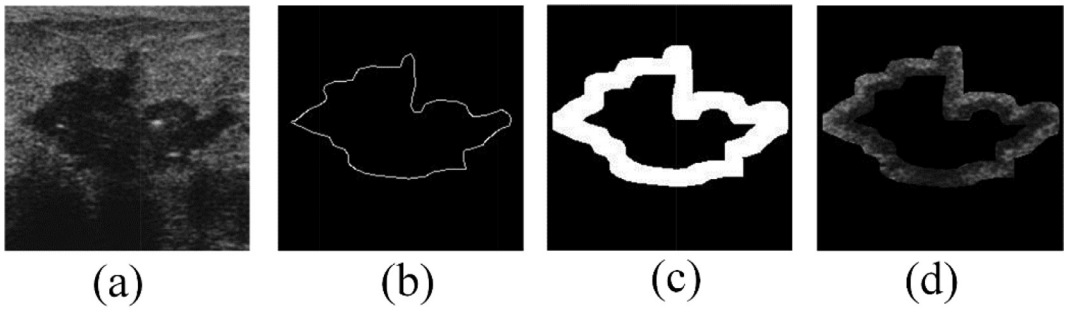

Fig. 4. Extracting the edge blur information of lesions: (a) Original breast ultrasound image; (b) Edge complexity image of the lesion; (c) Mask image of the edge of the lesion area; (d) Edge blur image of the lesion

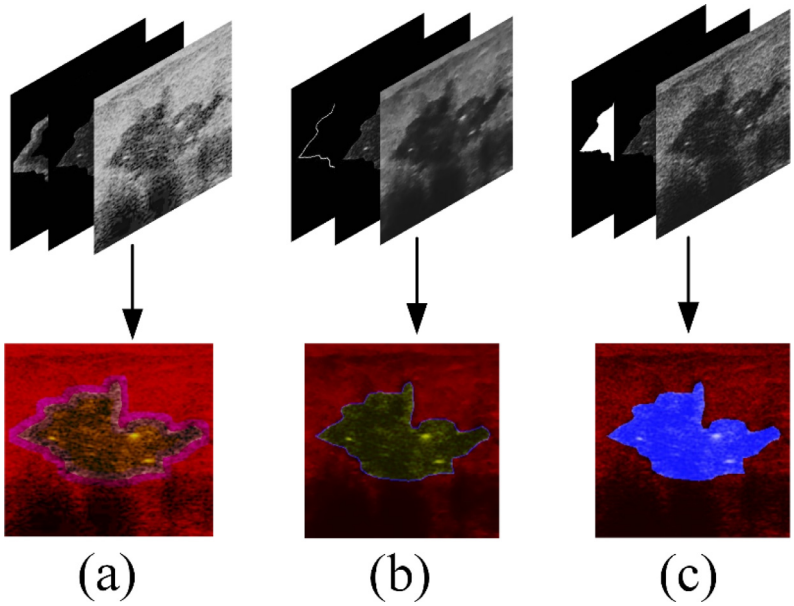

Fig. 5. Schematic illustration of image fusion. (a) Fusion image I is the RGB channel fusion of the fuzzy enhancement image, lesion image (LI), and edge blur image; (b) Fusion image II is the RGB channel fusion of the bilateral filter image, LI, and edge complexity image; (c) Fusion image III is the RGB channel fusion of the original image, LI, and mask image

blur image, which only contains the edge blur information of the lesion, as shown in Fig. 4.

2.2. Multi-channel image fusion based on decomposed images

Image fusion is widely used in target detection [19,20] and the detection of lesion locations from medical images [21]. In our study, using the above five characteristic images in combination with the original and mask images, multichannel fusion was conducted to fully utilize the cross information between different decomposed images to enhance the features that are used to differentiate between benign and malignant breast lesions. Image fusion is shown in Fig. 5.

In fusion image I (Fig. 5(a)), the contrast of the image is enhanced by the fuzzy enhancement image, which could weaken the interference of other tissues in the classification. Further, the edge blur image is used to strengthen the importance of edge ambiguity in discriminating between benign and malignant tumors. In fusion image II (Fig. 5(b)), bilateral filtering and edge complexity images are added, which can not only reduce the disturbance of speckle noise in the original breast ultrasound image, but also highlight the complexity of the tumor edge, and compensate for the loss of information in fusion image I caused by image enhancement. The third type of fusion image can show the original features of the breast tumor image and the morphological features of the lesion area more clearly, such as aspect ratio and the area of the lesion. Finally, because the lesion area of a breast ultrasound image contains rich features, such as echo, calcification, to distinguish between benign and malignant tumors, we added the lesion image as one of the RGB channels in the three types of fusion images.

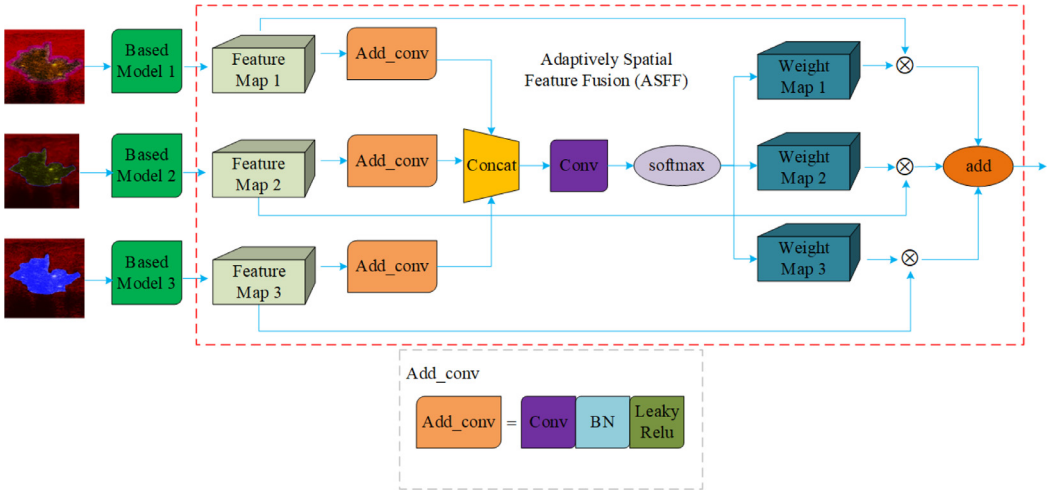

Fig. 6. Adaptive spatial feature fusion network. After deep learning features are extracted from the base model, the adaptive spatial feature fusion network is used for feature fusion

Experiments were conducted using VGG, ResNet, and densely connected CNN (DensNet) models to select the base model suitable for extracting each type of deep learning feature from the fused image. The VGG models are composed of several convolution blocks and a classification layer. In our experiment, the last convolution layer and the classification layer were fine-tuned. For the ResNet and DensNet models, only the classification layer is fine-tuned to prevent overfitting, making it suitable for detecting breast ultrasound images.

2.3. Multi-feature fusion based on an adaptive network

Feature fusion has been widely used in image classification [12,14,15], data fusion [22,23], and other fields [24]. However, general feature fusion usually adopts feature concatenation or simple feature addition. When the number of features is large, concatenating features will produce a large amount of computational overhead, and there will be more redundant information. When using simple feature addition, if there are discrimination conflicts between different feature maps, classification errors may be produced. In [25], multiple weight networks were trained to fuse the feature maps of different scales. This method solves the problem of error discrimination caused by the simple addition of features and has less computational overhead. Because the feature maps extracted from different models in our experiment have the same size, the adaptive spatial feature fusion network that is proposed in [25] was simplified. Only one weight network was used to satisfy the fusion of feature maps from different models in this study. The adaptive spatial feature fusion network used in this work is shown in Fig. 6. *BN* represents batch normalization, which was proposed by Ioffe [26]. It is significantly useful for accelerating deep network training by reducing the internal covariate shift. *Con* represents convolution.

First, after selecting the base model for various fusion images, the feature maps are extracted. A convolution module including a convolution layer, batch normalization layer, and activation function is used to calculate the feature maps, and the results are concatenated. Then, we use the softmax function to obtain three weight maps after convolution, which correspond to the three feature maps extracted from the base model. Finally, the weight map is multiplied by its corresponding feature map. Because the scale of the results is consistent, we can add them directly. As the adaptive network is a part of the entire classification network, the adaptive network can dynamically optimize the weight parameters of different feature maps through backpropagation to improve the classification ability of the entire model when training the classification network.

Another advantage of adaptive spatial feature fusion is its interpretability. That is, it can be explained using a formula derivation. When the adaptive spatial feature fusion network is not added, we can simplify the gradient calculation formula of a pixel (i, j) in the feature map as

$$\frac{\partial L}{\partial x_{ij}} = \frac{\partial L}{\partial y_{ij}^1} + \frac{\partial L}{\partial y_{ij}^2} + \frac{\partial L}{\partial y_{ij}^3}, \tag{2}$$

where L is the loss function and $\frac{\partial L}{\partial x_{ij}}$ is the gradient of pixels (i,j) during backpropagation. $\frac{\partial L}{\partial y_{ij}^1}$, $\frac{\partial L}{\partial y_{ij}^2}$, and $\frac{\partial L}{\partial y_{ij}^3}$ are the backpropagation gradients of each model. When the discrimination of a point is contradictory between different base models, this simple addition-based fusion method interferes with backpropagation. By adding an adaptive spatial feature fusion network, equation (2) can be expressed as

$$\frac{\partial L}{\partial x_{ij}} = \alpha_{ij}^1 \cdot \frac{\partial L}{\partial y_{ij}^1} + \alpha_{ij}^2 \cdot \frac{\partial L}{\partial y_{ij}^2} + \alpha_{ij}^3 \cdot \frac{\partial L}{\partial y_{ij}^3}, \tag{3}$$

Equation 3 adds trainable weight parameters based on equation 2. These parameters can achieve the optimal values with the entire network in the backpropagation process. Thus, the inaccuracy of artificial weight assignments can be avoided.

3. Results

3.1. Dataset and processing

As listed in Table 1, the dataset used in this study consists of four parts, which include 1328 breast ultrasound images, 707

Table 1
Dataset description

		Benign	Malignant	Total
Dataset 1	BUSI	437	210	647
Dataset 2	OMI	100	148	248
Dataset 3	Dataset B	110	53	163
Dataset 4	Hospital	60	210	270
	Total	707	621	1328

Table 2
Evaluation index.

Index	Definition	Explanation
Sensitivity / Recall	$\frac{TP}{TP+FN}$	The proportion of true positive samples that are correctly classified.
Specificity	$\frac{TN}{TN+FP}$	The proportion of true negative samples that are correctly classified.
Precision	$\frac{TP}{TP+FP}$	The proportion of correctly classified samples among all the samples identified as positive.
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$	The proportion of correctly classified samples in the total number of samples.
F1 score	$\frac{2TP}{2TP+FP+FN}$	Harmonic value of sensitivity and accuracy.

TP: True Positive, the number of positive samples that were accurately predicted as positive.
TN: True Negative, the number of negative samples that were accurately predicted as negative.
FP: False Positive, the number of negative samples that were predicted as positive samples.
FN: False Negative, the number of positive samples that were predicted as negative samples.

Table 3
Classification results of fusion image I using different deep learning models. The values in bold font indicate the best index values.

Model	Accuracy	Precision	Specificity	Sensitivity (Recall)	F1 Score
VGG11	**0.9296**	**0.9826**	**0.9840**	**0.9090**	**0.9419**
VGG16	0.9246	0.9549	0.9621	0.8842	0.9155
VGG19	0.8995	0.9388	0.9442	0.8473	0.8870
Resnet101	0.9196	0.9546	0.9579	0.9075	0.9294
Densnet121	0.9095	0.9355	0.9412	0.9005	0.9171
Densnet161	0.8744	0.9482	0.9596	0.8301	0.8837

benign cases, and 621 malignant cases. Datasets 1 [27], 2 [28], and 3 [6] are all public datasets, and dataset 4 is from the imaging department of the First Affiliated Hospital of Shantou University, which was obtained by a senior radiologist. This study randomly divided the dataset into three parts, i.e., training, validation, and test sets. The training set accounted for 70% of the dataset, and the validation and test sets each accounted for approximately 15%.

For the dataset, image processing was used to obtain fuzzy enhancement, bilateral filter, edge complexity, edge blur, and lesion images. Then, three types of fusion images are obtained through RGB channel fusion. Finally, to avoid overfitting and to learn more image information, in this study. the training set was flipped to expand the training dataset.

3.2. Evaluation index

To select the best base model for extracting deep learning features from each type of fusion image and evaluate the superiority of the proposed method, this study used sensitivity/recall, specificity, precision, accuracy, and F1 score to evaluate the classification results; these parameters are explained in Table 2. Moreover, we also used the receiver operating characteristic (ROC) curve and area under the ROC curve (AUC) value as evaluation indexes of different classification methods. The X- and Y-axes of the plane where the ROC curve is located are the false positive rate and true positive rate, respectively. Because the ROC curves are difficult to compare intuitively, the method is generally compared using the AUC values.

3.3. Base model selection

In this experiment, the fine-tuned VGG11 [30], VGG16 [30], VGG19 [30], ResNet101 [31], DensNet121 [32], DensNet161 [32] models and transfer learning were used to classify the three types of fusion images (as shown in Fig. 5), which were composed of decomposed images through RGB channel fusion. The experimental results for fusion images I, II, and III are listed in Tables 3, 4, and 5, respectively.

From Table 3, we can observe that when VGG11 is used for classifying fusion image I, the values of accuracy, precision, specificity, sensitivity/recall, and F1 score are 0.9296, 0.9826, 0.9840, 0.9090, and 0.9419, respectively, which are higher than the corresponding classification results of other models.

In addition, it can be observed from Table 4 that, when the VGG16 model is used as the transfer learning classifier, the values of accuracy, precision, specificity, sensitivity/recall, and F1 score are 0.9397, 0.9672, 0.9643, 0.9318, and 0.9482, respectively. When compared to other models, excluding the specificity that is marginally lower than that of the VGG19 model, the other indexes reflect that the VGG16 model has a higher classification ability for fusion image II. Therefore, this study used the convolution layer of VGG16 as the base model of fusion image II to extract deep learning features.

From Table 5, it can be observed that when VGG19 was used to classify fusion image III, the values of accuracy, precision, specificity, sensitivity/recall, and F1 score are 0.9246, 0.9579, 0.9703, 0.9022, and 0.9290, respectively, which are higher than the index values of other classification models. Therefore, the convolution layer in VGG19 was used as the base model for fusion image III.

3.4. Adaptive spatial feature fusion

For the three types of fusion images, the deep learning feature maps were extracted using their respective base models, which were selected in the former experiment. Then, the adaptive spatial feature network was trained to obtain three weight maps. Next, the weight map was multiplied with their corresponding deep-learning-based feature map, and the classification layer was used for classification. The proposed method (Method 8) was compared with several classification methods and the results are listed in Table 6. Method 1 indicates the method where VGG16 was used to classify the original breast ultrasound image. Methods 2, 3, and 4 represent the methods where the best deep learning model was selected for each type of fusion image. Fusion image III is also the image fusion method proposed in [15]. Method 5 uses the weighted-average method used in [15] to fuse deep learning features. Methods 6 and 7 fused deep learning features by using feature addition and feature concatenation, respectively.

Table 4
Classification results of fusion image II using different deep learning models. The values in bold font indicate the best index values.

	Accuracy	Precision	Specificity	Sensitivity (Recall)	F1 Score
VGG11	0.9046	0.8790	0.9020	0.8495	0.8614
VGG16	**0.9397**	**0.9672**	0.9643	**0.9318**	**0.9482**
VGG19	0.9296	0.9650	**0.9672**	0.9160	0.9389
Resnet101	0.8643	0.9100	0.9302	0.7643	0.8188
Densnet121	0.8643	0.9286	0.9329	0.7806	0.8445
Densnet161	0.8794	0.9143	0.9199	0.8460	0.8741

Table 5
Classification results of fusion image III using different deep learning models. The values in bold indicate the best index values

	Accuracy	Precision	Specificity	Sensitivity (Recall)	F1 Score
VGG11	0.9095	0.9461	0.9516	0.8990	0.9213
VGG16	0.8995	0.9573	0.9704	0.8152	0.8788
VGG19	**0.9246**	**0.9579**	**0.9703**	**0.9022**	**0.9290**
Resnet101	0.9045	0.9510	0.9621	0.8779	0.9118
Densnet121	0.9196	0.9500	0.9611	0.9019	0.9240
Densnet161	0.8794	0.9376	0.9397	0.8531	0.8894

Table 6
Comparison of different classification methods. The values in bold font indicate the best index values. Moreover, BUSI is an abbreviation for Breast Ultrasound Image.

		Accuracy	Precision	Specificity	Sensitivity (Recall)	F1 Score
Method 1	Original BUSI (VGG16)	0.8492	0.9088	0.9105	0.8395	0.8716
Method 2	Fusion image I (VGG11)	0.9296	**0.9826**	**0.9840**	0.9090	0.9419
Method 3	Fusion image II (VGG16)	0.9397	0.9672	0.9643	0.9318	0.9482
Method 4	Fusion image III (VGG19)	0.9246	0.9579	0.9703	0.9022	0.9290
Method 5	Weighted- Average[15]	0.9296	0.9719	0.9723	0.9020	0.9348
Method 6	Addition	0.9246	0.9594	0.9665	0.9006	0.9270
Method 7	Concatenation	0.9347	0.9084	0.9270	0.9383	0.9169
Method 8	Proposed method	**0.9548**	0.9811	0.9833	**0.9392**	**0.9571**

By comparing the classification results listed in Table 6, it can be observed that the accuracy, precision, specificity, sensitivity/recall, and F1 score obtained by using the proposed method based on adaptive spatial feature fusion are 0.9548, 0.9811, 0.9833, 0.9392, and 0.9571, respectively. The precision and specificity values are marginally lower than the scores of 0.9826 and 0.9840, respectively, which were obtained for fusion image I. The other indexes are higher than those of the other methods. Moreover, by comparing the ROC curves and AUC values of different classification models in Fig. 7, it can be determined that the AUC value obtained by using the adaptive feature fusion method is 0.9883, which is higher than that of the other methods.

4. Discussion

Different deep-learning-based transfer models show different classification abilities for different types of fusion images when selecting the base model. VGG11 showed better deep learning feature extraction ability than other deep learning models for fusion image I. This is because the fuzzy enhanced image used in fusion image I is processed by fuzzy enhancement, which reduces the complexity of the original breast ultrasound image. Therefore, the model with fewer convolution layers can extract features better. For fusion image II, the bilateral filtering image was added to weaken the speckle noise of the original ultrasound image while retaining the necessary information. With regard to image complexity, the complexity of the bilateral filtering image is higher than that of the fuzzy enhancement image, but lower than that of the original ultrasound image. Therefore, VGG16 with a moderate number of convolution layers achieved a better classification effect. The original image was used in fusion image III. Its complexity is higher than that of the fuzzy enhanced and bilateral filtering images. Therefore, VGG19 with more convolution layers can extract the deep learning features of fusion image III more effectively.

Further, when compared to ResNet or DenseNet models, choosing VGG models as the base model for transfer learning has a better classification effect for different fused images. This is because, this study aimed to classify benign and malignant breast ultrasound images, and the complexity of ultrasound images is significantly lower than that of the 1000 categories of natural images used in the pretraining model. When deeper models such as ResNet or DenseNet models are used, the speckle noise or irrelevant regions in breast ultrasound images have a more severe impact on the classification, which leads to a decline in the classification effect.

In recent years, adversarial machine learning (AML) has been gradually applied to train more robust image classification models [33,34]. A variety of proposed attack models have been developed to improve the interpretability of deep learning systems [35–37]. However, the breast ultrasound tumor image used in this study is different from a natural image, which contains a large amount of inherent speckle noise and is not suitable for use in AML. Moreover, the adaptive spatial feature fusion method proposed in this paper trains the weight and classification models together, which can automatically obtain more accurate weights with lesser computation cost when compared to artificial weight or simple addition methods [12–15].

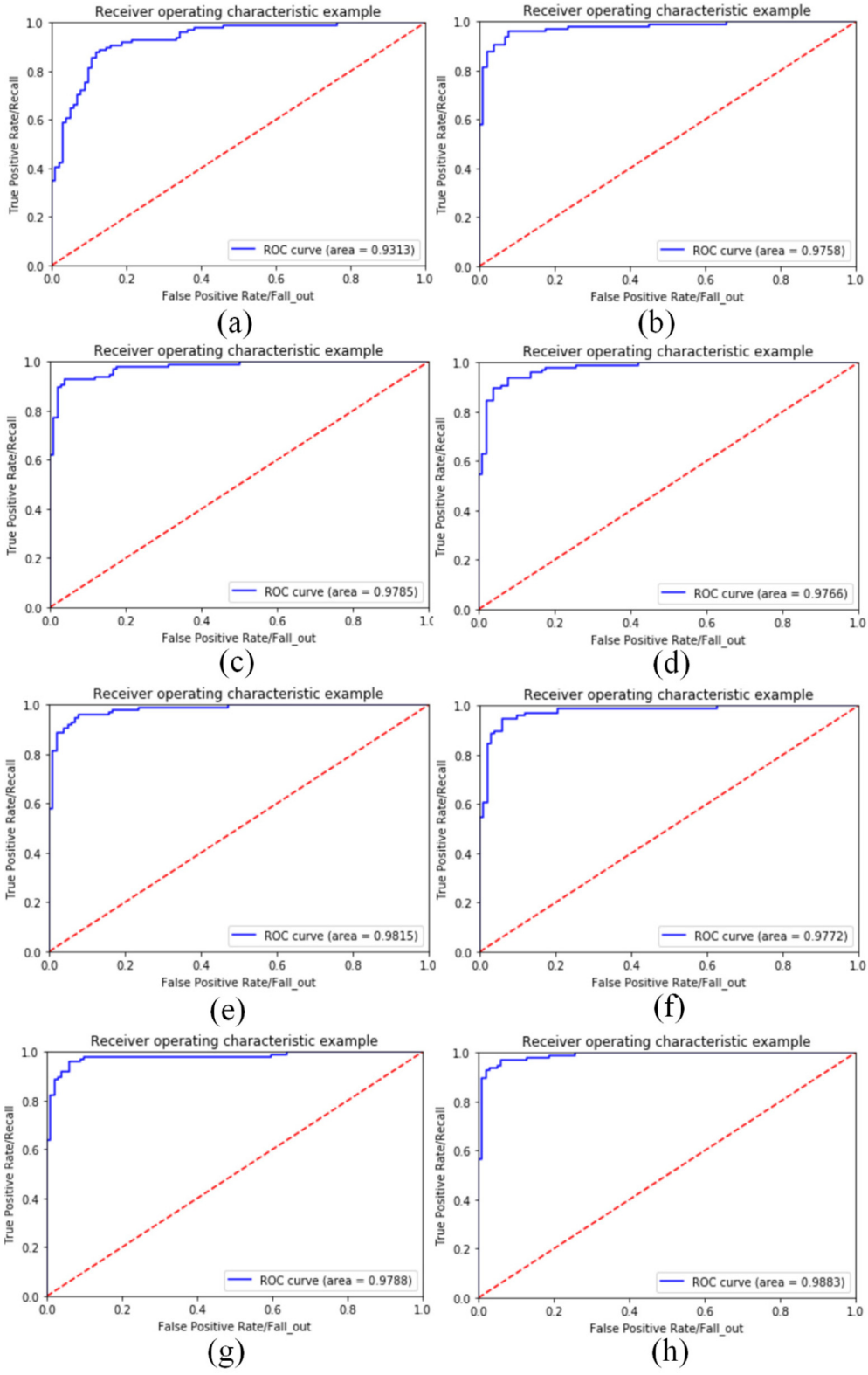

Fig. 7. Receiver operating characteristic (ROC) curves and area under ROC curve values of the classification method based on the results listed in Table 6. Here, (a)-(h) correspond to methods 1-8, respectively

5. Conclusion

CAD detection methods can assist doctors in the clinical diagnosis of breast ultrasound tumor images and other medical images [38–40]. Information fusion may also be implemented to provide a concise diagnosis of medical conditions [41]. In our research, first, based on the original ultrasound image, fuzzy enhancement and bilateral filtering were used to obtain fuzzy enhanced and bilateral filtering images, respectively. At the same time, according to the clinical experience of the doctor, this study used the original and mask images to decompose the image and obtain the edge complexity, lesion, and edge blur images; then, based on the

above seven types of images, three types of fusion images were obtained by RGB channel fusion. The lesion image was included in each type of fusion image because of its rich clinical features. Then, the fine-tuned VGG11, VGG16, VGG19, ResNet101, DensNet121, and DensNet161 models were used to select the base model of each type of fused image through transfer learning. For the deep learning features extracted from each type of fusion image based on the base model, this paper proposes the use of adaptive spatial feature fusion to fuse different deep learning features, and use an ANN to classify the fused features. The experimental results show that the proposed classification method based on adaptive spatial feature fusion can assist doctors more effectively in clinical diagnosis when compared to other methods.

Declaration of Competing Interest

The authors declare that there is no conflict of interests.

Acknowledgment

Ethical approval was not required for this research. This work was supported by the National Natural Science Foundation of China [grant number 82071992], Basic and Applied Basic Research Foundation of Guangdong Province [grant number 2020B1515120061], the Guangdong Province University Priority Field (Artificial Intelligence) Project [grant number 2019KZDZX1013], National Key R&D Program of China [grant number 2020YFC0122103], the Key Project of Guangdong Province Science & Technology Plan [grant number 2015B020233018]; and the Research Project of Shantou University, China [grant number NTF17016]. The authors declare that there are no conflicts of interest.

References

[1] F. Bray, J. Ferlay, I. Soerjomataram, R.L. Siegel, L.A. Torre, A. Jemal, Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries, CA: a cancer journal for clinicians 68 (6) (2018) 394–424.
[2] https://www.iarc.who.int/faq/latest-global-cancer-data-2020-qa/, (19 April 2012).
[3] F. Islami, A. Goding Sauer, K.D. Miller, R.L. Siegel, S.A. Fedewa, E.J. Jacobs, A.Jemal ..., Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States, CA: a cancer journal for clinicians 68 (1) (2018) 31–54.
[4] J. Anitha, J.D. Peter, Mammogram segmentation using maximal cell strength updation in cellular automata, Medical & biological engineering & computing 53 (2015) 737–749.
[5] W.A. Berg, Z. Zhang, D. Lehrer, R.A. Jong, E.D. Pisano, … G. Gabrielli, Detection of breast cancer with addition of annual screening ultrasound or a single screening MRI to mammography in women with elevated breast cancer risk, Journal of the American Medical Association 307 (13) (2012) 1394–1404.
[6] H. Yuzhou, Y. Guo, W. Yuanyuan, Y. Jinhua, L. Jiawei, Z. Shichong, C. Cai, Automatic tumor segmentation in breast ultrasound images using a dilated fully convolutional network combined with an active contour model, Medical Physics 46 (1) (2018) 215–228.
[7] M.H. Yap, G. Pons, J. Martí, S. Ganau, M. Sentís, R. Zwiggelaar, … R. Martí, Automated Breast Ultrasound Lesions Detection Using Convolutional Neural Networks, IEEE Journal of Biomedical and Health Informatics 22 (4) (2018) 1218–1226.
[8] W.A. Berg, L. Gutierrez, M.S. NessAiver, W.B. Carter, M. hargavan, R.S. Lewis, O.B. Ioffe, Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer, Radiology 233 (3) (2004) 830–849.
[9] M.J.G. Calas, R.M.V.R. Almeida, B. Gutfilen, W.C.A. Pereira, Intraobserver interpretation of breast 680 ultrasonography following the BI-RADS classification, European Journal of Radiology 74 (2010) 525–528.
[10] J.M.H. Timmers, H.J. van Doorne-Nagtegaal, A.L.M. Verbeek, G.J. den Heeten, A dedicated BIRADS training programme: effect on the inter-observer variation among screening radiologists, European Journal of Radiology 81 (9) (2012) 2184–2188.
[11] C. Dromain, B. Boyer, R. Ferre, S. Canale, S. Delaloge, C. Balleyguier, Computed-aided diagnosis (CAD) in the detection of breast cancer, European journal of radiology 82 (3) (2013) 417–423.
[12] T. Xiao, L. Liu, K. Li, W. Qin, S. Yu, Z. Li, Comparison of transferred deep neural networks in ultrasonic breast masses discrimination, BioMed research international (2018) (2018), doi:10.1155/2018/4605191.
[13] J. Virmani, R. Agarwal, Deep feature extraction and classification of breast ultrasound images, Multimedia Tools and Applications 79 (37) (2020) 27257–27292.
[14] 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 journal of biomedical and health informatics 24 (4) (2019) 984–993.
[15] W.K. Moon, Y.W. Lee, H.H. Ke, S.H. Lee, C.S. Huang, R.F. Chang, Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks, Computer methods and programs in biomedicine 190 (2020), doi:10.1016/j.cmpb.2020.105361.
[16] Z. Zhuang, Y. Kang, A.N. Joseph Raj, Y. Yuan, W. Ding, S. Qiu, Breast ultrasound lesion classification based on image decomposition and transfer learning, Medical Physics 47 (12) (2020) 6257–6269.
[17] P.D. Sathya, R. Kalyani, V.P. Sakthivel, Color image segmentation using Kapur, Otsu and Minimum Cross Entropy functions based on Exchange Market Algorithm, Expert Systems with Applications 172 (2021), doi:10.1016/j.eswa.2021.114636.
[18] X. Qi, L. Zhang, Y. Chen, Y. Pi, Y. Chen, Q. Lv, Z. Yi, Automated diagnosis of breast ultrasonography images using deep neural networks, Medical image analysis 52 (2019) 185–198.
[19] S. Liu and Z. Liu, Multi-channel CNN-based object detection for enhanced situation awareness, (2017), arXiv preprint arXiv:1712.00075.
[20] P. Yu, Y. Zhao, J. Zhang, X. Xie, Pedestrian detection using multi-channel visual feature fusion by learning deep quality model, Journal of Visual Communication and Image Representation 63 (2019), doi:10.1016/j.jvcir.2019.102579.
[21] J.C. Benson, M.L. Carlson, L. Yin, J.I. Lane, Cholesteatoma Localization Using Fused Diffusion-Weighted Images and Thin-Slice T2 Weighted Images, The Laryngoscope (2020), doi:10.1002/lary.29222.
[22] A. Shakya, M. Biswas, M. Pal, CNN-based fusion and classification of SAR and Optical data, International Journal of Remote Sensing 41 (22) (2020) 8839–8861.
[23] C. Tan, Y. Sun, G. Li, G. Jiang, D. Chen, H. Liu, Research on gesture recognition of smart data fusion features in the IoT, Neural Computing and Applications 32 (22) (2019) 1–13.
[24] H. Dong, K. Song, Y. He, J. Xu, Y. Yan, Q. Meng, PGA-Net: Pyramid feature fusion and global context attention network for automated surface defect detection, IEEE Transactions on Industrial Informatics 16 (12) (2019) 7448–7458.
[25] S. Liu, D. Huang, and Y. Wang, Learning spatial fusion for single-shot object detection, (2019), arXiv preprint arXiv:1911.09516.
[26] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, International conference on machine learning 37 (2015) 448–456.
[27] W. Al-Dhabyani, M. Gomaa, H. Khaled, A. Fahmy, Dataset of breast ultrasound images, Data in brief 28 (2020), doi:10.1016/j.dib.2019.104863.
[28] http://www.onlinemedicalimages.com/index.php/en/component/record/, (19 April 2012).
[29] K. Drukker, M.L. Giger, C.J. Vyborny, E.B. Mendelson, Computerized detection and classification of cancer on breast ultrasound[1], Academic radiology 11 (5) (2004) 526–535.
[30] K. Simonyan, and A. Zisserman, Very deep convolutional networks for large-scale image recognition, (2014), arXiv preprint arXiv:1409.1556.
[31] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.
[32] G. Huang, Z. Liu, L. Van Der Maaten, K.Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708.
[33] G. R. Machado, E. Silva, and R. R. Goldschmidt, Adversarial Machine Learning in Image Classification: A Survey Towards the Defender's Perspective, (2020), arXiv preprint arXiv:2009.03728.
[34] W. Woods, J. Chen, C. Teuscher, Adversarial explanations for understanding image classification decisions and improved neural network robustness, Nature Machine Intelligence 1 (11) (2019) 508–516.
[35] X. Zhang, N. Wang, H. Shen, S. Ji, X. Luo, T. Wang, Interpretable deep learning under fire, 29th {USENIX} Security Symposium ({USENIX} Security 20), 2020.
[36] Y. Ji, X. Zhang, S. Ji, X. Luo, T. Wang, Model-reuse attacks on deep learning systems, in: Proceedings of the 2018 ACM SIGSAC Conference on Computer and Communications Security, 2018, pp. 349–363.
[37] ... R. Pang, H. Shen, X. Zhang, S. Ji, Y. Vorobeychik, X. Luo, T. Wang, A tale of evil twins: Adversarial inputs versus poisoned models, in: Proceedings of the 2020 ACM SIGSAC Conference on Computer and Communications Security, 2020, pp. 85–99.
[38] K.K.L. Wong, Z. Sun, J.Y. Tu, S.G. Worthley, J. Mazumdar, D. Abbott, Medical image diagnostics based on computer-aided flow analysis using magnetic resonance images, Computerized Medical Imaging and Graphics 36 (7) (2012) 527–541.
[39] K.K.L. Wong, Z. Sun, J.Y. Tu, Medical imaging and computer-aided flow analysis of a heart with atrial septal defect, Journal of Mechanics in Medicine and Biology 12 (5) (2012) 1–28.
[40] K.K.L. Wong, G. Fortino, D. Abbott, Deep learning-based cardiovascular image diagnosis: A promising challenge, Future Generation Computer Systems 110 (2020) 802–811.
[41] X. Liu, M. Zhao, A. Liu, K.K.L. Wong, Adjusting forwarder nodes and duty cycle using packet aggregation routing for body sensor networks, Information Fusion 53 (2020) 183–195.