

# Leukocyte Detection in Blood Smear Image Based on Improved YOLOv3

Jing Li <sup>1+</sup>, Junmin Wu <sup>2</sup>

<sup>1</sup> School of Software Engineering, University of Science and Technology of China, China

<sup>2</sup> School of Computer Science and Technology, University of Science and Technology of China, China

**Abstract.** The classification of leukocytes is an important indicator for the detection of a variety of blood diseases in routine blood tests. Although there have been many papers studying the detection of WBCs (white blood cell) or classification independently, few papers consider them together. In this paper, we propose an end-to-end white blood cell localization and classification method based on improved YOLOv3. Firstly, we utilize the k-means clustering algorithm to generate the anchor boxes suitable for WBC. Secondly, to identify white blood cells of different sizes, we use multi-scale predictions with YOLOv3 network structure. Experimental results on both the LISC dataset [1] and a dataset of 7500[2] leukocyte smear images of five categories demonstrate that the proposed improved-YOLOv3 can achieve efficient detection performance in terms of accuracy. The recognition accuracy of the two data sets is reached respectively 96.4% and 95.5%.

**Keywords:** leukocyte, detection, k-means clustering algorithm, multi-scale predictions.

## 1. Introduction

In recent years, with the rapid development of computer technology, CAD (computer aided diagnosis, CAD) technology has achieved rapid development in some medical developed countries, especially in the field involving medical imaging. Among them, the segmentation and classification of leukocyte in blood smear are research hotspots. Because microscopic examination of blood smears is of great value in the diagnosis and evaluation of various blood diseases such as leukemic. In general, leukocytes can be divided into five categories, eosinophilic, basophilic, neutrophilic, lymphocytes, and monocytes.

In recent years, researchers have proposed methods to classify leukocytes using computer image recognition. These methods are generally summarized into three parts: firstly, the blood smear made successfully is placed under a microscope and the microscopic image is captured. Secondly, the image segmentation algorithm is used to segment the microscopic image and obtain the corresponding single leukocyte image. Finally, the leukocyte image is classified by classification algorithm. With the development of deep learning, computer vision has also made a breakthrough with great influence. Through the deep learning model, the computer can extract features through training, so as to realize the classification of images, and with the continuous improvement of the model, the classification accuracy is also improving. Therefore, the application of computer vision in medical image processing has become a new trend, which has research significance and practical value.

In this paper, we propose an end-to-end model of leukocyte smear image object detection based on improved YOLOv3, accurately locating and identifying five types of leukocytes. We evaluate our model on two datasets. They are the LISC dataset [1] and a large leukocyte dataset. In this large dataset, leukocytes were subdivided into 12 categories, and there were 1500 pictures in each class of the training set. We select 7500

---

<sup>+</sup> Corresponding author. Tel.: + 18896516019.  
E-mail address: sa517157@mail.ustc.edu.cn.

pictures of leukocytes as the five-category dataset we needed, and manually marked the position information of leukocytes.

The main contributions of this work are as follows: we first use the improved YOLOv3 model, the end-to-end model, for leukocyte localization and classification. The advantage over the previous segmentation first and classification later model is that the effect of segmentation accuracy on classification is eliminated. K-means clustering algorithm is used to generate anchor boxes of leukocytes. Our model improves the multi-scale detection of YOLOv3, extending the original 3-scale detection to 4-scale detection. Our model performs well both on small and large datasets. This achieved 96.4% accuracy on the LISC dataset of 274 images and 95.5% accuracy on the five-category dataset of 7,500 images.

The rest of the paper is organized as follows: First, in Section 2, a review of the literature. Then, the proposed improved YOLOv3 model described in detail in Section 3. Finally, in Section 4, the conclusion of this research is presented.

## 2. Related Work and Background

### 2.1. Related Work

In the past few decades, many scholars have studied the classification of leukocytes by image recognition, focusing on the algorithm of cell segmentation and recognition. Cell segmentation is one of the most challenging tasks in medical image processing because of the complex biological appearance of cells, inconsistent staining and light changes [3].

Based on flow cytometry data provided by Beckman-Coulter [4], support vector machine (SVM) classifier is used to cluster the parameter data. In this study, the results show that, the accuracy rate for the dataset of 100 images is 86.6%. The disadvantage of this method is that the leukocyte pictures can only be classified into four categories and this dataset cannot generate images of WBCs for further image analysis and validation if there are changes in in-class staining, shape, cell illumination, or overlapping cells.

In order to solve this problem, a method to divide WBCs into five categories by using standard extreme learning machine (ELM) technology and fast correlation vector machine (Fast-RVM) is proposed in [5]. The cells were segmented by ELM method, and the recognition characteristics were established according to the threshold technology. The system was trained and tested using ELM and fast RVM classifier. The results show that the accuracy rate is 80%. Obviously, although this method can effectively reduce the effects brought about by illumination and staining, the accuracy is far from enough.

Nowadays, more and more artificial intelligence methods are used to leukocyte classification.

An automatic detection and classification technique for peripheral blood WBCs based on [6] was proposed. The relationship between red, blue and morphological operations was used to detect WBCs from the microscope images. Support vector machine (SVM) classifier used particle size characteristics (pair rotation invariant co-occurrence local binary pattern and PRICoLBP feature) to classify basophilic and eosinophilic cells from other WBCs. Then, convolutional neural network (CNNs) is used to extract high level features from WBCs. For classification, random forests were used to identify other types of leukocytes: lymphocytes, monocytes, and neutrophils. The classification accuracy using these features was 92.6%. However, some cells were not detected correctly.

Changhun proposed a CNN-based WBC classification model named W-Net [7] to accurately recognize WBC types. After a series of pre-processing such as cropping, resizing and normalizing, the images were extracted features and classified by W-Net. W-Net achieved an average accuracy of 97% on a dataset, obtained from The Catholic University of Korea, that includes 6,562 real images of the five WBC types. This model was also evaluated on the LISC public data which contains 254 images of five types and the average accuracy was 96%.

### 2.2. YOLOv3

YOLOv3 backbone evolved from Yolov2's darknet-19 to darknet-53, deepening the network layers and introducing cross-layer addition in ResNet. The structure of YOLOv3 is shown in Fig.1. YOLOv3 borrows from the idea of feature pyramid networks [8], where a small size feature map detects a large size object, while

a large size feature map detects a small size object. YOLOv3 outputs 3 feature graphs in total. The first feature graph is sampled 32 times, the second is sampled 16 times, and the third is sampled 8 times. The output dimension of the feature map is  $N \times N \times [3 \times (4 + 1 + C)]$ , and  $N \times N$  is the number of grid points of the output feature map. There are three Anchor boxes in total in each scale.

Following is the structure of YOLOv3.

	Type	Filters	Size	Output	
1	Convolutional	32	3 3	256 256	
	Convolutional	64	3 3/2	128 128	
	Convolutional	32	1 1		
	Convolutional	64	3 3		
	Residual			128 128	
	Convolutional	128	3 3/2	64 64	
	2	Convolutional	64	1 1	
		Convolutional	128	3 3	
		Residual			64 64
		Convolutional	256	3 3/2	32 32
	8	Convolutional	128	1 1	
		Convolutional	256	3 3	
Residual				32 32	
Convolutional		512	3 3/2	16 16	
8	Convolutional	256	1 1		
	Convolutional	512	3 3		
	Residual			16 16	
	Convolutional	1024	3 3/2	8 8	
4	Convolutional	512	1 1		
	Convolutional	1024	3 3		
	Residual			8 8	
	Avgpool		Global		
Connected		1000			
Softmax					

Fig. 1: YOLOv3 structure.

### 3. Materials and Method

#### 3.1. Feature Extraction Using Depth Residual Network

In the deep neural network model, the ability to extract features degrades as the number of layers increases. Residual neural Networks [9] offers a solution. For a stacking structure (made up of several layers), the characteristics learned when the input is  $x$  are denoted as  $H(x)$ . Now we hope that it can learn the residual  $F(x) = H(x) - x$ , so the original learning feature is  $F(x) + x$ . This is because residual learning is easier than learning the original features directly. ResNet can solve this problem by utilizing skip connections or shortcuts to jump over layers. Typical ResNet models are implemented with double or triple layer skips. ResNet is a CNN widely used in image processing.

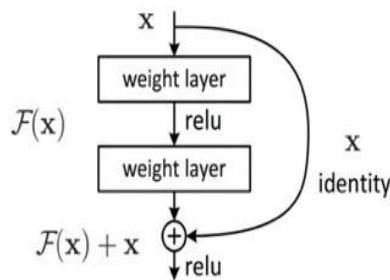


Fig. 2: Schematic diagram of residual block.

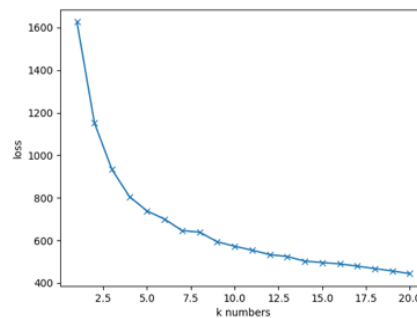


Fig. 3: The relationship between the number of anchor boxes and the k-means loss function.

In this paper, some modifications of the structure of YOLOv3 has been made. The proposed deep residual network structure is shown in Fig.2. In order to better extract the shallow features of the image, the depth of the residual network is increased after four times sampling. We use three residual blocks instead of the convolutional set before every prediction.

### 3.2. Using the Anchor Box to Predict the Leukocyte Boundary

The concept of anchor box was first proposed by Faster R-CNN[10], and then it was widely used in target detection models such as SSD, YOLOv2 and YOLOv3.

In this paper, k-means clustering algorithm will be used to generate anchor boxes suitable for leukocytes. But the distance measured to the centre of the cluster is no longer the standard Euclidean distance. We combine the distance formula with the IOU.

$$D(box, centroid) = 1 - IOU(box, centroid)$$

We found that after k was greater than 9, the loss function started to flatten out, so we still selected 9 anchor boxes. The relationship between the number of anchor boxes and the k-means loss function is shown in Fig.3.

### 3.3. Multi-Scale Networks Predict Leukocytes

Improved-YOLOv3 directly predicted the position and category of WBC by regression on the feature maps of multiple scales. The convolutional layer with different discrimination is divided into three branches, and WBC prediction was performed independently for each branch. Specifically, each branch is equipped with multiple convolutional layers, and for branches with resolution of  $8 \times 8$ ,  $16 \times 16$ ,  $32 \times 32$ , line up sampling is performed at twice the size. In order to enhance the representation ability of the feature pyramid, the up-sampling features and the corresponding size feature diagrams provided by the depth residual network in the previous section are concatenated. Therefore, the detection can be performed with contextual semantic information in each branch, while the three detection branches share features extracted from the residual network.

During the prediction phase, a three-dimensional tensor was predicted for the incoming blood smear image, including the leukocyte boundary box, confidence, and WBC class. Improved-YOLOv3 divides the characteristic graph into  $N \times N$  grids (different scales of characteristic graphs have different sizes of N), and its tensor can be expressed as  $N \times N \times [B \times (4 + 1 + 5)]$ , namely 4 boundary box offsets, 1 confidence and 5 leukocyte types. Different from YOLOv3, each grid no longer predicts three different boundary boxes. Through the leukocyte anchor boxes obtained before, B represents the number of candidate boxes predicted for each branch, corresponding to 2 32-fold samples, 4 16-fold samples and 2 8-fold samples.

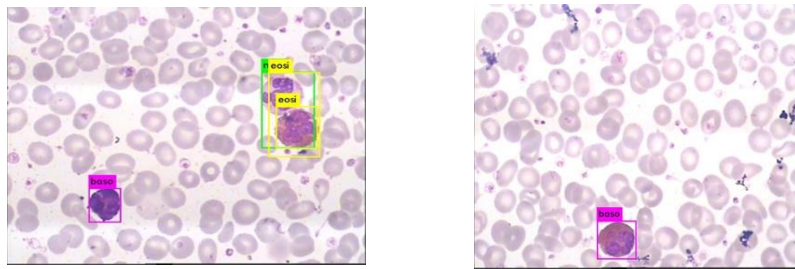
## 4. Experimental Results and Analysis

The experimental results of the proposed work and other related studies are shown in the table.1 below. The Improved-YOLOv3 is 0.4 percent more accurate than Jung et al. in LISC datasets. Meanwhile, the 91.5% accuracy obtained by training directly with the YOLOv3 network proved that Improved-YOLOv3 performed better. Moreover, our model also performed well on 7,500 five-category blood smears, with an accuracy of 95.5%

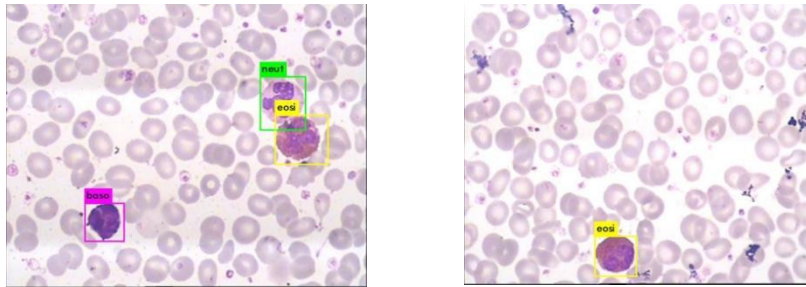
Table. 1: The performance of the proposed approach and existing approach on WBC detection

Models	Database	Baso	Eosi	Neut	Lymp	Mono	Accuracy
Agaian et al.2016[11]	ALL-IDBI	/	/	/	/	/	94%
Jung et al.2019[7]	LISC	96%	98%	100%	98%	88%	96%
YOLOv3	LISC	100%	85.7%	94.1%	88.9%	85.7%	91.5%
Improved-YOLOv3	LISC	100%	92.9%	100%	94.4%	92.9%	96.4%
Improved-YOLOv3	7500	99.0%	93.0%	97.5%	88.6%	99.3%	95.5%

Compared with the original model, we effectively solved the problem of inaccurate selection of candidate boxes and improved the accuracy of identifying and classifying leukocytes. As shown in Fig.4(a) below, multiple candidate boxes repeated the prediction and the candidate boxes inaccurately located eosinophilic leukocytes. Fig.4(b) wrongly predicted eosinophilic leukocytes into basophilic leukocytes. Improved-YOLOv3 improved the accuracy of the candidate box and improved the classification accuracy of WBC, as shown in Fig.5(a) and Fig.5(b) below.



4(a) 4(b)  
Fig. 4: The prediction of the YOLOv3 model.



5(a) 5(b)  
Fig. 5: The prediction of the Improved-YOLOv3 model.

## 5. Conclusions

An end-to-end white blood cell localization and classification method based on improved YOLOv3 was discussed in this paper. YOLOv3, a one-stage model, was used for the first time in leukocyte detection missions. Moreover, the accuracy of recognition on LISC dataset and a data set of 7500 images reached 96.4% and 95.5% respectively. The experimental result shows that the proposed method is superior to the existing method for the classification of leukocytes.

In the future, we will continue to study the accuracy and speed of leukocyte recognition and extend the current framework to evaluate additional data sets captured under various imaging conditions.

## 6. References

- [1] Rezatofghi, S. H., and Soltanian-Zadeh. Automatic recognition of five types of white blood cells in peripheral blood. *CMI and Graphics*. 35(4):333–343,2011.
- [2] Wang, Q., Bi, S., Sun, M., Wang, Y., Wang, D., Yang, S. Deep learning approach to peripheral leukocyte recognition. *PloS One*, 14(6): e0218808. <https://doi.org/10.1371/journal.pone.0218808>,2019.
- [3] H.T. Madhloom, S.A. Kareem, H. Ariffin, A.A. Zaidan, H.O. Alanazi and B.B. Zaidan. An Automated White Blood Cell Nucleus Localization and Segmentation using Image Arithmetic and Automatic Threshold. *Journal of Applied Sciences*, 10: 959-966,2010.
- [4] Adjouadi M, Zong N, Ayala M. Multidimensional pattern recognition and classification of white blood cells using support vector machines. *Particle & Particle Systems Characterization*. 22(2):107-18,2005.
- [5] Ravikumar S. Image segmentation and classification of white blood cells with the extreme learning machine and the fast relevance vector machine. *Artificial cells, nanomedicine, and biotechnology*. 44(3):985-9,2016.
- [6] Zhao J, Zhang M, Zhou Z, Chu J, Cao F. Automatic detection and classification of leukocytes using convolutional neural networks. *Medical & biological engineering & computing*. 55(8):1287-301,2017.
- [7] Jung, C., Abuhamad, M., Alikhanov, J., Mohaisen, A., Han, K., & Nyang, D. W-Net: A CNN-based Architecture for White Blood Cells Image Classification. *arXiv preprint arXiv:1910.01091*,2019.
- [8] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie. Feature pyramid networks for object detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2117–2125, 2017.

- [9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of CVPR, pages 770–778. [arXiv.org/abs/1512.03385](https://arxiv.org/abs/1512.03385),2016.
- [10] S. Ren, K. He, R. Girshick, and J. Sun. Faster R-CNN: Towards real-time object detection with region proposal networks. In NIPS, 2015.
- [11] Aгаian, S., Madhukar, M., & Chronopoulos, A. T. A new acute leukaemia- automated classification system. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*. 6(3), 303-314,2016.