Medical Tumor Image Classification Based on Few-Shot Learning

Wenyan Wang, Yongtao Li, Kun Lu, Jun Zhang, Peng Chen, Ke Yan, and Bing Wang

***Abstract*—As a high mortality disease, cancer seriously affects people's life and well-being. Reliance on pathologists to assess disease progression from pathological images is inaccurate and burdensome. Computer aided diagnosis (CAD) system can effectively assist diagnosis and make more credible decisions. However, a large number of labeled medical images that contribute to improve the accuracy of machine learning algorithm, especially for deep learning in CAD, are difficult to collect. Therefore, in this work, an improved few-shot learning method is proposed for medical image recognition. In addition, to make full use of the limited feature information in one or more samples, a feature fusion strategy is involved in our model. On the dataset of BreakHis and skin lesions, the experimental results show that our model achieved the classification accuracy of 91.22% and 71.20% respectively when only 10 labeled samples are given, which is superior to other state-of-the-art methods.**

***Index Terms*—Computer-aided diagnosis systems, few-shot learning, health care, medical image.**

I. Introduction

IMAGING examination is an important auxiliary means for tumor diagnosis and treatment [1]. Computed tomography (CT), magnetic resonance imaging (MRI), positron emission tomography (PET), mammography, ultrasound, and X-rays are the basic types of clinical images, and they can be used to find cancer, track cancer progression and judge the efficacy of cancer treatment. In this process, professional doctors can usually observe important tumor information from a few-shot different medical images [2], [3], [4]. However, the increase in patients leads to a greater workload for physicians and may therefore affect the accuracy of judgment. With the development of computer, computer aided diagnosis system (CAD) is proposed to assist physicians in making decisions and improve the accuracy of diagnosis [5], [6]. In the past two years, CAD have become even more essential, since the outbreak of COVID-19 has not only increased the burden of doctors' diagnosis, but also made it difficult to predict the severity of infection in patients with different cancers [7], [8], [9]. However, the calculation methods used to classify lesions in diagnostic systems usually rely on a large number of annotation data, which is opposite to the image sample size in the real world [10], [11]. Therefore, it is desirable to develop a medical image classification model based on limited learnable samples.

In computer aided diagnosis system, decision tree (DT), artificial neural network (ANN), support vector machine (SVM), etc. are widely used to distinguish different diseases [11], [12], [13], [14]. For example, Tania et al. proposed an improved support vector machine with hybrid kernel function for breast cancer classification, and achieved the classification accuracy of 95.78% [15]. Zhang et al. reported 92.06% recognition accuracy for breast cancer biopsy images classification by using a support vector machine (SVM) with local binary patterns and contrast measurement (LBP/C) descriptors [10]. Spanhol et al. published a breast cancer histopathology image dataset named BreakHis, and used various most advanced texture descriptors such as local binary pattern (LBP) and local phase quantization (LPQ) to classify different types of breast cancer, and their reported classification accuracy was between 80% and 85% [11]. However, these methods are relies on hand-craft features designed by human experts based on their prior knowledge, which makes it challenging for non-experts to design CAD systems using machine learning techniques.

In recent years, deep learning methods have been widely used in the field of computer vision and achieved great success [16], [17], [18], [19]. Spanhol et al. used DeCAF feature and Alexnet to classify skin damage and achieved an accuracy of 85% [20]. Wei et al. expanded the BreakHis dataset 14 times through offline data enhancement, and then applied them to the BiCNN model. As a result, they achieved 97% recognition performance on tens of thousands of images [21]. It can be seen that these deep learning-based methods are more effective when the number of available annotated samples in the training phase is large. When the training sample size is limited, the performance of the model may also decline.

Manuscript received 25 January 2022; revised 4 April 2023; accepted 28 May 2023. Date of publication 9 June 2023; date of current version 8 August 2024. This work was supported in part by the National Natural Science Foundation of China under Grants 62172004, 62072002, and 61872004, in part by Anhui Province Collaborative Innovation Project under Grants GXXT-2022-050 and GXXT-2022-053, and in part by the Educational Commission of Anhui Province under Grant 2022AH050336. *(Corresponding authors: Bing Wang; Peng Chen.)*

Wenyan Wang and Kun Lu are with the School of Electrical and Information Engineering, Anhui University of Technology, Ma'anshan, Anhui 243032, China (e-mail: wenyanwang9203@gmail.com; kunlu0819@gmail.com).

Bing Wang is with the School of Electrical and Information Engineering and Wuhu Technology and Innovation Research Institute, Anhui University of Technology, Ma'anshan, Anhui 243032, China (e-mail: wangbing@ustc.edu).

Yongtao Li is with the School of Materials Science and Engineering, Anhui University of Technology, Ma'anshan, Anhui 243032, China (e-mail: liyongtao@ahut.edu.cn).

Jun Zhang and Peng Chen are with the Co-Innovation Center for Information Supply & Assurance Technology, Anhui University, Hefei, Anhui 230032, China (e-mail: 00568@ahu.edu.cn; pchen@ahu.edu.cn).

Ke Yan is with the Department of the Built Environment, National University of Singapore, Singapore 117566 (e-mail: yanke@nus.edu.sg).

Digital Object Identifier 10.1109/TCBB.2023.3282226

However, human beings can quickly learn and identifying objects from a small number of samples based on prior experience and knowledge. In order to integrate this performance into the deep learning model, the concept of few-shot learning is proposed in recent years [22]. Moreover, to classify medical tumor images, Singh et al. presented a meta-learning-based "MetaMed" approach with low computational and image availability, and three publicly accessible medical data sets are used to evaluate the performance of their proposed model. However, this method only designed a CAD system for breast cancer to identify three rare benign tumors while ignoring the recognition of common malignant tumors in the real world, which is inconsistent with the actual clinical application [23].

In disease diagnosis, tumor type discrimination is closely related to the follow-up treatment plan and life safety. To solve this problem, a computational method based on few-shot learning is proposed in this work to identify benign and malignant tumors in breast and skin diseases. In addition to using an efficient network WRN28-10 to accelerate model training, a feature fusion strategy is also designed to compensate for the limited sample information. And then, the breast tumor images with different magnification factors are applied to the few-shot learning model to identify the tumor type more accurately. The experimental results show that the improved model is superior to other existing few-shot learning methods for break and skin lesion classification. However, the few-shot learning method is currently only suitable for image data, and there is no relevant application for clinical and genetic data processing in the medical field. With the development of natural language processing models, it is possible to apply our idea to text data processing in the furture.

II. Methods

A. Few-Shot Learning

With the development of deep learning methods, the performance of the image-based classification algorithm has achieved unprecedented success [24]. However, as the precondition of model training based on standard deep learning, collecting a large number of specific image instances such as medical tumor images, and labeling them has gradually developed into a new research issue [25], [26]. While human beings can quickly obtain the ability to distinguish new objects from a bunch of labeled samples. few-shot learning attempts to bridge this gap by imitating the human learning paradigm [27], [28], [29], [30].

In the few-shot learning method, the database is also divided into training and test sets. But the training set in this method is usually composed of a few samples with known labels, which is also called the support set. At the same time, an open database for classification, such as MiniImagnet, is added as the basic dataset $X_{base} := \{x_i, y_i\}_{i=1}^{K_{base}}$, in which each category contains thousands of labeled images, where x_i represents raw features of sample i, and y_i indicates its associated one-hot encoded label. In the learning process of few-shot, to obtain as much a priori knowledge as possible, the few-shot classification model is first trained on a basic public dataset X_{base}. Cross entropy loss is taken as the objective function. After several iterations, a feature extractor f_θ with trained parameters is obtained. Then, several

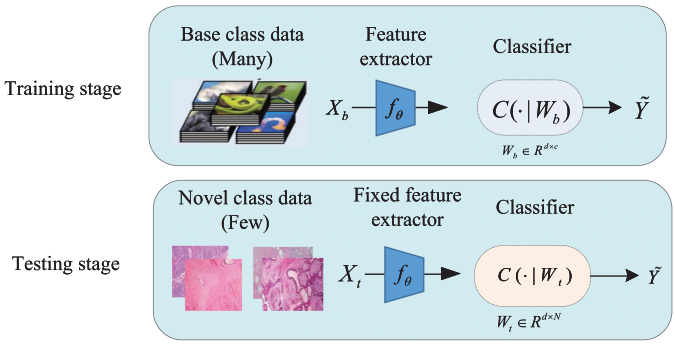


Fig. 1. The training and testing process of our proposed few-shot learning.

new labeled instances (the support set), which are selected from the database of specific tasks $X_N := \{x_i, y_i\}_{i=1}^{K_N}$ and disjoint with the testing classes, are given to fine-tune the classifier of the trained model. At this time, a small number of samples with unknown labels (the query set) $X_Q := \{x_i, y_i\}_{i=1}^{K_Q}$, which also from the specific database, such as breast cancer, are used to evaluate the performance of the model. Fig. 1 shows the training and testing process of the few-shot learning. For a K-way N-shot task, it represents randomly sampling N support samples from K different classes respectively. Typically, N is between 1 and 30, and $Q = 15$ [31], [32], [33].

Similar to the standard deep learning method based on convolutional neural network, the goal of the few-shot learning method in the training stage is to train many parameters of the feature extractor f_θ by using a large number of base classes data, and therefore obtain a priori knowledge model. In the test phase, the parameters of the f_θ are fixed, and some novel class data are only used to train a new classifier for identifying unknown target categories [23]. In order to use the information of query samples, a mutual information loss I_a is involved into the objective function of classifier training [27]. Different from the inductive method that uses the distance between query and support samples to distinguish the category of query samples, the few-shot learning method based on mutual information has been proved to have better classification performance, in which the transductive information maximization (TIM) is introduced as a loss function for parameter optimization of the new classifier, parametrized by weight matrix $W \in R^{K\times d}$ [27], [33], [34], [35], [36], [37], [38]. The training loss in the test phase is defined as

$$L = \gamma \cdot CE - I_a(X_Q; Y_Q) \tag{1}$$

where $CE = -\frac{1}{|S|}\sum_{i\in S}\sum_{n=1}^{N} y_{in} log(p_{in}), I_a(X_Q; Y_Q) = -\sum_{k=1}^{K} p_k log(p_k) + \alpha \frac{1}{|Q|}\sum_{i\in Q}\sum_{k=1}^{K} p_{ik} log(p_{ik})$, where $p_{ik} = P(Y = k \mid X = x_i; W, \theta)$ and $p_k = \frac{1}{Q}\sum_{i\in Q} p_{ik}$. The conditional entropy weight α and the cross-entropy weights γ are both set to 0.1.

B. Wideres Network and Feature Fusion

In recent years, the development of deep learning has promoted the optimization of backbone structure and performance improvement, because the backbone is an important core and component of the convolutional neural network (CNN)

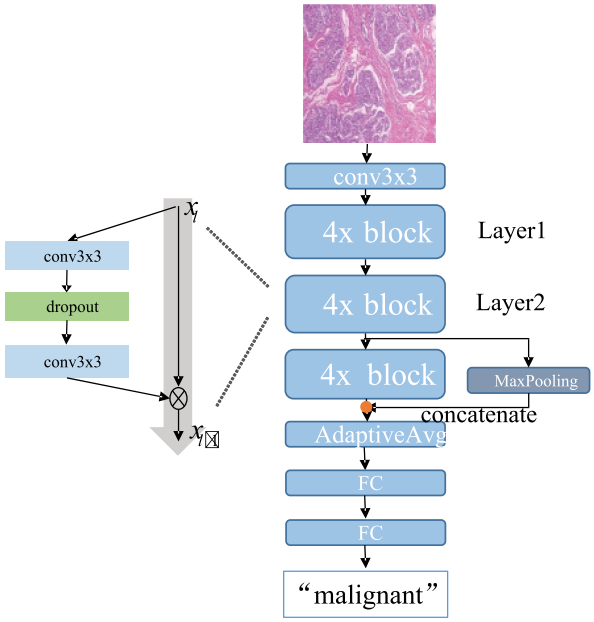


Fig. 2. The detailed network structure of *WRN28-10*. conv3x3 refers to a convolution layer with a kernel size equal to 3. 4x means the same convolution block is stacked 4 times. FC and AdaptiveAvg represent fully connection layer and adaptive average pooling. Orange dot refers to features concat.

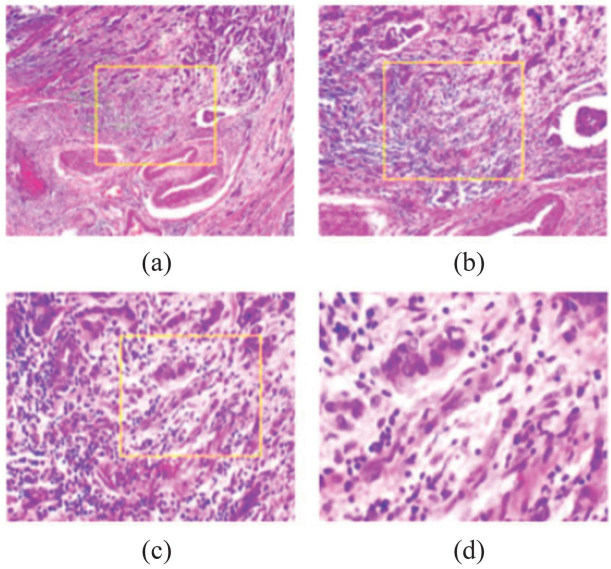


Fig. 3. A slide of breast malignant tumor (stained with HE) seen in different magnification factors: (a) 40X, (b) 100X, (c) 200X, and (d) 400X. The highlighted rectangle (added manually for illustrative purposes only) is the region of interest selected by the pathologist that will be detailed in the next higher magnification factor.

model [39], [40]. At the same time, improving the training speed of the model and then applying the CNN model to the actual production and life has gradually received extensive attention [41], [42]. In 2016, He et al. proposed the residual block with skip layer connection, and then various network variants should be shipped out [43]. From the perspective of improving the width and depth of the original ResNet, wide residual networks (WRN) were proposed by zagoruyko et al. in 2016, which has achieved advanced performance on CIFAR, SVHN, COCO, and important improvement on Imagenet [42].

WRN improves the prediction performance by using a shallow and wider (dimension or channel) channel on each convolution layer and adopts dropout regularization instead of batch normalization (BN) to avoid overfitting of the model [42]. Compared with the number of channels in ResNet, K is used to represent the multiple of network widening of WRN. In this work, we use a backbone *WRN28-10* with K of 10 and a depth of 28 as a feature extractor f_θ. In addition, to obtain more comprehensive feature information of small samples, a feature fusion strategy is involved in the backbone. Specially, the last two convolution blocks of the model are fused by increasing the number of neurons in the fully connected layer of the penultimate layer of the model. Here, max-pooling is adopted to unify the feature map size of different blocks and the fusion type is channel concatenation. The detailed network structure of *WRN28-10* is shown in Fig. 2.

III. Results and Discussion

A. Datasets

1) BreakHis Dataset: The BreakHis dataset contains 7,909 labeled histopathological images of 82 anonymous patients, including 2,480 benign tumor images and 5,429 malignant tumor images at four different magnification factors (40X, 100X, 200X, and 400X), as shown in Fig. 3 [21]. Histologically benign lesions are those that do not meet any malignant criteria, such as obvious cell atypia, mitosis, basement membrane rupture, and metastasis. Under normal circumstances, benign tumors are relatively harmless, grow slowly, and are restricted. A malignant tumor is synonymous with cancer where lesions can invade and destroy adjacent structures (locally aggressive), and spread to distant places (metastatic) leading to death. The scale of all histopathological images is 3 channels RGB, each color channel has a depth of 8 bits, and is fixed at 700×460 pixels. Benign tumors include four subtypes, namely adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenoma (TA), while malignant tumors include four subtypes which are ductal carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC), and papillary carcinoma (PC). The distribution of benign and malignant tumors in these subtypes is shown in Table I.

2) ISIC-2018 Skin Lesion Dataset: The ISIC-2018 Skin Lesions dataset encompasses a diverse range of dermoscopy images, originating from distinct anatomical sites and exhibiting 7 types of dermoscopic patterns, namely melanoma (HEL), melanocytic nevus (NV), basal cell carcinoma (BCC), actinic keratosis (AKIEC), benign keratosis (BKL), dermatofibroma (DF), and vascular lesion (VASC), as illustrated in Fig. 4, with a total of 10,015 images [38]. They are three-channel (RGB) images stored in JPEG format with the size of 600×450. The data distribution of these diseases is shown in Table II, which reflects a modified "real world" setting. Specifically, benign lesions are more than malignant lesions, but the proportion of malignant tumors is too high.

3) Implementation Details: In this work, all experiments are implemented on the NVIDIA Tesla V100 GPU hardware

TABLE I
IMAGE DISTRIBUTION BY MAGNIFICATION FACTOR AND CLASS

	Benign					Malignant				
Magnification	A	F	TA	PT	Total	DC	LC	MC	PC	Total
40x	114	253	109	149	598	864	156	205	145	1,370
100x	113	260	121	150	614	903	170	222	142	1,437
200x	111	264	108	140	594	896	163	196	135	1,390
400x	106	237	115	130	562	788	137	169	138	1,232
Total	444	1,014	453	569	2,368	3,451	626	792	560	5,429
# Patients	4	10	3	7	24	38	5	9	6	58

TABLE II
IMAGE DISTRIBUTION BY BENIGN AND MALIGNANT SUBTYPES OF DISEASES

Benign						Malignant		
DF	VASC	AKIEC	BKL	NV	Total	BCC	MEL	Total
115	142	327	1,099	6,705	8,388	514	1,113	1,627

TABLE III
THE ACCURACY OF OUR PROPOSED METHOD BOTH ON THE BREAKHISAND ISIC-2018 DATASETS

Dataset	1 shot	3 shots	5 shots	10 shots
BreakHis-40x	81.94 ± 1.24	88.34 ± 0.69	89.69 ± 0.55	**91.22 ± 0.50**
BreakHis-100x	79.53 ± 1.43	88.84 ± 0.65	89.51 ± 0.57	**91.31 ± 0.52**
BreakHis-200x	82.94 ± 1.33	89.94 ± 0.61	91.12 ± 0.54	**91.87 ± 0.45**
BreakHis-400x	79.41 ± 1.47	86.14 ± 0.65	87.34 ± 0.56	**88.08 ± 0.54**
ISIC-2018	63.49 ± 0.32	71.39 ± 0.24	74.72 ± 0.20	**77.20 ± 0.18**

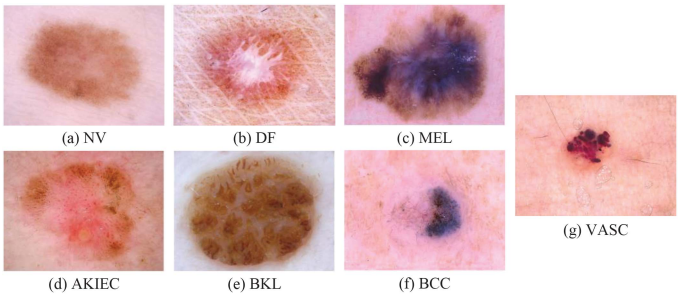


Fig. 4. Image samples of seven typical skin lesions.

platform with CUDA 10.1 and Pytorch 1.8.1. An SDG optimizer is adopted with a batch size of 256, the initial learning rate of 0.1, a weight decay of $1e$-3, and the total epochs is 90. These super parameters are determined based on similar work previously. The mini-ImageNet dataset is a subset of 100 classes selected from ILSVRC-12. There are 600 images in each class, which are divided into training, validation, and test meta sets with 64, 16, and 20 classes respectively [44]. The training set on mini-ImageNet is used to train a feature extractor, and 30% of the patients are in the BreakHis dataset and there are 30% of the images in the ISIC 2018 Skin lesions dataset are used to evaluate the performance of models followed by previous works [10], [20], [21], [45]. The final prediction accuracy is the average of the results of 600 repeated experiments, each containing 15 query samples.

B. Classification Performance of Improved Network

In order to apply the few-shot learning method to medical imaging recognition, this work proposes an improved few-shot classification model for breast cancer and melanoma recognition. The prediction results of our model on the BreakHis and ISIC-2018 datasets are listed in Table III.

As shown in Table III, the method proposed in this work can well distinguish benign and malignant tumors, especially for breast cancer images. Specifically, when the support sample is 1, the prediction accuracy of the model is 81.94%, 79.53%, 82.94%, 79.41%, and 63.49% respectively on BreakHis-40x, BreakHis-100x, BreakHis-200x, BreakHis-400x, and ISIC-2018 datasets. When the number of support samples for each new class increases to 3, the prediction accuracy of the model increases by 6.4%, 9.31%, 7%, 6.73%, and 7.9% respectively, reaching 88.34%, 88.84%, 89.94%, 86.14% and 71.39%. As the number of supporting samples increases, such as 5 and 10, the prediction performance of the model can be further improved. For example, when $K_s = 10$, the classification performance of the model is 91.22%, 91.31%, 91.87%, 88.08%, and 77.20%, respectively. However, it is worth noting that the improvement of model performance decreases with the increase of image magnification, which indicates that the information is lost with the image receptive field decreases, and it is therefore reasonable to think that there is an optimal combination between the magnification and the visual field of the image. According to the results of this work, the magnification of 200x is more useful

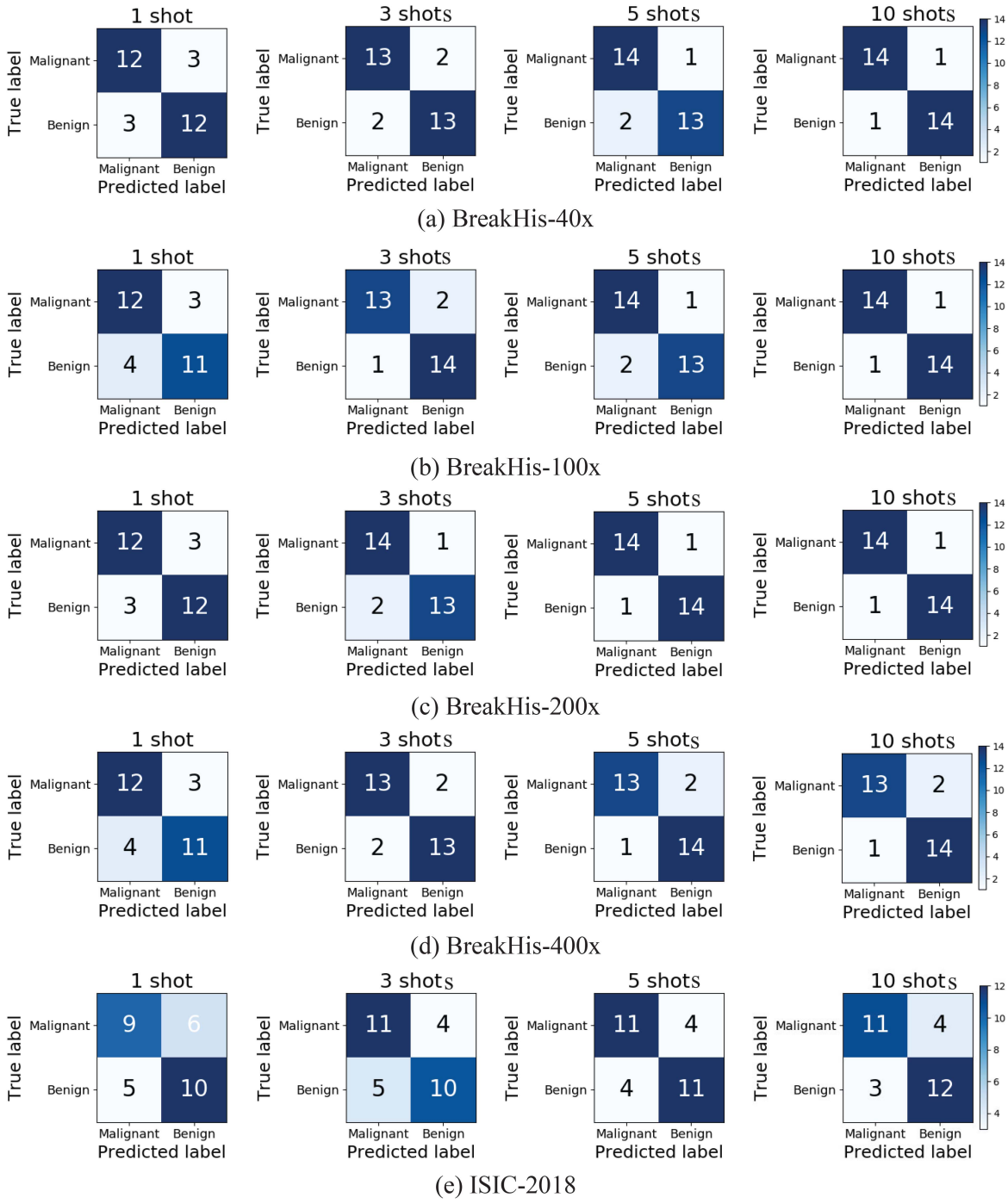


Fig. 5. Confusion matrices of the prediction results on the BreakHis-40x, BreakHis-100x, BreakHis-200x, BreakHis-400x, and ISIC-2018 datasets respectively.

for identifying tumor types. It can be seen from Fig. 5 that when K_s is equal to 1, 3, 5, and 10 respectively, there are 6, 4, 3, and 2 query samples with error predictions on BreakHis-40x, the error predicttions of query samples are 7, 3, 3, and 2 on BreakHis-100x, the error predictions of 6, 3, 2, and 2 query samples are on BreakHis-200x, the error predictions of 7, 4, 3, and 3 query samples are on BreakHis-400x, and there are 11, 9, 8, and 7 error predictions of query samples on ISIC-2018. Obviously, the few-shot learning-based model proposed in this work is effective with only a few samples participate in the learning.

C. Classification Performance of Feature Fusion

In convolutional neural network, it has been proved that the shallow features have the basic information of the target, such as the shape, texture and color etc., and with the deepening of the network layers, more semantic information is collected. However, due to the existence of the pooling layer, the image receptive field becomes larger, and the image details are gradually lost. To compensate for this information loss, a feature fusion module is introduced.

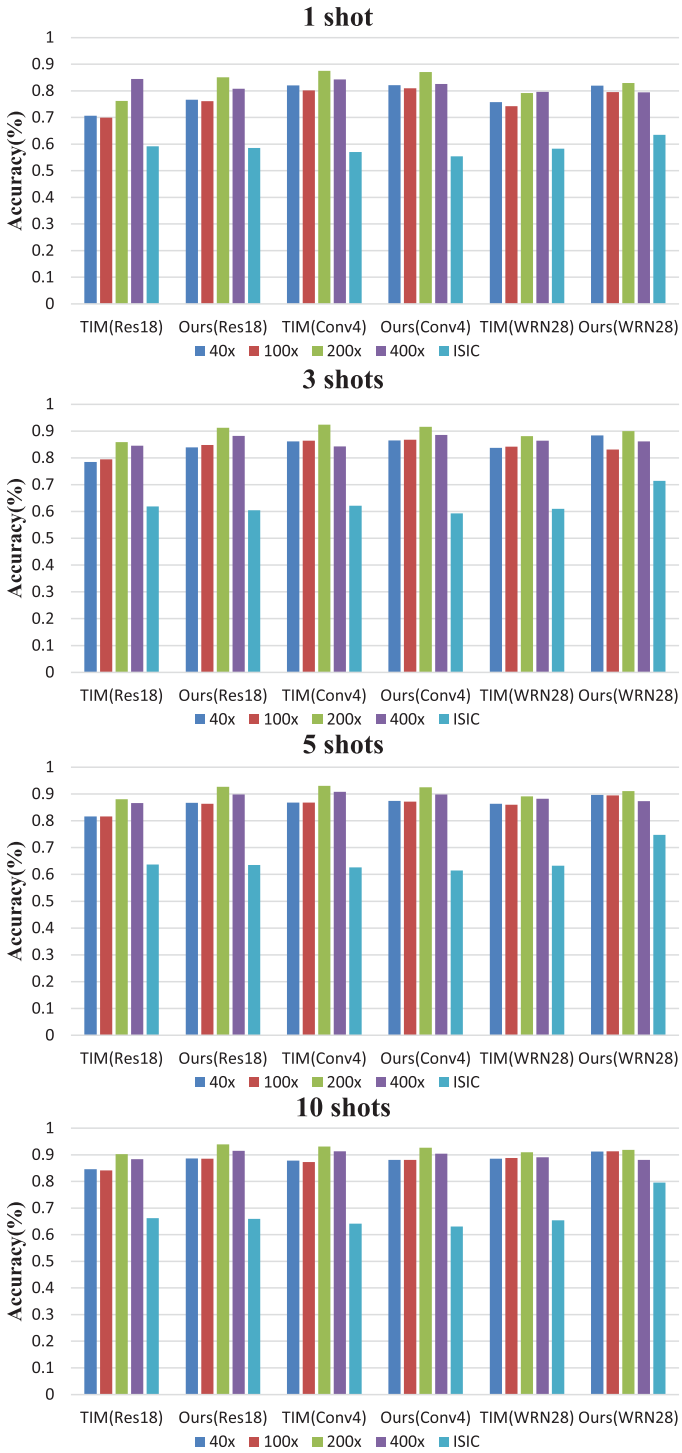


Fig. 6. The recognition accuracy comparison for our proposed method with the baseline on different backbones.

To verify the effectiveness of the feature fusion strategy proposed in this work, we compared the classification accuracy of different backbones with the baseline model, i.e. TIM, with support samples of 1, 3, 5, and 10, as shown in Fig. 6. Here, the same hyperparameter settings are adopted. It can be seen that the feature fusion strategy proposed in this work achieves the great performance gain on the BreakHis-40x, BreakHis-100x, and ISIC-2018 datasets. For example, when the support samples are 1, 3, 5, and 10 respectively,our WRN28-10 is 6.19%, 4.59%, 3.37%, 2.64% higher than the baseline model TIM on the BreakHis-40x dataset, 5.21%, 10.41%, 11.51%, and 11.83% higher on the ISIC-2018 dataset.Therefore, with the dimension of description sample features increasing, the samples can be distinguished more accurately, which also indicates that the feature fusion strategy proposed in this work is effective. Although the Conv4 network performs better on some datasets, such as BreakHis-200X and BreakHis-400X, its performance on ISIC-2018 datasets was far behind that of the WRN28-10 network. As shown in Fig. 6, our WRN28-10 has obtained the best and most balanced accuracy distribution on the five datasets. In other words, our WRN28-10 model achieved the best generalization performance.

D. Comparison With Other State-of-the-Art Methods

In order to further verify the effectiveness of our proposed method, several state-of-the-art few-shot learning methods are compared when support samples K_s is 1, 3, 5, and 10, respectively. Specifically, two inductive methods based on statistical rules which are Simple-shot and DeepEMDv2 [46], [48], and two transduction-based methods, i.e. PT-MAP [47] and TIM [27], are considered for comparison. As shown in Table IV, it can be seen that with the increase of the number of supported samples, the performance of all models based on few-shot learning has been further improved, but the performance gains are limited. When the support samples K_s are 1, 3, 5, and 10, respectively, the model proposed in this work achieves an accuracy of 81.94% to 91.22% on BreakHis-40x, 79.53% to 91.31% on BreakHis-100x, 82.94% to 91.87% onBreakHis-200x, 79.41% to 88.08% on BreakHis-400x, and 63.49% to 77.20% on ISIC-2018 dataset. In addition, the prediction accuracy of our proposed method is better than other state-of-the-art few-shot learning methods on the BreakHis-40x, BreakHis-100x, and ISIC-2018datasets. For example, when the support samples K_s are 1, 3, 5, and 10, respectively, the classification performance of our model are 9.91%, 7.11%, 4.86%, and 0.68% higher than Simple-shot, and 3.59%, 3.09%, 2.53%, and 4.79% higher than PT-MAP, 17.55%, 9.48%, 5.53%, 3.2% higher than DeepEMDv2, and 6.19%, 4.59%, 3.37%, 2.64% higher than TIM on BreakHis-40x dataset.

In addition, to verify that our model can achieve similar performance to the model based on large samples under the condition of small samples, Breakhis and ISIC-2018 datasets with large samples are therefore selected for foe-shot learning in this work. Specifically, to compare the difference in classification performance between few-shot learning and multiple-sample learning methods in this work, several deep learning methods are compared, such as AlexNet, MIL + CNN, deCAF + CaffeNet, and the results are shown in Table V. Spanhol et al. [21] chose the existing convolutional neural network model AlexNet as a classifier and reported an accuracy ranging from 80.7% to 85.6% on the BreakHis dataset. In the following work, they converted the CaffeNet model into a feature extractor and used logistic regression as the basic classifier, and achieving accuracy ranging

TABLE IV
PERFORMANCE COMPARISON OF DIFFERENT FEW-SHOT LEARNING METHODS

Dataset	Method	Model	1 shot	3 shots	5 shots	10 shots
BreakHis-40x	Simple-shot [46]	WRN28-10	72.03 ± 0.25	81.23 ± 0.17	84.83 ± 0.14	86.43 ± 0.13
	PT-MAP [47]	WRN28-10	78.35 ± 0.35	85.25 ± 0.20	87.16 ± 0.16	90.54 ± 0.12
	DeepEMDv2 [48]	ResNet12	64.39 ± 0.88	78.86 ± 0.79	84.16 ± 0.63	88.02 ± 0.49
	TIM [27]	WRN28-10	75.75 ± 1.27	83.75 ± 0.90	86.32 ± 0.68	88.58 ± 0.56
	ours	WRN28-10	**81.94 ± 1.24**	**88.34 ± 0.69**	**89.69 ± 0.55**	**91.22 ± 0.50**
BreakHis-100x	Simple-shot [46]	WRN28-10	71.90 ± 0.25	81.18 ± 0.17	85.01 ± 0.14	86.62 ± 0.13
	PT-MAP [47]	WRN28-10	75.92 ± 0.35	85.58 ± 0.20	87.46 ± 0.15	89.64 ± 0.13
	DeepEMDv2 [48]	ResNet12	64.57 ± 0.90	79.23 ± 0.74	83.88 ± 0.62	88.23 ± 0.50
	TIM [27]	WRN28-10	74.23 ± 1.42	84.15 ± 0.88	85.99 ± 0.72	88.78 ± 0.58
	ours	WRN28-10	**79.53 ± 1.43**	**88.84 ± 0.65**	**89.51 ± 0.57**	**91.31 ± 0.52**
BreakHis-200x	Simple-shot [46]	WRN28-10	79.16 ± 0.23	87.63 ± 0.15	**91.48 ± 0.12**	92.93 ± 0.10
	PT-MAP [47]	WRN28-10	**83.69 ± 0.31**	89.29 ± 0.16	90.49 ± 0.13	92.53 ± 0.11
	DeepEMDv2 [48]	ResNet12	69.97 ± 0.85	85.15 ± 0.70	89.77 ± 0.54	**93.81 ± 0.39**
	TIM [27]	WRN28-10	79.19 ± 1.29	88.11 ± 0.66	89.16 ± 0.62	90.94 ± 0.48
	ours	WRN28-10	82.94 ± 1.33	**89.94 ± 0.61**	91.12 ± 0.54	91.87 ± 0.45
BreakHis-400x	Simple-shot [46]	WRN28-10	79.95 ± 0.23	87.78 ± 0.15	**91.43 ± 0.11**	92.84 ± 0.10
	PT-MAP [47]	WRN28-10	**81.21 ± 0.32**	**88.57** ± 0.18	91.28 ± 0.14	92.37 ± 0.11
	DeepEMDv2 [48]	ResNet12	74.02 ± 0.93	87.66 ± 0.65	91.14 ± 0.48	**94.12 ± 0.37**
	TIM [27]	WRN28-10	79.63 ± 1.39	86.37 ± 0.72	88.22 ± 0.56	89.06 ± 0.55
	ours	WRN28-10	79.41 ± 1.47	86.14 ± 0.65	87.34 ± 0.56	88.08 ± 0.54
ISIC-2018	Simple-shot [46]	WRN28-10	58.91 ± 0.25	65.09 ± 0.22	67.73 ± 0.20	70.85 ± 0.17
	PT-MAP [47]	WRN28-10	59.28 ± 0.32	64.40 ± 0.27	67.50 ± 0.23	69.67 ± 0.20
	DeepEMDv2 [48]	ResNet12	57.14 ± 1.04	63.29 ± 0.91	65.01 ± 0.83	66.96 ± 0.69
	TIM [27]	WRN28-10	58.28 ± 1.10	60.98 ± 0.98	63.21 ± 0.86	65.37 ± 0.84
	ours	WRN28-10	**63.49 ± 0.32**	**71.39 ± 0.24**	**74.72 ± 0.2**	**77.20 ± 0.18**

TABLE V
CLASSIFICATION ACCURACY OF OUR FEW-SHOT LEARNING METHOD AND OTHER DEEP LEARNING APPROACHS BASED ON A LARGE NUMBER OF SAMPLES

Method	Accuracy				
	BreakHis-40x (1,404)	BreakHis-100x (1,460)	BreakHis-200x (1,410)	BreakHis-400x (1,258)	ISIC-2018 (9,432)
AlexNet [49]	85.60 ± 4.80	83.50 ± 3.90	82.70 ± 1.70	80.70 ± 2.90	-
MIL + CNN [50]	89.52	89.06	88.84	87.67	-
deCAF + CaffeNet [20]	83.00 ± 2.60	84.60 ± 5.00	84.00 ± 2.80	81.10 ± 3.90	-
GoogLeNet [51]	**94.82**	**94.38**	**94.67**	**93.49**	-
ResNet18 [43]	-	-	-	-	86.79
ResNet152d [43]	-	-	-	-	88.16
VGG16 [39]	-	-	-	-	**88.76**
	BreakHis-40x (10)	BreakHis-100x (10)	BreakHis-200x (10)	BreakHis-400x (10)	ISIC-2018 (10)
Ours	91.22 ± 0.50	91.31 ± 0.52	91.87 ± 0.45	88.08 ± 0.54	77.20 ± 0.18

The content in parentheses, such as (1,604) and (10), represents the number of training samples used in the large sample-based model or our few-shot learning method.

from 81.1% to 84.6% [47]. In the work of Das et al. a standard GoogLeNet network architecture and a majority voting-based strategy were used for slide level diagnosis. As a result, they reported accuracy ranging from 93.49% to 94.82% [52]. It can be seen that the model proposed in this work achieves accuracy ranging from 88.08% to 91.87% on the BreakHis dataset when only 10 samples are needed for each new class, which is better than other previous deep learning-based works, such as AlexNet, MIL + CNN, and deCAF + CaffeNet. On the ISIC-2018 dataset, the recognition performance of our proposed model reaches 77.20%, which is lower than the results of transfer learning for classical convolutional neural networks, such as ResNet and VGG. However, the samples used in this work are much smaller than those used in transfer learning. Generally, the results shown in Tables IV and V both illustrate the effectiveness and feasibility of our proposed method.

E. Discussion

It has been proved that the feature fusion strategy can improve the performance of the model. To explore the impact of different feature fusion methods on model performance, three

TABLE VI
ABLATION EXPERIMENTS ON FEATURE FUSION

Dataset	Layer1	Layer2	1 shot	3 shots	5 shots	10 shots
BreakHis-40x	✗	✗	75.75 ± 1.27	83.75 ± 0.90	86.32 ± 0.68	88.58 ± 0.56
	✓	✗	79.88 ± 1.41	86.72 ± 0.69	87.53 ± 0.57	88.81 ± 0.53
	✗	✓	**81.94 ± 1.24**	**88.34 ± 0.69**	**89.69 ± 0.55**	**91.22 ± 0.50**
	✓	✓	81.16 ± 1.26	88.00 ± 0.74	89.69 ± 0.58	91.16 ± 0.52
BreakHis-100x	✗	✗	74.23 ± 1.42	84.15 ± 0.88	85.99 ± 0.72	88.78 ± 0.58
	✓	✗	79.39 ± 1.51	86.74 ± 0.77	87.82 ± 0.54	88.83 ± 0.53
	✗	✓	79.53 ± 1.43	**88.84 ± 0.65**	89.51 ± 0.57	91.31 ± 0.52
	✓	✓	**80.07 ± 1.38**	88.44 ± 0.74	**89.61 ± 0.57**	**91.47 ± 0.51**
BreakHis-200x	✗	✗	79.19 ± 1.29	88.11 ± 0.66	89.16 ± 0.62	90.94 ± 0.48
	✓	✗	85.96 ± 1.26	**91.10 ± 0.48**	**91.91 ± 0.47**	**92.18 ± 0.43**
	✗	✓	82.94 ± 1.33	89.94 ± 0.61	91.12 ± 0.54	91.87 ± 0.45
	✓	✓	**83.34 ± 1.35**	90.38 ± 0.63	91.52 ± 0.53	91.93 ± 0.45
BreakHis-400x	✗	✗	79.63 ± 1.39	86.37 ± 0.72	88.22 ± 0.56	89.06 ± 0.55
	✓	✗	76.01 ± 1.53	84.93 ± 0.92	88.31 ± 0.56	89.24 ± 0.51
	✗	✓	79.41 ± 1.47	86.14 ± 0.65	87.34 ± 0.56	88.08 ± 0.54
	✓	✓	**80.46 ± 1.49**	**87.73 ± 0.62**	**88.59 ± 0.55**	**89.41 ± 0.53**
ISIC-2018	✗	✗	58.28 ± 1.10	60.98 ± 0.98	63.21 ± 0.86	65.37 ± 0.84
	✓	✗	51.89 ± 0.84	54.81 ± 0.90	55.90 ± 0.85	59.17 ± 0.82
	✗	✓	**63.49 ± 0.32**	**71.39 ± 0.24**	**74.72 ± 0.20**	**77.20 ± 0.18**
	✓	✓	58.92 ± 1.18	62.70 ± 1.01	63.63 ± 0.93	65.89 ± 0.83

"Layer 1" refers to increasing the number of neurons in the fully connected layer of the penultimate layer of the model by fusing the first convolution block of the model. "Layer2" indicates that the output of the second 4x module is also used as the input of the adaptive full connection layer.

fusion schemes are designed, and their results are shown in Table VI. It can be found that compared with the baseline model, when the output of the first convolution block named Layer1 is also used as the input of the full connection layer, the model recognition accuracy is improved by 3.8%, 2.97%, 1.21% and 0.23% respectively on the Breakhis-40x data set When K_s is 1, 3, 5 and 10. And the classification accuracy of Breakhis-100x data set is also improved by 5.16%, 2.59%, 1.83% and 0.05%. Similarly, the improvement of model performance is effective on Breakhis-200x, which shows that Layer1 is a desirable fusion method. When the second convolution block is concatented and input to the full connection layer, the performance of the model is further improved, and the best recognition results are obtained on Breakhis-40x and ISIC-2018. However, when these two convolution blocks are connected at the same time, the performance of the model is only optimal on breakhis-100x and breakhis-400x. The potential reason is that the serious information redundancy caused by too many features affects the performance of the model. Therefore, considering the classification performance of different fusion modes on different data sets, the fusion method of concatented the second convolution feature selected in this work is the best.

IV. CONCLUSION

Breast cancer and skin lesions are two common diseases with high mortality. To reduce the cost of diagnosis for patients and make medical diagnosis less error-prone, this work proposes a breast tumor recognition method based on few-shot learning. At the same time, skin lesion images are also discriminated to evaluate the generalization ability of our improved model. In the few-shot learning method, a feature map fusion strategy at different levels is involved in the model to expand the dimension of feature expression of limited sample information. As a result, the effectiveness of this method is proved by us. At the same time, when only a small number of labeled samples are given, such as 1, 3, 5, and 10, the experimental results demonstrate that our improved model has the most state-of-the-art classification performance, which is comparable to the deep learning model based on a large number of learnable samples.

Although the few-shot learning method proposed in this work can greatly reduce the demand for labeled training samples, its recognition accuracy is still to be improved compared with the standard convolutional neural network. In the future work, we attempt to make better use of other prior knowledge or unlabeled data to train the model.

REFERENCES

[1] T. Aoki et al., "Narrow-band imaging examination of microvascular architecture of subcapsular hepatic tumors," *J. Surg. Res.*, vol. 261, pp. 51–57, 2021.
[2] K. Okada et al., "Dual-energy computed tomography for evaluation of breast cancer: Value of virtual monoenergetic images reconstructed with a noise-reduced monoenergetic reconstruction algorithm," *Japanese J. Radiol.*, vol. 38, no. 2, pp. 154–164, 2019.
[3] A. Fogelstrom, F. Hallen, and K. Pekkari, "Computed tomography diagnosed first time diverticulitis and colorectal cancer," *Int. J. Colorectal Dis.*, vol. 35, pp. 1895–1901, 2020.
[4] F. Connolly, J. Roehl, C. Guthke, O. Wengert, J. Valdueza, and S. Schreiber, "Emergency room use of "fast-track" ultrasound in acute stroke: An observational study," *Ultrasound Med. Biol.*, vol. 45, no. 5, pp. 1103–1111, 2019.

[5] J. Tang, R. M. Rangayyan, J. Xu, I. E. Naqa, and Y. Yang, "Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances," *IEEE Trans. Inf. Technol. Biomed.*, vol. 13, no. 2, pp. 236–251, Mar. 2009.

[6] A. Jalalian, S. Mashohor, H. R. Mahmud, M. I. B. Saripan, A. R. bin Ramli, and B. Karasfi, "Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: A review," *Clin. Imag.*, vol. 37, no. 3, pp. 420–426, 2013.

[7] Q. Song et al., "Risk and outcome of breakthrough COVID-19 infections in vaccinated patients with cancer: Real-world evidence from the national covid cohort collaborative," *J. Clin. Oncol.*, vol. 40, no. 13, pp. 1414–1427, 2022. [Online]. Available: https://doi.org/10.1200/JCO.21.02419

[8] Y. Liu et al., "Intrapleural nano-immunotherapy promotes innate and adaptive immune responses to enhance anti-PD-L1 therapy for malignant pleural effusion," *Nature Nanotechnol.*, vol. 17, no. 2, pp. 206–216, 2022.

[9] Z. Lu, S. Rajan, Q. Song, Y. Zhao, and B. Lu, "3D scaffold-free microlivers with drug metabolic function generated by lineage-reprogrammed hepatocytes from human fibroblasts," *Biomaterials*, vol. 269, no. 5, 2021, Art. no. 120668.

[10] Y. Zhang, B. Zhang, F. Coenen, J. Xiao, and W. Lu, "One-class kernel subspace ensemble for medical image classification," *EURASIP J. Adv. Signal Process.*, vol. 2014, pp. 1–13, 2014.

[11] F. A. Spanhol, L. Oliveira, C. Petitjean, and L. Heutte, "A dataset for breast cancer histopathological image classification," *IEEE Trans. Biomed. Eng.*, vol. 63, pp. 1455–1462, Jul. 2016.

[12] K. Yan and H. Lu, "An extended genetic algorithm based gene selection framework for cancer diagnosis," in *Proc. IEEE 9th Int. Conf. Inf. Technol. Med. Educ.*, 2018, pp. 43–47.

[13] K. Yan, H.-L. Cheng, Z. Ji, X. Zhang, and H. Lu, "Accelerating smooth molecular surface calculation," *J. Math. Biol.*, vol. 76, pp. 779–793, 2018.

[14] K. Yan, B. Wang, H. Cheng, Z. Ji, J. Huang, and Z. Gao, "Molecular skin surface-based transformation visualization between biological macromolecules," *J. Healthcare Eng.*, vol. 2017, 2017, Art. no. 4818604.

[15] F. A. Tania and P. Chandra Shill, "An improved support vector machine for disease diagnosis," in *Proc. IEEE 22nd Int. Conf. Comput. Inf. Technol.*, 2019, pp. 1–6.

[16] X. Zhou, W. Liang, K. Wang, and S. Shimizu, "Multi-modality behavioral influence analysis for personalized recommendations in health social media environment," *IEEE Trans. Computat. Social Syst.*, vol. 6, no. 5, pp. 888–897, Oct. 2019.

[17] X. Zhou, X. Xu, W. Liang, Z. Zeng, and Z. Yan, "Deep-learning-enhanced multitarget detection for end-edge-cloud surveillance in smart IoT," *IEEE Internet Things J.*, vol. 8, no. 16, pp. 12588–12596, Aug. 2021.

[18] X. Zhou, Y. A. Li, and W. Liang, "CNN-RNN based intelligent recommendation for online medical pre-diagnosis support," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 3, pp. 912–921, May/Jun. 2021.

[19] X. Zhou, W. Liang, K. I.-K. Wang, R. Huang, and Q. Jin, "Academic influence aware and multidimensional network analysis for research collaboration navigation based on scholarly Big Data," *IEEE Trans. Emerg. Topics Comput.*, vol. 9, no. 1, pp. 246–257, First Quarter 2021.

[20] 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.*, 2017, pp. 1868–1873.

[21] B. Wei, Z. Han, X. He, and Y. Yin, "Deep learning model based breast cancer histopathological image classification," in *Proc. IEEE 2nd Int. Conf. Cloud Comput. Big Data Anal.*, 2017, pp. 348–353.

[22] W. Liang, Y. Hu, X. Zhou, Y. Pan, and K. I.-K. Wang, "Variational few-shot learning for microservice-oriented intrusion detection in distributed industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 18, no. 8, pp. 5087–5095, Aug. 2022.

[23] R. Singh, V. Bharti, V. Purohit, A. Kumar, A. Singh, and S. Singh, "Metamed: Few-shot medical image classification using gradient-based meta-learning," *Pattern Recognit.*, vol. 120, 2021, Art. no. 108111.

[24] X. Zhou, W. Liang, K. I.-K. Wang, and L. T. Yang, "Deep correlation mining based on hierarchical hybrid networks for heterogeneous Big Data recommendations," *IEEE Trans. Computat. Social Syst.*, vol. 8, no. 1, pp. 171–178, 2020.

[25] E. Miller, N. Matsakis, and P. Viola, "Learning from one example through shared densities on transforms," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2000, pp. 464–471.

[26] X. Zhou, Y. Hu, J. Wu, W. Liang, J. Ma, and Q. Jin, "Distribution bias aware collaborative generative adversarial network for imbalanced deep learning in industrial IoT," *IEEE Trans. Ind. Informat.*, vol. 19, no. 1, pp. 570–580, Jan. 2023.

[27] M. Boudiaf, I. M. Ziko, J. Rony, J. Dolz, P. Piantanida, and I. B. Ayed, "Transductive information maximization for few-shot learning," 2020, *arXiv: 2008.11297*.

[28] X. Zhou, W. Liang, S. Shimizu, J. Ma, and Q. Jin, "Siamese neural network based few-shot learning for anomaly detection in industrial cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 17, no. 8, pp. 5790–5798, Aug. 2021.

[29] X. Zhou et al., "Intelligent small object detection for digital twin in smart manufacturing with industrial cyber-physical systems," *IEEE Trans. Ind. Informat.*, vol. 18, no. 2, pp. 1377–1386, Feb. 2022.

[30] H. Li, D. Eigen, S. F. Dodge, M. D. Zeiler, and X. Wang, "Finding task-relevant features for few-shot learning by category traversal," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2019, pp. 1–10.

[31] Q. Cai, Y. Pan, T. Yao, C. C. Yan, and T. Mei, "Memory matching networks for one-shot image recognition," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 4080–4088.

[32] F. Sung, Y. Yang, L. Zhang, T. Xiang, P. H. S. Torr, and T. M. Hospedales, "Learning to compare: Relation network for few-shot learning," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 1199–1208.

[33] Z. Ji, X. Chai, Y. Yu, Y. Pang, and Z. Zhang, "Improved prototypical networks for few-shot learning," *Pattern Recognit. Lett.*, vol. 140, pp. 81–87, 2020.

[34] P. T. de Boer, D. P. Kroese, S. Mannor, and R. Y. Rubinstein, "A tutorial on the cross-entropy method," *Ann. Operations Res.*, vol. 134, pp. 19–67, 2005.

[35] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[36] J. Liu, L. Song, and Y. Qin, "Prototype rectification for few-shot learning," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 741–756.

[37] Y. Ma et al., "Transductive relation-propagation network for few-shot learning," in *Proc. Int. Joint Conf. Artif. Intell.*, 2020, pp. 804–810.

[38] J. Zou, X. Ma, C. Zhong, and Y. Zhang, "Dermoscopic image analysis for ISIC challenge2018," 2018, *arXiv:1807.08948*.

[39] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2015, *arXiv:1409.1556*.

[40] M. Sandler, A. G. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 4510–4520.

[41] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 6848–6856.

[42] S. Zagoruyko and N. Komodakis, "Wide residual networks," 2016, *arXiv:1605.07146*.

[43] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*2016, pp. 770–778.

[44] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," *Int. J. Comput. Vis.*, vol. 115, pp. 211–252, 2015.

[45] M. Simon, E. Rodner, and J. Denzler, "ImageNet pre-trained models with batch normalization," 2016, *arXiv:1612.01452*.

[46] Y. Wang, W.-L. Chao, K. Q. Weinberger, and L. van der Maaten, "SimpleShot: Revisiting nearest-neighbor classification for few-shot learning," 2019, *arXiv: 1911.04623*.

[47] Y. Hu, V. Gripon, and S. Pateux, "Leveraging the feature distribution in transfer-based few-shot learning," in *Proc. Int. Conf. Artif. Neural Netw.*, 2021, pp. 487–499.

[48] C. Zhang, Y. Cai, G. Lin, and C. Shen, "DeepEMD: Differentiable earth mover's distance for few-shot learning," 2020, *arXiv:2003.06777*.

[49] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Commun. ACM*, vol. 60, pp. 84–90, 2012.

[50] K. Das, S. Conjeti, A. G. Roy, J. Chatterjee, and D. Sheet, "Multiple instance learning of deep convolutional neural networks for breast histopathology whole slide classification," in *Proc. IEEE 15th Int. Symp. Biomed. Imag.*, 2018, pp. 578–581.

[51] C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.

[52] K. Das, S. P. K. Karri, A. G. Roy, J. Chatterjee, and D. Sheet, "Classifying histopathology whole-slides using fusion of decisions from deep convolutional network on a collection of random multi-views at multi-magnification," in *Proc. IEEE 14th Int. Symp. Biomed. Imag.*, 2017, pp. 1024–1027.

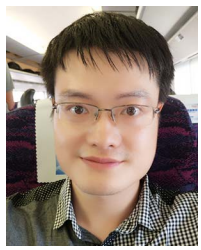
Wenyan Wang received the MS degree from the School of Electrical Information and Engineering, Anhui University of Technology, in 2018, and the PhD degree from the School of Metallurgical Engineering. She is currently serving as a postdoctoral teacher in Anhui University of Technology. Her research interests include deep learning, intelligent inspection, data mining, and bioinformatics.

Yongtao Li received the PhD degree in physical electronics from Fudan University (FDU), in 2011, and the Postdoctoral degree in energy storage materials from Fudan University. He is currently a professor of materials science and director of Low-Carbon New Materials Research Center, Anhui University of Technology. He joined the Anhui University of Technology and further worked with prof. E. Akiba/ Hai-Wen Li research group, Kyushu University, Japan as a visiting professor. His research focuses on the design and applications of nanostructured metal materials as energy storage media, electrodes and solid electrolytes.

Kun Lu is currently working toward the PhD degree with the School of Management Science and Engineering, Anhui University of Technology, Ma'anshan, China. His research interests include machine learning, data mining, and bioinformatics.

Jun Zhang received the PhD degree from the University of Science and Technology of China. He served as a postdoc fellow with the University of Louisville, USA, from 2009-2011. He focuses on deep learning with application to bioinformatics, Cheminformatics and computer vision etc. He has published more than 40 papers in international conferences and journals. Currently, he is an associate professor with the School of Electrical Engineering and Automation, Anhui University, 230601 Hefei, China.

Peng Chen received the PhD degree from the University of Science and Technology of China. He served as a senior research associate with the City University of Hong Kong, in 2006, as a postdoc fellow with Howard University, USA from 2008-2009, as a research fellow with Nanyang Technological University, Singapore from 2009-2010, and as a postdoc fellow with the King Abdullah University of Science and Technology (KAUST), Saudi Arabia from 2012-2014. He is a professor with the School of Computer Science & Technology and Institute of Physical Science & Information Technology, Anhui University, Hefei, China. He speciliases in machine learning and data mining with applications to bioinformatics, drug discovery, computer vision, etc. He has published about 100 high quality referred papersin international conferences and journals.



Ke Yan (Member, IEEE) received the bachelor's and PhD degrees in computer science from the School of Computing (SoC), National University of Singapore (NUS), Singapore, in 2006 and 2012, respectively. He worked with the Masdar Institute of Science and Technology (MIST) campus, Khalifa University, Abu Dhabi, UAE as a post-doctoral researcher from 2012 to 2014. He was with China Jiliang University, Hangzhou, China and currently an assistant professor of College ofDesign and Engineering (CDE), NUS. He has been the leading/corresponding guest editor of special issues for multiple IEEE journals, such as *IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB)*, *IEEE Transactions on Industrial Informatics (TII)* and *Building and Environment (BAE)*. He is actively engaged in cross-discipline research fields, including machine learning, artificial intelligence, cyber intelligence, applied mathemat-ics, sustainability, applied energy and etc. He has published more than 60 full length papers with highly ranked confer-ences and journals, such as International Conference on Autonomous Agents and Multi-Agent Systems (AAMAS), Association for the Advancement of Artificial Intelligence (AAAI), *IEEE Transactions on Industrial Informatics (TII)*, *IEEE Transactions on Sustainable Energy (TSE)*, *IEEE Transactions on Systems, Man and Cybernetics: Systems (SMCA)*, *IEEE Access*, *Applied Energy (AE)*, *Future Gen-eration Computer Systems (FGCS)*, *Energy and Buildings (ENB)*, *Building and Environment (BAE)* and *International Journal of Refrigeration (JIJR)*.

Bing Wang (Senior Member, IEEE) received the PhD degree from the University of Science and Technology of China, Hefei, China, in 2006. He worked as a senior research associate with the City University of Hong Kong, 2006-2007, and a postdocoral fellow with the University of Louisville and Vanderbilt University, USA, from 2008 to 2012. Currently, he is serving as a full professor with the School of Electrical and Information Engineering, Anhui University of Technology, Ma'anshan, China. He has more than 100 publications, and more than 1,000 citations. His research interests mainly focus on machine learning, image processing, computational biology, and chemoinformatics.