Improved detection performance in blood cell count by an attention-guided deep learning method

ZHENGFEN JIANG,[1] XIN LIU,[2,3] ⓘ ZHUANGZHI YAN,[1,4] ⓘ WENTING GU,[1] AND JIEHUI JIANG[1]

[1]*School of Communication & Information Engineering, Shanghai University, Shanghai 200444, China*
[2]*Academy for Engineering & Technology, Fudan University, Shanghai 200433, China*
[3]*xinliu.c@gmail.com*
[4]*zzyan@shu.edu.cn*

Abstract: Blood cell count plays an important role in the field of clinical medical diagnosis. To effectively automate the counting of blood cells, recently, the deep-learning-based detection method represented by the YOLO has been proposed and used successfully. Nevertheless, the YOLO detection method has difficulties in insufficient positioning of the bounding boxes and in distinguishing overlapping objects. To overcome the limitations, we propose a new deep-learning-based method, termed Attention-YOLO, which is achieved by adding the channel attention mechanism and the spatial attention mechanism to the feature extraction network. By using the filtered and weighted feature vector to replace the original feature vector for residual fusion, Attention-YOLO can help the network to improve the detection accuracy. The experimental results indicate that compared to the standard YOLO network, the Attention-YOLO can achieve a better detection performance in blood cell count without introducing too many additional parameters, where the recognition accuracy of cells (RBCs, WBCs, and platelets) has an improvement of 6.70%, 2.13%, and 10.44%, respectively, and the mean Average Precision (mAP) has an improvement of 7.10%.

1. Introduction

Blood cell count is an important part of the blood test. The accurate cell count is help for detecting potential diseases and related lesions, and determining the type and development of the patient's diseases. The three common types of blood cells are red blood cells (RBCs), white blood cells (WBCs), and platelets. The main function of RBCs is to transport oxygen, WBCs play an immune role, and platelets play a role in the process of hemostasis [1]. Generally, the number of these blood cells is huge. The traditional manual counting process is very cumbersome, time-consuming, and error-prone. Furthermore, the accuracy rate will also be affected by the status and experience of the clinical laboratory analysts [2].

In the automatic counting process of blood cells, there are usually two different methods, i.e., the image processing methods and the deep-learning (DL)-based methods. In the early stage, the image processing methods are widely used, e.g., the automatic cell recognition technology based on Hough transform [3–5], and the watershed method based on distance transformation [6], etc. However, these image processing methods have some problems in cell detection, for example, the accuracy of detecting cells cannot meet the requirements in areas with high cell overlap. In recent years, the DL methods have been successfully applied to different medical applications [7], e.g., the multi-class diagnosis of Alzheimer's disease [8], automatic segmentation of the left ventricle in cardiac MRI [9], and lung nodule detection in lung CT [10], etc. Inspired by DL, the DL-based detection methods [11–19] have been proposed and used for blood cell detection and counting. By using convolutional neural network (CNN) as classifier, Habibzadeh *et al.* propose a blood

cell classification method, which can automatically classify WBCs into one of five types from cytological images [11]. However, the speed of this method is slow. By employing a hierarchical RBC patch extraction method followed by a shape-invariant RBC patch normalization technique, Xu *et al.* reduce the computational cost during both the training and learning procedures and classify the RBC shape from sickle cell of disease patients [14]. Furthermore, in 2019, with the help of YOLO (you only look once) [20], Alam and Islam automatically identify and count RBCs, WBCs, and platelets, with the detection speed less than a second [18]. Nevertheless, YOLO method has difficulties in distinguishing overlapping objects and positioning bounding box. Together, in cell experiments, blood cells may touch or overlap each other, or appear as clusters in the image, which make that YOLO detection method is difficult to accurately detect each cell.

To overcome the limitations, in this work, we propose a new DL-based method, termed as Attention-YOLO, which is achieved by adding the channel attention mechanism and the spatial attention mechanism to the feature extraction network. Referring to [21], by using the filtered and weighted feature vector to replace the original feature vector for residual fusion, Attention-YOLO is helpful for further improving the detection accuracy of the network. To evaluate the performance of the proposed method, a series of experimental data are used. The experimental results show that compared to the standard YOLO detection method, Attention-YOLO can achieve better performance in blood cell count without introducing too many additional parameters, where the recognition accuracy of cells (RBCs, WBCs, and platelets) has an improvement of 6.70%, 2.13%, and 10.44%, respectively, and the mean Average Precision (mAP) has an improvement of 7.10%.

The outline of this paper is as follows. In section 2, the methods and materials are detailed. In section 3, the experimental results are shown. Finally, we discuss the results and draw conclusions in section 4.

2. Methods and materials

2.1. YOLO network structure and detection principle

YOLO treats the detection task as a regression problem [22] and has been widely used in image processing fields. Recently, many versions of YOLO (e.g., YOLOv1, YOLOv2, and YOLOv3) have also been proposed. Compared to YOLOv1 and YOLOv2, YOLOv3 [23] is with the following advantages: 1) The improved classification performance on complex datasets. 2) The increased amount of information in the feature map. 3) The deeper network layers. So, in this work, YOLOv3 is chosen and used.

Briefly, YOLOv3 uses Darknet-53 as the feature extraction network [23], and adopts the method similar to the feature pyramid network [24]. It can directly use the original input images and annotations for training. As a result, it saves computing resources and time-consuming. When performing the detection task, first input the image into the feature extraction network, and then the extracted feature vectors are sent to a structure similar to the feature pyramid network, and the grid cell is obtained at 3 scales. Furthermore, each grid cell predicts 3 bounding boxes, and each bounding box predicts a vector P, as follows:

$$P = (t_x + t_y + t_w + t_h) + P_0 + (P_1 + P_2 + \cdots + P_n) \tag{1}$$

with

$$P_0 = P_r(Object) \times IOU_{pred}^{truth} \tag{2}$$

where t_x, t_y, t_w, t_h is the coordinates related to the bounding box. $P_r(Object)$ represents the probability that the object is in the prediction box. IOU reflects the accuracy of the object's position. Finally, the non-maximum suppression is performed on the generated prediction to obtain the final prediction result.

2.2. Attention-YOLO network structure and detection principle

It is worth mentioning that in YOLOv3 detection processes, each region in entire feature map is treated equally. That is, each area is considered with the same contribution to the final detection. However, in experiments, blood cells may touch or overlap each other, or appear as clusters in the image. In other words, there are often complex and rich contextual information around cells to be detected in the image. Weighting the features of the target area can more clearly show the parts that the detection network focuses on, and better distinguish the objects to be detected.

Based on the above considerations, to improve the detection accuracy, the attention mechanism module is introduced into the network. Generally speaking, channel attention mechanism can filter and weight features in channel dimension, which is helpful for improving the detection performance. The spatial attention mechanism can model features on the spatial relationship, which can supplement the positional relationship information that the channel attention mechanism cannot obtain [25]. Considering that: 1) Convolutional Block Attention Module (CBAM) [21] combines both space and channel attention mechanisms, which allows quickly find important feature from numerous features, suppresses the irrelevant or unimportant features, and improve the efficiency and accuracy of important features processing; 2) CBAM can be conveniently embedded in network structures, and the end-to-end training can be carried out without changing the original network structure, in this work, CBAM is chosen and introduced into YOLOv3 network.

In detail, as shown in Fig. 1, the input feature maps are subjected to global maximum pooling and global average pooling operations, and then sent to the multilayer perceptron (MLP) for channel information filtering. After that, the MLP output features based on the weighting element-wise for generating the channel attention feature map. Sequentially, the maximum pooling and average pooling are performed along the channel dimension. Furthermore, the output of the two operations is combined and a feature descriptor is obtained. Finally, the convolution operation is used to encode, and the spatial attention map is obtained.
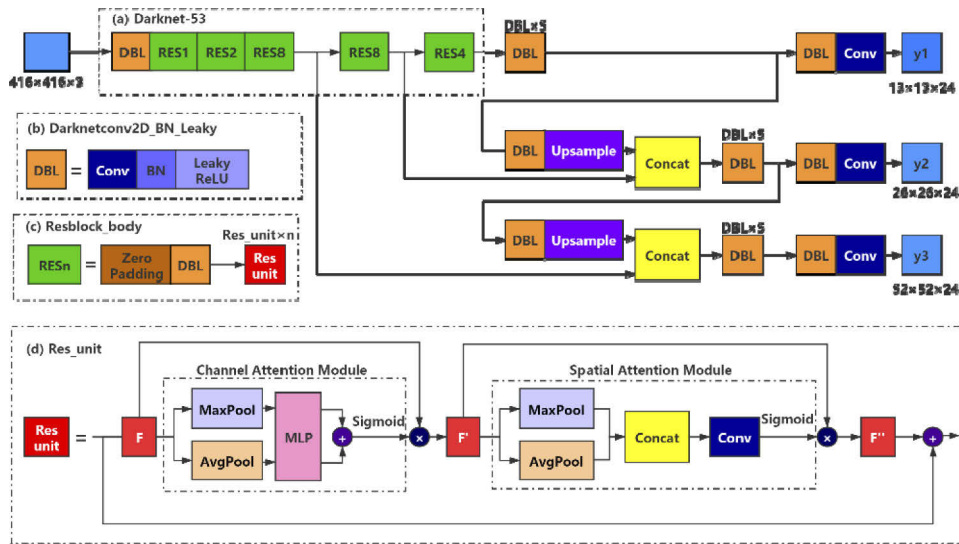

Fig. 1. The proposed Attention-YOLO network structure, which is achieved by adding the channel attention mechanism and the spatial attention mechanism into YOLOv3 network. (a) The network structure of Darknet-53, which mainly consists of the DBL and Resblock module. (b) DBL module, which consists of the convolution, batch normalization, and leaky ReLU. (c) Resblock module which consists of DBL and *n* Res_unit. (d) Res_unit. Here, a channel and spatial attention module are introduced into the residual module.

It should also be pointed out that the original implementation of the YOLO configuration is trained for 80 different classes. When using YOLO for blood count, the number of classes is changed from 80 to 3 (RBCs, WBCs, and platelets). Corresponding, the number of filters in the final convolutional layer is changed to 24. Referring to [18], the number is calculated by $NF = NA \times (NC + 5)$ with NA being the number of anchor boxes and NC being the number of classes. Here, NA and NC are set to 3, respectively.

2.3. Blood cell dataset

The dataset used for this work is a public blood cell dataset (Blood Cell Count Dataset, BCCD), which is available on https://github.com/Shenggan/BCCD_Dataset. BCCD contains 364 cell images and 4,888 annotations (RBCs, WBCs, and platelets). The image resolution is 640×480 pixels. The choice of these datasets is guided by the fact that they are open source and fully accessible to the research community and the general public.

It should be pointed out that in a BCCD dataset, some images contain RBCs, but the provided annotation file does not contain the corresponding RBCs. The mismatches may affect the accuracy of the evaluation. To address the problem, in this work, based on the previously provided annotations, we manually annotate the dataset again, which is achieved by LabelImg image annotation software (https://github.com/tzutalin/labelImg). Here, the annotation operation follows two principles: 1) For the local cells at the edge of the image, if the area is less than 30% of the entire cell, it will not be labeled; 2) For the highly adherent cells if the cell overlap exceeds 80%, only one cell is marked. Figure 2 shows some labeled images manually. The new dataset is available from Ref. [26].
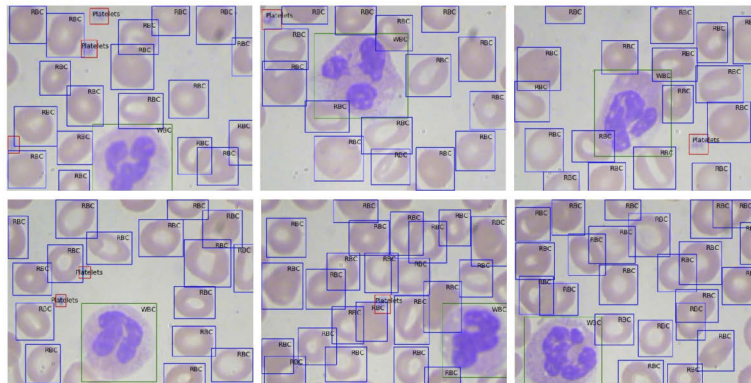

Fig. 2. The cell images and the corresponding annotations in a BCCD dataset.

2.4. Training

Considering that Attention-YOLO is a data-driven method, sufficient training samples are important for training. Based on the consideration, in this work, the Mosaic data enhancement method is used to generate large amounts of training data. By randomly zooming, cropping, stitching operations, the data augmentation increases the variability of the input image, making the detection model more robust to images in different environments. Figure 3 shows the generated image data by the augmentation method.

In this work, the BCCD dataset contains 364 images. We randomly selected 292 annotated blood cell images as the training set, 15 images as the validation set, and 57 images as the testing set. The principle of setting the training/validation/testing images is that the ratio of the number of training set to the sum of testing set and verification set is ~4:1, and the ratio of the number of testing set
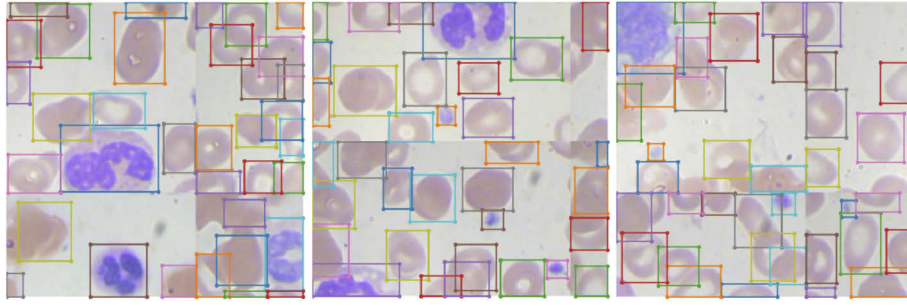
Fig. 3. The image data generated by the Mosaic data enhancement method.

to verification set is ~4:1. The input size of the network is fixed at 416×416. In the training phase, the stochastic gradient descent is used with momentum=0.9. The weight attenuation coefficient is set to 0.0005. Figure 4 shows the learning curve of Attention-YOLO according to the loss function. As expected, as the epoch increases, the loss function gradually decreases. Considering that the minimum loss can be obtained at epoch 393, here, we use the epoch for testing purposes. Here, the training is performed on a server, equipped with an NVidia Tesla V100 GPU (16 GB RAM), 2 Intel Xeon Gold 6130 (2.1GHZ), and 192 G DDR4 REG ECC. The software environment required include torch = 1.1.0, torchvision = 0.2.3, keras = 2.2.4, scipy = 1.2.1, tqdm = 4.31.1, matplotlib = 3.0.3, numpy = 1.16.3, tensorboard = 1.14.0, terminaltables = 3.1.0.
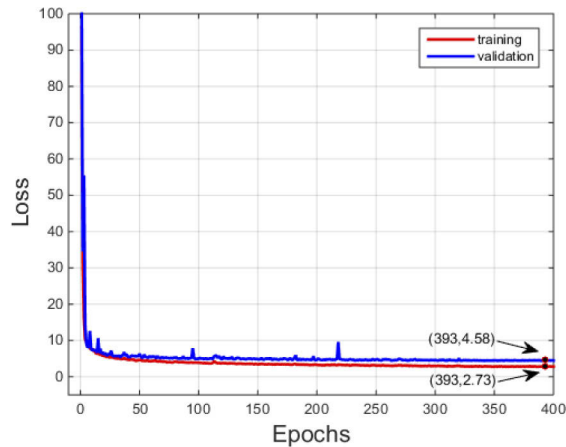

Fig. 4. Learning curve of Attention-YOLO for blood cell recognition.

Finally, Fig. 5 shows the flowchart of the proposed method. Once the model is trained, the Attention-YOLO network returns the number of three labels (RBCs, WBCs, and platelets) for each input image, i.e., the number of the detected cells. In addition, it should be pointed out that since the dataset is not large, the leave-N cross-validation is used to evaluate the performance of the proposed method. Considering that the number of validation set is 15, we choose N=15 when using the leave-N cross-validation. Correspondingly, a total of 24 (364/15) trials are trained and then used to test in this work.
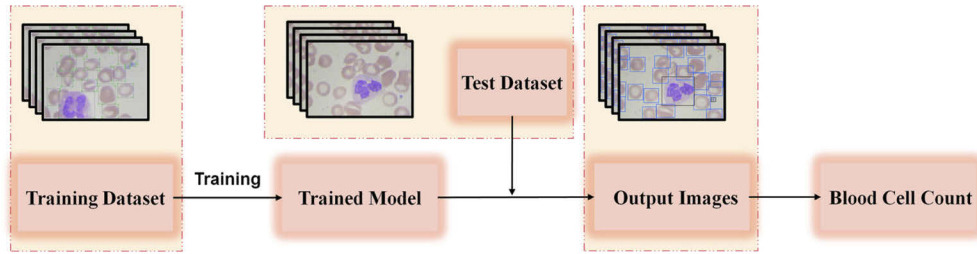
Fig. 5. The flowchart of blood cell recognition and counting based on the proposed Attention-YOLO method.

3. Results

3.1. Testing with a BCCD dataset

When Attention-YOLO method is used to detect RBCs, WBCs, and platelets, the confidence threshold plays an important role. The redundant and false predictions can be further eliminated by choosing an appropriate confidence threshold. Referring to [18], here, the optimal confidence threshold is determined by calculating the average absolute error (AAE) between the truth cell number and the estimated cell number. In this work, the error is computed using

$$\varepsilon^{\alpha} = \frac{1}{K} \sum_{i=1}^{K} \left| \mathbb{N}_{truth}^{(i)} - \mathbb{N}_{estimated}^{(i)} \right| \tag{3}$$

where $\mathbb{N}_{true}$ depicts the actual cell number, and $\mathbb{N}_{estimated}$ depicts the predicted cell number. ε^{α} represents the AAE value with α being cell type (RBCs, WBCs, or platelets). K is the size of the testing set with 57 used in this experiment.

As mentioned above, a total of 24 trials are trained and then used to test in this work when using the leave-N cross-validation. As an example, Table 1 summarizes the calculated AAE values, selected from the 6th trial of all 24 trials. The results show that for RBCs, when the confidence threshold is 55%, the obtained AAE value is the smallest. As a result, the confidence threshold is set to 55% for detecting RBCs. Correspondingly, when detecting WBC and platelets, the confidence thresholds are set to 40% and 25%, respectively.

Table 1. The AAE of RBCs, WBCs, and platelets with different confidence thresholds.

Threshold (%)	RBCs	WBCs	Platelets
20	4.211	0.035	0.053
25	3.860	0.035	**0.018**
30	3.351	0.018	0.035
35	2.842	0.018	0.053
40	2.211	**0**	0.123
45	1.526	0.035	0.158
50	0.649	0.035	0.175
55	**0.018**	0.035	0.228
60	1.105	0.053	0.333

Figure 6 shows the detection results of RBCs based on the obtained confidence threshold (55%), selected from the 6th trial of all 24 trials. The 1st row of Fig. 6 shows the detection results obtained by the standard YOLO method. In contrast, the detection results from the proposed Attention-YOLO method are shown in the 2nd row of Fig. 6. Similarly, Fig. 7 and Fig. 8

compare the detection results of WBCs (confidence threshold=40%) and platelets (confidence threshold=25%), selected from the 6th trial of all 24 trials. The results are obtained by YOLO and Attention-YOLO methods. The experimental results indicate that the standard YOLO method is prone to missed detections and false detections (see the black arrows in Figs. 6–8). This may be caused by the cell touch or overlap each other, or appear as clusters. In contrast, when using the proposed Attention-YOLO, the detection performance is improved.
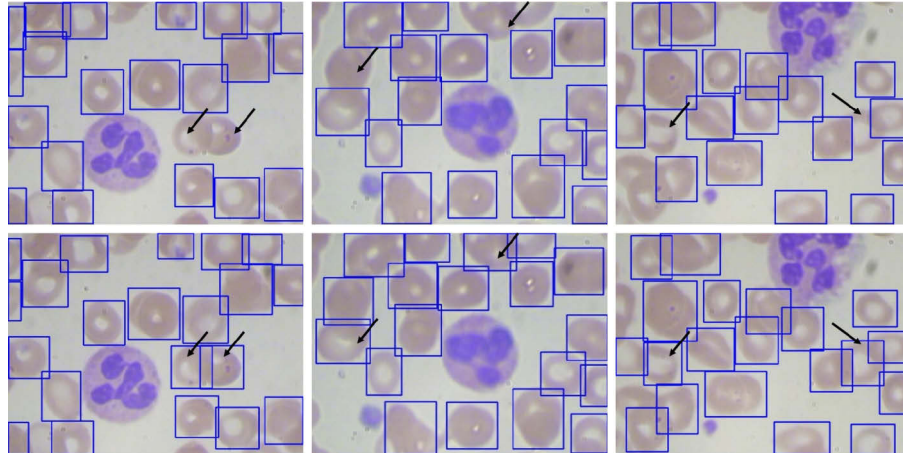

Fig. 6. The detection results of RBCs with confidence threshold=55%, selected from the 6th trial of all 24 trials. The 1st row depicts the detection of the standard YOLO method, selected randomly 3 images from all 57 testing images. In contrast, the 2nd row depicts the detection of the proposed Attention-YOLO method.
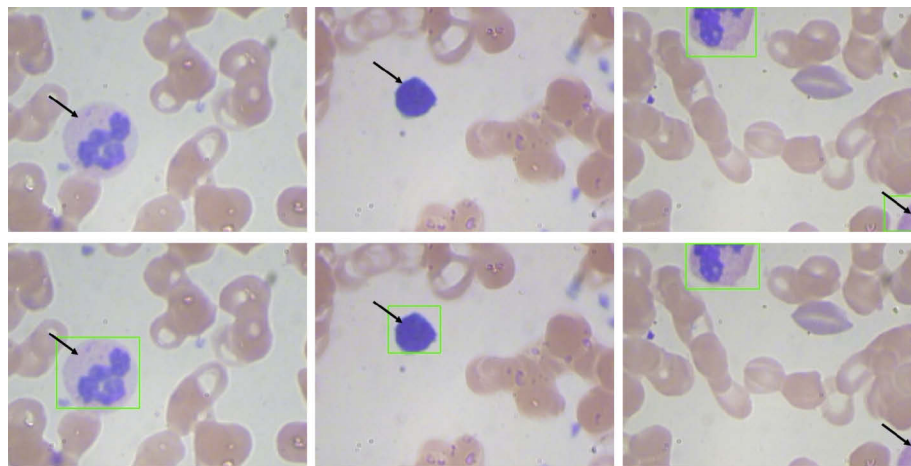

Fig. 7. The detection results of WBCs with confidence threshold=40%, selected from the 6th trial of all 24 trials. The 1st row depicts the detection of YOLO method. In contrast, the 2nd row depicts the detection of Attention-YOLO method.

Table 2 compares the estimated number of cells obtained by the standard YOLO method and the proposed Attention-YOLO method, selected from the 6th trial of all 24 trials. Further, Table 3 quantitatively compare the detection performance obtained by the above two methods. In this work, the detection performance is quantitatively evaluated by recognition accuracy
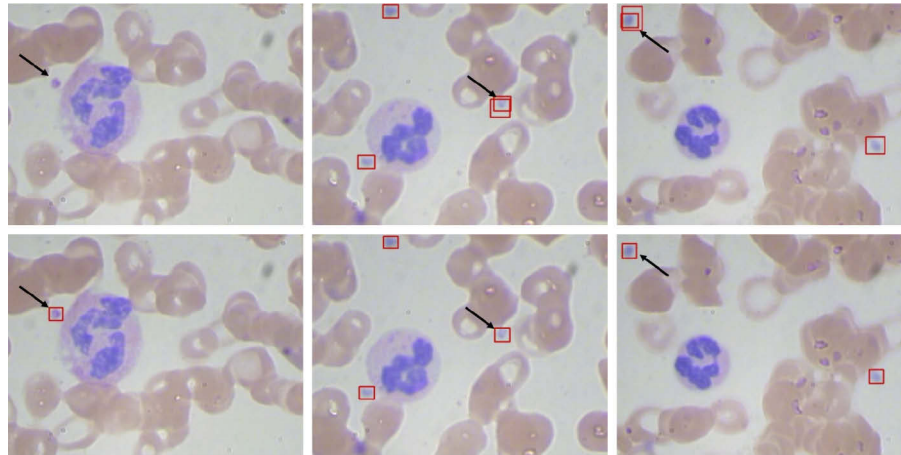
Fig. 8. The detection results of platelets with confidence threshold=25%, selected from the 6th trial of all 24 trials.. The 1st row depicts the detection of YOLO method. In contrast, the 2nd row depicts the detection of Attention-YOLO method.

(RA) [18] and mAP [27]. Note that the obtained RA and mAP summarized in Table 3 are obtained by calculating the mean from all 24 trials. The results indicate that the detection performance of Attention-YOLO is superior to the standard YOLO method, in terms of RA and mAP, whether they are used to detect RBCs, WBCs, or platelets. For the detection of RBCs, the recognition accuracy of YOLO is 90.74%, while the recognition accuracy of Attention-YOLO is improved to 97.44%. When detecting WBCs, the recognition accuracy of YOLO is 97.33%, and the recognition accuracy of Attention-YOLO is further improved to 99.46%. Especially when Attention-YOLO is used to detect platelets, the recognition accuracy is 96.99%. Comparably, the recognition accuracy of YOLO is dropped to 86.55%, where there exist multiple and missed detection. The results from mAP further validate the conclusion, where mAP is increased from 0.872 (YOLO) to 0.943 (Attention-YOLO), indicating a higher detection performance obtained by Attention-YOLO.

Table 2. Comparison of the estimated number of cells obtained by the standard YOLO method and the proposed Attention-YOLO method, selected from the 6th trial of all 24 trials.

	RBCs	WBCs	Platelets
The truth number	1033	57	69
The estimated number of YOLO	970	58	77
The estimated number of Attention-YOLO	1032	57	70

Table 3. Comparison of the detection performance of the standard YOLO method and the proposed Attention-YOLO method, obtained by calculating the mean from all 24 trials.

	RA(%)			mAP
	RBCs	WBCs	Platelets	
YOLO	90.74	97.33	86.55	0.872
Attention-YOLO	97.44	99.46	96.99	0.943

3.2. Testing with an RBCS dataset

To demonstrate the robustness of Attention-YOLO model, another dataset [28] is used to test. Briefly, the new dataset contains 30 images, where the number of RBCs accounts for the majority, and the number of WBCs and platelets is small. Correspondingly, here, the dataset is termed as RBCS (red blood cells). In this case, the trained model based on a BCCD dataset is directly applied to the new dataset. The confidence threshold for this case is determined also by calculating the AAE, similar to that determined in a BCCD dataset. The obtained confidence thresholds are 0.2 (RBCs), 0.2 (WBCs), and 0.1 (platelets), respectively. Since an RBCS dataset contains 30 images, the number of tested datasets is 30 in this case. The corresponding detection results are shown in Fig. 9. Furthermore, Table 4 summarizes the obtained recognition accuracy. The experimental results indicate that as expected, Attention-YOLO can effectively identify RBCs, WBCs, and platelets, even if it is applied to a different dataset. In addition, we can also find that in this case, the 100% recognition accuracy is obtained in identifying WBCs. But, it is hardly to hold the 100% recognition accuracy when Attention-YOLO is applied to the other blood cell dataset.
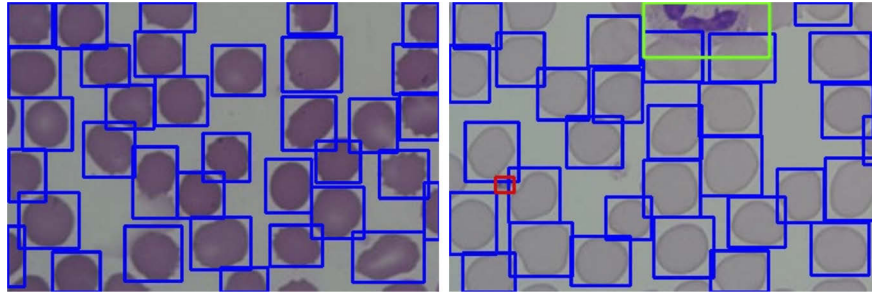

Fig. 9. The detection results obtained Attention-YOLO with an RBCS dataset. The detected RBCs, WBCs, and platelets are framed by blue, green, and red rectangles, respectively.

Table 4. The recognition accuracy of Attention-YOLO with an RBCS dataset.

	RBCs	WBCs	Platelets
The truth number	796	15	10
The estimated number	766	15	9
The recognition accuracy	96.23%	100%	90.00%

4. Discussion and conclusion

Accurate cell counting is important in medical image analysis. In clinical applications, generally, various types of cells are manually counted, leading to a huge workload. The DL-based detection method, e.g., YOLO, can automatically identify and count RBCs, WBCs, and platelets. However, the current YOLO method has difficulties in distinguishing overlapping objects and positioning the bounding box. The purpose of this paper is to improve cell detection accuracy, which is achieved by adding channel attention and spatial attention mechanisms to the feature extraction network.

It can be seen from the experimental results that compared with the standard YOLO method, the proposed Attention-YOLO method achieves the better performance in detecting RBCs, WBCs, and platelets, in terms of RA and mAP (Figs. 6–8, Table 3), where the RA of RBCs, WBCs, and platelets has an improvement of 6.70%, 2.13%, and 10.44%, respectively, and mAP has an

improvement of 7.10%. In addition, the results from another dataset further demonstrate the robustness of Attention-YOLO model, where RBC, WBC, and platelets can also be effectively identified (Fig. 9 and Table 4).

It should be noted that for the DL-based method, a large amount of training data can improve the generalization ability and robustness of the model. In this work, the training set has only 292 images. In addition, although the Attention-YOLO model can detect other datasets, the RA has declined, which may be caused by the insufficient training datasets. On the other hand, here, only blood cell datasets are used to train and test. In clinical applications, the types of cells are more complex and diverse, so different types of cell data should be collected and used. Moreover, in this work, the confidence threshold is determined by calculating the average absolute error. The better methods for determining the confidence threshold may further improve the detection performance of the proposed method. Furthermore, considering that the dataset used in this work is imbalanced, the detection performance of the Attention-YOLO method might be further improved by thoroughly processing these imbalanced data. Finally, due to the addition of the attention mechanism, the detection time of Attention-YOLO has increased by about 20 ms compared to YOLO. These related works will be investigated in our future study.

In conclusion, by adding the channel attention mechanism and the spatial attention mechanism to the feature extraction network, we improve the detection performance of cells (RBCs, WBCs, and platelets) with higher recognition accuracy (97.44%, 99.46%, 96.99%) and mAP (0.943). In the future, we will apply the proposed to detect the more complex cell, e.g., neuron.

Disclosures. The authors declare that there are no conflicts of interest related to this article.

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 method," Comput. Math. Method. Med. **2014**, 1–17 (2014).
2. S. Acharjee, S. Chakrabartty, M. I. Alam, N. Dey, V. Santhi, and A. S. Ashour, "A semiautomated approach using GUI for the detection of red blood cells," *2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT)* (2016), pp. 525–529.
3. Z. Ejaz, C. A. Ahmad, and H. Justin, "Automatic red blood cell detection and counting system using Hough transform," Am. J. Pharm. Sci. **5**(8), 267–271 (2018).
4. V. Acharya and P. Kumar, "Identification and red blood cell automated counting from blood smear images using computer-aided system," Med. Biol. Eng. Comput. **56**(3), 483–489 (2018).
5. S. M. Mazalan, N. H. Mahmood, and M. A. A. Razak, "Automated red blood cells counting in peripheral blood smear image using circular Hough transform," *2013 International Conference on Artificial Intelligence, Modelling and Simulation (AIMS)* (2013), pp. 320–424.
6. J. Hari, A. S. Prasad, and S. K. Rao, "Separation and counting of blood cells using geometrical features and distance transformed watershed," *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPR)* (2014), pp. 77–82.
7. G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. W. M. van der Laak, B. van Ginneken, and C. I. Sanchez, "A survey on deep learning in medical image analysis," Med. Image. Anal. **42**, 60–88 (2017).
8. S. Q. Liu, S. D. Liu, W. D. Cai, H. Y. Che, S. Pujol, R. Kikinis, D. G. Feng, and M. J. Fulham, "Multimodal neuroimaging feature learning for multiclass diagnosis of Alzheimer's disease," IEEE. Trans. Biomed. Eng. **62**(4), 1132–1140 (2015).
9. M. R. Avendi, A. Kheradvar, and H. Jafarkhani, "A combined deep-learning and deformable-model approach to fully automatic segmentation of the left ventricle in cardiac MRI," Med. Image. Anal. **30**, 108–119 (2016).
10. Q. Dou, H. Chen, L. Q. Yu, J. Qin, and P. A. Heng, "Multilevel contextual 3-d CNNs for false positive reduction in pulmonary nodule detection," IEEE. Trans. Biomed. Eng. **64**(7), 1558–1567 (2017).
11. M. Habibzadeh, A. Krzyzak, and T. Fevens, "White blood cell differential counts using convolutional neural networks for low resolution images," Artif. Intell. Soft. Comput. **7895**, 263–274 (2013).
12. M. Habibzadeh, M. Jannesari, Z. Rezaei, H. Baharvand, and M. Totonchi, "Automatic white blood cell classification using pre-trained deep learning models: Resnet and Inception," *2017 International Conference on Machine Vision (ICMV)* (2017), pp. 1–8.

13. J. W. Choi, Y. Ku, B. W. Yoo, J. A. Kim, D. S. Lee, Y. J. Chai, H. J. Kong, and H. C. Kim, "White blood cell differential count of maturation stages in bone marrow smear using dual-stage convolutional neural networks," PLoS One **12**(12), e0189259 (2017).

14. M. J. Xu, D. P. Papageorgiou, S. Z. Abidi, M. Dao, H. Zhao, and G. E. Karniadakis, "A deep convolutional neural network for classification of red blood cells in sickle cell anemia," Plos. Comput. Biol. **13**(10), e1005746 (2017).

15. X. Wang, T. F. Xu, J. Z. Zhang, S. N. Chen, and Y. Z. Zhang, "SO-YOLO based wbc detection with fourier ptychographic microscopy," IEEE Access. **6**, 51566–51576 (2018).

16. A. I. Shahin, Y. H. Guo, K. M. Amin, and A. A. Sharawi, "White blood cells identification system based on convolutional deep neural learning networks," Comput. Methods. Programs. Biomed. **168**, 69–80 (2019).

17. L. Alzubaidi, M. A. Fadhel, O. Al-Shamma, J. L. Zhang, and Y. Duan, "Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis," Electronics **9**(3), 427 (2020).

18. M. M. Alam and M. T. Islam, "Machine learning approach of automatic identification and counting of blood cells," Healthc. Technol. Lett. **6**(4), 103–108 (2019).

19. A. B. Chowdhury, J. Roberson, A. Hukkoo, S. Bodapati, and D. Cappelleri, "Automated complete blood cell count and malaria pathogen detection using convolution neural network," IEEE Robot. Autom. Lett. **5**(2), 1047–1054 (2020).

20. J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 21–26.

21. S. Woo, J. Park, J. Y. Lee, and J. Y. Lee, "CBAM: convolutional block attention module," arXiv:1807. 06521 (2018).

22. J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2016), pp. 779–788.

23. J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv:1804. 02767 (2018).

24. T. Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 936–944.

25. C. J. Xu, X. F. Wang, and Y. D. Yang, "Attention-YOLO:YOLO detection algorithm that introduces attention mechanism," Comput. Eng. and Appl. **55**(6), 13–23 (2019).

26. Z. F. Jiang, "A modified blood cell dataset based on Blood Cell Count Dataset," figshare (2020), https://doi.org/10.6084/m9.figshare.13018625.v4.

27. R. Padilla, S. L. Netto, and E. A. B. da Silva, "A survey on performance metrics for object-detection algorithms," *2020 International Conference on Systems, Signals and Image Processing (IWSSIP)* (2020), pp. 237–242.

28. O. Sarrafzadeh, A. M. Dehnavi, H. Rabbani, N. Ghane, and A. Talebi, "Circlet based framework for red blood cells segmentation and counting," *IEEE Workshop on Signal Processing Systems* (2015).