Received October 21, 2019, accepted November 7, 2019, date of publication November 15, 2019,
date of current version November 27, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2953775

Research on Urine Sediment Images Recognition Based on Deep Learning

QINGBO JI, XUN LI, ZHIYU QU, AND CHONG DAI
College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China

Corresponding author: Zhiyu Qu (quzhiyu@hrbeu.edu.cn)

This work was supported in part by the National Natural Science Foundation of China under Grant 61801143 and Grant 61671168, in part by the National Natural Science Foundation of Heilongjiang Province under Grant LH2019F005 and Grant LH2019F010, and in part by the Fundamental Research Funds for the Central Universities under Grant 3072019CF0801 and Grant 3072019CFM0802.

ABSTRACT Detection of urine sediment microscopic images of human urine samples plays an important part in vitro examination. Doctors usually use automatic urine sediment analyzer to assist manual examine. At present, automatic urine sediment analyzers mostly use traditional method of artificial feature extraction to recognize urine sediment images. However, traditional image processing methods based on the selection and combination of feature operators and classifiers require a lot of work and subjective experience for engineers in the implementation process. It's also difficult to deal with urine sediment images recognition tasks with large scale categories, and particles in some different categories are often confused in recognition using traditional image processing methods, such as red blood cells (RBCs) and white blood cells (WBCs). In this paper, a combination convolution neural network (CNN) recognition method with area feature algorithm is proposed. The disadvantage that CNN can weaken the area feature of input image is solved by area feature algorithm (AFA) proposed in this paper. The network models which use 300,000 urine sediment images for training can quickly and accurately recognize 10 categories of urine sediment images, and several confusing categories' recognition indexes are remarkably improved. The test accuracy in the test set reached 97%.

INDEX TERMS Automatic diagnosis, deep learning, image processing, urine sediment.

I. INTRODUCTION

The urine discharged by healthy people is ought to be clear and sterile, however certain sediments in the urine can reflect the corresponding diseases. For example, from the appearance of a great quantity of WBCs in the urine we can speculate that the patient is with urinary system inflammation. A large number of RBCs, calcium oxalate crystals (CAOXs), and hyaline casts (HYALs) shows that the patient is with urinary tract calculi and so on [1]–[3]. Therefore, microscopic examination of micro-particles in human urine analysis is one of the most important external diagnostic projects in nephrology and urology departments. Compared with the urine dry chemical method, urine sediment microscopy can directly reflect the corresponding symptoms of patients using microscopic images, and plays an irreplaceable role in urine analysis [4], [5].

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

For urine microscopic images are characterized by many categories of particles and the particles are tiny and unevenly distributed, urine microscopic examination requires experienced laboratory physicians to operate. Moreover, urine analysis has a large workload and is very time-consuming, and can be easily affected by the subjective will and working fatigue of laboratory physicians. With the continuous development of digital image technology, automatic image recognition has been widely used in the field of medicine. The traditional method based on "target segmentation + feature selection and extraction + classifier" has achieved certain achievements in recognition and classification of urine sediment images [6]–[9]. The performance of these methods mainly depends on the accuracy of target segmentation and the effectiveness of feature selection and combination. Compared with other general images, urine sediment images have fewer distinctive features visible to the human eye and high similarity among some categories. Thus it is difficult to break through the bottleneck of high-accuracy classification using traditional methods.

In recent years, with the rapid development of high performance computing hardwares as graphics processing unit (GPU), deep learning technology has made considerable progress. CNN [10] plays an important role in the field of image recognition. Compared with traditional methods of feature extraction, CNN can automatically extract more features than manual methods, and optimize the combination of features. Considering the characteristics of urine sediment image, we proposed a method combined AlexNet with area feature algorithm (AFA) which can quickly and accurately recognize urine sediment images. This method not only saves a lot of workload of artificial feature extraction, but also achieves higher recognition accuracy than traditional methods under the background of complex tasks.

The main contributions of this study are as below:

1) A combined network method based on AlexNet network and AFA is designed. We combine 3 CNNs models with an AFA module, which improves the accuracy of discrimination between certain confusing categories. The combined network achieves 96.75% accuracy and 6.8 ms recognition time for each segmented image.
2) An AFA method based on the area features of RBCs and WBCs is proposed, which combines the output of AFA with the output of CNN in order to solve the defect of CNN weakening cell area feature.
3) A large number of urine sediment images were manually labeled by professionals, thus produced a large-scale urine sediment data set which provided valid support for networks' training. Furthermore, we use transfer learning and data augmentation technology to tackle the problem of images number of some categories in the dataset is too small.

The rest of this paper is arranged as follows: Section 2 introduces the research status of traditional methods and other deep learning methods in urine sediment recognition. Section 3 presents the theory and structure of the method designed in this paper. Section 4 introduces the relevant experiments and results, and compares our results with the results of relevant papers. Finally, the conclusion of this paper is given.

II. RELATED WORK

A. CONVOLUTIONAL NEURAL NETWORK

Recently, deep learning has made great process in many areas, attracting more and more attention. CNN, which works well in the field of image recognition, is an important branch of deep learning and inspired by the cognitive mechanism of biological natural vision. In 1998, Yann LeCun put forward the classic LeNet-5 in [11] aiming at letter recognition in postal departments and achieved good results. Limited by computing power and network depth caused by gradient disappearance, the development of CNN lags behind the traditional image process methods represented by support vector machine (SVM) [12] and BP shallow network [13] for a long time. In 2012, Alex Krizhevsky *et al*. [14] put forward AlexNet which has deeper network level, its innovative training method of adding ReLU activation function and Dropout in the network greatly inhibited the gradient disappearance. And AlexNet won the championship in the ILSVRC competition and set off a research upsurge of CNN. VGG [15], ResNet [16], GoogLeNet [17]–[20] and other networks have deeper and deeper layers, which can cope with increasingly complex recognition tasks thus promotes the development of artificial intelligence technology.

In medical areas, CNN has also gained great progress. It has achieved milestone results in medical image recognition tasks such as diabetic retinopathy [21], skin cancer [22], optical coherence tomography (OCT) [23] and Chest X-rays images [24]. In recent years, Microsoft, Google, and other high-tech companies have also vigorously deployed and invested in research and application in the intelligent medical field. Thus we can see that deep learning is leading the trend of intelligent medicine in the future.

B. AUTOMATIC RECOGNITION OF URINE SEDIMENT IMAGES

The traditional automatic recognition methods, as the main methods of urine sediment images recognition, mainly relied on artificial feature extraction and classification. Sun *et al*. [25] proposed a new detector named aggregate channel features plus (ACF+) detector which is based on aggregate channel features (ACF) for urine sediment detection. Sun *et al*. [26] adopt ACF which are variant and discriminative, combing improved soft-cascade adaboost classifier for RBCs detection in urine sediment micrograph. Zheng *et al*. [27] obtained the feature of urine sediment images according to the chain code method. Jiang *et al*. [28] proposed a method for the segmentation of urine sediment images using the magnification of 20-fold microscopy based on Markov model. They selected sum average feature derived from the spatial gray co-occurrence matrix for the classification in the neighborhood window. Zhou and Zhou [29] proposed an easy-to-implement automatic urine sediment analysis system, which uses SVM classifier to distinguish urine sediment particles; Li *et al*. [30] proposed a method mainly focused on segmentation of urine particles and extraction of their texture features. After segmentation by watershed algorithm and combining Gabor filter with scattering transform, by robust feature description, it can not only keeps the invariance of scaling, rotation and translation, but also shows good performance in SVM classification process. Shen and Zhang [31] used AdaBoost algorithm and SVM algorithm to classify through Harr feature, and used cascade acceleration algorithm to improve the computing speed. Zhou *et al*. [32] combined LHS and D-GDCM with SVM and used them for classification. Avci *et al*. [33] proposed a new adaptive feature extraction theorem based on adaptive discrete wavelet energy method. After image denoising, contrast enhancement and segmentation in the pre-processing stage, features can be automatically extracted and classified by artificial neural network (ANN).

Although the traditional methods have gained some progress, they need a lot of work to find the suitable combination of features and classifier for urine sediment. At the same time, the complex characteristics of urine sediment particles themselves also add a lot of difficulties to recognition. Therefore, extracting features using deep learning methods provides a new way for recognizing urine sediment particles. People begin to use the deep learning methods to extract objects or recognize urine sediment particles. Aziz *et al*. [34] proposed a novel unsupervised method for extracting objects from urine sediment images and applied U-net for extracting these objects. Liang *et al*. [35] proposed feature pyramid network with denseNet (DFPN) method to solve the problem of category confusion in the urine sediment images, which can classify 7 categories. Li *et al*. [36] presented a deep learning method based on the LeNet-5 network for classifying recognisable shapes of urine sample images. Pan *et al*. [37] constructed a CNN model, and achieved 97% accuracy in the classification of 3 urine sediment categories, including RBCs, WBCs, and CAOXs. Kang *et al*. [38], [39] combined Faster RCNN with SSD to classify 7 categories of urine sediment particles. Although these methods have achieved satisfying results, these methods still have disadvantages such as confusion of some different categories in recognition. The automatic recognition of urine sediment images still has a long way to go.

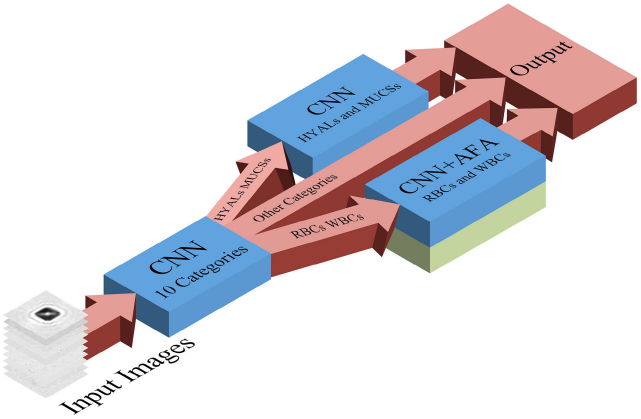


FIGURE 1. Structures of urine sediment images recognition method.

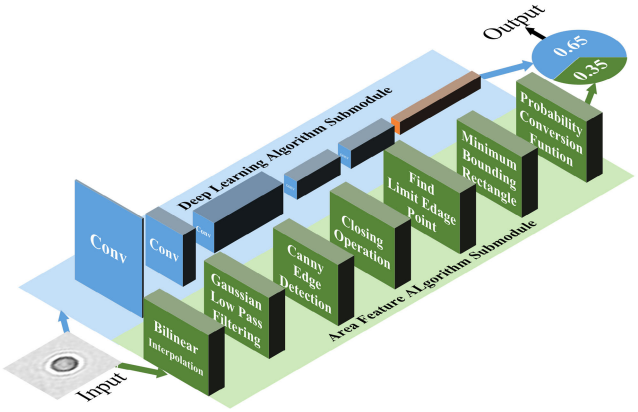


FIGURE 2. RBC-WBC secondary recognition module (as "CNN+AFA" module in Figure 1).

III. METHOD

A. STRUCTURES

Because of the physiological characteristics of the urinary system and the physical characteristics of the microphotographic equipments, the images of urine sediment have the following main characters:

1) Urine sediment particles are unevenly distributed within the capture range of the lens, and recognized urine sediment particles from different urine samples are often different. A large number of blank images without any urine sediment particles will appear in many samples, resulting in the waste of computing resources.
2) The large morphological differences between some particles of the same category, such as yeasts (BYSTs) in different states and HYALs, poses new challenges to the generalization performance of the network.
3) The morphological differences between samples of some different categories are small, such as RBCs and WBCs, HYALs and mucous filaments (MUCSs), etc. And the small size of urine sediment and the poor quality of lens imaging result in the insufficient number of features carried by small particles. It is difficult to classify them.

For characteristics (1), because the time of urine sediment particles segmentation is much shorter than that of traversing convolution and classification, the preprocessing operation of segmentation is used to segment the captured micro-image into several small images, whose edge length varies from tens of pixels to hundreds of pixels, and each small image contains only one particle. For characteristics (2), feature extraction is carried out by using the AlexNet model with an enough depth, and dropout in AlexNet model is also helpful to improve the generalization of the network. For characteristics (3), the urine sediment images recognition method as shown in Figure 1 is designed. This method consists of 3 parts: the main network module, the RBC-WBC secondary recognition module (including the RBC-WBC recognition network sub-module and AFA sub-module, as shown in Figure 2), and the HYAL-MUCS secondary recognition module (including only a HYAL-MUCS secondary recognition network). Firstly, the main network module recognizes the input images with 10 categories. If the images are recognized as RBCs or WBCs, the images will be sent to the corresponding RBC-WBC secondary recognition module for recognizing. If the images are recognized as HYALs or MUCSs, the images will be sent to the corresponding HYAL-MUCS secondary recognition module for recognizing. If the images are recognized as other than these 4 categories, the classification results will be output directly. Finally, the main network module combines 2 secondary recognition modules to output the complete classification results.

B. MAIN NETWORK MODULE

1) NETWORK MODEL

This paper uses AlexNet, which performs well in general classification tasks, as the basic model. Experiments show that although AlexNet has fewer layers than VGG-16, VGG-19, Inception and other networks, it can meet the requirements of accuracy for the classification task of urine sediment images. Besides, AlexNet will bring faster recognition speed and lower computational cost because of its relatively simple structure. In order to further eliminate the gradient descent and improve the effect of feature extraction at the bottom of the network, the pre-trained model is used to initialize the weight of the network by using the transfer learning technology.

2) TRANSFER LEARNING

Generally speaking, the complete training of a network model requires tens of thousands or even larger dataset for training. However, the collection and labeling of urine sediment dataset are very tedious, and some categories of particles occur less frequently. As a result, the existing datasets are not large enough to fit the networks with high accuracy for urine sediment recognition. Therefore, transfer learning technology is used for initializing the weight of the network with a pre-training model. The pre-training model is trained by the open source Image net [40] dataset, consisting of 14.2 million images labeled accurately for 1000 categories. During the training process, all weights in the network are frozen except the full connection layer, which is randomly initialized and updated in training.

3) DATA AUGMENTATION

Because the probability of various particles appeared in urine samples is different, some particles, such as BACTs, BYSTs, etc., are more likely to appear. So the number of them in the dataset is as high as tens of thousands. The probability of some particles appearing is small, such as HYALs, which has only over 100 images, and HYALs usually misclassifies with the MUCSs in the actual classification experiment, and whether they appear or not also has certain clinical significance. Therefore, in order to improve the recognition indexes of the networks, especially to reduce the probability of the misclassification between MUCSs and HYALs, data augmentation was performed in increasing the number of the particles of HYALs.

Common data augmentation options include flip, rotation, stretch, zoom in and out, random cropping, color change, and so on. Through data augmentation, the scale of datasets can be greatly increased, and the generalization of network can be improved. Experiments show that flip and rotation have a greater impact on the improvement of accuracy than other augmentation methods. Therefore, 90 degrees, 180 degrees, 270 degrees, and up-down or left-right flips of HYAL particles are carried out respectively, which is a total of five times the size of the original data set.

C. RBC-WBC SECONDARY RECOGNITION MODULE

When the urine sediment images are recognized as RBCs or WBCs by the main network module, they will be sent to this module for secondary recognition. The schematic diagram of the module structure is shown in Figure 2. The module is divided into 2 parts. The blue area is the sub-module of the RBC-WBC recognition network, and the green area is the sub-module of AFA.

1) RBC-WBC RECOGNITION NETWORK SUB-MODULE

In the RBC-WBC recognition network sub-module, the network structure is the same as the main network module except the number of output softmax nodes in the output layer. The training set is also augmented, and the network weight is initialized by transfer learning in the training process. In the RBC-WBC recognition network sub-module in the figure 2, each blue block represents the convolution part of the network, the orange block represents the last two full connection layers in the network, and the depth of the block represents the number of feature graphs (points). It can be seen that with the deepening of network computing, the number of feature graphs (points) is increasing, and the classification features of RBCs and WBCs can be fully extracted.

According to the morphological characteristics of RBCs and WBCs, the main consideration of classification between them is to extract their internal morphological features and area features. Because of photographic conditions and urine sample differences, the internal features of RBCs and WBCs are often not very obvious, which can result in confusion between RBCs and WBCs. Meanwhile, the CNN needs to normalize the images of different sizes in the recognition process, which leads to the further weakening of the area features of RBCs and WBCs. Therefore, the area features of RBCs and WBCs are considered as an auxiliary, to weighted fuse the output of CNN.

2) AFA SUB-MODULE

Through experiments, we design a set of methods suitable for urine sediment recognition and named them AFA, which can effectively extract the area features of cells. The AFA sub-module mainly includes the following steps:

1) Bilinear interpolation amplification. Firstly, the input image is enlarged 10 times by bilinear interpolation. Compared with size normalization of CNN, the size relationship of each particle is preserved, and the area feature of cells in the image is easier to extract. Since RBCs and WBCs images are mostly distributed between 20×20 and 60×60, it is very likely that images beyond 60×60 are misidentified by the main network module. Therefore, we add constraint on image size to this step to prevent images larger than 60×60 from being input into AFA.
2) Gauss low-pass filtering. The enlarged image is filtered by using 3×3 Gauss kernel to make the cell edge smoother.

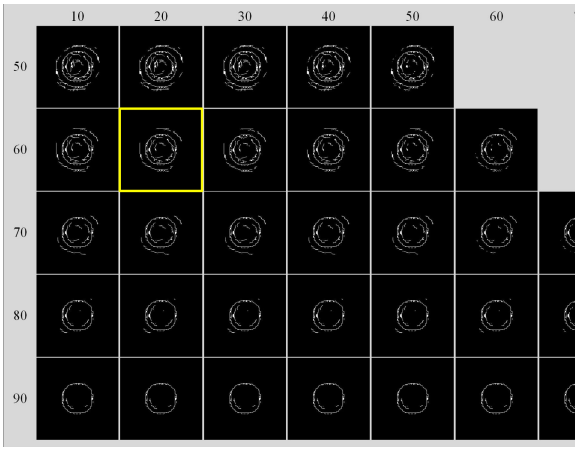


FIGURE 3. Different edge images under different upper and lower threshold limit. The column index represents the lower threshold, the row index represents the upper threshold, and the yellow circle represents the final choice.

3) Edge extraction. Canny edge extraction algorithm is used to extract the edge of the image. Canny edge is formed by applying lag threshold to pre-processed pixels, including two thresholds: upper and lower. A piexl is considered as an edge pixel if its gradient is greater than the upper threshold, and discarded if it is lower than the lower threshold. Figure 3 shows the different edge images under different upper and lower threshold limit. Considering the integrity of the main contour and the reduction of noise points, and combined with the general experience that the ratio of upper and lower threshold is about 3:1 in Canny detector, upper threshold of 60 and lower threshold of 20 are selected as thresholds.
4) Closed operation. The image was expanded first and then corroded by using 3×3 kernel to enhance the connectivity of cell edges.
5) Finding the limit point of edge. Find out the outermost pixels in the 4 directions (up, down, left, and right) of each cell edge, named as the limit point of edge.
6) Drawing the bounding rectangle of cells. 4 limit points of edge are used to draw the bounding rectangle parallel to the edge of the image to represent the area feature of the cells in the image, and then output the area of the rectangle. Figure 4 shows the process of extracting the bounding rectangle of RBCs and WBCs using AFA. It can be seen that for distorted or damaged cells, the general method of minimum bounding rectangle provided by OpenCV is not satisfactory. Because the edge is discrete, its rectangle is replaced by a series of small rectangles. Therefore, the extracted areas can not represent the real feature of cells. Meanwhile, the AFA sub-module can achieve good results.
7) Probability conversion function. In order to convert the output area of bounding rectangle into a probability value weighted with the softmax output value of CNN, referring to softmax the probability conversion function

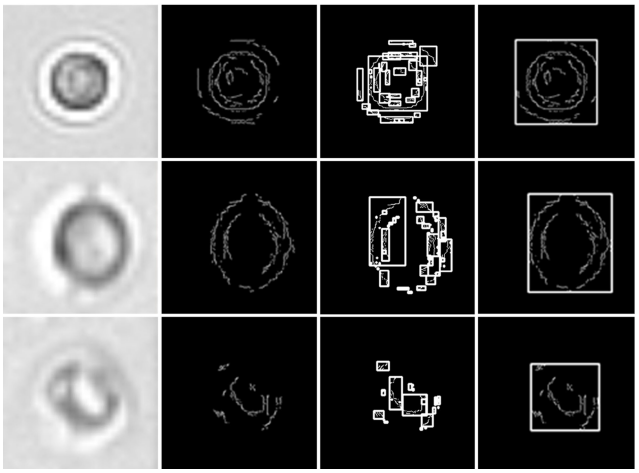


FIGURE 4. The process of generating the bounding rectangle, from left to right, is the bilinear interpolation enlarged image, the edge image processed by Canny algorithm, the minimum bounding rectangle using OpenCV (as a comparison), and the final generated bounding rectangle by our method.

is designed as follows:

$$Prob_{\mathrm{WBC}} = \frac{1}{1 + e^{-(\frac{Area - Threshold}{1000})}} \tag{1}$$

$$Prob_{\mathrm{RBC}} = 1 - \frac{1}{1 + e^{-(\frac{Area - Threshold}{1000})}} \tag{2}$$

where *Area* is the area of the bounding rectangle, and *Threshold* is the area threshold, which is set to 1200 by comparing the experiments. This function has the following good characteristics: firstly, its range of value is between 0 and 1, which is consistent with the output of probability value; moreover, its domain of definition can cover all areas, and when special cells appear, it will not throw out exceptions or errors; at the same time, near the threshold, its probability value is 0.5, which is the same as the probability of RBCs and WBCs, and conforms to the actual situation; The absolute value of the derivative near threshold is larger, and the absolute value of the derivative far from the threshold decreases gradually, which can enlarge the area feature of cells. The above characteristics can effectively improve the ability of the algorithm to extract cell area feature.

After the recognition of the RBC-WBC recognition network sub-module and AFA sub-module, the output probability of the two sub-modules is assigned to 0.65 and 0.35 weights, respectively, as the final output. Namely:

$$Output_{\mathrm{WBC}} = 0.65 \times Softmax_{\mathrm{CNN,WBC}} + 0.35 \times Prob_{\mathrm{AFA,WBC}} \tag{3}$$

$$Output_{\mathrm{RBC}} = 0.65 \times Softmax_{\mathrm{CNN,RBC}} + 0.35 \times Prob_{\mathrm{AFA,RBC}} \tag{4}$$

D. HYAL-MUCS SECONDARY RECOGNITION MODULE

The HYAL-MUCS secondary recognition module is composed of a 2-category network model. Its structure is identical

TABLE 1. Recognition results of main network module.

		Predict										Sen
		BACT	BYST	CAOX	HYAL	MUCS	RBC	SPRM	SQEP	WBC	WBCC	
Original	BACT	21313	9	0	0	17	0	3	0	1	0	1.00
	BYST	19	4303	55	0	3	19	0	35	53	50	0.95
	CAOX	2	22	3068	0	0	59	1	0	41	3	0.96
	HYAL	0	0	0	142	27	0	0	3	0	0	0.83
	MUCS	2	7	1	35	5310	0	0	45	0	0	0.98
	RBC	0	2	21	0	0	4448	0	0	621	0	0.87
	SPRM	7	2	1	0	2	0	173	0	0	0	0.94
	SQEP	1	30	0	18	50	0	0	5744	0	20	0.98
	WBC	5	41	19	0	0	161	1	0	6259	4	0.96
	WBCC	0	77	3	0	3	0	0	10	13	1034	0.91

to that of the RBC-WBC recognition network sub-module of RBC-WBC secondary recognition module.

IV. EXPERIMENTS

A. DATASET

The quality of datasets has a great influence on the recognition accuracy of the network. In order to make a dataset of image quality and quantity up to standard, hundreds of urine samples of patients were collected by micro-photography equipment. 650 high-definition images were collected from each sample and segmented. The magnification of microscopic equipment is 20 times, so the relationship of size between the particles in the image also indicates the relationship of size between the particles in practice. After that, it was labeled manually by a professional medical team, among which 10 categories of urine sediment with high quantity and certain medical significance were selected. They are bacterias (BACTs), BYSTs, CAOXs, HYALs, MUCSs, RBCs, sperms (SPRMs), squamous epithelial cells (SQEPs), WBCs, and white blood cell clusters (WBCCs).

Subsequently, according to the ratio of 4:1, images are randomly selected from the dataset and grouped into training set and test set. The training set is responsible for training the network, and the test set is responsible for evaluating the training effect of the network. Figure 5 shows the number distribution of images in the dataset.

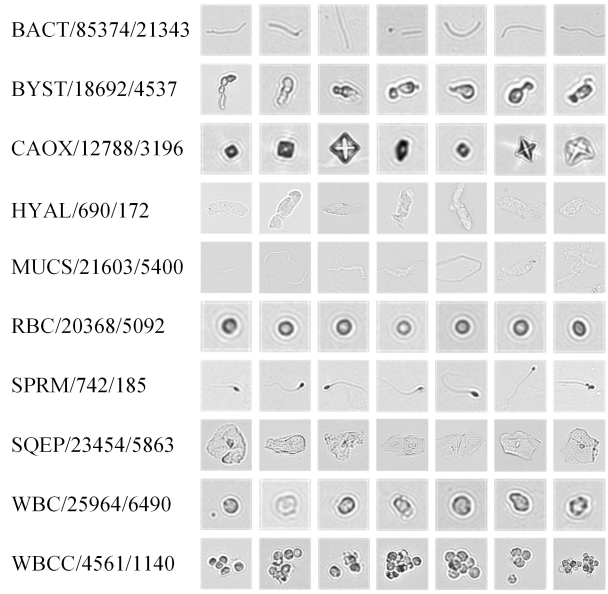


FIGURE 5. Urine sediment dataset. The left column represents the labels of each category, the number of training sets and the number of test sets.

B. EXPERIMENTS AND ANALYSIS

The training and testing of models were carried out on the CPU of Intel Core i7-7700K and GPU of Nvidia Quadro P2000. The network was built using the Tensorflow framework of Google. The learning rate was 0.001 and the epoch iteration was 20 times. Table 1 shows the test result confusion matrix of the main network module. As can be seen in the Table 1, there are many images of HYALs misclassified as MUCSs, and there are also many images of RBCs misclassified as WBCs.

In order to solve the misclassification between MUCSs and HYALs, RBCs and WBCs, the RBC-WBC recognition network sub-module in the RBC-WBC secondary recognition module and HYAL-MUCS secondary recognition module were further trained. Compared with the 10-category model, the 2-category model has a relatively simple recognition task and a higher accuracy of classification. According to the method of Section 3, the test set is used to test the recognition method consisting of the main network module, the RBC-WBC secondary recognition module and the HYAL-MUCS secondary recognition module. The confusion matrix is represented by Table 2 and the receiver operating characteristic (ROC) curve is shown by Figure 6. It can be seen that the combination network model is suitable for recognition of most categories of urine sediment particles.

Sensitivity, also called recall rate, is an important index to evaluate the performance of recognition tasks, which is

TABLE 2. Recognition results of the method combining main network module and 2 secondary recognition modules.

		Predict										Sen
		BACT	BYST	CAOX	HYAL	MUCS	RBC	SPRM	SQEP	WBC	WBCC	
Original	BACT	21313	9	0	1	16	0	3	0	1	0	1.00
	BYST	19	4303	55	0	4	22	0	35	50	50	0.95
	CAOX	2	22	3068	0	0	64	1	0	36	3	0.96
	HYAL	0	0	0	152	17	0	0	3	0	0	0.88
	MUCS	2	7	1	102	5243	0	0	45	0	0	0.97
	RBC	0	2	21	0	0	4552	0	0	517	0	0.89
	SPRM	7	2	1	0	2	0	173	0	0	0	0.94
	SQEP	1	30	0	23	45	0	0	5744	0	20	0.98
	WBC	5	41	19	0	0	319	1	0	6101	4	0.94
	WBCC	0	77	3	1	2	0	0	10	13	1034	0.91

TABLE 3. Comparison of recognition indicators of different categories by different methods.

Methods	RBCs Sensitivity	WBCs Sensitivity	HYALs Sensitivity	MUCSs Sensitivity
Main Network Module	87.35%	96.44%	82.56%	98.33%
Combination Network	89.40%	94.01%	88.37%	97.09

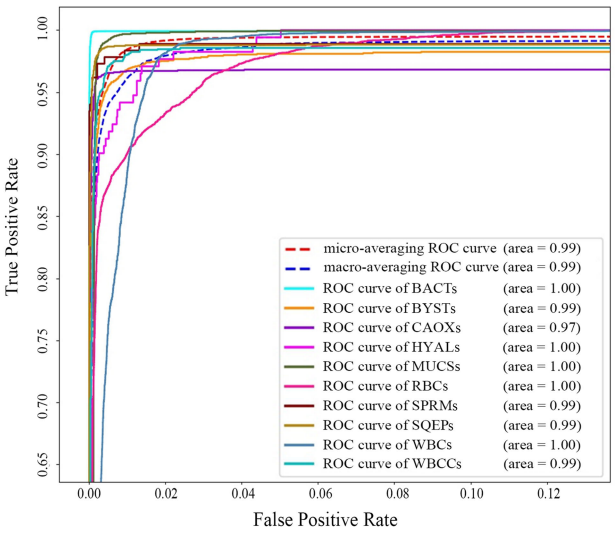


FIGURE 6. 10-category ROC curves of combination network.

defined as:

$$R = \frac{TP}{TP + FN} \tag{5}$$

where TP is true positive, and FN is false negative. Table 3 summarizes the changes of the sensitivity of RBCs, WBCs, HYALs, and MUCSs. It can be seen that the sensitivity of RBCs has increased from 87.35% to 89.40% and the sensitivity of HYALs has increased from 82.56% to 88.37% through this method. Although the sensitivity of WBCs and MUCSs decreased correspondingly, but its indicators are more balanced and performance is more stable. What we have done is to find a balance between sensitivity of different categories, sacrificing a small amount of sensitivity of WBCs and MUCSs (they are high enough) in exchange for greater sensitivity gain of RBCs and HYALs, so as to compensate for the less ideal recognition effect of RBCs and HYALs in the initial recognition of main network module.

Macro-averaging and micro-averaging are important indicators for evaluating multi-category recognition models. Macro-averaging refers to calculating index value, and then calculating the arithmetic average value for all categories. Assuming that the recognition task has n categories, the macro-average formula is as follows:

$$Macro_F = \frac{2 \times Macro_P \times Macro_R}{Macro_P + Macro_R} \tag{6}$$

and the definitions of *Macro* are as follow:

$$Macro_R = \frac{1}{n} \sum_{i=1}^{n} R_i \tag{7}$$

$$Macro_P = \frac{1}{n} \sum_{i=1}^{n} P_i \tag{8}$$

P is the precision, which is defined as:

$$P = \frac{TP}{TP + FP} \tag{9}$$

where FP is false positive.

Macro-averaging is to establish a global confusion matrix for each sample in the dataset without classification, and then calculate the corresponding indicators. The formula is as follows:

$$Micro_F = \frac{2 \times Micro_P \times Micro_R}{Micro_P + Micro_R} \tag{10}$$

TABLE 4. Comparison of accuracy and average time-consuming of methods from different papers.

Number	Methods	Dataset and categories	Accuracy	Average time-consuming per image
1	Our Method	10 categories of RBCs, WBCs, BACTs, etc.	96.75%	61ms
2 [37]	CNN	3 categories of RBCs, WBCs, and CAOXs	98%	
3 [36]	CNN (LeNet-5)	4 categories of RBCs, WBCs, Crystals, etc.	92.3%	
4 [35]	CNN (DFPN)	7 categories of RBCs, WBCs, Crystals, etc.	91.4%	201ms
5 [30]	Gabor+Scattering Transformation+SVM	3 categories of RBCs, WBCs, and Crystals	90%	
6 [31]	AdaBoost+SVM	6 categories of RBCs, WBCs, Casts, etc.	90%	
7 [38]	Faster RCNN+SSD	7 categories of RBCs, WBCs, Casts, etc.	84.1%	72ms

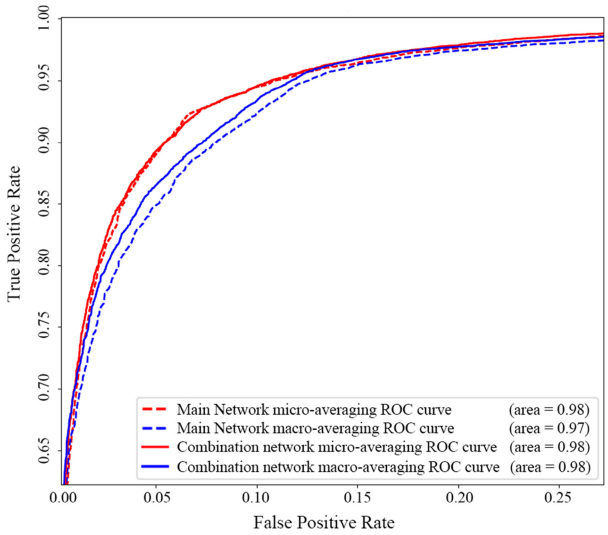


FIGURE 7. Comparison of Macro-averaging (blue) and Micro-averaging (red) between the main network module and the combination network (local), dotted line as the main network module, solid line as the combination network.

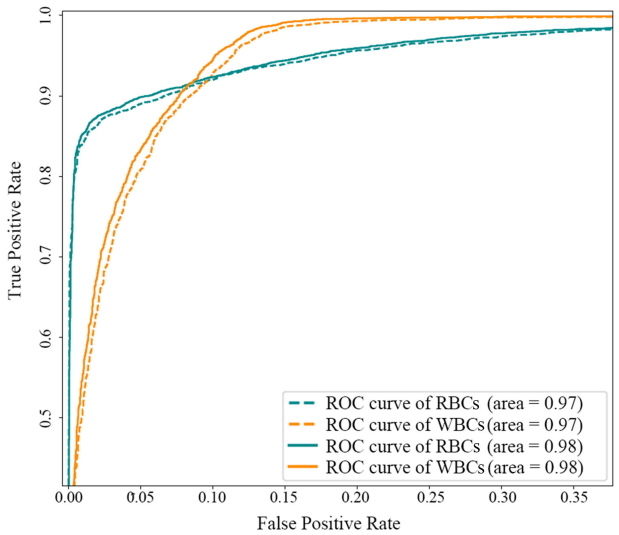


FIGURE 8. Comparison of RBCs (green) and WBCs (yellow) on the main network module (dotted line) and combination network (solid line) respectively.

and the definitions of *Micro* are as follow:

$$Micro_P = \frac{\sum_{i=1}^{n} TP_i}{\sum_{i=1}^{n} TP_i + \sum_{i=1}^{n} FP_i} \tag{11}$$

$$Micro_R = \frac{\sum_{i=1}^{n} TP_i}{\sum_{i=1}^{n} TP_i + \sum_{i=1}^{n} FN_i} \tag{12}$$

Compared with the main network module, the combination network also achieves good performance on the micro-averaging and macro-averaging of the 10-category tasks. From Figure 7, it can be seen that the combination network is superior to the main network module in most areas. In macro-averaging, the combination network is superior to the main network module in all areas.

Figure 8 separately shows the ROC curves of RBCs and WBCs in the main network module and the combination network respectively. From the Figure 8, we can see that the RBCs ROC curve of combination network is better than that of the main network module in most areas, while the WBCs ROC curve of combination network is better than that of the main network module in all areas, and the increase is obvious. It can be seen that the method proposed in this paper has a significant effect on improving the sensitivity of confusing categories in the task of urine sediment recognition.

Compared with other papers, the method proposed in this paper has obvious advantages, as shown in Table 4. It can be seen that the accuracy of the proposed method is slightly lower than that of the CNN method proposed in [37], and higher than that of the traditional method proposed in [30], [31] and CNN mothods proposed in [35], [36], [38]. The recognition task of this paper is 10 categories, and the methods proposed in [37] only have 3 categories. Compared with other methods, this method has the most categories, the highest complexity and the greater difficulty in recognition. It has more practical clinical application value. Under these conditions, the main reasons for the high accuracy of our method are as follows:

1) As mentioned above, this paper proposes a combination network structure, and improves the performance of the network by means of transfer learning and data augmentation. Because the 2-category network model receives less interference from other categories and has stronger pertinence than the 10-category network model, the secondary recognition module is used to correct the recognition results of the main network module.

2) The resize mechanism of CNN weakens or even eliminates the size features of cells. At the same time, CNN itself also has some size invariance. The AFA designed in this paper can restore the size features of cells and integrate them into the recognition results of CNN. In other words, CNN is used to extract the internal details of cells, and AFA is used to extract the area features of cells.
3) The general image processing algorithm has some limitations in dealing with damaged cells, so this paper improves the minimum bounding rectangle to better adapt to the urine sediment particles.
4) The urine sediment dataset produced and used in this paper contains more than 300,000 hand-labeled images, which is also the largest urine sediment in the current public papers. The high quality dataset directly affect the final test results of the network.

In terms of recognition time, the average recognition time of the main network module is 6.5 ms per image, and the average recognition time of the combination network is 6.8 ms per image. At the same time, the average recognition time of single frame containing multiple particles is 61 ms, which is faster than 72 ms per frame in [38] and 201 ms per frame in [35]. At the same time, the cost of this method is lower. The GPU price of TITAN X used in [38] and TITAN XP used in [35] is three times that of Quadro P2000 used in this paper.

V. CONCLUSION

Urine sediment microscopy detection is one of the main detection items in urology and other departments. It is of great significance for the diagnosis of related diseases. In order to solve the shortcomings of traditional recognition methods, such as heavy workload in the process of algorithm designed, confusion of some different categories in recognition, and low accuracy in recognition, this paper designs a urine sediment image recognition method based on deep learning. The method uses CNN to build a recognition model based on AlexNet. A 10-category network model is combined with 2 secondary recognition modules, and an AFA sub-module is proposed in the RBC-WBC secondary recognition module, which effectively solves the confusion between RBCs and WBCs, HYALs and MUCSs images. The method uses dataset containing 300,000 images to train the CNNs. The accuracy of the test set is 97% and the average recognition time is about 6 ms. In addition, the CNN has a strong expansibility, which can introduce more categories for recognition in practical application.

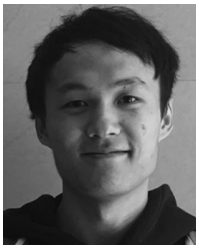
REFERENCES

[1] J. Lakatos, T. Bodor, Z. Zidarics, and J. Nagy, "Data processing of digital recordings of microscopic examination of urinary sediment," *Clin. Chim. Acta*, vol. 297, nos. 1–2, pp. 225–237, 2000.
[2] T. Kouri, G. Fogazzi, V. Gant, H. Hallander, W. Hofmann, and W. G. Guder, "European urinalysis guidelines," *Scand. J. Clin. Lab. Invest.*, vol. 60, no. 231, pp. 1–96, 2000.
[3] R. Davis, J. S. Jones, D. A. Barocas, E. P. Castle, E. K. Lang, R. J. Leveillee, E. M. Messing, S. D. Miller, A. C. Peterson, and T. M. T. Turk, "Diagnosis, evaluation and follow-up of asymptomatic microhematuria (AMH) in adults: AUA guideline," *J. Urol.*, vol. 188, no. 6S, pp. 2473–2481, Dec. 2012.
[4] F. D. İnce, H. Y. Ellidag, M. Koseoglu, N. Simsek, H. Yalçin, and M. O. Zengin, "The comparison of automated urine analyzers with manual microscopic examination for urinalysis automated urine analyzers and manual urinalysis," *Practical Lab. Med.*, vol. 5, pp. 14–20, Aug. 2016.
[5] A. M. Perazella, "The urine sediment as a biomarker of kidney disease," *Amer. J. Kidney Diseases*, vol. 66, no. 5, pp. 748–755, Nov. 2015.
[6] F. H. Deindoerfer, J. R. Gangwer, C. W. Laird, and R. R. Ringold, "'The Yellow IRIS,' urinalysis workstation–the first commercial application of 'automated intelligent microscopy,'" *Clin. Chem.*, vol. 31, no. 9, pp. 1491–1499, 1985.
[7] M. Ranzato, P. E. Taylor, J. M. House, R. C. Flagan, Y. LeCun, and P. Perona, "Automatic recognition of biological particles in microscopic images," *Pattern Recognit. Lett.*, vol. 28, no. 1, pp. 31–39, Jan. 2007.
[8] Y. Liang, B. Fang, J. Qian, L. Chen, C. Li, and Y. Liu, "False positive reduction in urinary particle recognition," *Expert Syst. Appl.*, vol. 36, no. 9, pp. 11429–11438, Nov. 2009.
[9] M. D. Almadhoun and A. El-Halees, "Automated recognition of urinary microscopic solid particles," *J. Med. Eng. Technol.*, vol. 38, no. 2, pp. 104–110, 2014.
[10] L. O. Chua and T. Roska, "The CNN paradigm," *IEEE Trans. Circuits Syst. I, Fundam. Theory Appl.*, vol. 40, no. 3, pp. 147–156, Mar. 1993.
[11] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
[12] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
[13] R. Hecht-Nielsen, "Theory of the backpropagation neural network," in *Neural Networks for Perception*. Amsterdam, The Netherlands: Elsevier, 1992, pp. 65–93.
[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, Jun. 2012, pp. 1097–1105.
[15] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," Apr. 2015, *arXiv:1409.1556*. [Online]. Available: https://arxiv.org/abs/1409.1556
[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 770–778.
[17] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Boston, MA, USA, Jun. 2015, pp. 1–9.
[18] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," Feb. 2015, *arXiv:1502.03167*. [Online]. Available: https://arxiv.org/abs/1502.03167
[19] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Seattle, WA, USA, Jun. 2016, pp. 2818–2826.
[20] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. A. Alemi, "Inception-v4, inception-resnet and the impact of residual connections on learning," in *Proc. 31st AAAI Conf. Artif. Intell.*, San Francisco, CA, USA, Feb. 2017, pp. 1–7.
[21] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, and J. Cuadros, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *J. Amer. Med. Assoc.*, vol. 316, no. 22, pp. 2402–2410, Dec. 2016.
[22] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Jan. 2017.
[23] D. S. Kermany, M. Goldbaum, W. Cai, C. C. Valentim, H, Liang, S. L. Baxter, A. McKeown, G. Yang, X. Wu, and F. Yan, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, Feb. 2018.

[24] P. Rajpurkar, J. Irvin, K. Zhu, B. Yang, H. Mehta, T. Duan, D. Ding, A. Bagul, C. Langlotz, and K. Shpanskaya, M. P. Lungren, and A. Y. Ng, "CheXNet: Radiologist-level pneumonia detection on chest X-rays with deep learning," Dec. 2017, *arXiv:1711.05225*. [Online]. Available: https://arxiv.org/abs/1711.05225
[25] Q. Sun, S. Yang, C. Sun, and W. Yang, "Exploiting aggregate channel features for urine sediment detection," *Multimedia Tools Appl.*, vol. 78, no. 17, pp. 23883–23895, Sep. 2019.
[26] Q. Sun, S. Yang, C. Sun, and W. Yang, "An automatic method for red blood cells detection in urine sediment micrograph," in *Proc. 33rd Youth Academic Annu. Conf. Chin. Assoc. Automat. (YAC)*, Nanjing, China, May 2018, pp. 241–245.
[27] X.-H. Zheng, X.-M. Zhou, and M.-J. Zheng, "Research on urinary micro-particles image feature extraction and analysis method," in *Proc. Int. Conf. Image, Video Signal Process.*, Shanghai, China, Feb. 2019, pp. 63–67.
[28] X. Jiang, F. Chen, Q. Chen, M. Si, and W. Wang, "Texture segmentation of urinary sediment image based on a weighted Gaussian mixture model with Markov random fields," in *Proc. 7th Int. Conf. Bioinf. Biomed. Sci.*, Shenzhen, China, Jun. 2018, pp. 82–87.
[29] Y. Zhou and H. Zhou, "Automatic classification and recognition of particles in urinary sediment images," in *Computer, Informatics, Cybernetics and Applications*. Chongqing, China: Springer, Nov. 2012, pp. 1071–1078.
[30] C. Li, Y. Y. Tang, H. Luo, and X. Zheng, "Join Gabor and scattering transform for urine sediment particle texture analysis," in *Proc. IEEE 2nd Int. Conf. Cybern. (CYBCONF)*, Gdynia, Poland, Jun. 2015, pp. 410–415.
[31] M.-L. Shen and R. Zhang, "Urine sediment recognition method based on SVM and AdaBoost," in *Proc. Int. Conf. Comput. Intell. Softw. Eng. (CiSE)*, Wuhan, China, Dec. 2009, pp. 1–4.
[32] X. Zhou, X. Xiao, and C. Ma, "A study of automatic recognition and counting system of urine-sediment visual components," in *Proc. 3rd Int. Conf. Biomed. Eng. Inform.* Yantai, China: Yantai Univ., vols. 1–7, Oct. 2010, pp. 78–81.
[33] D. Avci, M. K. Leblebicioglu, M. Poyraz, and E. Dogantekin, "A new method based on adaptive discrete wavelet entropy energy and neural network classifier (ADWEENN) for recognition of urine cells from microscopic images independent of rotation and scaling," *J. Med. Syst.*, vol. 38, no. 2, p. 7, Feb. 2014.
[34] A. Aziz, H. Pande, B. Cheluvaraju, and T. R. Dastidar, "Improved extraction of objects from urine microscopy images with unsupervised thresholding and supervised U-Net techniques," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Salt Lake City, UT, USA, Jun. 2018, pp. 2230–2238.
[35] Y. Liang, Z. Tang, M. Yan, and J. Liu, "Object detection based on deep learning for urine sediment examination," *Biocybern. Biomed. Eng.*, vol. 38, no. 3, pp. 661–670, 2018.
[36] T. Li, D. Jin, C. Du, X. Cao, H. Chen, J. Yan, N. Chen, Z. Chen, Z. Feng, and S. Liu, "The image-based analysis and classification of urine sediments using a LeNet-5 neural network," *Comput. Methods Biomech. Biomed. Eng., Imag. Vis.*, pp. 1–6, May 2019. [Online]. Available: https://www_tandfonline.xilesou.top/doi/abs/10.1080/21681163.2019.1608307
[37] J. Pan, C. Jiang, and T. Zhu, "Classification of urine sediment based on convolution neural network," *AIP Conf.*, vol. 1955, Apr. 2018, Art. no. 040176.
[38] R. Kang, Y. Liang, C. Lian, and Y. Mao, "CNN-based automatic urinary particles recognition," Mar. 2018, *arXiv:1803.02699*. [Online]. Available: https://arxiv.org/abs/1803.02699
[39] Y. Liang, R. Kang, C. Lian, and Y. Mao, "An end-to-end system for automatic urinary particle recognition with convolutional neural network," *J. Med. Syst.*, vol. 42, no. 9, p. 165, 2018.
[40] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A large-scale hierarchical image database," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Miami, FL, USA, Jun. 2009, pp. 248–255.

QINGBO JI was born in Harbin, Heilongjiang, China, in 1975. She received the B.S., M.S., and Ph.D. degrees in communication engineering from Harbin Engineering University, China, in 1998, 2004, and 2008, respectively.

She is currently an Associate Professor with the College of Information and Communication Engineering, Harbin Engineering University. Her major research interests include image processing, image recognition, deep learning, and embedded systems.



XUN LI was born in Jinan, Shandong, China, in 1995. He received the bachelor's degree in communication engineering from Northeastern University, Qinhuangdao, China, in 2017. He is currently pursuing the degree in electronic and communication engineering with the College of Information and Communication Engineering, Harbin Engineering University.

ZHIYU QU was born in Changchun, Jilin, China, in 1983. She received the Ph.D. degree in communication and information system from Harbin Engineering University, China, 2008. In 2008, she joined the Nanjing Research Institute of Electronic Technology, where she has been studying radar signal processing and passive target tracking. She is currently an Associate Professor with Harbin Engineering University. Her current research interests include wide band signals detection, high precision passive direction finding, and spatial spectrum estimation.

CHONG DAI was born in Yingcheng, Xiaogan, Hubei, China, in 1994. He received the B.S. degree in communication engineering from the Heilongjiang University of Science and Technology, China, in 2017. He is currently pursuing the master's degree with the College of Information and Communication Engineering, Harbin Engineering University.