# A deep learning model for detection of leukocytes under various interference factors

Meiyu Li[1], Lei Li[2], Shuang Song[3], Peng Ge[3], Hanshan Zhang[4], Lu Lu[5], Xiaoxiang Liu[6], Fang Zheng[1], Cong Lin[6]*, Shijie Zhang[7]* and Xuguo Sun[1]*

1 School of Medical Laboratory, Tianjin Medical University, Tianjin, China.
2 Clinical Laboratory, Tianjin Chest Hospital, No.261 Taierzhuangnan Road, Jinnan District, Tianjin 300350, PR China.
3 Department of Laboratory, Key Laboratory of Cancer Prevention and Therapy, National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin Clinical Research Center for Cancer, Tianjin 300060, P.R. China.
4 The Australian National University, Canberra, Australia.
5 Institute of Disaster Medicine, Tianjin University, Tianjin, China.
6 School of Intelligent Systems Science and Engineering, Jinan University, Zhuhai, China.
7 Department of Pharmacology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin, China.

*Corresponding author:
Cong Lin: School of Intelligent Systems Science and Engineering, Jinan University, Zhuhai, China. Email: conglin@jnu.edu.cn
Shijie Zhang: PhD, Department of Pharmacology, School of Basic Medical Sciences, Tianjin Medical University, Tianjin, China. Email: shijie.zhang@tmu.edu.cn
Xuguo Sun: Professor, School of Medical Laboratory, Tianjin Medical University, No.1 Guangdong Road, Hexi District, Tianjin 300203, PR China. E-mail: sunxuguo@tmu.edu.cn

# Supplementary materials

**Evaluation Criteria**

The main metric for evaluating the overall performance of our leukocyte detector is mean average precision (mAP) [http://mscoco.org/], which is a popular performance indictor in object detection domain. Given a set of pre-set IoU thresholds, the mAP is the mean value of average precision across all leukocyte types to be detected, which can be expressed as follows:

$$\text{mAP} = \frac{1}{|C|}\sum_{c\in C} AP_c \ , \qquad (1)$$

where $C$ is a label set of all leukocyte types in dataset, $|C|$ is the order or number of items of set $C$, and $AP_c$ is average precision across different an IoU thresholds set $U$ for leukocyte type $c$, which is a function of $TP_c^u$ and $FP_c^u$:

$$AP_c = \frac{1}{|U|}\sum_{u\in U} \frac{TP_c^u}{TP_c^u+FP_c^u} \ . \qquad (2)$$

Generally, the higher mAP value suggests the higher of recognition accuracy of the model for leukocytes in the image. However, the variance between APs of leukocyte types should not be too high. Otherwise, the accuracies of different types are imbalanced.

The mean average recall (mAR), which is another commonly used metric, is also employed as auxiliary indicator in our experiments. The average recall (AR) is mathematically expressed as:

$$AR_c = \frac{1}{|U|}\sum_{u\in U} \frac{TP_c^u}{TP_c^u+FN_c^u} \ . \qquad (3)$$

The mAR is the average recall (AR) of multiple types of leukocytes in the image, and the higher value represents the lower missed detection rate of leukocytes recognition by the model.

The specific meanings of $TP_c^u$, $FP_c^u$ and $FN_c^u$ are given as follows:

$TP_c^u$: The number of true positive detections which the detected bounding boxes exhibit a higher IoU overlapping than threshold $u$. True positive detections are objects that are correctly

classified as the groundtruth type $c$.

$FP_c^u$: The number of false positive detections that come from two type of false positive cases. One type of the false positive is the detected bounding boxes that exhibit higher IoU overlappings than threshold $u$, however, are incorrectly classified into other types rather than the groundtruth type $c$. Another type of the false positive is the detected bounding boxes that contain no underlying object or merely contain parts of the objects with a lower IoU overlapping than threshold $u$.

$FN_c^u$: The number of false negative detections, which is the number of groundtruth objects with no detection or with no covering bounding box shows a higher IoU overlapping than threshold $u$. In other words, the detector misses the targets.

A detection, regardless being correct or incorrect, is determined by whether its confidence is higher than the pre-set confidence threshold. Through all our experiments, the confidence threshold is pre-set empirically to 0.05; In the MS-COCO criteria, the IoU thresholds are a set of percentage values from the range of [0.5,0.95] with a step of 0.05; while in the more relaxed Pascal VOC criteria, the IoU is by default set to be 0.5.

# Supplementary Tables

**Supplementary Table 1:** Detailed transforms list of data augmentation and their parameters.

| Type of Transform | Parameters | Occurrence Probability |
|---|---|---|
| Flip | - | 0.5 |
| RandomRotate90 | - | 0.5 |
| RGBShift | 10 (Shift limit for each channel) | ~0.1 |
| HueSaturationValue | 20 (Shift limit for each channel) | ~0.1 |
| GaussNoise | - | ~0.05 |
| RandomBrightnessContrst | [0.1, 0.3] variation range | ~0.1 |
| RandomGamma | - | ~0.1 |

**Supplementary Table 2:** The number of images that interfere with leukocyte detection in our training set.

| Types | BG | EG | L | M | NG |
|---|---|---|---|---|---|
| Color casts on blood cell smear images | 5 | 129 | 58 | 36 | 2 |
| Low illuminative intensity of blood cell smear images | 23 | 68 | 31 | 44 | 2 |
| Giant platelets | 58 | 88 | 65 | 63 | 109 |
| Incorrect imaging focal length of cell images | 21 | 33 | 31 | 18 | 11 |
| Images containing dyes or other impurities | 47 | 109 | 269 | 217 | 246 |
| Overlapping leukocytes | 2 | 22 | 12 | 19 | 82 |
| Degenerated leukocytes | 3 | 11 | 4 | 2 | 2 |
| Excessively high phosphate buffer solution pH >6.8 | 90 | 235 | 82 | 62 | 77 |
| Excessively low phosphate buffer solution pH<6.4 | 4 | 117 | 76 | 32 | 9 |

NG: neutrophil, BG: basophil, EG：eosinophil, L：lymphocyte, M：monocyte

**Supplementary Table 3：** Multiple leukocytes of our training set.

| Types | One | Two | Three | Four | Five | Six | Seven | Ratio of multiple leukocytes |
|---|---|---|---|---|---|---|---|---|
| Numbers | 3365 | 933 | 294 | 76 | 24 | 5 | 2 | 39.64% |

**Supplementary Table 4：** Cascade R-CNN and ensemble models correctly detect the number of pictures of leukocytes in different scenarios.

| **Models** | **pH >6.8** | **Overlapping** | **Incomplete** | **Impurities** |
|---|---|---|---|---|
| Cascade R-CNN | 412/447 | 25/38 | 84/86 | 156/170 |
| Ensemble | 426/447 | 36/38 | 85/86 | 164/170 |

a/b: a represents the number of pictures correctly detected by the model in each scene; b represents the total number of pictures in each scene.