IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received February 16, 2022, accepted March 18, 2022, date of publication March 31, 2022, date of current version April 7, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3163822

Breast Cancer Classification From Histopathological Images Using Resolution Adaptive Network

YIPING ZHOU, CAN ZHANG, AND SHAOSHUAI GAO, (Member, IEEE)
School of Electronic, Electrical and Communication Engineering, University of Chinese Academy of Sciences, Beijing 101408, China

Corresponding author: Yiping Zhou (zhouyiping14@mails.ucas.ac.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61571416 and Grant 61271282, and in part by the Award Foundation of Chinese Academy of Sciences under Grant 2017-6-17.

ABSTRACT A histopathological analysis performed by pathologists plays a key role in the diagnosis of breast cancer. A novel approach based on an image processing technique is proposed to help pathologists efficiently produce accurate diagnoses, that is composed of two modules, namely, anomaly detection with a support vector machine (ADSVM) method and a resolution adaptive network (RANet) model. The ADSVM method screens mislabeled patches to improve the training performance of the RANet model. In the RANet model, subnetworks with variable resolutions and depths are utilized to classify images according to the classification difficulty. This adaptive mechanism potentially increases the computational efficiency and prediction accuracy. The proposed RANet-ADSVM approach is evaluated using two public datasets: the BreaKHis and BACH 2018 datasets. Binary and multiclass classifications of patient and image levels at different magnification factors are conducted on the BreaKHis dataset. The best accuracies of 98.83% and 99.14% are obtained for the binary classification at 200× magnification at the patient and image levels, respectively. For the BACH 2018 dataset, binary and multiclass classifications on patch and image levels are performed. Experimental results show that the best accuracies for multiclass and binary classifications at the image level are 97.75% and 99.25%, respectively. Additionally, comparative experiments are performed and indicated that the proposed approach achieves significant improvements in both the classification accuracy and computational efficiency. Compared with similar networks (ResNet and DenseNet), the computational time is reduced by approximately 50%.

INDEX TERMS Breast cancer image classification, resolution adaptive network, anomaly detection, support vector machine, convolutional neural network.

I. INTRODUCTION

Breast cancer is the leading cause of morbidity and mortality in women worldwide. According to a study reported by the International Agency for Research on Cancer (IARC) [1], approximately 9.2 million new cancer cases in women were reported in 2020, of which 2.29 million were breast cancer cases. Breast cancer accounted for 15.5% of all 4.4 million cancer-related deaths in women. Early diagnosis is critical for increasing the survival rate and improving the quality of life of the patients [2]. In breast cancer diagnosis, a histopathological analysis of breast tissue biopsy images is essential.

The associate editor coordinating the review of this manuscript and approving it for publication was Francesco Piccialli.

Breast tissues are classified into four types: normal tissues, benign lesions, *in situ* carcinomas, and invasive carcinomas. Benign lesions occur in the normal tissue of the breast parenchyma and are unrelated to malignant carcinogenesis. *In situ* carcinomas and invasive carcinomas are two types of malignant breast cancer. The cancerous cells are limited inside the mammary ductal-lobular system in *in situ* tissue, whereas the cells spread beyond the structure in invasive tissue. Pathologists analyze the microstructure in images of a biopsy stained with hematoxylin and eosin (H&E) to grade and stage the tissue [3]. Image translation, rotation, and scaling operations are required during the analysis. This diagnostic process requires well-experienced specialists and considerable time and effort. Due to the complexity and

variability among histopathological images, the average diagnostic accuracy of pathologists is only 75% [4].

During the last decade, different machine-learning methods have been applied to breast cancer classification from histopathological images [5]–[7]. Early researchers mostly used traditional machine-learning methods and evaluated their methods using private datasets with a small sample size [8]–[10]. As a result, these methods do not meet the expectation in clinical practice. Recently, convolutional neural networks (CNNs) have been widely used for breast cancer classification [11]–[13]. Methods based on classic networks (e.g., ResNet [14] and DenseNet [15]) have been proposed and have yielded remarkable results.

However, in the previous CNN models, the original image is commonly subdivided into patches for follow-up CNN processing. These patches were marked with the same label as that of the corresponding image. Because malignant images may contain benign issues, this operation may cause the CNN to be trained on mislabeled patches, which may reduce the performance of the model. In addition, complex CNN models are utilized to achieve accurate classification results, which reduces the computational efficiency.

In this study, a deep learning approach based on an anomaly detection with a support vector machine (ADSVM) method and a resolution adaptive network (RANet) model is proposed for breast cancer classification. The main contributions of our work are summarized below.

1) A patch screening method based on ADSVM is proposed. In the method, a soft-margin SVM classifier is trained for binary classification (benign vs. malignant). The trained SVM classifier is used to recognize the benign patches, which are mislabeled as malignant, subdivided from malignant images. As a result, the ADSVM method reduces the number of mislabeled patches and increases the follow-up training performance.
2) A RANet model [16] is proposed for breast cancer classification from histopathological images. The RANet model is composed of subnetworks with adaptive input resolution feature maps to reduce the computational cost. Using different subnetworks, the images are classified based on distinguishing difficulty. The ''easy'' images are recognized by a low-resolution subnetwork with few convolutional layers, while the ''hard'' images are processed by a high-resolution subnetwork. This adaptive mechanism increases the computational efficiency based on the premise of prediction accuracy.
3) The image-level and patient-level evaluations are performed on the BreaKHis dataset for binary and multiclass classification with different magnifications.
4) The patch-level and image-level evaluations are performed on the BACH 2018 dataset for binary and multiclass classification.
5) The recognition performance and computational cost are compared with similar studies. Our approach outperforms other models in terms of classification accuracy and computational efficiency.

The remaining parts of the manuscript are organized as follows: Section II discusses related works, the datasets used in this study are described in Section III, the proposed approach is presented in detail in Section IV, Section V presents the classification experiments and results, and Section VI concludes the paper.

II. RELATED WORKS

As a result of advances in machine learning and image processing techniques, a significant increase in interest in the computer-aided diagnosis of breast cancer has been noted in the last decade [5]–[7]. Early researchers mainly used traditional machine-learning methods for breast cancer classification [8]–[10]. These approaches were always composed of two modules: a handcrafted feature extraction phase and a classifier. The feature extraction methods (e.g., scale invariant feature extraction (SIFT), local binary patterning (LBP), and local phase quantization (LPQ)) were used to extract local features from the input image. These features were then applied to a classifier (e.g., SVM, principle component analysis (PCA), random forest (RF), etc.) for binary classification (benign vs. malignant). These approaches were evaluated using a private dataset with a small sample size. As a result, these methods do not retain their accuracy and robustness in clinical practice.

Recently, convolutional neural networks (CNNs) have been widely used for breast cancer classification using well-known public datasets and have achieved satisfactory performances [11]–[13]. Spanhol *et al.* [17] proposed a method based on a pre-trained AlexNet model. The method was evaluated for patient-level and image-level classification of images with different magnifications in a popular database named BreakHis. At the patient level, the accuracy ranged from 84% to 90%. At the image level, the accuracy ranged from 80.8% to 85.6%. Dalal *et al.* [18] proposed a CNN-based approach for automatic binary and multiclass classification of the BreaKHis dataset. The CNN-based approach was compared with traditional machine learning approaches based on handcrafted features. The experimental results showed that the CNN-based approach achieved the best accuracy of 83.31% to 88.23% for multiclass classification. Although these CNN-based methods have achieved better performance than traditional machine learning methods, their classification performance has room for improvement.

With the developments in CNN models, classic networks (e.g., ResNet and DenseNet) have been proposed for image classification and have yielded remarkable results. Man *et al.* [19] constructed a DenseNet121-based network for the classification of breast cancer images in the BreaKHis dataset. The method achieved accuracies ranging from 85.16% to 96.32% for patient-level classification and 85.20% to 99.13% for image-level classification. Wang *et al.* [20] proposed another deep learning approach based on the FE-BkCapsNet model.

This approach combined the CNN model (focused on semantics) and CapsNet (focused on spatial features). The results showed accuracies ranging from 92.71% to 94.52% for classification of images in the BreaKHis dataset. However, these CNN models use a large number of convolutional layers and parameters to achieve accurate results, which increases computing resources and time.

Similar studies have been conducted using other public datasets. Teresa *et al.* [21] presented a Bioimaging2015 dataset and conducted two approaches based on a CNN and a CNN + SVM model for binary and multiclass classifications. High-resolution images were divided into 512 × 512 pixels patches to reduce the computational resources. The results showed an accuracy of 83.3% for binary classification and 77.8% for four-class classification. Moreover, Aresta *et al.* [22] introduced the BACH 2018 dataset, which contains more images than the Bioimaging2015 dataset. The highest classification accuracy of 87% was achieved using the DenseNet121 model. Similarly, Nawaz *et al.* [23] proposed an AlexNet-based approach to breast cancer classification. The experimental results showed that the approach obtained a patch-wise accuracy of 75.73% and an image-wise accuracy of 81.25% for multiclass classification of the BACH 2018 dataset. In addition, Yao *et al.* [24] developed a parallel structure that combined a CNN and a recurrent neural network (RNN). Different approaches have been evaluated using the BACH 2018 dataset. The highest overall accuracy was 92% for multiclass classification. In general, the classification accuracy and computational efficiency still require improvement for both the Bioimaging2015 and BACH 2018 datasets.

III. MATERIALS

The proposed approach was evaluated using two public datasets: the BreaKHis and BACH 2018 datasets. The datasets used in this study are described in detail below.

A. BREAKHIS DATASET

The BreaKHis dataset was developed by Spanhol *et al* [25]. It consists of 7909 breast cancer histological images from 82 patients, who were enrolled in a clinical study in 2014. Images were generated from breast tissue biopsy slides stained with hematoxylin and eosin, using an Olympus BX-50 system microscope and a Samsung Digital Color Camera SCC-131AN. They are captured in 3-channel RGB color with a 700 × 460 pixels resolution at 40×, 100×, 200×, and 400× magnification. Each sample was diagnosed by an experienced pathologist and confirmed by complementary examinations. Images were classified into benign and malignant categories. Each category has four subclasses: adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenoma (TA) for benign lesions; and ductal carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC), and papillary carcinoma (PC) for malignant cancer. Table 1 summarizes the statistics for the BreaKHis dataset. Fig. 1 shows images of some samples at 200× magnification.

TABLE 1. Statistics for the BreaKHis dataset.

Classes	Sub-class	Patients Number	Magnification Factor 40×	100×	200×	400×	Total
Benign	A	4	114	113	111	106	444
	F	10	253	260	264	237	1014
	TA	3	109	121	108	115	453
	PT	7	149	150	140	130	569
Malignant	DC	38	864	903	896	788	3451
	LC	5	156	170	163	137	626
	MC	9	205	222	196	169	792
	PC	6	145	142	135	138	560
Total	-	82	1995	2081	2013	1820	7909

B. BACH 2018 DATASET

The BACH 2018 dataset was constructed by Teresa *et al.* [22]. It comprises breast cancer histological images available from the Breast Cancer 2018 Grand Challenge. Images were obtained from breast tissue biopsy slides and stained with hematoxylin and eosin. These images were acquired under the same conditions using a Leica DM 20000 LED microscope and a Leica ICC50 high-definition camera. Each image had a size of 2048 × 1536 pixels and was captured at 200× magnification and a pixel scale of 0.42 μm × 0.42 μm. The entire dataset contained 400 images collected from different patients in Covilh and Porto. These images were labeled by two medical experts and classified into four categories of the same size: normal tissue, benign lesion, *in situ* carcinoma, and invasive carcinoma. Fig. 2 shows sample images of these four categories.

IV. METHODOLOGY

Fig. 3 shows the schematic of the proposed approach. This approach consists of three parts.

1) *Pre-processing:* First, the input images were normalized to solve the inter-image variability caused by the staining and image acquisition process. Second, the original images were subdivided into small patches using a patch extraction approach to reduce the complexity of the model. Third, a data augmentation method was used to enlarge the dataset for model training.
2) *Patch screening:* An ADSVM algorithm was used to screen the mislabeled patches in the malignant images. SURFs were extracted from the patches and then the features were encoded using the LLC method to conduct an SVM classifier.
3) *RANet classification:* Patches were used to train a RANet model for cancer classification. RANet selects a resolution adaptive subnetwork for the input image based on the recognition difficulty.

A. IMAGE NORMALIZATION

Histological images were acquired under different conditions (including different staining protocols, staining materials,

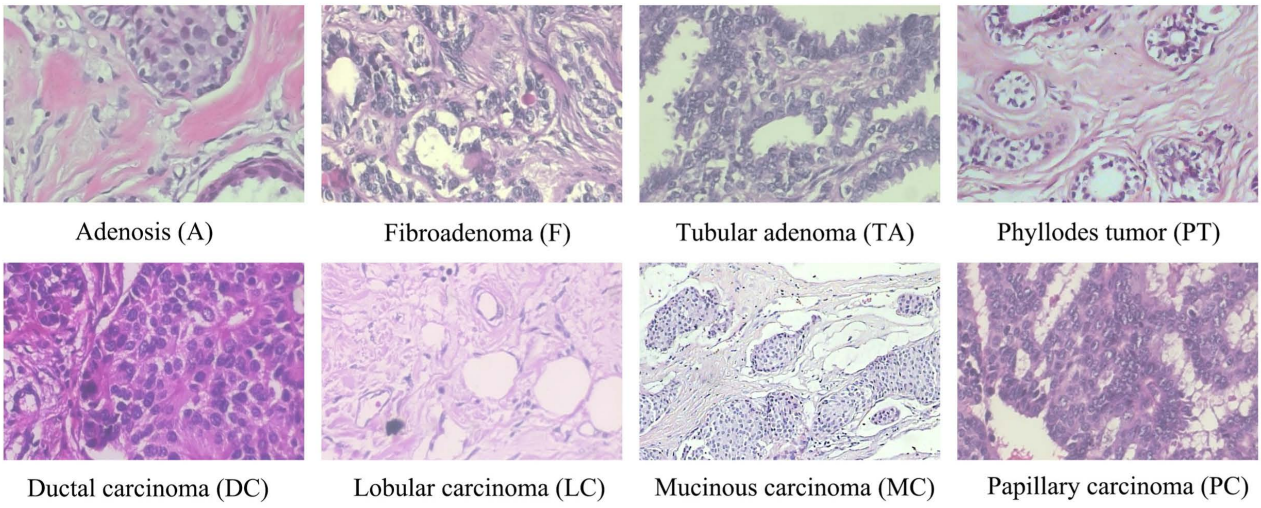

FIGURE 1. Images of some samples at 200× magnification in the BreaKHis dataset. The top row shows the four types of benign tumors and the bottom row shows the malignant tumors.

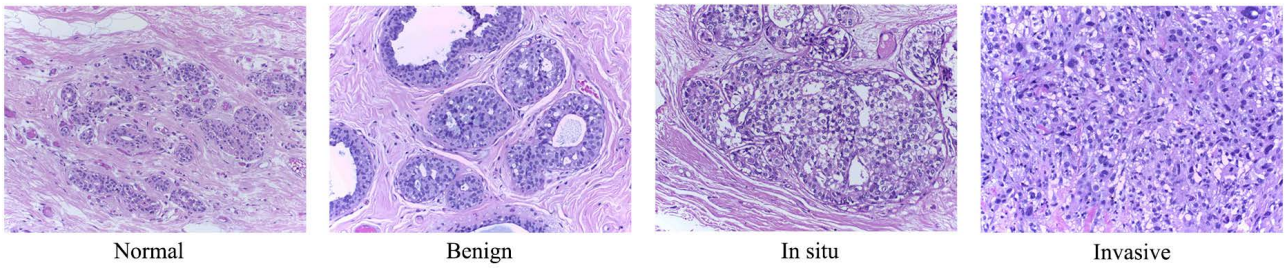

FIGURE 2. Samples images of four categories in the BACH. 2018 dataset.

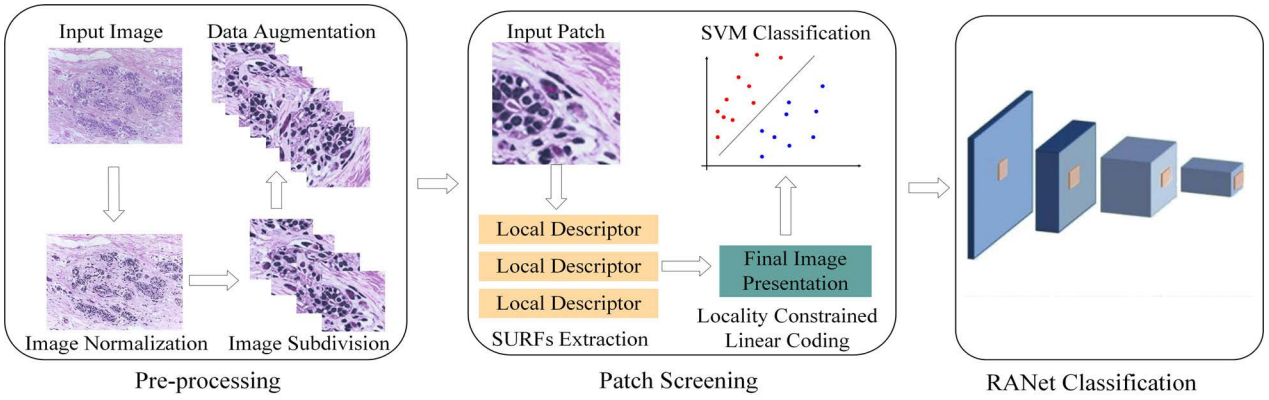

FIGURE 3. The schematic illustration of the proposed classification approach for breast cancer histopathological images.

and acquisition times); thus they, presented color differences. These differences may cause overfitting in the subsequent model training process. Image normalization potentially eliminates color variability, thereby improving the robustness and accuracy [26]–[28].

In the present study, the images were normalized using a novel approach. The normalization algorithm is composed of two procedures based on Retinex theory [29] and the histogram equalization technique [30]. The Retinex method divides the illuminance and reflection components using a logarithmic algorithm. These components were then filtered using a function that simulates human vision. Thus, the low-frequency part of the illuminance component decreased, whereas the high-frequency part of the reflection component

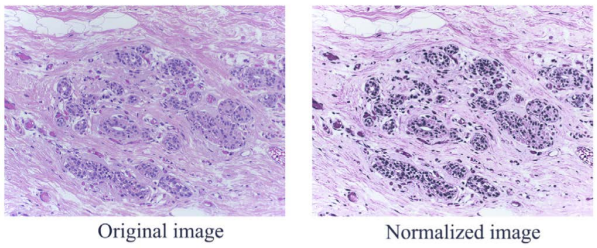

FIGURE 4. The performance of image normalization. Left is the original image and right is the normalized image.

was enhanced. In addition, a histogram equalization method was used for image normalization. A level adaptive histogram equalization (LAHE) method was applied to each color component. Histogram equalization was performed on the patches subdivided from the original image to avoid over-enhancement. Histograms were clipped to adjust the color component. Fig. 4 shows the image normalization results.

B. IMAGE SUBDIVISION

Extensive computational resources and a long time are required when a CNN model is trained on high-resolution images. The high-resolution images were subdivided into patches for subsequent model training to improve the computing efficiency. Image level classification results were obtained by combining the corresponding patch level results.

The input images should be subdivided into patches with appropriate sizes that contain sufficient tissue information for the diagnosis to achieve great classification performance. Small patches with a size of 32×32 pixels or 64×64 pixels do not provide sufficient tissue information for classification [17], [21]. Meanwhile, considering the structure of the RANet model, the size of subdivided patches should be set to 112×112, 224×224, or 448×448 pixels, suggesting that the base feature maps with four different scales (56×56, 28×28, 14×14, and 7×7 pixels) can be easily generated from the input patches using a convolution layer (7×7 conv) and a max pooling layer (3×3 max pool) with different strides of 1 or 2. The classification performance on different patch sizes was evaluated using the BreaKHis and BACH 2018 datasets. Experimental results show that the patch with a size of 224×224 pixels achieved the overall best performance for image classification in the BreaKHis and the BACH 2018 datasets. Therefore, the images were subdivided into patches with a size of 224×224 pixels in our study.

Additionally, data imbalance affects the performance of the CNN model, and the classification results are biased toward the major category. An adaptive patch number approach is proposed to reduce the influence of data imbalance. For the k^{th} category in n classes, the number of patches subdivided per image N_k is [19]:

$$N_k = \left\lfloor \frac{\sum_{i=1}^{n} x_i}{n \times x_k} \times \alpha \right\rfloor \tag{1}$$

where x_k refers the image number of the k^{th} category and α is a control parameter.

After the size and the number of the patches were determined, a sliding window process was used to select the N_k patches from each input image. Because the breast cancer diagnosis mainly focuses on nuclear features, the patches are chosen based on the nuclear density, and thus the first N_k patches with high entropy were selected.

C. DATA AUGMENTATION

The data augmentation approach may extend the image dataset from the original dataset to avoid overfitting [31], [32]. Regarding the analysis process for histopathological images, the data augmentation methods based on geometric transform were used to increase the quantity of data. The geometric transform includes rotation, flipping, and shifting which maintain the histopathological characteristic of the tissues without altering the classification results [33], [34]. The classification performance of these methods (including rotation with different degrees, flipping, and shifting) was evaluated using the BreaKHis and the BACH 2018 datasets to select appropriate transform methods for data augmentation. The experimental results show that the rotation + shifting + flipping method achieved the best performance.

Therefore, each patch was randomly rotated from 0 to 90 degrees to generated four new patches initially, and then shifted horizontally and vertically by a random pixel within 56 pixels (the size of the stride for the sliding windows process in the image subdivision method). Next, the rotated and shifted patches were flipped horizontally and vertically. Therefore, 36 patches were generated from each input patch.

D. PATCH SCREENING

Because the histology images in most datasets are classified at the image-level, the patches obtained from the image subdivision and data augmentation methods are marked with the same label as the whole image. However, a malignant image may contain benign and normal tissues, and thus, some patches are mislabeled. This mislabeling reduces the recognition performance of the follow-up model. A mislabeled patch screening method based on anomaly detection using a support vector machine (ADSVM) is proposed to solve this problem.

Fig. 5 shows a schematic of the ADSVM method. It consists of three parts: speeded-up robust feature (SURF) extraction, locality constrained linear coding (LLC), and SVM classification. In the proposed method, the SURFs are extracted from the patches and encoded using the locality constrained linear coding (LLC) method. Then, the encoded features are used to train a soft-margin SVM classifier for

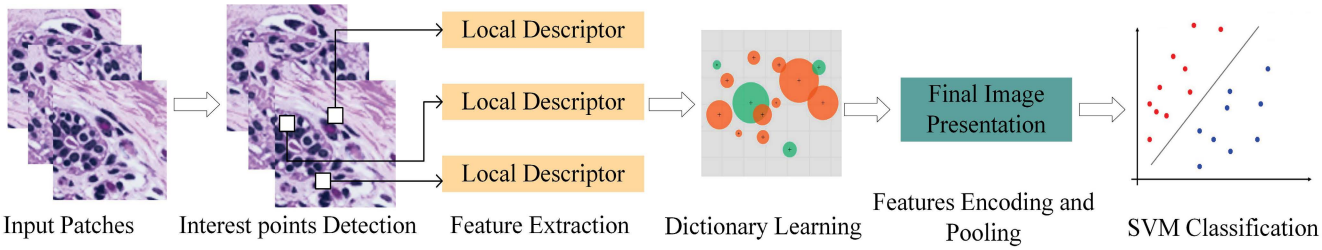

FIGURE 5. The schematic of patch screening based on the ADSVM method.

binary classification (benign vs. malignant). Finally, the trained SVM classifier is used to recognized the benign patches, which are mislabeled as malignant, subdivided from malignant images. These mislabeled patches are replaced by new patches extracted from the corresponding image with malignant results. As a result, the ADSVM method reduces the number of mislabeled patches and increases the follow-up training performance.

1) SPEED-UP ROBUST FEATURE EXTRACTION

SURF [35] is a fast and robust local feature extraction method based on the scale invariant feature transform (SIFT) method. The method comprises three main parts: scale-space construction, local feature point detection, and feature point description. The specific implementation process is described as follows:

1) Constructing a scale-space representation of the image,
2) Creating a Hessian matrix and generating points of interest for feature extraction,
3) Detecting the position and orientation of points of interest, and
4) Constructing the descriptor for points of interest.

2) LOCALITY CONSTRAINED LINEAR CODING

The local feature descriptors applied to SVM classifiers must have a fixed length, whereas the number of SURFs extracted from the image varies. Thus, feature coding is an essential procedure that must be conducted between SUEF extraction and SVM classification. The LLC method was used to map the features into an image presentation.

LLC is a simple and effective feature coding method [36]. It projects the local feature descriptors into a local-coordinate space. Then the image presentation is obtained by maximum pooling of the projection results.

Let $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_N] \in \mathbb{R}^{D\times N}$ denote the local feature descriptors. The LLC method utilizes the following criteria to obtain the codes $\mathbf{C} = [\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N] \in \mathbb{R}^{M\times N}$ [36]:

$$\begin{aligned} &\min_{\mathbf{C}} \sum_{i=1}^{N} \|\mathbf{x}_i - \mathbf{B}\mathbf{c}_i\|^2 + \lambda \|\mathbf{d}_i \otimes \mathbf{c}_i\|^2 \\ &s.t. \quad \mathbf{1}^T\mathbf{c}_i = 1, \quad \forall i \end{aligned} \tag{2}$$

where $\mathbf{B} = [\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_M] \in \mathbb{R}^{D\times M}$ is the codebook presenting the local-coordinate space, $\otimes$ denotes the element-wise multiplication, and $\mathbf{d}_i \in \mathbb{R}^M$ is the locality adaptor [36]:

$$\mathbf{d}_i = \exp\left(\frac{\text{dist}(\mathbf{x}_i, \mathbf{B})}{\sigma}\right) \tag{3}$$

where $\text{dist}(\mathbf{x}_i, \mathbf{B}) = [\text{dist}(\mathbf{x}_i, \mathbf{b}_1), \ldots, \text{dist}(\mathbf{x}_i, \mathbf{b}_M)]^T$ denotes the Euclidean distance between the $\mathbf{x}_i$ and $\mathbf{B}$; σ is a parameter that controls the distance weight for $\mathbf{d}_i$.

A simplified approximate calculation for the LLC method was conducted by ignoring the second part of Equation (2). The K-means method was also employed to accelerate the process. For the descriptor $\mathbf{x}_i$, only the K-nearest neighbors were used to calculate the code coefficients, whereas the other coefficients were set to 0. Then, the codes were maximally pooled to generate the final image presentation.

3) SUPPORT VECTOR MACHINE CLASSIFICATION

Using the final image presentation as the input, a linear SVM classifier for binary classification was trained. The trained SVM classifier then recognized and screened the non-malignant patches from the malignant images. However, the screening process may lead to a data imbalance. The screened patch was replaced with a new patch extracted from the corresponding image with a malignant result to solve this problem.

E. RESOLUTION ADAPTIVE NETWORK

Recently, researchers have mainly exploited the CNN model architecture in network depth or width to achieve outperformance, which increases the model complexity and computational resources. A RANet model based on a multi-scale dense connection (MSDNet) [37] was proposed for classification to increase the computational efficiency without sacrificing accuracy.

In the RANet model, input images are classified using different subnetworks according to classification difficulty. The "easy" images are processed by sub-networks that use low-resolution feature maps and consist of few convolutional layers. These subnetworks produce classification results in much less time than that of the deeper networks used in classic models (e.g., ResNet [14] and DenseNet [15]). The "hard" images are predicted by deeper sub-networks using higher resolution feature maps. Compared with the networks in classic models, these subnetworks use more convolutional

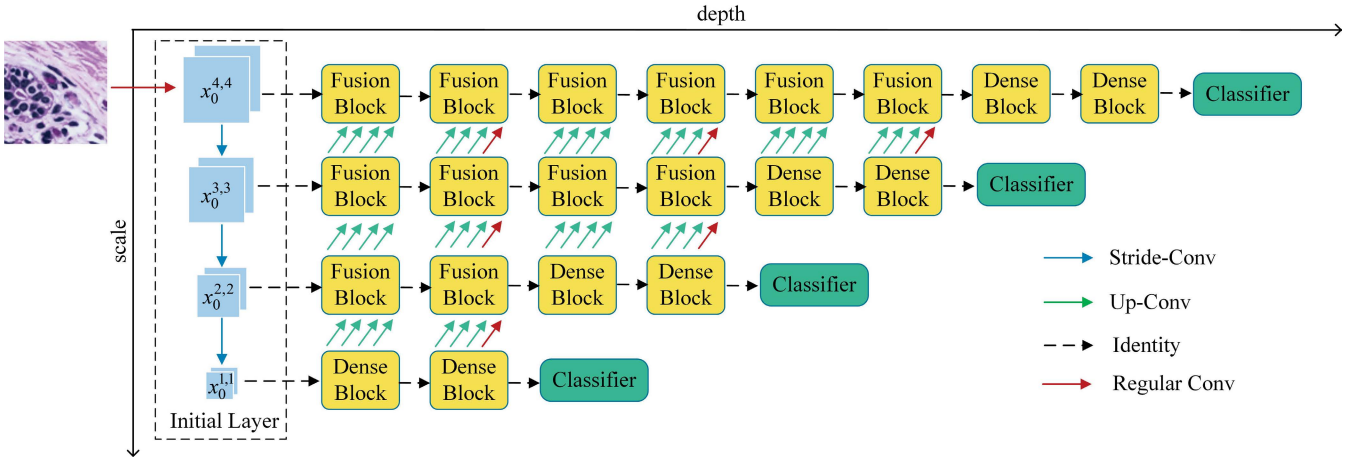

FIGURE 6. The structure of the RANet model.

layers to increase the prediction accuracy. Although the operation increases the computational time of "hard" images, the computational time for the whole model is reduced because most images are "easy" images. Consequently, the RANet model increases the computational efficiency based on the premise of prediction accuracy.

The global structure of the RANet model is shown in Fig. 6. In the illustration, the model is composed of an *initial layer* and subnetworks in S scales. $\mathbf{x}_0^{s,h}$ denotes the feature maps obtained from the input image at scale s for subnetwork h. The prediction process is first implemented via subnetwork 1 with the feature maps $\mathbf{x}_0^{1,h}$ as input. If subnetwork 1 fails to obtain a convincing result, subnetwork 2 that works on higher-resolution feature maps is used for classification. This process is repeated until a convincing classification result is achieved or the last subnetwork is utilized.

The RANet architecture includes the important components described below.

Initial layer: In the *initial layer*, the base feature maps are obtained from the input image. The highest-resolution feature maps are generated from a *Regular-Conv* layer [15], composed of a *Bottleneck* layer [15] and a regular convolution layer. These two layers consist of a *Batch Normalization* (BN) layer [38], a *ReLU* layer [39], and a convolution layer. Small-scale feature maps are generated from the larger-scale feature maps by a *Strided-Conv* layer, which is similar to a *Regular-Conv* layer. Here, the stride of the second convolution layer was 2.

Subnetwork 1: Subnetwork 1 works on the coarsest-scale feature map $\mathbf{x}_0^{1,h}$. As shown in Fig. 7 (a), subnetwork 1 consists of *Dense Blocks* [15] with l layers. In each *Dense Block,* the output feature maps $\mathbf{x}_i^{1,h}$ from the i-th layer are reused in subnetwork 2. Subnetwork 1 is perceived as DenseNet [15] utilizing coarsest-scale feature maps.

Subnetwork s: Subnetwork s ($s > 1$) utilizes the base feature maps $\mathbf{x}_0^{s,h}$ and lower-resolution feature maps from subnetwork $(s-1)$ for classification. As shown in Fig. 7 (b, c), *Dense Blocks* with feature fusion are called *Fusion Blocks.*

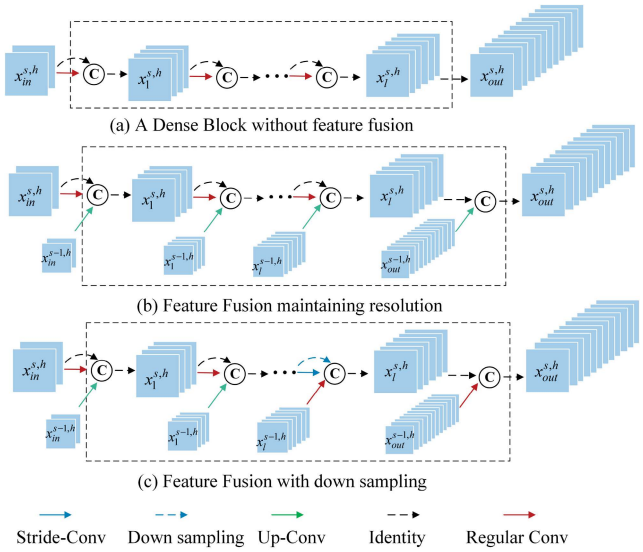

FIGURE 7. The dense block (a) and fusion blocks (b, c) in the RANet model.

The *Fusion Block* processes feature maps from two sources.

Two different feature fusion methods are used. The method shown in Fig. 7 (b) maintains the same output feature map resolution as the input features, while the other method shown in Fig. 7 (c) reduces the output resolution using a *Strided-Conv* layer. As shown in Fig. 7 (b), the new feature maps $\mathbf{x}_1^{s,h}$ are first obtained from the input feature maps $\mathbf{x}_{in}^{s,h}$ using a *Regular-Conv* layer. Then, the input feature maps $\mathbf{x}_{in}^{s-1,h}$ from the lower sub-network are handled by an *Up-Conv* layer, consisting of a *Regular-Conv* layer and an up-sampling bilinear interpolation. Feature maps with the same resolution were concatenated using dense connections.

Another *Fusion Block* in Fig. 7 (c) reduces the feature resolution using a *Strided-Conv* layer at the end of the block, and the feature maps generated from the previous layers in

the block are concatenated after a down-sampling operation. Meanwhile, the feature maps from the lower subnetwork are processed using a *Regular-Conv* layer. Finally, the two categories of feature maps are concatenated by a dense connection.

Transition layer: Transition layers are utilized to reduce the number of feature maps between *Dense Blocks,* thus reducing the computational cost. A *transition layer* consists of a *BN* layer, a *ReLU* layer, and a 1×1 convolution layer. In Fig. 6, the *transition layers* are omitted to simplify the flowchart.

Classifiers and loss function: Four classifiers are placed at the end block of each subnetwork shown in Fig. 6. The output of the k-th classifier is defined as follows [16]:

$$\mathbf{p}^k = [p_1^k, \ldots, p_C^k]^T \in \mathbb{R}^C \tag{4}$$

where $p_c^k \in [0, 1]$ is the prediction probability for the c-th category. Prediction results were obtained only if the prediction probability p_c^k was larger than ε. Here ε denotes a threshold that adjusts the tradeoff between the accuracy and efficiency. Additionally, *softmax* classifiers were used, and the final results were predictions from the first classifier with the prediction confidence reaching the threshold. The classifiers used the following criteria [16]:

$$k^* = \min \left\{ k | \max_c p_c^k \geq \varepsilon \right\} \tag{5}$$

$$\widehat{y} \in arg \max_c p_c^{k^*} \tag{6}$$

Each classifier used a cross-entropy loss function. The overall loss function is the weighted cumulative loss of all classifiers.

V. EXPERIMENTS AND RESULTS

A. EXPERIMENTAL SETUP

As shown in Fig. 3, the image classification process is composed of a preprocessing phase, patch screening phase, and RANet classification phase.

In the preprocessing phase, the input images were first normalized using the Retinex and LAHE methods. Then, the normalized images were subdivided into patches using the sliding window approach. For the BreaKHis dataset, each image with a resolution of 700×460 pixels was divided into 45 patches of 224×224 pixels. Here, the sliding stride was set to 56 and the parameter α was set to 32. Then, the number of patches subdivided per image N_k was calculated. The first N_k patches with high entropy were selected for subsequent processing. The same operations were performed on the BACH 2018 dataset. Images with 2040×1536 pixels were subdivided into 792 patches, and the parameter α was set to 640.

The patch screening phase consisted of SURF extraction, LLC, and SVM classification. The sample was first processed using SURF extraction to obtain 64-dimensional feature vectors. For LLC, the K-means method was used to cluster the 20 nearest neighbors ($K = 20$), and then a max-pooling operation was conducted to pool the sparse codes. A codebook with 1500 visual words was also generated using the K-means method. The SVM classifier was applied using the Libsvm library [40].

The RANet architecture is composed of an initial layer and subnetworks in 4 scales. In the *initial layer*, feature maps with 56×56 pixels were generated from input patches of 224×224 pixels utilizing a convolution layer (7×7 conv, stride $= 2$) and a max pooling layer (3×3 max pool, stride $= 2$) firstly. Then base feature maps with four different scales (56×56, 28×28, 14×14, and 7×7) for corresponding subnetworks were generated using a *Regular-Conv* layer or a *Strided-Conv* layer. The numbers of base feature channels for the four scales were 32, 64, 64, and 128, respectively.

The four subnetworks consisted of 2, 4, 6, and 8 blocks with growth rates of 64, 32, 32, and 16, respectively. The layer number in each block was 8. The stride of the last convolution layer (3×3 conv) was 2 in the even Fusion Block. This is to reduce the feature resolution. The stride of the other convolution layers in subnetworks was 1. A compression factor of 0.25 was applied to each Fusion Block. Thus, 75% of the feature maps were generated from the current subnetwork, and the others were generated from the previous low-resolution subnetwork. The transition layers were utilized to reduce the number of feature maps between blocks by 50%. In addition, the classification layer consisted of a global average pooling layer (7×7 conv, stride $= 1$), a fully connected layer and a softmax layer. In addition, the threshold ε was set to 0.8, indicating that the classification results were obtained only when the prediction confidence p_c^k was larger than 0.8.

Moreover, the RANet model was trained using the stochastic gradient descent (SGD) optimizer with a batch size of 64. The RANet model was pretrained on the ImageNet dataset [41]. The model was trained for 100 epochs, with an initial learning rate of 0.01 and the learning rate was decreased by a factor of 10 after 50 epochs. The momentum and weight decay were set to 0.9 and 0.0001, respectively.

B. RESULTS FOR THE BREAKHIS DATASET

As shown in Table 1, images from the BreaKHis dataset were classified into eight categories within two classes (benign and malignant). Additionally, the images were classified at the patient level. Thus, we evaluated the performance of the breast cancer classification approach for multiclass and binary classes at the image and patient levels, respectively. Meanwhile, we conducted the analysis with different classification systems based on the RANet model with and without these modules to evaluate the effects of the new algorithm modules (data balance and ADSVM) described in the previous sections.

Fig. 8 shows the training and validation accuracy curves for the eight-class classification at different magnification levels. Accordingly, the training and validation processes were affected by magnification factors. The best training and validation accuracies were achieved at $200\times$ magnification.

Tables 2 and 3 show the testing accuracy for multiclass classification and binary classification, respectively.

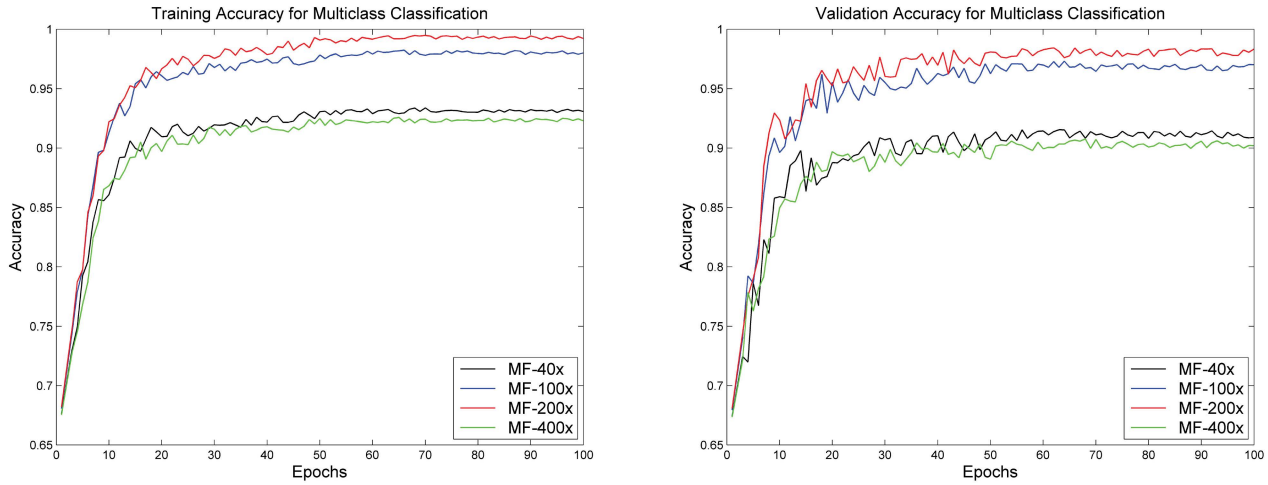

FIGURE 8. The training and validation accuracy curves for multiclass classification with different magnification factors in the BreaKHis dataset.

TABLE 2. Breast cancer classification accuracy for multiclass (8 classes) in the BreaKHis dataset (with and without data balance, with and without ADSVM). The best results are in bold.

Accuracy (100%)	Method	Magnification factor			
		40×	100×	200×	400×
Patient level	RANet	70.12 ± 3.3	72.70 ± 2.6	73.85 ± 2.4	69.62 ± 3.6
	RANet + Data Balance	84.74 ± 2.2	91.26 ± 1.3	91.74 ± 1.2	86.13 ± 1.9
	RANet + Data Balance + ADSVM	**91.21 ± 1.0**	**96.90 ± 0.6**	**97.43 ± 0.4**	**88.43 ± 1.3**
Image level	RANet	71.57 ± 2.5	72.34 ± 2.8	73.50 ± 2.7	69.45 ± 3.2
	RANet + Data Balance	83.86 ± 2.1	91.84 ± 1.2	91.73 ± 1.3	86.05 ± 1.8
	RANet + Data Balance + ADSVM	**91.14 ± 1.1**	**96.83 ± 0.7**	**98.05 ± 0.3**	**90.30 ± 1.2**

TABLE 3. Breast cancer classification accuracy for binary classes (benign vs. malignant) in the BreaKHis dataset (with and without data balance, with and without ADSVM). The best results are in bold.

Accuracy (100%)	Method	Magnification factor			
		40×	100×	200×	400×
Patient level	RANet	80.03 ± 1.8	84.34 ± 1.6	85.27 ± 1.5	79.49 ± 2.2
	RANet + Data Balance	88.93 ± 1.7	95.14 ± 0.8	96.08 ± 0.7	91.10 ± 1.1
	RANet + Data Balance + ADSVM	**94.16 ± 0.9**	**97.91 ± 0.4**	**98.83 ± 0.3**	**92.64 ± 0.9**
Image level	RANet	79.84 ± 2.0	84.17 ± 1.5	84.23 ± 1.6	80.04 ± 1.9
	RANet + Data Balance	89.65 ± 1.2	94.59 ± 0.8	96.38 ± 0.7	91.47 ± 1.2
	RANet + Data Balance + ADSVM	**94.43 ± 0.8**	**98.31 ± 0.3**	**99.14 ± 0.2**	**93.35 ± 0.9**

In addition, other metrics, including precision, recall, and $F1_{score}$ for binary classification, are also shown in Table 4. As shown in Tables 2 and 3, the data balance method has affected the system performance. In the image-level evaluation, the average classification accuracy for the multiclass analysis with the data balance approach was approximately 17% better than that of the model without it. The improvement in the binary classification was 11%. Meanwhile, the accuracies at the patient level were improved by approximately 16% and 10% for multiclass and binary classification, respectively. In addition, other metrics showed that the data balance method improved the performance of the classification system. A potential explanation for this finding is that the data balance method removes the uneven distribution problem in the BreaKHis dataset and increases the quality of the training data.

Additionally, according to Tables 2 and 3, the ADSVM method improved the accuracy and other metrics for breast cancer classification. For multiclass classification, the accuracies increased by approximately 6% at the patient and image levels with magnification factors of 40×, 100×, and 200×. The improvement was approximately 3% at 400×

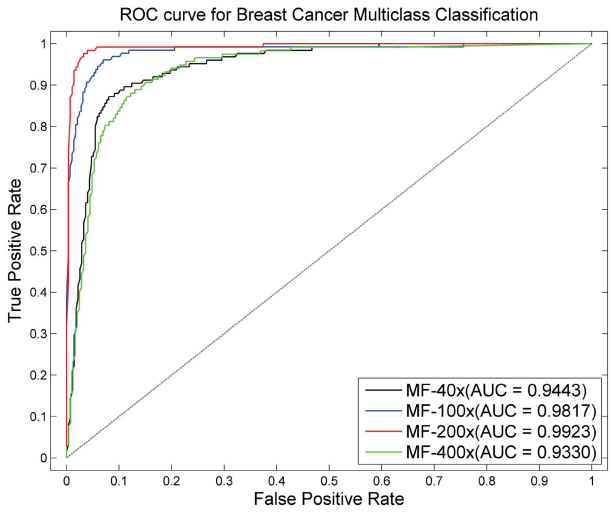

FIGURE 9. The ROC curves with AUC for multiclass classification with different magnification factors in the BreaKHis dataset.

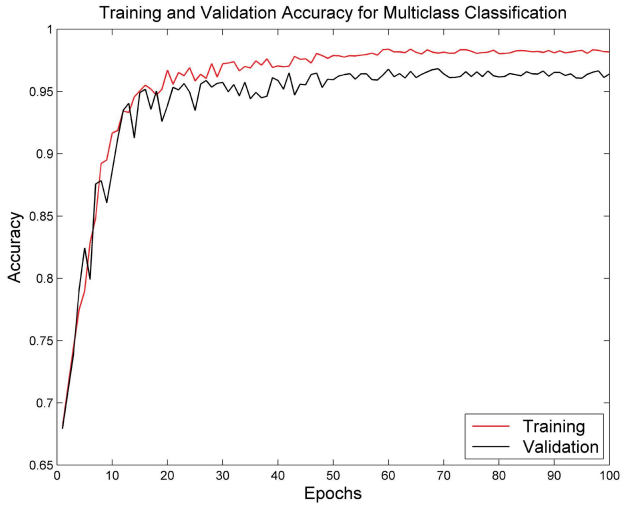

FIGURE 10. The training and validation accuracy curves for multiclass classification in the BACH 2018 dataset.

magnification. Meanwhile, the accuracy of the binary classification was improved by approximately 1.7% and 4% at 400× and other magnification levels, respectively. The images captured at 400× magnification contained fewer benign tissues than the images captured at other magnifications; thus, the number of mislabeled patches screened by the ADSVM approach was less than that screened at other resolutions.

In general, for the image-level evaluation, the best accuracies of 98.05% and 99.14% for multiclass and binary classification were obtained at 200× magnification, respectively. The best precision of 98.62%, the best recall of 99.05%, and the best $F1_{score}$ of 98.83% were achieved at the same magnification. Meanwhile, the approach achieved the best accuracies of 97.43% and 98.83% at 200× magnification in the patient-level evaluation for multiclass and binary classifications, respectively. In addition, the worst classification performance was obtained at a magnification of 400×, because the subdivided patches with 224×224 pixels from the images at this resolution do not contain sufficient information for CNN model training.

Additionally, the performance of the RANet + ADSVM model was analyzed by constructing receiver operating characteristic (ROC) curves. Fig. 9 shows the ROC curves for multiclass classification of images captured at different magnifications. The area under curve (AUC) is also shown in Fig. 9. Accordingly, the method achieved the best AUC of 0.9923, meaning that the best classification performance was obtained at 200× magnification.

C. RESULTS FOR THE BACH 2018 DATASET

For the BACH 2018 dataset, the classification performance was evaluated for multiclass and binary classes based on two criteria: the image level and patch level [22], [34]. Additionally, different metrics were used to estimate the effect of the ADSVM module on the classification system. During

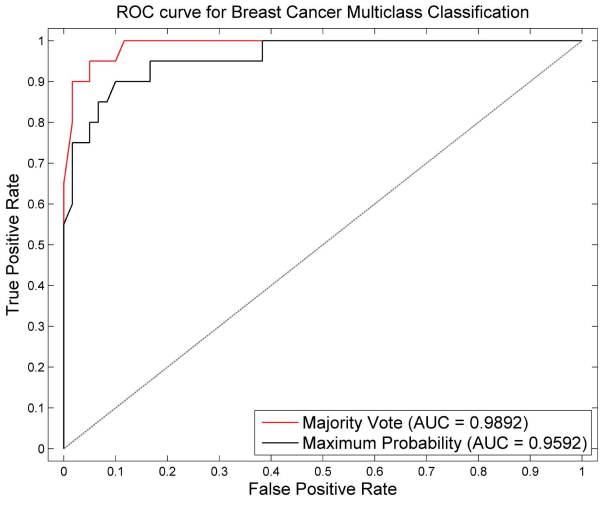

FIGURE 11. The ROC curves with AUC for multiclass classification in the BACH 2018 dataset.

processing, each image was subdivided into 70 patches of 224×224 pixels for prediction. The classification result for the original whole image was generated by combining the results obtained from patches. Additionally, fusion methods affect the classification results for the entire image. The performances of the majority vote and maximum probability fusion methods were evaluated. For the majority vote fusion method, the image classification result was determined by the label with the highest number within the subdivided patches. In the maximum probability fusion method, the image label was determined by the maximum probability among the subdivided patches.

The training and validation curves of the accuracy for the four-class classification are shown in Fig. 10. Table 5 shows the testing performance for multiclass and binary classifi-

TABLE 4. Breast cancer classification metrics (precision, recall, and F1score) for binary classes (benign vs. malignant) in the BreaKHis dataset (with and without data balance, with and without ADSVM). The best results are in bold.

Metric (100%)	Method	Magnification factor			
		40×	100×	200×	400×
Precision	RANet	76.42 ± 2.3	82.26 ± 1.6	82.45 ± 1.6	77.53 ± 2.2
	RANet + Data Balance	88.43 ± 1.3	92.75 ± 0.9	95.61 ± 0.6	89.81 ± 1.2
	RANet + Data Balance + ADSVM	**93.03 ± 0.8**	**98.12 ± 0.4**	**98.62 ± 0.3**	**92.27 ± 0.9**
Recall	RANet	81.13 ± 1.8	85.04 ± 1.6	86.22 ± 1.5	80.87 ± 1.8
	RANet + Data Balance	90.76 ± 1.1	95.41 ± 0.7	96.52 ± 0.6	91.85 ± 1.0
	RANet + Data Balance + ADSVM	**94.21 ± 0.7**	**97.80 ± 0.5**	**99.05 ± 0.1**	**92.76 ± 0.8**
$F1_{score}$	RANet	79.97 ± 1.9	83.38 ± 1.6	83.62 ± 1.5	78.06 ± 2.0
	RANet + Data Balance	89.56 ± 1.3	94.35 ± 0.9	96.07 ± 0.7	91.36 ± 1.1
	RANet + Data Balance + ADSVM	**93.56 ± 0.8**	**98.02 ± 0.3**	**98.83 ± 0.3**	**92.45 ± 0.9**

TABLE 5. Metrics (accuracy, precision, recall, and F1score) for multiclass (4 classes) and binary classification at patch level in the BACH 2018 dataset (with and without ADSVM). The best results are in bold.

Classification task	Method	Accuracy (100%)	Precision (100%)	Recall (100%)	$F1_{score}$ (100%)
Multiclass	RANet	91.80 ± 0.5	91.91 ± 0.5	91.80 ± 0.5	91.79 ± 0.5
	RANet + ADSVM	**96.37 ± 0.3**	**96.38 ± 0.3**	**96.37 ± 0.3**	**96.36 ± 0.3**
Binary	RANet	95.71 ± 0.3	94.68 ± 0.3	96.92 ± 0.2	95.78 ± 0.3
	RANet + ADSVM	**97.43 ± 0.2**	**96.32 ± 0.3**	**98.64 ± 0.1**	**97.46 ± 0.2**

TABLE 6. Metrics (accuracy, precision, recall, and F1score) for multiclass (4 classes) and binary classification at image level in the BACH 2018 dataset (majority vote and maximum probability). The best results are in bold.

Classification task	Voting method	Accuracy (100%)	Precision (100%)	Recall (100%)	$F1_{score}$ (100%)
Multiclass	Majority Vote	**97.75 ± 0.2**	**97.93 ± 0.2**	**97.75 ± 0.2**	**97.76 ± 0.2**
	Maximum probability	94.50 ± 0.3	95.23 ± 0.3	94.50 ± 0.3	94.56 ± 0.3
Binary	Majority Vote	**99.25 ± 0.1**	**99.01 ± 0.1**	**99.50 ± 0.1**	**99.25 ± 0.1**
	Maximum probability	96.50 ± 0.2	97.45 ± 0.2	95.50 ± 0.1	96.44 ± 0.2

cation at the patch level. Table 6 shows the performance for multiclass and binary classification at the image level. As shown in Table 5, the best accuracies of 96.37% and 97.43% were achieved for multiclass and binary classification at the patch level, respectively. Similar to the performance of the model in analyzing the BreaKHis dataset, the accuracy was reduced by 4.6% and 1.7% for multiclass and binary classifications without the ADSVM method, respectively. In addition, the best precision of 96.38%, the best recall of 96.37%, and the best $F1_{score}$ of 96.36% were obtained for multiclass classification. For binary classification, the best precision reached 96.32%, the best recall reached 98.64%, and the best $F1_{score}$ reached 97.46%.

As shown in Table 6, for the image-level evaluation, the classification system with the majority vote method achieved better performance than that with the maximum probability method. The highest accuracies of 97.75% and 99.25% were obtained using the majority vote method for multiclass and binary classifications, respectively. This accuracy was 3% better than that of the maximum probability method for both multiclass and binary classifications. In general, the model achieved the highest precision of 97.93% and 99.01% for multiclass and binary classifications, respectively. Meanwhile, the best recall of 97.75% and 99.50%, and the best F1score of 97.76% and 99.25% were achieved for multiclass and binary classifications, respectively. Additionally, Fig. 11 shows the ROC curves for multiclass classification using these voting methods. A higher AUC of 0.9892 was obtained using the majority vote method.

D. COMPARISON WITH OTHER CNN MODELS

Different approaches based on other CNN models were applied to the two datasets to evaluate the classification accuracy and computational efficiency of the proposed approach. The results of the comparison between the pro-

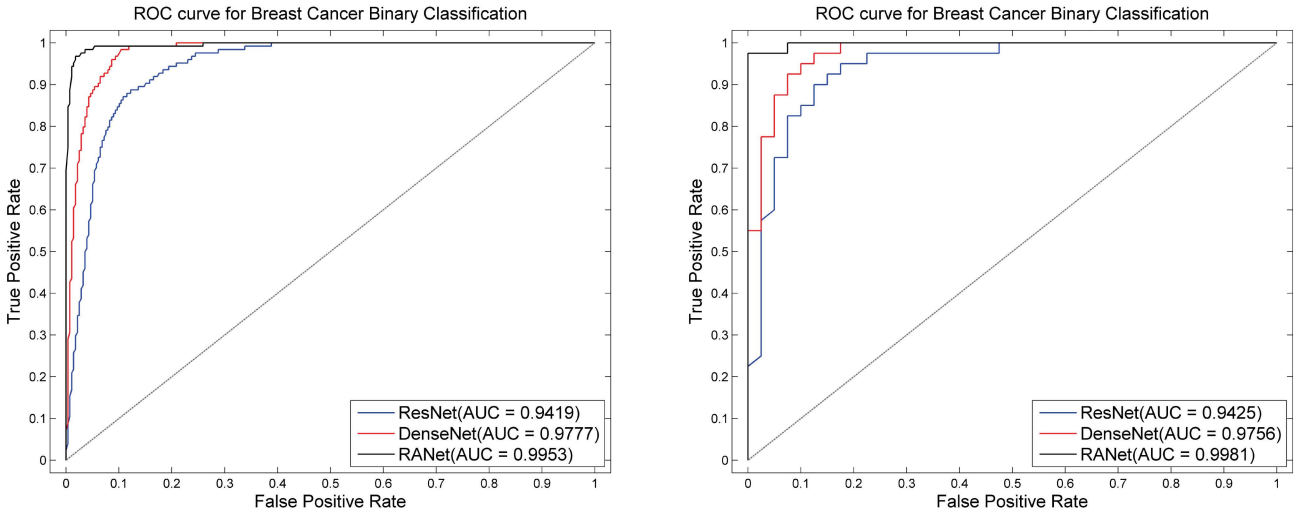

FIGURE 12. The ROC curves of different CNN models with AUC for binary classification. (a): ROC curves in the BreaKHis dataset. (b) ROC curves in the BACH 2018 dataset.

TABLE 7. Classification performance and computational time of different CNN models (ResNet101, DenseNet121, and RANet) in two datasets. The best results are in bold.

Dataset	Model	Accuracy (100%)	Precision (100%)	Recall (100%)	AUC	Gini Coefficient	Time per image(s)	Time per patch(s)
BreaKHis	ResNet101	91.43 ± 0.6	90.91 ± 0.7	91.32 ± 0.6	0.9419	0.6112	0.9216	0.0768
	DenseNet121	96.74 ± 0.3	96.23 ± 0.4	96.55 ± 0.3	0.9777	0.6607	1.2348	0.1029
	RANet	**99.14 ± 0.2**	**98.62 ± 0.3**	**99.05 ± 0.1**	**0.9953**	**0.6851**	**0.5676**	**0.0473**
BACH 2018	ResNet101	91.53 ± 0.5	91.24 ± 0.5	91.78 ± 0.4	0.9425	0.4425	5.4042	0.0772
	DenseNet121	96.38 ± 0.3	96.16 ± 0.3	96.45 ± 0.3	0.9756	0.4756	7.2385	0.1034
	RANet	**99.25 ± 0.2**	**99.01 ± 0.1**	**99.50 ± 0.1**	**0.9981**	**0.4981**	**3.1574**	**0.0451**

posed approach and the approaches based on ResNet101 [14] and DenseNet121 [15] are shown in Table 7. Additionally, Fig 12 shows the comparison of the ROC curves.

For the BreaKHis dataset, these approaches were evaluated for binary classification at the image level using a magnification factor of 200×. Our approach achieved the highest accuracy of 99.14% compared with the accuracies of 96.74% and 91.43% for DenseNet121 and ResNet101, respectively. Meanwhile, our approach achieved the best precision of 98.62%, the best recall of 99.05%, the best AUC of 0.9953 and the best Gini coefficient of 0.6851 compared with the other models. In addition, the computational time per image using our approach was 0.57s, which was less than that of DenseNet121 (1.23 s) and ResNet101 (0.92 s).

For the BACH 2018 dataset, our approach achieved 99.25% accuracy for binary classification at the image level, which was better than the accuracy of 96.38% for DenseNet121 and 91.53% for ResNet101. Meanwhile, a precision of 99.01%, a recall of 99.50%, an AUC of 0.9981 and a Gini coefficient of 0.4981 were achieved using our approach, which were better than those of the other models. In addition, the computational time per image using our approach was 3.16 s, which was less than that of DenseNet121 (7.24 s) and ResNet101 (5.40 s).

In general, compared with the DenseNet121 and ResNet101 models, our RANet + ADSVM approach significantly improves both the classification accuracy and computational efficiency.

E. COMPARISON WITH SIMILAR WORKS

In this section, we compare the performance of the RANet + ADSVM approach with methods proposed in recent studies. Table 8 shows the accuracy for multiclass classification at the image level and patient level with different magnification factors for the BreaKHis dataset. Our approach achieved the best performance for classification at the image and patient levels with magnification factors of 100× and 200×. However, the classification accuracies of our approach were 94.16% and 94.43% at the patient level and image level at 40× magnification, respectively, which were less than the accuracies of 96.32% and 99.13% obtained with the DenseNet121-AnoGAN approach [19]. A potential explanation for this finding is that patches from images captured at 40× magnification represent larger breast tissues than those from images captured at other magnifications. The patches divided from malignant images may contain more benign tissues, and these patches may be misclassified by the low-resolution subnetwork in the RANet model. Moreover, the approach based

TABLE 8. Performance comparison with similar works for binary classification in the BreaKHis dataset. The best results are in bold.

Accuracy (100%)	Model	Year	Magnification factor			
			40×	100×	200×	400×
Patient level	AlexNet [17]	2016	88.6 ± 5.6	84.5 ± 2.4	85.3 ± 3.8	81.7 ± 4.9
	DeCAF [43]	2017	84.0 ± 6.9	83.9 ± 5.9	86.3 ± 3.5	82.1 ± 2.4
	Sequential framework [44]	2018	94.71 ± 0.88	95.9 ± 4.2	96.76 ± 1.09	89.11 ± 0.12
	PFTAS + NPMIL [45]	2019	92.1 ± 5.9	89.1 ± 5.2	87.2 ± 4.3	82.7 ± 3.0
	DenseNet121-AnoGAN [19]	2020	**96.32 ± 1.3**	95.98 ± 0.9	86.91 ± 2.0	85.16 ± 1.3
	RANet+ADSVM		94.16 ± 0.9	**97.91 ± 0.4**	**98.83 ± 0.3**	**92.64 ± 0.9**
Image level	AlexNet [17]	2016	89.6 ± 6.5	85.0 ± 4.8	84.0 ± 3.2	80.8 ± 3.1
	DeCAF [43]	2017	84.6 ± 2.9	84.8 ± 4.2	84.2 ± 1.7	81.6 ± 3.7
	New designed CNN [46]	2018	82 ± 2.8	86.2 ± 4.6	84.6 ± 3	84 ± 4
	PFTAS + NPMIL [45]	2019	87.8 ± 5.6	85.6 ± 4.3	80.8 ± 2.8	82.9 ± 4.1
	DenseNet121-AnoGAN [19]	2020	**99.13 ± 0.2**	96.39 ± 0.7	86.38 ± 1.2	85.20 ± 2.1
	FE-BkCapsNet [20]	2021	92.71 ± 0.16	94.52 ± 0.11	94.03 ± 0.25	**93.54 ± 0.24**
	RANet+ADSVM		94.43 ± 0.8	**98.31 ± 0.3**	**99.14 ± 0.2**	93.35 ± 0.9

TABLE 9. Performance comparison with similar works for multiclass classification in the BACH 2018 dataset. The best results are in bold.

	Model	Year	Accuracy (100%)	Precision (100%)	Specificity (100%)
Image level	AlexNet [23]	2018	81.25	-	-
	DenseNet161 [22]	2019	87	-	95.75
	ResNet + G-loss [42]	2019	91	93	91
	Parallel CNNs [24]	2020	92	90	92
	RANet+ADSVM		**97.75**	**97.93**	**98.21**

on the FE-BkCapsNet model [20] achieved an accuracy of 93.54% for image-level classification at 400× magnification, which was better than the accuracy of 93.35% using our approach. In our approach, images are subdivided into 224 × 224 pixels patches, which may not contain sufficient information for the follow-up classification at 400× magnification. However, in general, our approach achieved the best overall accuracy compared with the other approaches.

The image level classification performance using the BACH 2018 dataset is presented in Table 9. Compared with other similar studies [22]–[24], [42], our approach achieved the highest accuracy of 97.75%, precision of 97.93%, and specificity of 98.21%.

VI. CONCLUSION

In this study, we propose a breast cancer classification approach based on the RANet model and the ADSVM method. The ADSVM method screens mislabeled patches from malignant images to improve classification accuracy. In addition, the RANet model utilizes resolution adaptive subnetworks with variable depths for different images. This approach reduces the computational resources and the time based on the premise of the classification accuracy. Our approach was evaluated using the BreaKHis and BACH 2018 datasets. For the BreaKHis dataset, the best accuracies of 98.83% and 99.14% were achieved for binary classification at 200× magnification at the patient and image levels, respectively. Meanwhile, the approach achieved the best accuracies of 97.75% and 99.25% for multiclass and binary image-level classification of the BACH 2018 dataset, respectively. In addition, the computational efficiency was evaluated for the two datasets. Experimental results show that the proposed approach reduced the computational cost by approximately half compared with the ResNet101 and DenseNet121 models. Compared with similar studies of breast cancer image classification, our approach achieved the best performance in classifying images from the BreaKHis dataset at 100× and 200× magnification. The performance for the BACH 2018 dataset was also better than that described in other studies. In general, the RANet+ADSVM approach shows a significant improvement in performance in terms of classification accuracy and computational efficiency, which may help pathologists produce rapid and accurate diagnoses.

Although the proposed approach achieved outperformance at magnification factors of 100× and 200×, the classification performance required improvement at 400× magnification. Future studies will evaluate the effect of the patch size on high-resolution image classification. Moreover, different-resolution feature maps were used separately in our approach. In future studies, we will develop a new CNN model combining different resolution feature maps extracted from images captured at different magnifications from one patient to generate prediction results. This analysis will allow us to directly conduct the classification approach at the patient level and may increase the performance of patient-level classification. Additionally, new feature extraction method which can replicate the features as derived from the RANet model will be studied to future improve the classification performance and computational efficiency.

ACKNOWLEDGMENT

The authors would like to thank Fabio A. Spanhol for providing the BreaKHis dataset and Araújo Teresa for providing the BACH 2018 dataset.

REFERENCES

[1] J. Ferlay *et al.*, "Cancer statistics for the year 2020: An overview," *Int. J. Cancer*, vol. 149, pp. 778–789, 2021.
[2] R. A. Smith, V. Cokkinides, and H. J. Eyre, "American cancer society guidelines for the early detection of cancer, 2004," *CA, A Cancer J. Clinicians*, vol. 54, no. 1, pp. 41–52, Jan. 2004.
[3] C. W. Elston and I. O. Ellis, "Pathological prognostic factors in breast cancer. I. The value of histological grade in breast cancer: Experience from a large study with long-term follow-up," *Histopathology*, vol. 41, p. 151, Nov. 2002.
[4] J. G. Elmore, "Diagnostic concordance among pathologists interpreting breast biopsy specimens," *Jama*, vol. 313, no. 11, pp. 1122–1132, 2015.
[5] A. Sohail, A. Khan, N. Wahab, A. Zameer, and S. Khan, "A multi-phase deep CNN based mitosis detection framework for breast cancer histopathological images," *Sci. Rep.*, vol. 11, no. 1, p. 6215, Dec. 2021.
[6] Q. Juppet, F. De Martino, E. Marcandalli, M. Weigert, O. Burri, M. Unser, C. Brisken, and D. Sage, "Deep learning enables individual xenograft cell classification in histological images by analysis of contextual features," *J. Mammary Gland Biol. Neoplasia*, vol. 26, no. 2, pp. 101–112, Jun. 2021.
[7] I. Hirra, M. Ahmad, A. Hussain, M. U. Ashraf, I. A. Saeed, S. F. Qadri, A. M. Alghamdi, and A. S. Alfakeeh, "Breast cancer classification from histopathological images using patch-based deep learning modeling," *IEEE Access*, vol. 9, pp. 24273–24287, 2021.
[8] Y. M. George, H. H. Zayed, M. I. Roushdy, and B. M. Elbagoury, "Remote computer-aided breast cancer detection and diagnosis system based on cytological images," *IEEE Syst. J.*, vol. 8, no. 3, pp. 949–964, Sep. 2014.
[9] M. Kowal, P. Filipczuk, A. Obuchowicz, J. Korbicz, and R. Monczak, "Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images," *Comput. Biol. Med.*, vol. 43, no. 10, pp. 1563–1572, 2013.
[10] P. Filipczuk, T. Fevens, A. Krzyzak, and R. Monczak, "Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies," *IEEE Trans. Med. Imag.*, vol. 32, no. 12, pp. 2169–2178, Dec. 2013.
[11] Y. Xue, J. Ye, Q. Zhou, L. R. Long, S. Antani, Z. Xue, C. Cornwell, R. Zaino, K. C. Cheng, and X. Huang, "Selective synthetic augmentation with HistoGAN for improved histopathology image classification," *Med. Image Anal.*, vol. 67, Jan. 2021, Art. no. 101816.
[12] N. Dif, A. M. Oualid, Z. Elberrichi, M. Lebbah, and H. Azzag, "Transfer learning from synthetic labels for histopathological images classification," *Appl. Intell.*, vol. 52, pp. 358–377, Jan. 2021.
[13] K. C. Burak, M. K. Baykan, and H. Uuz, "A new deep convolutional neural network model for classifying breast cancer histopathological images and the hyperparameter optimisation of the proposed model," *J. Supercomputing*, vol. 77, pp. 1–17, Jan. 2021.
[14] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
[15] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4700–4708.
[16] L. Yang, Y. Han, X. Chen, S. Song, J. Dai, and G. Huang, "Resolution adaptive networks for efficient inference," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2020, pp. 2369–2378.
[17] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using convolutional neural networks," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 2560–2567.
[18] D. Bardou, K. Zhang, and S. M. Ahmad, "Classification of breast cancer based on histology images using convolutional neural networks," *IEEE Access*, vol. 6, pp. 24680–24693, 2018.
[19] R. Man, P. Yang, and B. Xu, "Classification of breast cancer histopathological images using discriminative patches screened by generative adversarial networks," *IEEE Access*, vol. 8, pp. 155362–155377, 2020.
[20] P. Wang, J. Wang, Y. Li, P. Li, L. Li, and M. Jiang, "Automatic classification of breast cancer histopathological images based on deep feature fusion and enhanced routing," *Biomed. Signal Process. Control*, vol. 65, Mar. 2021, Art. no. 102341.
[21] T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, A. Polónia, and A. Campilho, "Classification of breast cancer histology images using convolutional neural networks," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0177544.
[22] G. Aresta, T. Araújo, S. Kwok, S. S. Chennamsetty, M. Safwan, V. Alex, and B. Marami, "BACH: Grand challenge on breast cancer histology images," *Med. Image Anal.*, vol. 56, pp. 122–139, Aug. 2019.
[23] W. Nawaz, S. Ahmed, A. Tahir, and H. A. Khan, "Classification of breast cancer histology images using alexnet," in *Proc. Int. Conf. Image Anal. Recognit.*, 2018, pp. 869–876.
[24] H. Yao, X. Zhang, X. Zhou, and S. Liu, "Parallel structure deep neural network using CNN and RNN with an attention mechanism for breast cancer histology image classification," *Cancers*, vol. 11, no. 12, p. 1901, Nov. 2019.
[25] F. A. Spanhol, L. S. Oliveira, L. Heutte, and C. Petitjean, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1455–1462, Jul. 2016.
[26] A. Vahadane, T. Peng, A. Sethi, S. Albarqouni, L. Wang, M. Baust, K. Steiger, A. M. Schlitter, I. Esposito, and N. Navab, "Structure-preserving color normalization and sparse stain separation for histological images," *IEEE Trans. Med. Imag.*, vol. 35, no. 8, pp. 1962–1971, Aug. 2016.
[27] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, C. Schmitt, and N. E. Thomas, "A method for normalizing histology slides for quantitative analysis," in *Proc. IEEE Int. Symp. Biomed. Imaging: From Nano to Macro*, Boston, MA, USA, Jun. 2009.
[28] E. Reinhard, M. Adhikhmin, B. Gooch, and P. Shirley, "Color transfer between images," *IEEE Comput. Graph. Appl.*, vol. 21, no. 5, pp. 34–41, Sep./Oct. 2002.
[29] E. H. Land and J. J. McCann, "Lightness and Retinex theory," *J. Opt. Soc. Amer.*, vol. 61, no. 1, pp. 1–11, 1971.
[30] I. Fondón, A. Sarmiento, A. I. García, M. Silvestre, C. Eloy, A. Polónia, and P. Aguiar, "Automatic classification of tissue malignancy for breast carcinoma diagnosis," *Comput. Biol. Med.*, vol. 96, pp. 41–51, May 2018.
[31] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, and S. Li, "Breast cancer multi-classification from histopathological images with structured deep learning model," *Sci. Rep.*, vol. 7, no. 1, p. 4172, 2017.
[32] G. Murtaza, A. W. A. Wahab, G. Raza, and L. Shuib, "A tree-based multiclassification of breast tumor histopathology images through deep learning," *Comput. Med. Imag. Graph.*, vol. 89, Apr. 2021, Art. no. 101870.
[33] M. Z. Alom, C. Yakopcic, M. S. Nasrin, T. M. Taha, and V. K. Asari, "Breast cancer classification from histopathological images with inception recurrent residual convolutional neural network," *J. Digit. Imag.*, vol. 32, no. 4, pp. 605–617, Aug. 2019.
[34] H. Elmannai, M. Hamdi, and A. AlGarni, "Deep learning models combining for breast cancer histopathology image classification," *Int. J. Comput. Intell. Syst.*, vol. 14, no. 1, p. 1003, 2021.

[35] A. Bosch, A. Zisserman, and X. Munoz, "Image classification using random forests and ferns," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
[36] J. Wang, J. Yang, K. Yu, F. Lv, T. Huang, and Y. Gong, "Locality-constrained linear coding for image classification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, San Francisco, CA, USA, Jun. 2010, pp. 13–18.
[37] C. Zhang, G. Li, and S. Du, "Multi-scale dense networks for hyperspectral remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 11, pp. 9201–9222, Nov. 2019.
[38] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 448–456.
[39] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines Vinod Nair," in *Proc. 27th Int. Conf. Mach. Learn. (ICML)*, Haifa, Israel, Jun. 2010, pp. 807–814.
[40] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.
[41] D. Jia, D. Wei, R. Socher, L. J. Li, L. Kai, and F. F. Li, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 248–255.
[42] Z. Meng, Z. Zhao, and F. Su, "Multi-classification of breast cancer histology images by using gravitation loss," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, May 2019.
[43] F. A. Spanhol, L. S. Oliveira, P. R. Cavalin, C. Petitjean, and L. Heutte, "Deep features for breast cancer histopathological image classification," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2017, pp. 1868–1873.
[44] V. Gupta and A. Bhavsar, "Sequential modeling of deep features for breast cancer histopathological image classification," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW)*, Jun. 2018, pp. 2254–2261.
[45] P. J. Sudharshan, C. Petitjean, F. Spanhol, L. E. Oliveira, L. Heutte, and P. Honeine, "Multiple instance learning for histopathological breast cancer image classification," *Expert Syst. Appl.*, vol. 117, pp. 103–111, Mar. 2019.
[46] K. Kumar and A. C. S. Rao, "Breast cancer classification of image using convolutional neural network," in *Proc. 4th Int. Conf. Recent Adv. Inf. Technol. (RAIT)*, Mar. 2018, pp. 1–6.

YIPING ZHOU received the B.S. degree in electronic information science and technology from Tsinghua University, Beijing, China, in 2014. He is currently pursuing the Ph.D. degree in communication and information system with the University of Chinese Academy of Sciences, Beijing. His research interests include the field of artificial intelligence, deep learning, machine learning, image processing, and medical engineering.

CAN ZHANG received the B.S. and M.S. degrees from the University of Science and Technology of China, Hefei, China. She is currently a Professor with the University of Chinese Academy of Sciences. Her research interests include video processing and communication, mobile wireless communication, wireless security, and information processing.

SHAOSHUAI GAO (Member, IEEE) received the B.S. degree from Tianjin University, Tianjin, China, in 1998, and the Ph.D. degree in signal and information processing from the Graduate University of Chinese Academy of Sciences, Beijing, China. He conducted postgraduate research on error resilient video coding with Nanyang Technological University, Singapore, from 2003 to 2004, and he was a Researcher with the National Institute of Standards and Technology, USA, from 2004 to 2009. He has been a Professor with the University of Chinese Academy of Sciences, since 2009. His research interests include wireless communication, visible light communication, video processing, and communication.