IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

SPECIAL SECTION ON MISSION CRITICAL SENSORS AND SENSOR NETWORKS (MC-SSN)

Received July 10, 2018, accepted August 8, 2018, date of publication August 23, 2018, date of current version October 8, 2018.

Digital Object Identifier 10.1109/ACCESS.2018.2865541

SO-YOLO Based WBC Detection With Fourier Ptychographic Microscopy

XING WANG[1], TINGFA XU[1,2], JIZHOU ZHANG[1], SINING CHEN[1], AND YIZHOU ZHANG.[1]

[1]Image Engineering and Video Technology Laboratory, School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China
[2]Key Laboratory of Photoelectronic Imaging Technology and System, Ministry of Education of China, Beijing 100081, China

Corresponding author: Tingfa Xu (ciom_xtf1@bit.edu.cn)

This work was supported in part by the Major Science Instrument Program of the National Natural Science Foundation of China under Grant 61527802 and in part by the General Program of National Natural Science Foundation of China under Grant 61371132 and Grant 61471043.

ABSTRACT White blood cells (WBC) play a significant role in human immune system, so WBC detection is a meaningful work. In this paper, we propose a novel framework which combines Fourier ptychographic microscopy (FPM) and SO-you only look once (YOLO) for WCB detection. FPM is a recently developed microscope technology which can achieve high-resolution, wide-field, and quantitative phase imaging at the same time without mechanical moving and phase measurement. With the FPM, we can get high resolution, wide field-of-view blood cell images at one time. After obtaining high-resolution images, we propose SO-YOLO as our detection method. As a convolutional neural network, the SO-YOLO outperforms the state-of-the-art detection methods both in accuracy and speed. The proposed framework in this paper can be used in clinical diagnose, because the experiment setup is flexible and the detection is fast. The data set in this paper is made with FPM by ourselves. Our data set contains 1000 color high resolution images and all the WBC has been detected and counted by a human expert. The detection results by the experts are regarded as the ground truth. The experiment results show that the SO-YOLO performs well in detecting small objects compared with other methods.

INDEX TERMS White blood cells, detection, Fourier ptychographic microscopy, convolutional neural network, SO-YOLO, small objects.

I. INTRODUCTION

Imaging has become an essential component in many fields of bio-medical research and clinical practice. Numerous modalities of medical image such as PET, MRI, CT, or microscopy make quantitative analysis possible, which can help doctors in researching, diagnosing, monitoring, and treating medical disorders. In medical field, the analysis of white blood cells(WBC) is of vital importance for diagnosing diseases, for WBC play a significant role in body's immune system [1]. Usually, the proportion of WBC reveals important diagnostic information which can help doctors make correct treatment to their patients and observe curative effect. Therefore the detection of WBC is a meaningful task [2]. However, WBC in different stages of maturation have different shape, texture and density, so the detection of WBC remains a challenging problem [3]. In the past, the detection of WBC was completed by hematology experts, which is unfortunately a tedious, time consuming, subjective and error-prone work. Nowadays, with the development of image processing technology, WBC automatic detection is achieved and the detection is becoming better and faster.

No matter what method is adopted for WBC detection, the basic step is the acquisition of WBC images [4]. Usually, in order to get clear cell micrograph, we use conventional microscope with 20× objective lens. However, the field-of-view(FOV) of 20× objective lens is so small that mechanical scanning of a specimen for detecting and counting an adequate number of WBC is necessary. For a blood smell, the mechanical moving is unfavorable to the accuracy of WBC detection because it's so subjective and error-prone. In addition, when capturing images, the blood smells are not always uniform in thickness, which leads to the out-of-focus images in the thick area. So we need to repeat the image acquisition process after properly focusing the objective, which is time-consuming and boring. The above problems imply that there exist many drawbacks in the acquisition of blood cell images with standard microscope.

Here, we introduce Fourier Ptychographic Microscopy (FPM) [5] as a solution that can figure out the problems existing objectively in the acquisition of blood cell images. What's more, there are other advantages about FPM. FPM is a recently developed computational technique [6] which can increase the numerical aperture(NA) of a microscope system to acquire a high resolution, wide FOV image with a sequence of low-resolution images with no need of mechanical scanning. FPM requires minor modifications to conventional microscopy setup, only replacing the common light source with a programmable light emitting diode(LED) array and adding a Charge Coupled Device(CCD) camera. By sequentially turning on single LED element in the matrix, the thin specimen is illuminated with oblique plane waves from different angles so that we can capture the corresponding low-resolution images. Since the illumination source can be regarded as plane waves, the spectrum of the specimen is on the back focal plane of the objective lens. Different illuminating angles correspond to different positions in fourier domain. Though the finite numerical aperture (NA) of the objective lens acts as a low-pass filter which limits the resolution of the system, the spectrum of the specimen illuminated by LED with different angles can carry some frequency components [7]which are beyond the NA of the objective lens. By iteratively stitching them together in fourier domain, a high resolution, wide FOV image can be got. Besides, the depth of focus of the microscope is extended from $80\mu m$ to $300\mu m$ with FPM procedure, which provides a large tolerance to the unevenness of blood smells. Due to the modification to the conventional microscope is minor, FPM offers a flexible and low-cost approach compared to the whole slide imaging methods, which usually involve an expensive precision mechanical stage.

With so many advantages, FPM has been applied in many biological fields such as hematology [8], pathology [9], [10], etc and non-biological fields like quantitive phase imaging [11] within a few years. There are also many improvements of FPM lately in imaging performance [12]–[14], implementation methods [15]–[17] and imaging mode [18]–[20]. These applications and modifications show the great potential of FPM in biomedical observation and clinical diagnosis.

After obtaining high resolution, wide FOV blood cell images, the detection of WBC becomes easier. In the past years, several work has been done in the field of WBC detection. Shitong and Min [4] proposed a detection algorithm based on Fuzzy Cellular Neural Networks(FNCC) for WBC detection. The detection algorithm combines the advantages of threshold segmentation followed by mathematical morphology and the fuzzy logic method. It detects successfully only one white blood cell in an image, but for images which contain several white blood cells it doesn't show the detection results. What's more, the detection results are influenced by the definition of iteration number, so how to decide the proper iteration number is a challenging problem. By using a different approach, [21] considers the automatic detection of WBC as a multi-ellipse detection problem. The approach firstly segments the WBC using the DEM algorithm, then it gets the edge map from the segmented image. After that, it starts the ellipse detector based on DE over the edge map while saving best ellipse. Finally, by defining parameter values for each ellipse, the WBC can be identified. This approach can detect several WBC in one image, but it doesn't have robustness for large images. WBC in large images can't be well detected, for the background is complex. As we all know, the segment of WBC is an important step for WBC detection. In [22], an automatic segmentation technique of white blood cells nuclei is proposed for the detection and counter of WBC. The technique is based on gray scale contrast enhancement and filtering, in which minimum segment size is implemented to remove false objects. This method performs well without the condition that the nucleus of WBC is lobulated. This disadvantage limits the detection accuracy greatly and leads to incorrect count of WBC's total number.

Nowadays, the vast emergence of convolutional neural network (CNN) has created an impressive performance in object detection and classification [23], [24]. Compared to the detection methods without CNN, there are obvious improvements in accuracy and robustness. The region proposal network with convolutional neural network(RCNN) was first proposed by Girshick *et al*. [25] for object detection. For an image, 2000 region of proposals are extracted with selective search in RCNN. Then it needs to perform a forward CNN to realize feature extraction for each region of proposal obtained from the selective search, which makes R-CNN computationally expensive. Then the author proposed Fast-RCNN [26], which is many times faster than RCNN both in training and testing. A special layer named ROI layer is proposed in Fast-RCNN, which can enhance the performance. However, Fast-RCNN still uses selective search to extract region of proposals in an image, which will limit the processing speed. In [27] Ren *et al*. proposed Faster-RCNN , which replaced the selective search with Region Proposal Network(RPN) for proposals abstraction. So the processing speed is improved.

Faster-RCNN has achieved a good detection result. However, due to the 2-stage detection deep learning models that it belongs to, the detection speed is hard to further increase. Redmon *et al.* come up with a totally new idea, named You Only Look Once (YOLO) [28], which belongs to 1-stage detection models. As a groundbreaking method which describes detection as a unified, end-to-end regression problem in the field of object detection, YOLO is named after processing a whole image at one time to obtain the position and classification simultaneously. The training and testing of YOLO are all conducted in a single network, so the optimization can be carried out end-to-end directly on detection performance. Here we propose a CNN-based SO-YOLO framework for the detection of WBC, which are small objects. The SO-YOLO framework is on the base of YOLOv2, which is the improved model of YOLO. By using a novel, multi-scale training method with prior work in YOLO, YOLOv2 outperforms state-of-the-art methods like Faster

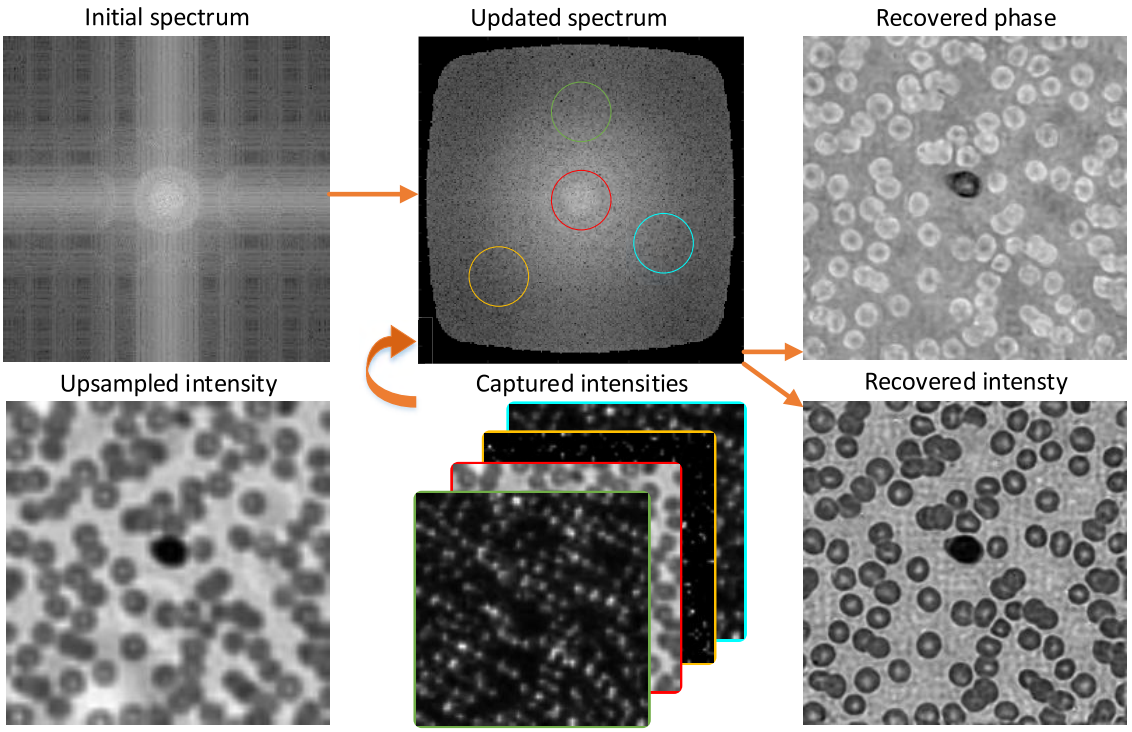


FIGURE 1. The reconstruction procedure of FPM.

R-CNN and SSD while still running faster. Our network, SO-YOLO, not only keeps YOLOv2's advantage of fast detection speed but also has high accuracy for the detection of WBC, which are small objects in high-resolution images that are reconstructed by FPM.

In this paper, we propose the combination of FPM and SO-YOLO for WBC detection. The overall framework is shown in Figure 3. FPM is an excellent method for lab research and clinical detection because of its low-cost and simple experimental equipment. Then we conduct WBC detection on base of the reconstruction results of FPM with the high accuracy and speed method, SO-YOLO. The experiment results verified that the combination of FPM and SO-YOLO for WBC detection is of good effectiveness.

The rest of the paper is organized as follows. In section 2, we describe our proposed method in detail. The framework of our method is presented, which contains FPM and SO-YOLO. In section 3, we present a series of experiment results and analysis. Finally, conclusion is given in section 4.

II. THEORIES AND METHODS

A. RECONSTRUCTION PROCESS OF FPM

As a recently developed optical microscopy, FPM combines the concepts of ptychography [29], [30], synthetic aperture, and phase retrieval [31]–[33]. With only intensity images capturing, phase images can be got by applying phase retrieval methods. Figure 1 is the reconstruction procedure of FPM. First, initialize the high-resolution complex amplitude distribution with the amplitude of the LR image corresponding to the vertically incident plane wave. The oblique plane waves with wave vectors r_m $(m = 1, 2, \cdots, M)$ illuminate a thin specimen $o(l)$, where M is the total number of the LEDs, $l = (x, y)$ represents the coordinate in the spatial domain. Then the exit field of the specimen is $n(l) = o(l)\exp(i2\pi r_m l)$. So the optical field exiting the lens is $\mathcal{F}\{n(l)\} = O(r - r_m)$. Here we define $r = (r_u, r_v)$ as the coordinate in the fourier domain, $O(r) = \mathcal{F}\{o(l)\}$ as the Fourier spectrum of the sample. The complex amplitude φ in the detector is the light wave that transmits to the detector. So it is the convolution of exit wave and the spatially invariant point spread function p(r) of the microscope system, i.e. $\varphi = n(l) \otimes p(l)$. In the fourier domain:

$$\begin{aligned} F(\varphi) &= F\{n(l) \otimes p(l)\} \\ &= F\{n(l)\} * F\{p(l)\} \\ &= O(r - r_m)P(r) \\ &= S(r - r_m) \qquad (1) \\ S(r - r_m) &= O(r - r_m)P(r) \qquad (2) \end{aligned}$$

where $P(r) = \mathcal{F}\{p(l)\}$ is the pupil function of the image system. S is the initialized high-resolution spectrum that we want to get.

Second, corresponding to a certain illuminating angles, a low-resolution complex amplitude distribution, which is called the target complex amplitude distribution, can be captured by the interception of a certain sub-aperture of the initial high-resolution spectrum with the pupil function of the objective lens. The following is how producing the target

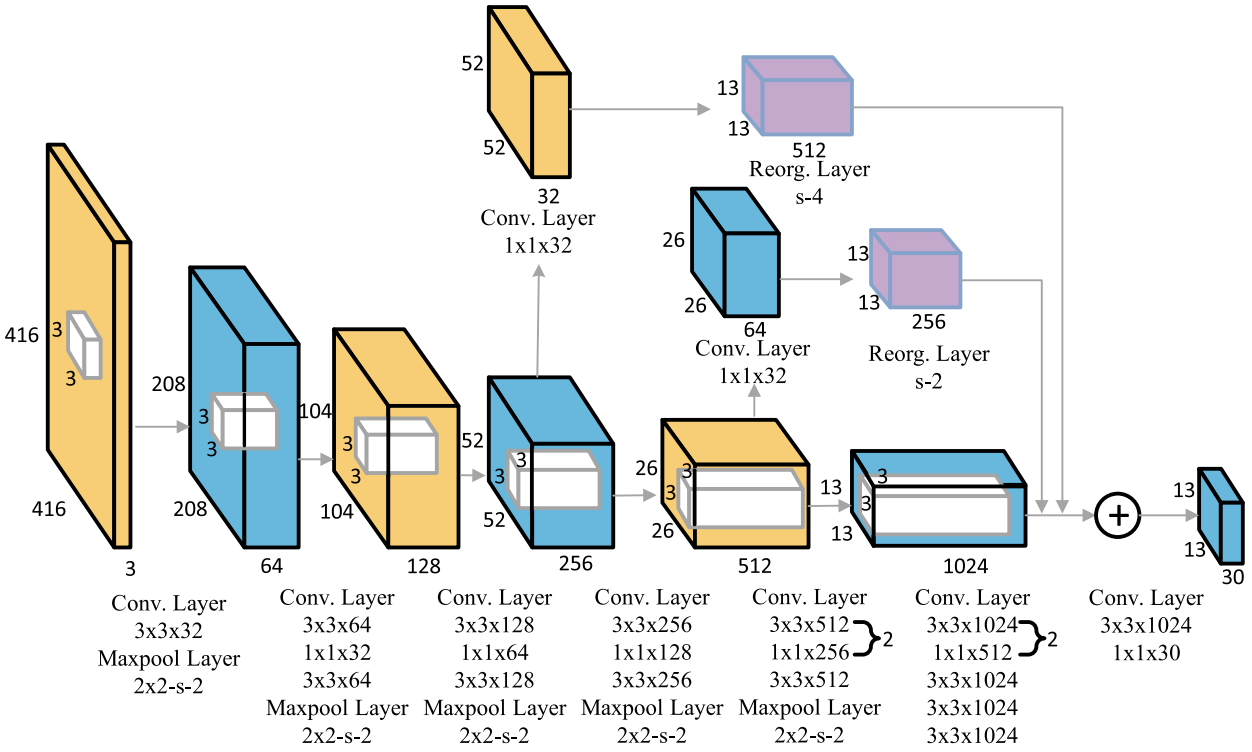


FIGURE 2. The Architecture of SO-YOLO.

complex amplitude distribution:

$$S^e(r - r_m) = S(r - r_m)P(r) \tag{3}$$

$$s^e(l) = \mathcal{F}^{-1}\left\{S^e(r - r_m)\right\} \tag{4}$$

where the superscript e represents the target spectrum and target complex amplitude distribution that wait for update.

Third, maintain the phase of the target complex amplitude image unvaried and update the amplitude part of the target complex amplitude image with the actual intensity measured at the corresponding illumination angle. The amplitude update formula is

$$s^u(l) = \sqrt{I_m^c(l)} \cdot \frac{s^e(l)}{|s^e(l)|} \tag{5}$$

where $s^u(l)$ is the updated target complex amplitude distribution, $I_m^c(l)$ is the actual measurement of the intensity corresponding to the certain LED. The superscript c indicates the captured image.

Fourth, capture the spectrum of the updated target complex amplitude image using the Fourier transform and replace the corresponding subregion of the high-resolution Fourier space. The update formula is

$$S^u(r - r_m) = \mathcal{F}\{s^u(l)\}P(r) \tag{6}$$

Fifth, repeat the above steps for all illuminating angles. Finally, steps 2-5 are repeated until the solution converges.

B. DETECTION METHOD BASED ON SO-YOLO

1) PRINCIPLE OF YOLO AND YOLOV2

In today's detection systems, better detection performance is often accompanied by a larger neural network or the integration of multiple detection models. The goal of YOLO is high-precision real-time monitoring, so it is expected to improve positioning errors and low recall without increasing the network and reducing accuracy. Different from the 2-stages detection method, YOLO's network uses features from the whole image to predict each bounding box. After being resized, the input images are divided into $N * N$ grid cells. Each grid cell is responsible for the object whose center falls into it and predicts M bounding boxes. In YOLO, there are 5 predictions for each bounding box, which are x, y, w, h and confidence. x and y are the coordinates of the bounding box's center which is relative to the bounds of the grid cell. w and h are the width and height of the box relative to the full image. Confidence is defined as $\Pr(Object) \times IOU_{pred}^{truth}$, where IOU is the intersection over union between the predicted box and any ground truth box. In YOLOv2, which is the improved model of YOLO, each bounding box predicts C conditional class probabilities, in addition to the above five parameters. Here C is the number of object types that we detect. So, when testing , we can get the product as follows,

$$\Pr(Class_i|Object) \times \Pr(Object) \times IOU_{pred}^{truth} = \Pr(Class_i) \times IOU_{pred}^{truth} \tag{7}$$

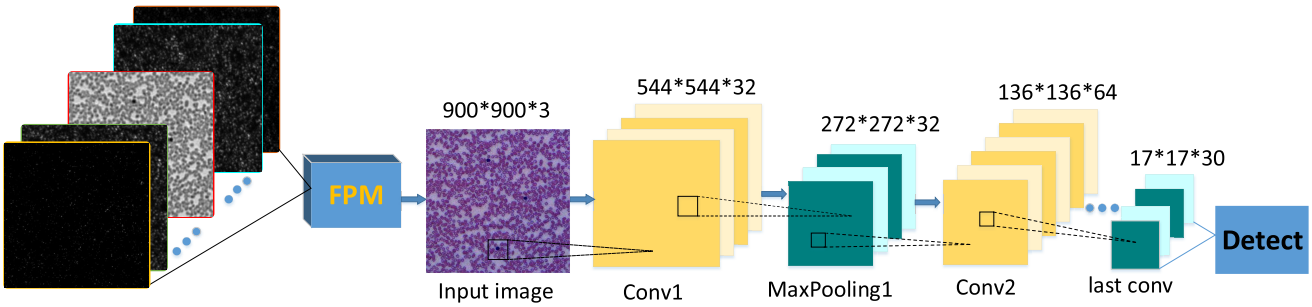


FIGURE 3. The overall framework.

Every bounding box can get a class-specific confidence score, which contains both the class and location information. Then after the network, we can get one bounding box for each object. The multi-part loss function in YOLO includes coordinates, size and class information of bounding boxes:

$$\begin{aligned}
&\lambda_{coord}\sum_{i=0}^{N^2}\sum_{j=0}^{M}\Im_{ij}^{obj}[(x_i-\hat{x}_i)^2+(y_i-\hat{y}_i)^2]\\
&+\lambda_{coord}\sum_{i=0}^{N^2}\sum_{j=0}^{M}\Im_{ij}^{obj}[(\sqrt{w_i}-\sqrt{\hat{w}_i})^2+(\sqrt{h_i}-\sqrt{\hat{h}_i})^2]\\
&+\sum_{i=0}^{N^2}\sum_{j=0}^{M}\Im_{ij}^{obj}[(C_i-\hat{C}_i)^2]\\
&+\sum_{i=0}^{N^2}\Im_{i}^{obj}\sum_{c\in classes}(p_i(c)-\hat{p}_i(c))^2 \qquad (8)
\end{aligned}$$

Inspired by GoogLeNet model, YOLO's network architecture contains 24 convolutional layers and 2 fully connected layers. 1 × 1 convolutional layers are adopted followed by 3 × 3 convolutional layers to reduce the parameter number. But the fully connected layers in the last two layers of the network limit the size of input images. So the fully connected layers are replaced with convolutional layers in YOLOv2. As the improved version of YOLO, a new classification model is used as the base of YOLOv2, which contains 19 convolutional layers and 5 maxpooling layers. When training for detection, the last convolutional layer is replaced with three 3 × 3 convolutional layers which all have 1024 filters.

2) WBC DETECTION WITH SO-YOLO

WBC are small objects in blood cell images so WBC detection is difficult. Many detection methods such as Faster-RCNN and YOLO, perform badly in small objects detection. With some considerable improvements in the original framework of YOLOv2, our proposed SO-YOLO can solve this problem perfectly. The SO-YOLO network structure is shown in Figure 2.

One of the modification in our work is the scale that we change when training. In order to detect well for small objects, the size that the input images are resized to before entering the network is enlarged. The framework of SO-YOLO downsamples by a factor of 32, so we pull from the following multiples of 32: 448, 480,..., 768. Compared to YOLOv2, the size in our detection method is improved a lot. Therefore, the largest option is 768 × 768 and the smallest is 448 × 448. When training, the network is resized to a random dimension every 10 batches. The multi-scale training makes our network be robust to running on images of different sizes.

As shown in Figure 2, SO-YOLO network consists of 26 convolutional layers and 5 maxpooling layers. All the convolutional layers employ a kernel of 3 × 3 or 1 × 1 with a padding size of 1. The addition of convolutional layers compared to the original network can improve the abstract of image feature. The downsampling is conducted by maxpooling layers so we can get feature maps of different dimension size, which contain different feature information of the image and our small detection objects, WBC. As we know, the larger the size of the feature map, the more detailed white blood cell features we can get. The smaller the size of feature map, the more semantic information we can get. Therefore, in order to detect small objects better, we concatenates 52 × 52 features, 26 × 26 features and 13 × 13 features with passthrough layers. By stacking four adjacent features into different channels, we turn the 52 × 52 × 32 feature map into a 13 × 13 × 512 feature map. By stacking adjacent features into different channels, we turn the 26 × 26 × 64 feature map into a 13 × 13 × 256 feature map. Then we concatenate the 13 × 13 × 512 feature map and the 13 × 13 × 256 feature map with the original 13 × 13 × 1024 feature map, which turn out 13 × 13 × 1792 feature map. Then the expanded feature map acts as the input of the last convolutional layer. The combination of multi-resolution features makes our network detect better for small objects.

When an image is passed through the network of SO-YOLO, it extracts global features based on the entire image. The division of the image into NxN grid cells is an important step for the detector. Big objects may occupy several cells, so the detection will be accuracy and easy. But for small objects, such as our detection targets, WBC, there may be some difficulties in detecting them. The size of one white

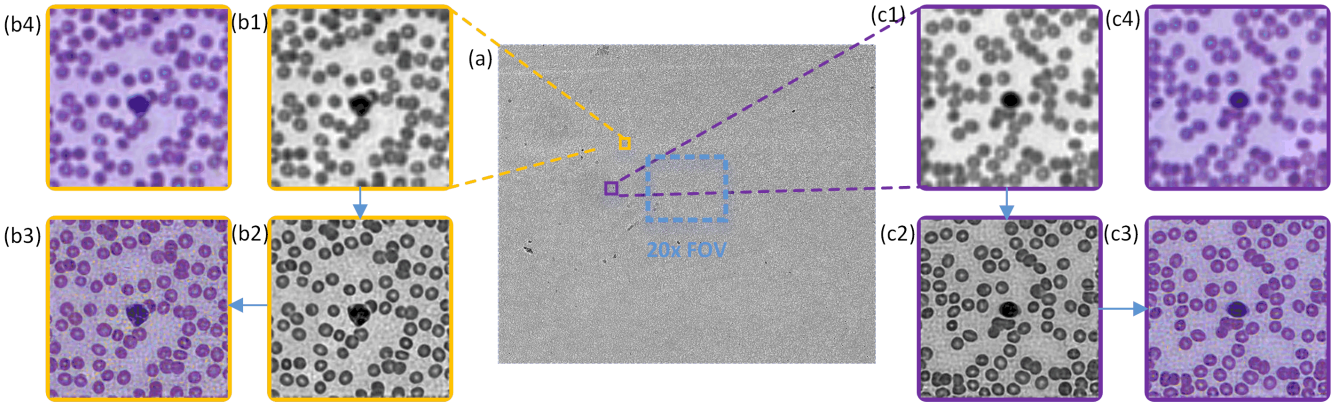


FIGURE 4. (b1) and (c1) are enlarged region of interest. (b2) and (c2) are monochrome high resolution images which are the reconstruction results of FPM. (b3) and (c3) are color high resolution images. (b4) and (c4) are color low resolution images that are synthesized from monochrome low resolution images of three channels.

blood cell compared to the high resolution blood cell image is so small that it can't occupy one cell, so some WBC may be missed in the detection. In order to improve detection accuracy, we propose to increase the number of divided grid cells when detecting. The original network divides the image into 13×13 grid cells. We divide the input image into 15×15, 17×17 and 19×19 gird cells respectively and run the detector. Then by doing experiments for different division, we can find that the more we divide the image, the more accurate the detection is. So we choose 17×17 as a good tradeoff between detection speed and detection accuracy.

When detecting, the size of input image is 544×544. After the network of SO-YOLO, it is downsampled by a factor of 32, so the size becomes 17×17. The input image is divided into 17×17 grid cells. Each grid cell is responsible for the object whose center falls into it and predicts K bounding boxes. There are several parameters for every box, which contain x, y, w, h and $class$. Here $class$ indicates the number of object class that we detect. The location prediction of the bounding box is relative to the grid cell. What's more, for stability, logistic action is used to constrain the network's prediction between 0 and 1. 5 coordinates are predicted for each bounding box in the network, x_p, y_p, w_p, h_p, c_p. If the cell grid is offset from the top left corner of the image by (x_c, y_c) and the width and height of the bounding box prior is w_r, h_r, then we can get the predictions as following,

$$\begin{aligned} x &= s(x_p) + x_c \\ y &= s(y_p) + y_c \\ w &= w_r \cdot e^{w_p} \\ h &= h_r \cdot e^{h_p} \end{aligned} \tag{9}$$

What's more, the confidence is,

$$P_r(object) * IOU(b, object) = s(c_p) \tag{10}$$

Here b represents the predicted box. The above location prediction make the network more stable. In our network, $K = 5$, and we only detect WBC, so $class = 1$. Therefore, the output of the network is $N \times N \times K \times (5 + class) = 17 \times 17 \times 5 \times (5 + 1) = 17 \times 17 \times 30$. By predicting $17 \times 17 \times 5 = 1445$ boxes, we can detect any white blood cell with a optimized box.

Figure 3 shows the overall framework of our system. Our system takes advantage of both FPM and SO-YOLO network. For a blood cell smear which needs testing, we firstly obtain its wide field-of-view, high resolution image with FPM, then the image is delivered to SO-YOLO network for detection. The whole process is not only fast but also accurate. What's more, the cost of our system is low because our setup is only the minor modification of the conventional microscope. The above advantages make our system great potential in the field of clinical medicine, hematopathology, etc.

III. EXPERIMENTAL RESULTS AND COMPARISON

A. EXPERIMENTAL SETUP

The FPM setup is a conventional microscope which replaces its common light source with a programmable light emitting diode(LED) array.

As shown in Figure 5, there is a real experimental system that we built for the work. We choose the Nikon ECLIPSE Ni-U microscope as the experimental microscope, which has CFI60 infinity optical system. The objective lens of the microscope can correct chromatic aberration across the wave length from 435nm to 850nm. Here we use the four times objective(0.13NA) to capture the low resolution images. Connected directly to the microscope is a pre-photographic science grade CMOS(sCMOS) camera named Andor Zyla 5.5. The Zyla 5.5 uniquely offers both Rolling and true Global Shutter exposure modes. The high sensitivity and wide dynamic range are especially beneficial to our image capture process since only one LED is lightened at a time. The light source of the microscope, which is 100mm away from the sample, is replaced with a 16×16 LED array of which only 13×13 elements are used in our experiment. the distance between each LED element is 8mm. The LED array is controlled by an Ardruino circuit board which is connected

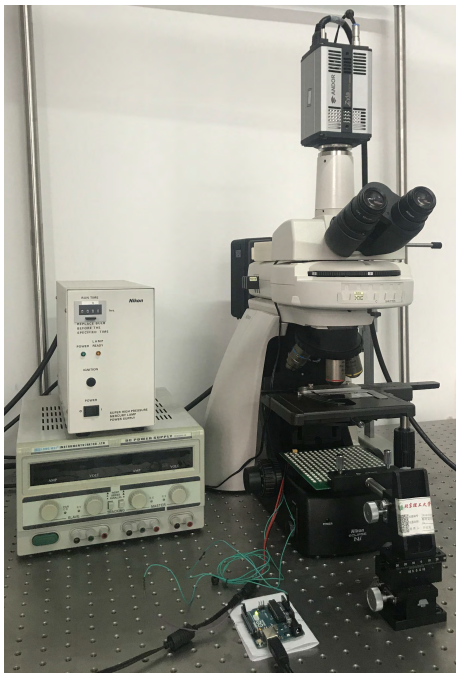


FIGURE 5. The experiment setup of FPM.

to the computer so that the code can be written to the circuit board.

B. DATA SET MAKING

1) IMAGE CAPTURE

A high-resolution image(HRI) is the reconstruction result of FPM with the input of 169 low-resolution images(LRIs) in our experimental system. The training and testing images for SO-YOLO are color images so we need to capture images under illumination of three colors sequentially. That's to say, if we want to get a color HRI for training, we need to capture 169 × 3 LRIs. During image capture, The LED-array board are controlled directly by computer with MATLAB. We made a GUI window to decide the path of the obtained images, set the exposure time and control the beginning of the image capture process, etc. After clicking the '' start capturing '' button, the 169 LEDs are lightened one by one with red light, the images under which are saved in ''Red'' file. After that, the light of LEDs is changed to green automatically and a ''Green'' file with 169 images can be got. Similarly, we can obtain a ''Blue'' file in which the 169 images are captured under blue light. So far, an image capture progress is done. Repeat the progress until we get enough pictures.

2) DATASET MAKING WITH FPM

As shown in Figure 4, the images of datasets are color HRIs that are synthesized from monochrome HRIs of three channels, which are obtained with FPM progress sequentially. 1000 color HRIs are got after the above progress. We prepare our own data set according to the architecture of VOC datasets. We labelled these images with 'labelImg', an image annotation tool with which the annotations containing the bounding box information of the ground truth can be obtained. Then we converted the ground truth annotation acquired by 'labelImg' to the format accepted by SO-YOLO. So far, the dataset has been made, which contains 1000 images and the corresponding label files.

C. EXPERIMENTAL RESULTS AND COMPARISON

The images we used for detecting are the FPM reconstruction results, which are clearer than the images that are captured under the same NA objective lens of the conventional microscope. It's meaningless to do experiments with images that are not the reconstructions of FPM because the wide FOV and the high resolution can't be obtained simultaneously with conventional microscope. In order to compare and well present the performance of different detecting methods, we first do experiments with 300 × 300 resolution images. Due to the performance of SO-YOLO is same as that of YOLOv2 for low resolution blood cell images, so the comparison between YOLOv2 and SO-YOLO is temporarily ignored here and we just present the detecting results of three methods, a DE-based method(DED) in [21], a method based on automatic segmentation(AS) in [22] and our proposed SO-YOLO. The intuitive comparison of the detection results with the three methods are demonstrated in Figure 5. (a1)-(a5) are original color images with different forms of WBC. The green boxes are ground truth. Figure (b1)-(b5) show the detecting results with DED. Each red circle indicates a white blood cell that has been detected. Figure (c1)-(c5) present the detecting results with the black block in the white background. Figure (d1)-(d5) show SO-YOLO's detecting results, a bounding box selecting each white blood cell with the class name of the detecting object. We can see that SO-YOLO has good robustness for different shapes of WBC. What's more, it can detect the whole cells without selecting part of the cell like DED and AS as shown in (b3) and (c3). The completeness degree of the detected WBC is important for classification task because the features extracted from the detected WBC have a strong influence on the performance of the classifier. We can see easily that SO-YOLO is better than the other two methods. In addition, For the white blood cell whose nucleus is lobulated like figure (a1) and figure (a5), SO-YOLO can correctly detect the cell as one target while the number that DED or AS counts is more than one.

Since the field-of-view of the 4× objective lens is wider than that of the 20× objective lens, the resolution of the images that we use for detecting is large. So the ability of detecting WBC in large images is significant. In order to present well the detection results, we segment the high resolution images into 1000 × 1000 images and show the results respectively. Figure7 shows the detecting results with the four methods. (a1)-(a5) are original images with ground truth. (b1)-(b5) are detecting results with DED. None of WBC is detected with this method indicates that it can't detect small objects in large images. (c1)-(c5) are AS's detection results. We can see that as images get bigger, the background become more complicated. This method mistakenly detects part of

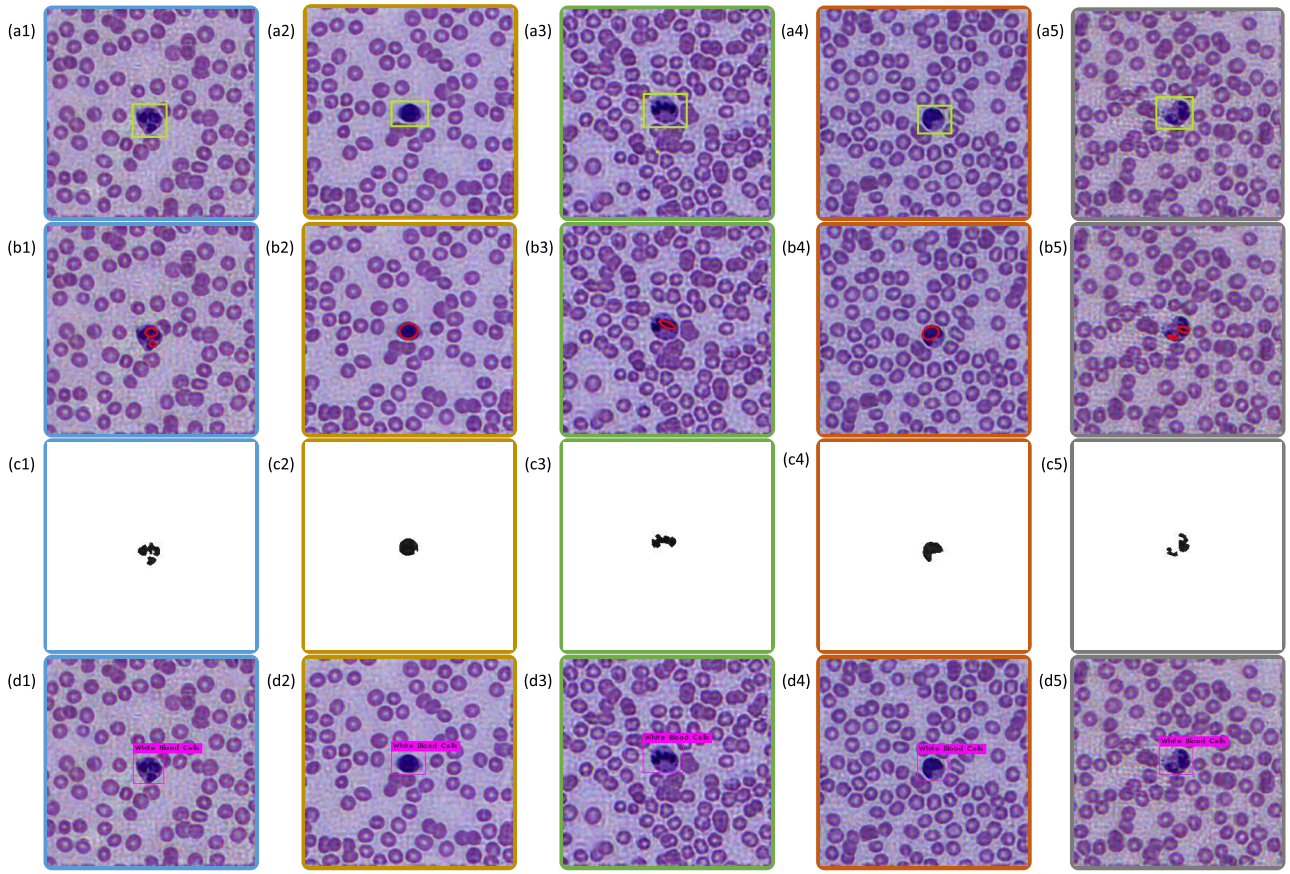


FIGURE 6. 300 × 300 original images for detecting and the the detection results with three methods. (a1)-(a5) Original images with ground truth. (b1)-(b5) Detecting results with DED. (c1)-(c5) Detecting results with AS. (d1)-(d5) Detecting results with SO-YOLO.

TABLE 1. comparison of WBC detecting results on different scales of images between FPM+DED, FPM+AS, FPM+YOLO and FPM+SO-YOLO.

image size	detection	FPM+DED	FPM+AS	FPM+YOLO	FPM+SO-YOLO
300x300	**PR**	0.91	0.93	0.98	**1.00**
	RC	1.00	1.00	1.00	**1.00**
1000x1000	**PR**	0	0.47	0.85	**0.97**
	RC	0	1.00	0.91	**1.00**

the background as WBC. (d1)-(d5) and (e1)-(e5) present the detecting results with YOLOv2 and SO-YOLO respectively. We can see that the detecting boxes of YOLOv2 is too large for the detecting WBC while the detecting boxes of SO-YOLO is suitable. As we know, the features extracted from the detecting WBC affect the performance of the classifier, so the more suitable the detecting boxes, the better the detection effect. (d2) and (e2) show that SO-YOLO has higher detection recall rate. The small WBC which is near the edge of the images can be well detected with SO-YOLO. What's more, from (d3)-(e3) and (d4)-(e4) we can see that the accuracy of SO-YOLO is higher than that of YOLO. SO-YOLO can distinguish between WBC and other objects in the background which is similar with WBC. (d5) and (e5) show that for the WBC which are close SO-YOLO is able to accurately detect them. The experiment results demonstrate that SO-YOLO has good robustness and accuracy for detecting small objects.

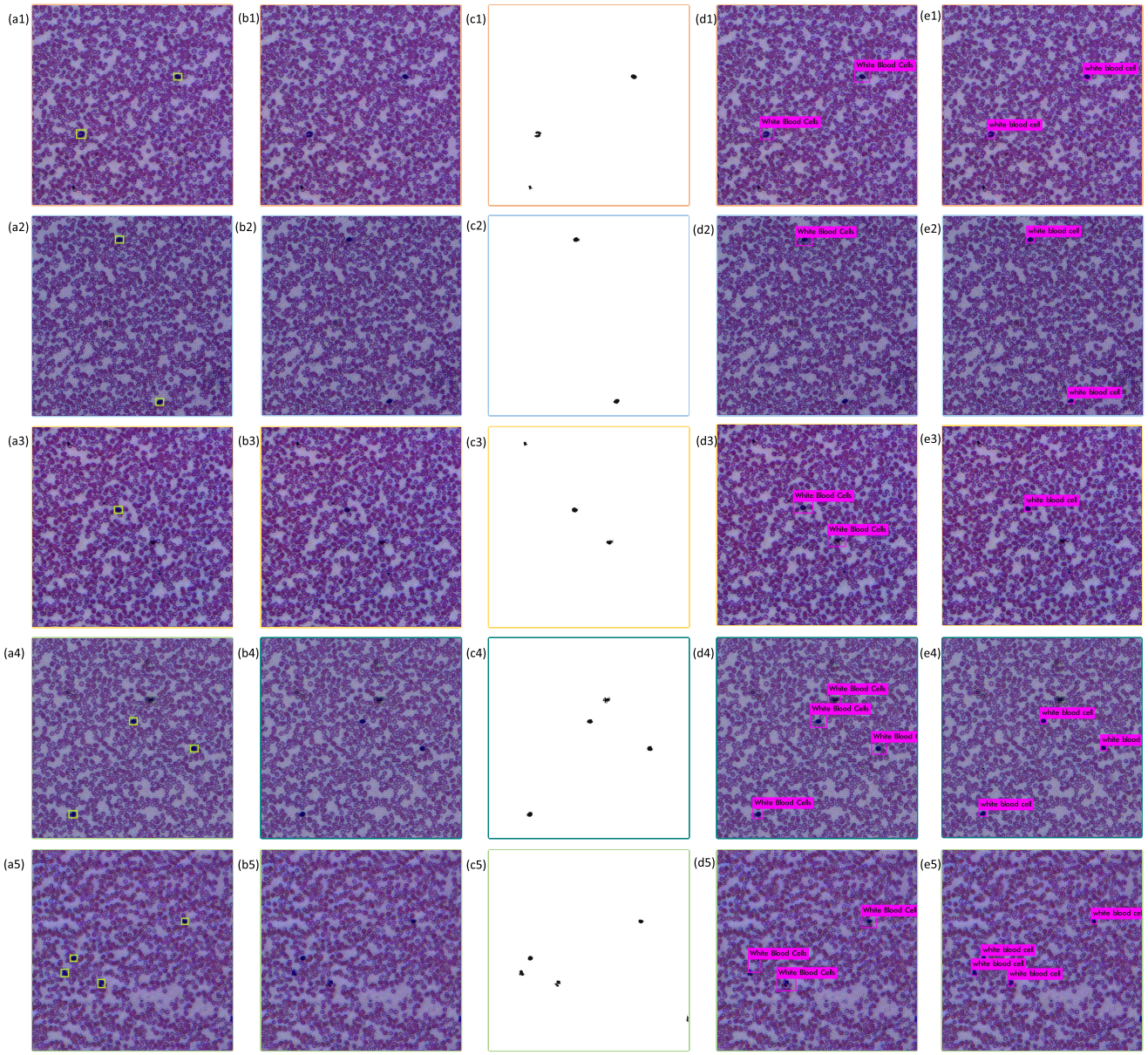


FIGURE 7. 1000 × 1000 original images for detecting and the the detection results with four methods. (a1)-(a5) Original images with ground truth. (b1)-(b5) Detecting results with DED. (c1)-(c5) Detecting results with AS. (d1)-(d5) Detecting results with YOLO. (e1)-(e5) Detecting results with SO-YOLO.

The detection performance of the four FPM-based methods is summarized in Table 1. Here PR is precision rate, RC is recall rate. The definition of PR and RC are as follows,

$$PR = \frac{Tp}{Tp + Fp} \tag{11}$$

$$RC = \frac{Tp}{Tp + Fn} \tag{12}$$

Here, Tp, Fp and Fn are short for true positive, false positive, and false negative respectively. Tp means detecting white blood cells correctly as white blood cells. Fp means detecting other objects in the images as white blood cells mistakenly. Fn means the white blood cells which are missed in the detection. There are 1000 images that we obtained through the reconstruction process of FPM for detecting. The detection images include 300 × 300 resolution images and 1000 × 1000 resolution images. The WBC in the images have been detected and counted by a human expert. So the detection results by expert are regarded as the ground truth for all detection images. From Table 1 we can see that when detecting WBC in small images, DED and AS can detect well. Our proposed method can get 100% precision rate and 100% recall rate. What's more, our detection speed is fast. When detecting WBC in high resolution images, the advantages of SO-YOLO are revealed. All the WBC can be detected with a fast speed and the detection accuracy is high. The precision

and recall rate have been greatly improved compared to YOLO. However, DED can't detect any WBC because they are so small compared to the whole images. Due to the complicated background of large images, the precision of AS is also low.

IV. CONCLUSION

In this paper, we propose a complete system for WBC detection which includes FPM as image acquisition method and SO-YOLO as detection method. The contrast results between our method and other methods show that SO-YOLO has good accuracy and robustness in WBC detection. What's more, our detection speed is fast so that it has great advantages for clinical detection. The combination of FPM and SO-YOLO can achieve fast detection over an image which contains all information of a blood smear without mechanical moving. Our experiment setup is flexible and low-cost because FPM only needs minor modifications of conventional microscope. All the advantages of our detection system make promise in further applications.

ACKNOWLEDGMENT

The authors sincerely acknowledge the editor and the anonymous reviewers for their valuable suggestions.

REFERENCES

[1] Y. M. Alomari, S. N. H. S. Abdullah, R. Z. Azma, and K. Omar, "Automatic detection and quantification of WBCs and RBCs using iterative structured circle detection algorithm," *Comput. Math. Methods Med.*, vol. 2014, Apr. 2014, Art. no. 979302.

[2] C. Di Ruberto, A. Loddo, and L. Putzu, "A leucocytes count system from blood smear images," *Mach. Vis. Appl.*, vol. 27, no. 8, pp. 1151–1160, 2016.

[3] L. B. Dorini, R. Minetto, and N. J. Leite, "Semiautomatic white blood cell segmentation based on multiscale analysis," *IEEE J. Biomed. Health Informat.*, vol. 17, no. 1, pp. 250–256, Jan. 2013.

[4] W. Shitong and W. Min, "A new detection algorithm (NDA) based on fuzzy cellular neural networks for white blood cell detection," *IEEE Trans. Inf. Technol. Biomed.*, vol. 10, no. 1, pp. 5–10, Jan. 2006.

[5] G. Zheng, R. Horstmeyer, and C. Yang, "Wide-field, high-resolution Fourier ptychographic microscopy," *Nature Photon.*, vol. 7, pp. 739–745, Jul. 2013.

[6] F. Zhao, L. Wei, and H. Chen, "Optimal time allocation for wireless information and power transfer in wireless powered communication systems," *IEEE Trans. Veh. Technol.*, vol. 65, no. 3, pp. 1830–1835, Mar. 2016.

[7] F. Zhao, H. Nie, and H. Chen, "Group buying spectrum auction algorithm for fractional frequency reuse cognitive cellular systems," *Ad Hoc Netw.*, vol. 58, pp. 239–246, Apr. 2016.

[8] J. Chung, X. Ou, R. P. Kulkarni, and C. Yang, "Counting white blood cells from a blood smear using Fourier ptychographic microscopy," *PLoS ONE*, vol. 10, no. 7, p. e0133489, 2015.

[9] R. Horstmeyer, X. Ou, G. Zheng, P. Willems, and C. Yang, "Digital pathology with Fourier ptychography," *Comput. Med. Imag. Graph.*, vol. 42, pp. 38–43, Nov. 2015.

[10] A. J. Williams *et al.*, "Fourier ptychographic microscopy for filtration-based circulating tumor cell enumeration and analysis," *J. Biomed. Opt.*, vol. 19, no. 6, p. 066007, 2014.

[11] X. Ou, R. Horstmeyer, C. Yang, and G. Zheng, "Quantitative phase imaging via Fourier ptychographic microscopy," *Opt. Lett.*, vol. 38, no. 22, pp. 4845–4848, 2013.

[12] S. Dong, Z. Bian, R. Shiradkar, and G. Zheng, "Sparsely sampled Fourier ptychography," *Opt. Express*, vol. 22, no. 5, p. 5455, 2014.

[13] S. Dong, R. Shiradkar, P. Nanda, and G. Zheng, "Spectral multiplexing and coherent-state decomposition in Fourier ptychographic imaging," *Biomed. Opt. Express*, vol. 5, no. 6, p. 1757, 2014.

[14] L. Tian, X. Li, K. Ramchandran, and L. Waller, "Multiplexed coded illumination for Fourier Ptychography with an LED array microscope," *Biomed. Opt. Express*, vol. 5, no. 6, pp. 1757–1767, 2014.

[15] S. Dong, K. Guo, P. Nanda, R. Shiradkar, and G. Zheng, "FPscope: A field-portable high-resolution microscope using a cellphone lens," *Biomed. Opt. Express*, vol. 5, no. 10, pp. 3305–3310, 2014.

[16] S. Dong *et al.*, "Aperture-scanning Fourier ptychography for 3D refocusing and super-resolution macroscopic imaging," *Opt. Express*, vol. 22, no. 11, pp. 13586–13599, 2014.

[17] S. Dong, P. Nanda, R. Shiradkar, K. Guo, and G. Zheng, "High-resolution fluorescence imaging via pattern-illuminated Fourier ptychography," *Opt. Express*, vol. 22, no. 17, pp. 20856–20870, 2014.

[18] L. Tian, J. Wang, and L. Waller, "3D differential phase-contrast microscopy with computational illumination using an LED array," *Opt. Lett.*, vol. 39, no. 5, pp. 1326–1329, 2014.

[19] L. Tian and L. Waller, "Quantitative differential phase contrast imaging in an LED array microscope," *Opt. Express*, vol. 23, no. 9, pp. 11394–11403, 2015.

[20] C. Zuo, J. Sun, S. Feng, Y. Hu, and Q. Chen, "Programmable colored illumination microscopy (PCIM): A practical and flexible optical staining approach for microscopic contrast enhancement," *Opt. Lasers Eng.*, vol. 78, pp. 35–47, Mar. 2016.

[21] E. Cuevas, M. Diaz, M. Manzanares, D. Zaldivar, and M. Perez, "An improved computer vision method for white blood cells detection," *Comput. Math. Methods Med.*, vol. 2013, Apr. 2013, Art. no. 137392.

[22] M. Mohamed, B. Far, and A. Guaily, "An efficient technique for white blood cells nuclei automatic segmentation," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2012, pp. 220–225.

[23] Q. Lu, C. Liu, Z. Jiang, A. Men, and B. Yang, "G-CNN: Object detection via grid convolutional neural network," *IEEE Access*, vol. 5, no. 99, pp. 24023–24031, 2017.

[24] K. Muhammad, J. Ahmad, I. Mehmood, S. Rho, and S. W. Baik, "Convolutional neural networks based fire detection in surveillance videos," *IEEE Access*, vol. 6, no. 99, pp. 18174–18183, 2018.

[25] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2014, pp. 580–587.

[26] R. Girshick, "Fast R-CNN," in *Proc. Int. Conf. Comput. Vis. (ICCV)* Dec. 2015, pp. 1440–1448.

[27] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2015, pp. 91–99.

[28] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 779–788.

[29] J. M. Rodenburg, "Ptychography and related diffractive imaging methods," *Adv. Imag. Electron Phys.*, vol. 150, no. 7, pp. 87–184, 2008.

[30] A. M. Maiden, J. M. Rodenburg, and M. J. Humphry, "Optical ptychography: A practical implementation with useful resolution," *Opt. Lett.*, vol. 35, no. 15, pp. 2585–2587, 2010.

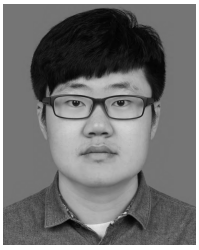
[31] J. R. Fienup, "Phase retrieval algorithms: A comparison," *Appl. Opt.*, vol. 21, no. 15, pp. 2758–2769, 1982.

[32] A. M. Maiden and J. M. Rodenburg, "An improved ptychographical phase retrieval algorithm for diffractive imaging," *Ultramicroscopy*, vol. 109, no. 10, pp. 1256–1262, 2009.

[33] W. Huang, D. Liu, X. Zhang, Y. Zhang, and J. Zhu, "Analysis of a digital phase retrieval method for wave-front reconstruction," *Chin. Opt. Lett.*, vol. 9, no. 8, p. 080101, 2011.

XING WANG is currently pursuing the M.E. degree with the School of Optoelectronics, Beijing Institute of Technology, Beijing, China. Her research interests include computer vision and Fourier ptychographic microscopy.

TINGFA XU received the Ph.D. degree from the Changchun Institute of Optics, Fine Mechanics and Physics, Changchun, China, in 2004. He is currently a Professor with the School of Optoelectronics, Beijing Institute of Technology, Beijing, China. His research interests include optoelectronic imaging and detection and hyper-spectral remote sensing image processing.



JIZHOU ZHANG received the bachelor's degree from the School of Optics and Photonics, Beijing Institute of Technology, Beijing, China, in 2014, where he is currently pursuing the Ph.D. degree. His research interests include Fourier ptychographic microscopy, computational imaging, and machine learning.

SINING CHEN is currently pursuing the M.E. degree with the School of Optoelectronics, Beijing Institute of Technology, Beijing, China. Her research interests include computer imaging and Fourier ptychographic microscopy.

YIHZOU ZHANG received the master's degree in optical engineering from the Beijing Institute of Technology in 2015, where she is currently pursuing the Ph.D. degree with the Key Laboratory of Optoelectronics Imaging Technology and System of the Education Ministry of China, School of Optics and Photonics. Her research interests include Fourier ptychographic microscopy and infrared image processing.