SW-YOLO: Improved YOLOv5s algorithm for Blood Cell Detection

Yonglin Wu[1], Yinfeng Fang[2], Dongxu Gao[3], Hongwei Gao[1], Zhaojie Ju[4(✉)]

[1] School of Automation and Electrical Engineering, Shenyang Ligong University, Shenyang 110158, China
ghw1978@sohu.com

[2] School of Telecommunication Engineering, Hangzhou Dianzi University, Hangzhou 310000, China
yinfeng.fang@hdu.eud.cn

[3] School of Computing, University of Portsmouth, Portsmouth PO13HE, UK
dongxu.gao@port.ac.uk

[4] School of Computing, University of Portsmouth, Portsmouth PO13HE, UK
Zhaojie.Ju@port.ac.uk

Abstract. This paper proposes an improved target detection algorithm SW-YOLO based on the YOLOv5s framework to solve the problems of low detection accuracy, wrong detection and missed detection in blood cell detection tasks. To begin with, the end of the backbone network is fused with Swin Transformer to improve network feature extraction. Next, since blood cells are mostly small and medium-sized targets, resulting in poor detection of large cells, the network layer that identifies large cells is removed. In addition, the normal convolution in the PANet network is replaced with depth-separable convolution during the feature fusion process to ensure the accuracy and real-time detection while having better detection results for small targets. At last, the loss function of the prediction layer uses EIOU to solve the positive and negative sample imbalance problem of CIOU. Compared with existing target detection algorithms such as Faster-RCNN, YOLOv4 and YOLOv5s, SW-YOLO improves to 99.5%, 95.3% and 93.3% mAP on the blood cell dataset BCCD for white blood cells, red blood cells and platelets respectively. The experimental results are eximious and the algorithm is highly practical for blood cell detection.

Keywords: Blood Cells Testing, Swin Transformer, PAN, EIOU.

1 Introduction

Blood cell testing is a technique that analyses the cellular components of blood by means of specialist instruments. In medicine, the cells circulating in the blood are

*The authors would like to acknowledge the support from the National Natural Science Foundation of China (52075530).

** Corresponding author: Zhaojie Ju

divided into three categories: white blood cells, red blood cells and platelets [1]. When the results of these three types of cell count are abnormally high or low, the body may already be suffering from certain diseases. Accurate and efficient identification of these three types of cells can provide an overview of a patient's overall health status [2].

Previously, blood cell testing mainly used traditional image processing techniques, most classically manual photography and staining, followed by manual observation and classification with microscope [3], which required the testers to have extensive basic knowledge of cell morphology and to be skilled in morphological observation and discrimination through continuous study and repeated practice.

In recent years, as many efficient deep learning-based target detection algorithms have been proposed, they have also been widely applied to blood cell detection tasks. Current deep learning-based target detection algorithms can be broadly classified into two main categories. Two-stage algorithms, represented by Faster R-CNN [4], extract regions of interest (ROI) through a region proposal network (RPN) and then classify the regions of interest. This type of algorithm usually has high detection accuracy, but the detection speed is slow. Single-stage algorithms, represented by the SSD [5], RetinaNet [6] and YOLO [7-8] series, treat localisation and classification as a regression problem and achieve end-to-end detection. This class of algorithms is fast in detection, but low in detection accuracy. From Figure 1, we can see that the detection results with Faster R-CNN, SSD, and YOLO algorithms do have the problem of wrong detection and missed detection.

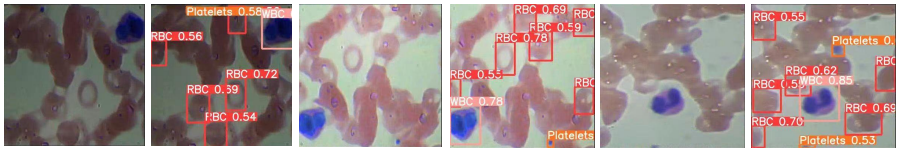


Fig. 1. Comparison of blood cells in target detection algorithms

This paper introduces SW-YOLO small target detection algorithm to address the issues of inadequate detection accuracy, incorrect detection and missed detection which are common in target detection algorithms for blood cell detection tasks. The precision of small-target blood cell identification is enhanced while maintaining real-time . This paper focuses on the following tasks:

(1) Enhancing feature extraction for small targets by connecting the end of the backbone network to the Swin Transformer [9].

(2) By removing the large target detection layer, the number of parameters in the network can be reduced, while improving detection performance.

(3) Replacing the normal convolution in the PANet network with depth-separable convolution, the detection accuracy and real-time performance are guaranteed while the detection of small targets is more effective.

(4) Replacing the loss function CIOU with the loss function EIOU [10] can solve the problem of positive and negative sample imbalance.

2 Related Work

2.1 YOLOv5s model

YOLOv5 is divided into four models: s, m, l and x by varying the depth and width of the model to obtain different numbers of convolution kernels [11], thus choosing the appropriate model according to one's needs. To facilitate the detection of small and medium-sized models such as blood cells, YOLOv5s, which has the smallest model, is used as the base framework in this paper. As shown in Figure 2, the network structure is divided into four main parts: input, backbone, neck and prediction.

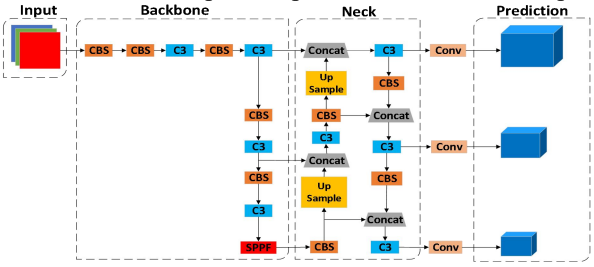
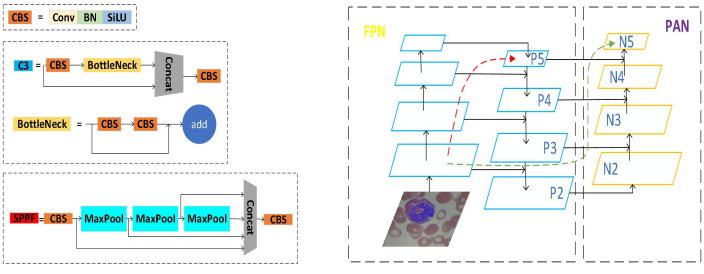


Fig. 2. YOLOv5s network architecture

Input. The role of the input is to scale, colour change and data enhancement of the input image. The processed image on the input side then enters the backbone network for feature information extraction. The current backbone networks in the field of target detection are generally ResNet [12], AmoebaNet [13], Xception [14] and other networks.

Backbone. The backbone network in YOLOv5s is mainly composed of the CBS module, C3 module and SPPF modules, as shown in Figure 3(a). The CBS is to speed up the inference of the model, the C3 module makes full use of the geometric information and resolution of the shallow features, the perceptual field and semantic information of the deep features.The SPPF module is able to convert feature maps into fixed size feature vectors regardless of the size of the input feature map, and it serves to improve the speed of generating candidate frames and reduce overfitting.



(a) (b)

Fig. 3. (a) Represents the main modules of YOLOv5s backbone network, (b) Representing the structure of the neck network

Neck. After the feature map is extracted, it enters the neck network, which is composed of FPN [15] (feature pyramid network) and PANet [16] (path aggregation network). The network structure is shown in Figure 3(b). The semantic features of the feature map are first passed from the top down by the FPN layer, and then the localization features are passed from the bottom up by the PANet, so that the shallow semantic features and the deep localization features can be fused and the feature fusion capability of the model can be improved.

Prediction. The key function of the output prediction network is the recognition and classification of the incoming feature maps. YOLOv5s uses a total of three different sizes for image detection, enabling object detection of small, medium and large targets respectively.The loss function for the prediction part consists of three main components: confidence loss L_{obj}, classification loss L_{cls} and loss of position of the target and predicted frames L_{box}. The confidence loss and classification loss can be obtained by using the cross-entropy loss function, where the classification loss function is expressed as shown in equation (1). The CIOU loss function is used to calculate the position loss of the target and prediction frames, as shown in equation (2), thus improving the regression speed of the prediction frame and the convergence speed of the network, and solving the bounding box overlap problem of YOLOv4. The total loss function is the sum of the three components of confidence loss, classification loss and position loss of the target and prediction frames. The expression is shown in equation (3).

$$L_{cls}(p,\ y) = -\frac{1}{N}\sum_{i=1}^{N}[y\log p + (1-y)\log(1-p)] \tag{1}$$

$$\begin{cases} L_{obj} = 1 - IoU + \frac{\rho^2(b,\ b^{gt})}{c^2} + \alpha v \\ v = \frac{4}{\pi^2}\left(\arctan\frac{w^{gt}}{h^{gt}} - \arctan\frac{w}{h}\right) \\ \alpha = \frac{v}{1-IoU(b,\ b^{gt})+v} \end{cases} \tag{2}$$

$$L_{total} = L_{obj} + L_{cls} + L_{box} \tag{3}$$

3 Proposed Method

The proposed SW-YOLO model, which focuses on the fusion of Swin Transformer, improved PANet structure and optimised loss function, is mainly based on YOLOv5s [13]. The SW-YOLO structure is shown in Figure 5(a).

3.1 Integration of the Swin Transformer module

The Swin Transformer network architecture is shown in Figure 4. Firstly, the input [H, W, 3] image is passed into the Patch Partition layer, where every 4×4 adjacent pixels are chunked into a patch and expanded along the channel direction, making the image dimension into [H/4, W/4, 48], and then the channel data of each pixel is

linearly transformed through the Linear Merging layer, making the image dimension becomes [H/4, W/4, C], while each sample is normalised in the feature dimension.

The Swin Transformer employs a window-based MSA module in lieu of the standard Multi-head self-attention module (MSA) in the conventional Transformer module, a window-based MSA is composed of a combination of window multi-head self-attention layers (W-MSA) and sliding window multi-head self-attention layers (SW-MSA). This technique splits the input image into multiple segments that do not overlap and limits the attention computation to each separate window. The backbone network with two Swin Transformer modules in the first, second and fourth stages and six Swin Transformer modules in the third stage, with the feature maps of the second, third, fourth and fourth stages being fused together. The output feature maps are extracted for 3 distinct grid sizes. This technology is capable of identifying targets of varying sizes within blood cells.

YOLOv5s can leverage Swin Transformer as its core network to gain more accurate image feature extraction.

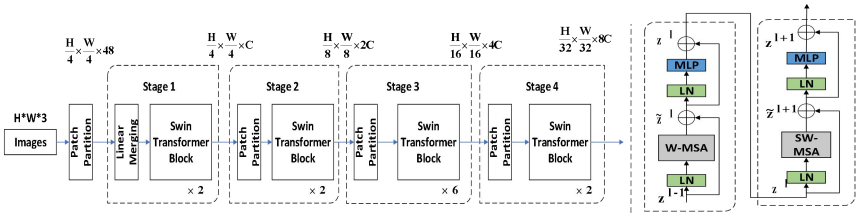


Fig. 4. Swin Transformer Network Architecture

3.2 Improved the PANet network architecture

Since the PANet network outputs predictions at three different scales for small, medium and large targets through 8x, 16x and 32x down-sampling respectively. In the BCCD dataset, for example, the image size is 1024×1024. If the image size is 640×640, the output prediction layer is 80×80, 40×40 and 20×20 after 8x, 16x and 32x down-sampling respectively. However, in the BCCD dataset, the target size is between 8×9 and 285×442, and small and medium-sized cells account for 95% of the total blood cells. When the image is down-sampled by 32 times, the target size of blood cells smaller than 102×102 in the image will be compressed into one pixel, resulting in low efficiency of the network in detecting small-sized cells. In order to reduce the number of network model parameters while improving the efficiency of target detection, this paper does not use the 32-fold down-sampling correlation network layer for network redundancy, thus improving the efficiency of detection of small-sized cells.

3.3 Improved convolutional structures in PAN

The top diagram in Figure 5(b) is a normal Conv-BN-SiLu module, where SiLu is the activation function. The lower diagram in Figure 5(b) is an over-parameterised schematic, which uses M identical parallel branches and then sums the outputs of all

branches before feeding them into the activation function. Over-parameterisation methods can increase the complexity of the model at training time while maintaining the same inference structure.

The depth-separable convolution is a method of over-parameterisation. As shown in Figure 5(c), it consists of two parts, the upper part based on Depthwise Convolution and the lower part on Pointwise Convolution. Depthwise Convolution is essentially a grouped convolution with the same number of groups g as the input channels. Pointwise Convolution, on the other hand, is a 1×1 convolution. The depth convolution module in the figure consists of three branches. The leftmost branch is a 1 × 1 convolution, the middle branch is an over-parameterised 3 × 3 convolution, i.e. k 3 × 3 convolutions, and the right-hand part is a jump connection containing the BN layer. Here both the 1×1 convolution and the 3×3 convolution are deep convolutions (the number of groupings g is equal to the number of input channels). The dot-convolution module in the figure consists of two branches. The left branch is an over-parameterised 1 × 1 convolution, consisting of k 1 × 1 convolutions. The right branch is a jump connection containing a BN layer. Replacing the normal convolution in the PANet network with a deeply separable convolution provides better detection of small targets while ensuring accuracy and real-time detection.

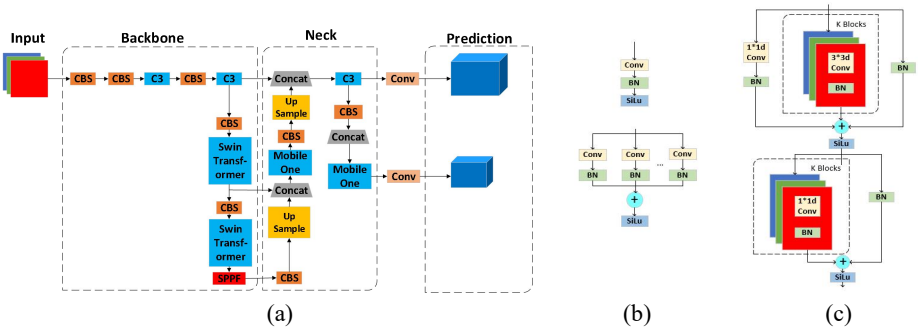


Fig. 5. (a) Representing the SW-YOLO structure, (b) Representing over-parameterised structures, (c) Representing depth-separable convolutional structure

3.4 Improved IOU loss function

As can be seen from the classification loss function in equation (1), when the prediction probability value ($p \geqslant 0.5$) is larger, the loss value of positive samples is smaller than that of negative samples, making it difficult for positive samples to affect the training parameters, thus affecting the detection accuracy. In order to solve the above problem, this paper adopts EIOU based on the cross-entropy loss function, the expression is shown in equation (4), the penalty term of EIOU is based on the penalty term of CIOU to split the influence factor of aspect ratio to calculate the length and width of target frame and anchor frame respectively, the loss function contains three parts: overlap loss, centre distance loss, width and height loss. The first two parts continue the approach in CIOU, but the width-height loss directly minimises the difference between the width and height of the target and anchor boxes, allowing for

faster convergence. The regression accuracy is higher and can solve the positive and negative sample imbalance problem. Therefore, EIOU Loss is chosen to replace CIOU Loss in this paper.

$$L_{EIOU} = L_{IOU} + L_{dis} + L_{asp} = 1 - IOU + \frac{\rho^2(b,b^{gt})}{c^2} + \frac{\rho^2(w,w^{gt})}{c_w^2} + \frac{\rho^2(h,h^{gt})}{c_h^2} \quad (4)$$

4 Experiments

4.1 Dataset

In this paper, we use the BCCD public dataset [17], which is a small-scale dataset of relatively homogeneous blood shapes for cell detection with 768 train images and 106 test images, generated by annotating the images in the dataset. The BCCD dataset contains a small number of white blood cells (WBC), platelets and a large number of red blood cells (RBC).

4.2 Experimental configuration

The experimental configuration: Ubuntu 9.4.0 as the operating system, Pytorch 1.7.1 as the deep learning framework, E-2136 as the CPU, 16 GB as the RAM, and Quadro P5000 as the GPU with 16 G of video memory.

The details regarding the training parameters for this paper are provided in Table 1.

Table 1. Training parameter settings for the BCCD dataset

Parameter Name	Parameter values	Parameter Name	Parameter values
Batch size	16	Epoch	133
Learning rate	0.01	Momentum	0.937
Box loss	0.05	Cls loss	0.5
Anchors	(10,13), (16,30), (33,23), (30,61), (62,45), (59,119)	Obj loss	1.0

4.3 Evaluation indicators

To verify the effectiveness of the proposed SW-YOLO blood cell detection algorithm, the SW-YOLO algorithm in this paper is compared with Faster R-CNN, YOLOv3, YOLOv4 and YOLOv5 networks to evaluate the algorithm detection performance in terms of experimental detection accuracy (Precision), recall and mean average precision (mAP). This is shown in equations (5) to (7).

$$\text{Precision} = \frac{TP}{TP+FP} \quad (5)$$

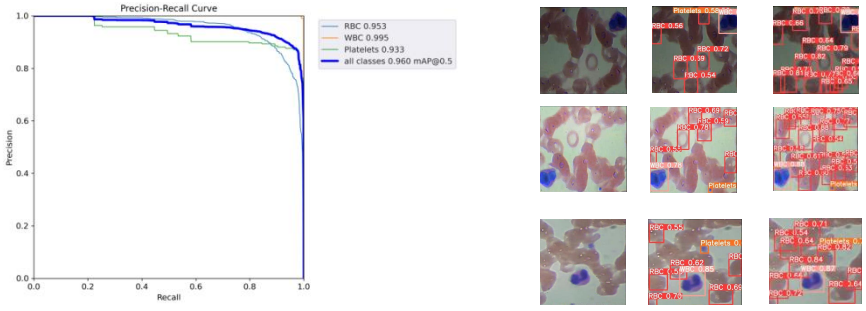
$$\text{Recall}=\frac{\text{TP}}{\text{TP+FN}} \tag{6}$$

$$\text{mAP}=\frac{1}{\text{N}}\int_0^1 \text{P}_\text{i}(\text{R})\,\text{dR} \tag{7}$$

TP denotes the number of correctly identified blood cell classes, FP denotes the number of incorrectly identified blood cell classes, FN denotes the number of incorrectly identified non-blood cell classes, and TN denotes the number of correctly identified non-blood cell classes. Therefore, Precision is the probability of correctly predicting blood cell classes as a percentage of predicted blood cell classes, and Recall is the probability of correctly predicting blood cell classes as a percentage of true blood cell classes. N in equation (7) represents the number of total blood cell categories, so mAP is the mean value of the average accuracy of all categories.

4.4 Results

The P-R curves of the SW-YOLO algorithm for the three classes of blood cells are shown in Figure 6(a), and the mAP values for these classes of blood cells can be obtained by calculating the area under the P-R curve.



(a) (b) (c) (d)

Fig. 6. (a) Representing the P-R curves of SW-YOLO, (b) Original image of blood cells, (c) Blood cell images under other detection algorithms, (d) Blood cell images under SW-YOLO detection algorithm

Figure 6(c) and Figure 6(d) represent the detection results of blood cells under the other algorithms and the SW-YOLO algorithm respectively. From the figures, we can see that our proposed SW-YOLO does solve the problem of missed detection of other algorithms. Preliminary proof that our algorithm is effective.

This paper evaluates the performance of the SW-YOLO model in comparison to the traditional single-stage and two-stage target detection models, and the results of the comparison are illustrated in Table 2.

Table 2. Comparative analysis experiments

Method	Input size	WBCS	RBCS	Platelets	mAP	FPS

Faster-RCNN[4]	1000×600	0.803	0.722	0.770	0.765	9.2
YOLOv3[7]	608×608	0.914	0.829	0.774	0.839	34.5
YOLOv4[8]	640×640	0.930	0.798	0.813	0.847	36.1
TE-YOLOF-B3[19]	416×416	0.987	0.873	0.898	0.919	43
ISE-YOLO[20]	416×416	0.965	0.927	0.896	0.857	34.5
YOLOv5[13]	640×640	0.977	0.838	0.873	0.896	56.7
SW-YOLO	640×640	0.995	0.953	0.933	0.960	43.4

Table 2 provides the detection speed as FPS (frames per second), which represents the amount of images that the algorithm can process in one second. Examining Table 2, it is clear that the SW-YOLO not only has reduced complexity but also has significant advantages in precision compared to the majority of target detection networks. The mAP metric showed that the SW-YOLO network model is 19.5, 12.1, 11.3, 4.1, 10.3 and 6.4 percentage points better than Faster-R-CNN [4], YOLOv3 [7], YOLOv4 [8], TE-YOLOF-B3 [18], ISE-YOLO [19] and YOLOv5 [13] respectively. The results of the comparative analysis demonstrate that this algorithm outperforms all the other models in terms of the mAP metric. The SW-YOLO have a higher detection rate and are faster by 0.4 and 9 FPS than the TE-YOLOF-B3 and ISE-YOLO algorithms, which are utilized to carry out detection on the publicly accessible blood cell dataset BCCD. The SW-YOLO show an increase in both speed and accuracy of detection on the blood cell data set compared to the prior two network models. The algorithm is not as speedy as YOLOv5, and the processing time of a single image increases by 0.07s with a frame rate of 43.4, which is sufficient to support real-time detection needs on hardware terminals while still improving the false and missed detections issue in the original network. The results of the experiment suggest that this algorithm has the best overall performance compared to other models, as it is able to increase the mAP while maintaining the detection speed.

The performance of the SW-YOLO model is compared with the latest YOLO detection model and the comparative results are shown in Table 3.

Table 3. Comparative experiments with the latest technology

Method	Input size	WBCS	RBCS	Platelets	mAP	FPS	F1_curve	percision
YOLOv7	640×640	0.995	0.954	0.928	0.959	9.1	0.92	0.864
YOLOv8	640×640	0.995	0.960	0.945	0.967	54.8	0.93	0.807
SW-YOLO	640×640	0.995	0.953	0.933	0.960	43.4	0.92	0.851

SW-YOLO compared to YOLOv7, although the precision is 1.3% lower than YOLOv7, its F1_curve is the same as YOLOv7 and the map is 0.1% higher than YOLOv7, and the detection speed is 31.3 FPS higher than YOLOv7.

SW-YOLO compared to YOLOv8, although F1_curve is 0.01% lower than YOLOv8, map is 0.7% lower than YOLOv8 and detection speed is 11.4 FPS lower than YOLOv8, SW-YOLO has a 4.4% higher precision than YOLOv8.

In summary, our improved SW-YOLO compares favourably with the most advanced YOLO technology, although some of its performances are inferior to theirs, some are better than YOLOv7 and YOLOv8, so it is further proof that SW-YOLO can meet the current requirements in terms of detection accuracy and real-time performance.

5 Comparison with other datasets

To determine the generality of the SW-YOLO algorithm, other datasets are tested in this paper. Firstly, the LIDC-IDRI dataset [20], which is the same medically relevant lung CT dataset. The annotation of the images result in 3144 train set images and 349 test set images. The another is dataset is TT100K [21], which is also a small to medium sized target dataset. The TT100K dataset is dedicated to the recognition of traffic signs and after annotating the images contained in the dataset, we have 8816 images in the train set and 1712 images in the test set. The performance metrics under both datasets are shown in Figure 7.

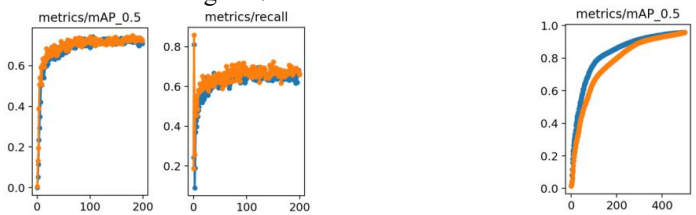


Fig. 7. (a) Comparison of the YOLOv5s algorithm and the SW-YOLO algorithm on the LIDC-IDRI dataset, (b) Comparison of the YOLOv5s algorithm and the SW-YOLO algorithm on the TT100K dataset. Where the blue line represents the original YOLOv5s algorithm and the orange line represents the SW-YOLO algorithm.

From Figure 7(a) we can see that the SW-YOLO algorithm is significantly more effective on the LIDC-IDRI dataset than YOLOv5s, and the experiments also show that SW-YOLO yields the mAP of 72%, which is higher than the 70% mAP of YOLOv5s. From Figure 7(b), we can see that the SW-YOLO algorithm does not work as well as YOLOv5s on the TT100K dataset, but it is not very bad either, and the experiments also show that the mAP of SW-YOLO is similar to that of YOLOv5s, both reaching 96%, so our proposed SW-YOLO is still effective.

6 Ablation Studies

In order to verify the effectiveness of each module and to analyze the impact of each module on the accuracy of the SW-YOLO algorithm, we conducted an ablation

experiment, using YOLOv5s is used as the base network to gradually add improvement points, setting the fuse Swin Transformer as A, only small and medium target blocks are detected as B, Replacing the ordinary convolution in PAN with the depth-separable convolution as C, and replacing CIOU loss function with EIOU loss function as D. The experimental results are shown in Table 4. According to the data from the ablation experiments in Table 4, the improvement of either module resulted in 4.1% mAP to 7.5% mAP increase over the YOLOv5s algorithm, F1_curve increased by 0.1% to 0.3% and precision increased by 0.3% to 6.4%. SW-YOLO show an average accuracy improvement of 6% over the original YOLOv5s network for the detection of 3 types of blood cells. It is highly practical.

Table 4. Ablation studies

Method	Input size	WBCS	RBCS	Platelets	mAP	F1_curve	percision
YOLOv5s [13]	640×640	0.977	0.838	0.873	0.896	0.90	0.858
0.93YOLOv5s+A	640×640	0.995	0.949	0.959	0.968	0.93	0.922
YOLOv5s +B	640×640	0.991	0.938	0.953	0.961	0.92	0.892
YOLOv5s +C	640×640	0.995	0.959	0.945	0.966	0.92	0.865
YOLOv5s +D	640×640	0.995	0.950	0.967	0.971	0.93	0.888
YOLOv5s +A+B+C	640×640	0.995	0.948	0.953	0.965	0.92	0.862
YOLOv5s +A+C+D	640×640	0.995	0.948	0.955	0.966	0.93	0.909
YOLOv5s +B+C+D	640×640	0.995	0.939	0.958	0.964	0.92	0.904
SW-YOLO	640×640	0.995	0.953	0.933	0.960	0.92	0.902

7 Conclusion

This paper proposes a SW-YOLO network model for BCCD blood cell detection. This paper first introduces the structure of YOLOv5s, followed by the SW-YOLO network structure, which includes fusing Swin Transformer, removing the large target detection structure, replacing the ordinary convolution in PAN with depth-separable convolution, and replacing the CIOU loss function with the EIOU loss function. In this paper, comparison experiments with other mainstream detection algorithms, comparison experiments with other datasets and ablation experiments are conducted respectively. The experimental results demonstrate that the SW-YOLO algorithm has excellent detection accuracy and real-time detection effect.

References

1. Chen, Y.: Study on the value of blood smear analysis in routine blood tests. China Pharmaceutical Guide. 16(01), 118-119 (2018).
2. Chan, L., Laverty, D., Smith,T.: Accurate measurement of peripheral blood mononuclear cell concentration using image cytometry to eliminate RBC-induced counting error. Journal of Immunological Methods. 388(01), 25-32 (2013).
3. Lehmann, T.M., Guld, M.O., Thies, C.: Content-based image retrieval in medical applications. Methods of information in medicine. 43(04), 354-361 (2004).
4. Ren, S.Q., He, K., Girshick, R.: Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. IEEE transactions on pattern analysis and machine intelligence. 39(6), 1137-1149 (2017).
5. Chen, D.H., Sun, S.R., Wang, Y.C.: Research on improved SSD algorithm for small target detection. Sensors and Microsystems. 42(03), 65-68 (2023).
6. Liu, J.C., Li, X.F., Liu, A.X.: Improved RetinaNet for UAV small target detection. Science Technology and Engineering. 23(01), 274-282 (2023).
7. Yan, W.J., Dai, J.H.: Traffic sign recognition based on improved YOLOv3. Journal of Wuhan Engineering Vocational Technology College. 35(01), 31-35 (2023).
8. Chen, Y.F., Yan, C.C., Zhou, C.: Improved YOLOv4-based vehicle detection for autonomous driving scenarios. Automation and Instrumentation. 38(01), 59-63 (2023).
9. Zheng, C.W., Lin, H.: A YOLOv5 helmet wearing detection method based on Swin Transformer. Computer Measurement and Control. 31(03), 15-21(2023).
10. Zhang, Y.F., Ren, W., Zhang, Z.: Focal and efficient IOU loss for accurate bounding box regression. Neurocomputing. 506, 146-157 (2022).
11. Ouyang, D., Huang, H., Li, J.: Improved yolov5s model for aerial image target detection algorithm. Fujian Computer. 39(05), 7-15 (2023).
12. He, K., Zhang, X., Ren, S.: Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 770-778. (2016).
13. Shah, S.A.R., Wu, W., Lu, Q.: AmoebaNet: An SDN-enabled network service for big data science. Journal of Network and Computer Applications. 119, 70-82 (2018).
14. Chollet, F.: Xception: Deep learning with depthwise separable convolutions. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1251-1258. (2017).
15. Lin, T.Y., Dollar, P., Girshick, R.: Feature pyramid networks for object detection. Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2117-2125. (2017).
16. Liu, S., Qi, L., Qin, H.: Path aggregation network for instance segmentation. Proceedings of the IEEE conference on computer vision and pattern recognition, 8759-8768. (2018).
17. https://github.com/Shenggan/BCCD_Dataset, last accessed 2018/2/24.
18. Xu, F., Li, X., Yang, H.: TE-YOLOF: Tiny and efficient YOLOF for blood cell detection. Biomedical Signal Processing and Control. 73, 103416 (2022).
19. Liu, C., Li, D., Huang, P.: ISE-YOLO: Improved Squeeze-and-Excitation Attention Module based YOLO for Blood Cells Detection. 2021 IEEE International Conference on Big Data, 3911-3916. (2021).
20. Lung Nodule Analysis 2016. Available at https://luna16.grand-challenge.org, last accessed 2020/10/20.
21. Huang, S.A.: Traffic sign detection based on improved YOLO model. Science and Technology Innovation. 18, 194-196 (2021).