



A deep learning method and device for bone marrow imaging cell detection

Jie Liu^{1#}, Ruize Yuan^{2,3#}, Yinhao Li^{2,3#}, Lin Zhou¹, Zhiqiang Zhang⁴, Jidong Yang⁴, Li Xiao^{2,3,5}

¹Department of Laboratory, The Seventh Medical Center of Chinese PLA General Hospital, Beijing, China; ²School of Computer and Control Engineering, University of Chinese Academy of Sciences, Beijing, China; ³Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China; ⁴Hanyuan Pharmaceutical Co., Ltd., Beijing, China; ⁵Ningbo Huamei Hospital, University of Chinese Academy of Sciences, Ningbo, China

Contributions: (I) Conception and design: L Xiao, J Liu; (II) Provision of study materials or patients: J Liu, L Zhou; (III) Collection and assembly of data: L Zhou, Z Zhang, J Yang; (IV) Methodology Development: L Xiao, R Yuan; (V) Data analysis and interpretation: Y Li, R Yuan, L Xiao; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Li Xiao. No. 6 Zhongguancun South Road, Haidian District, Beijing 100090, China. Email: andrew.lxiao@gmail.com or xiaoli@ict.ac.cn.

Background: Morphological analysis of bone marrow cells is considered as the gold standard for the diagnosis of leukemia. However, due to the diverse morphology of bone marrow cells, extensive experience and patience are needed for morphological examination. automatic diagnosis system through the comprehensive application of image analysis and pattern recognition technology is urgently needed to reduce work intensity, error probability and improves work efficiency.

Methods: In this article, we establish a new morphological diagnosis system for bone marrow cell detection based on the deep learning object detection framework. The model is based on the Faster Region-Convolutional Neural Network (R-CNN), a classical object detection model. The system automatically detects bone marrow cells and determines their types. As specimens have severe long-tail distribution, i.e., the frequency of different types of cells varies dramatically, we proposed a general score ranking loss to solve such a problem. The general score ranking loss considers the ranking relationship between positive and negative samples and optimizes the positive sample with a higher classification probability value.

Results: We verified this system with 70 bone marrow specimens of leukemia patients, which proved that it can realize intelligent recognition with high efficiency. The software is finally integrated into the microscope system to build an augmented reality system.

Conclusions: Clinical tests show that the response speed of the newly developed diagnostic system is faster than that of trained diagnostic experts.

Keywords: Morphological analysis; diagnosis of leukemia; deep learning; object detection

Submitted Jan 05, 2022. Accepted for publication Feb 18, 2022.

doi: [10.21037/atm-22-486](https://doi.org/10.21037/atm-22-486)

View this article at: <https://dx.doi.org/10.21037/atm-22-486>

Introduction

Leukemia is a malignant cancer in which abnormal white blood cells (leukemia cells) in bone marrow diffusely

proliferate. Leukemia cells replace normal bone marrow tissue and invade the surrounding blood, resulting in changes in the number and quality of peripheral white blood cells (1,2). Leukemia is a common malignant disease,

[^] ORCID: [0000-0002-5354-9633](https://orcid.org/0000-0002-5354-9633).

especially in children and adolescents. There are various methods used to detect leukemia, such as myelomorphology, cytochemical staining, and immunophenotyping (3-5). Among them, myelomorphology is considered as the gold standard for the diagnosis of leukemia. However, due to the diverse morphology of bone marrow cells, extensive experience and patience are needed for morphological examination. At present, there are some bone marrow cell image analysis systems that can be applied to morphometric analysis, bone marrow cytology inspection reports, and chromosome analysis reports, which greatly reduces work intensity and error probability, and improves work efficiency (6,7). However, advanced work is still needed an automatic classification, recognition, and position of bone marrow cell images and the classification and recognition of bone marrow diseased cells through the comprehensive application of image analysis and pattern recognition technology (8-10). Thus, constructing blood disease diagnosis equipment that integrates artificial intelligence (AI) and big data analysis functions can provide support for the accurate diagnosis of leukemia and improve the effectiveness of the medical service system.

AI uses sophisticated algorithms to learn features from a large volume of healthcare data, and then uses the obtained insights to assist clinical practice (11-13). Therefore, AI has the potential to revolutionize disease diagnosis and management by performing classification and reviewing of huge amounts of images rapidly. Deep convolutional neural network (CNN) has allowed for significant gains in the ability to classify images and detect objects in images (14-16). Based on this, Kermany *et al.* established a diagnostic tool for the screening of patients with common treatable blinding retinal diseases, which demonstrated comparable classification performance to that of human experts (17). Using CNN, mapping of the relationships between the image of bone marrow cells and classification can be established to realize the effective recognition of bone marrow morphology (18-20). Thus, establishing an efficient algorithm and training through a large number of images can break the dependence of bone marrow cell morphology analysis on human recognition and realize intelligent recognition. Recently, some pioneer works have been proposed to use deep learning algorithm to perform bone marrow cells detection on whole-slide images (21). Our work is to detect bone marrow cells on the field of view images under the microscope and successfully integrate the model into the software and hardware system of the augmented reality microscope. We present the

following article in accordance with the MDAR reporting checklist (available at <https://atm.amegroups.com/article/view/10.21037/atm-22-486/rc>).

Methods

Data collection

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Ethics Committee Of Seventh Medical Center of Chinese PLA General Hospital (No. 2022-25) and informed consent was taken from all individual participants.

We collected and labeled real bone marrow cell images under the microscope, stored the data and labeling information in a unified format, and randomly divided the training set and testing set. In this embodiment, 4,451 real images of bone marrow cells on the Hyde star HDS-BFS high-speed micro scanning image system were collected, with a resolution of 4,000×3,000. The labeling tool is used to mark the rectangular boxes of all cells in the picture, generate the coordinates of the vertices at the upper left corner and the lower right corner of the rectangular box in the pixel coordinate system, and label the corresponding categories. Labeling file generated by the label tool, where each<object></object> represents a target, that is, a cell, and <name></name> represents the category of the target, <xmin></xmin>, <ymin></ymin> represent the x-axis coordinates and y-axis coordinates of the vertices at the upper left corner, <xmax></xmax> and <ymax></ymax> represent the x-axis coordinates and y-axis coordinates of the vertices at the upper left corner, respectively.

Data preprocessing

During the preprocessing stage, we filtered the samples of the training set, removed the samples without labeling and non-labeled categories, and used categories with number of cell samples greater than 100 to construct the dataset. A total of 20 classification categories were obtained and summarized in *Table 1*, which were as follows: eosinophilic lobulated neutrophils, promyelocytes, primordial plasma cells, promyelocytes, megaloblastic erythrocytes, heterotypic lymphocytes, megaloblastic erythrocytes, neutrophils, neutrophils, neutrophils, immature lymphocytes, immature plasma cells, immature monocytes, late immature erythrocytes, primordial lymphocytes, mature lymphocytes,

Table 1 Summary of the number of different types of cells

Cell type	Training (count of cells)	Testing (count of cells)
Eosinophilic lobulated nuclei	168	55
Basophilic normoblast	185	52
Protoplasma cell	331	101
Promyelocyte	398	124
Megalocytocytes	464	146
Lymphocyte atypia	472	126
Neutrophils	629	200
Polychromatic erythroblast	720	213
Neutrophilic metamyelocyte	942	283
Neutrophilic myelocyte	982	279
Naive lymphocyte	1,090	368
Naive plasma cell	1,095	334
Neutrophilic granulocyte band form	1,104	347
Naive monocyte	1,259	359
Metarubricyte	1,378	427
Protolymphocyte	1,531	365
Mature lymphocyte	1,605	465
Myeloblast	2,137	638
Abnormal promyelocytic granulocytes	2,528	667
Protomonocyte	2,784	830
Sum	21,820	6,379

primordial granulocytes, abnormal promyelocytes, primordial monocytes. The training dataset was augmented by 5 data enhancement methods: image horizontal flip, vertical flip, image rotation, image translation, and image adding Gaussian noise.

Object detection model

The object detection model is based on the Faster Region-Convolutional Neural Network (R-CNN). Faster R-CNN is composed of the region proposal network (RPN) and Fast R-CNN, in which RPN is used to select candidate target boxes and Fast R-CNN is used for accurate target classification and regression.

The loss function of the classification task is as follows:

$$L_{cls} = -\frac{1}{N} \sum_{i=1}^N \log(p_i) \quad [1]$$

The classification task is a binary classification task, that is, the probability that the current prediction region belongs to the foreground or background.

The loss function of the regression task is as follows:

$$L_{reg} = \sum_{i \in (x,y,w,h)} smooth_{L1}(t_i^u - v_i) \quad [2]$$

$$smooth_{L1}(x) = \begin{cases} 0.5x^2|x| < 1 \\ |x| - 0.5 \text{ otherwise} \end{cases} \quad [3]$$

The regression task uses the loss of $smooth_{L1}$ function, where x, y, w, h are the coordinates of the center point of the candidate rectangular box and the length and width of the rectangular box. In order to ensure the translation invariance and the consistency of the length and width of coordinates, it is parameterized to generate 4-dimensional vectors t . The specific calculation method is as follows:

$$t_x = \frac{(x - x_a)}{w_a}, \quad t_y = \frac{(y - y_a)}{h_a} \quad [4]$$

$$t_w = \log\left(\frac{w}{w_a}\right), \quad t_h = \log\left(\frac{h}{h_a}\right) \quad [5]$$

$$t_x^* = \frac{(x^* - x_a)}{w_a}, \quad t_y^* = \frac{(y^* - y_a)}{h_a} \quad [6]$$

$$t_w^* = \log\left(\frac{w^*}{w_a}\right), \quad t_h^* = \log\left(\frac{h^*}{h_a}\right) \quad [7]$$

Where subscript a indicates the template box, and the superscript $*$ indicates the real rectangular box.

The complete loss function of the RPN is the weighted sum of the above 2 task loss functions, which is expressed as follows:

$$L(\{p_i\}, \{t_i\}) = \frac{1}{N_{cls}} \sum_i L_{cls}(p_i, p_i^*) + \lambda \frac{1}{N_{reg}} \sum_i p_i^* L_{reg}(t_i, t_i^*) \quad [8]$$

In the Fast R-CNN part, the Fast R-CNN uses the candidate bounding box predicted by the RPN to sample the features of the same size under the region of interest (ROI) pooling, specifically 7es. After flattening the feature map, the prediction results are obtained through a series of full connection layers. Classification prediction and regression prediction are carried out for the obtained feature vectors. The two parts are similar to the classification regression task

calculation method of the RPN. The classification task in the invention is a two-class classification task, that is, foreground cell and background. The regression task is the same as the method of calculating 4-dimensional regression parameters by the RPN. The specific loss function is expressed as follows:

$$L(p, u, t^u, v) = L_{cls}(p, u) + \lambda[u \geq 1]L_{loc}(t^u, v) \quad [9]$$

Where p is the softmax probability distribution predicted by the classifier, u is the real classification label corresponding to the target, t^u is the corresponding category predicted by the regressor corresponding to the candidate bounding box, $(t_x^u, t_y^u, t_w^u, t_h^u)$ are regression parameters, and v is the regression parameter corresponding to the real bounding box (v_x, v_y, v_w, v_h) .

Both the RPN and Fast R-CNN need to use the CNN to extract image features. In the structure of Fast R-CNN, these two parts share a feature extraction network.

Generalized average precision loss

Due to the difference in the occurrence frequency of different categories of bone marrow-like cells, we proposed a general score ranking loss to evaluate the positive and negative samples through the classification probability value, rather than using the heuristic sampling training method, which is more conducive to solving the problem of long tail data distribution. In the traditional ranking loss, the positive samples are generally ranked before the negative samples, without considering the ranking relationship between the two kinds of samples. There are two tasks to sort losses using scores. In the scoring task, the positive and negative samples are divided according to the threshold, and the model aims to make the classification probability value of the positive samples greater than the negative samples. The sorting task is constrained within the positive samples to make the positive samples with high confidence have a large classification probability value.

Firstly, specify that the Intersection-of-Union (IOU) of the prediction box and GTBox exceeds the threshold as a positive sample, otherwise it is a negative sample. Define the negative sample set as N and the positive sample set as P .

The score difference between the two prediction frames is calculated by:

$$\forall i, j, x_{\{i,j\}} = -(s(b_i; \theta) - s(b_j; \theta)) = -(s_i - s_j) \quad [10]$$

Where s is the score of the rectangular box, and we define the step function:

$$H(x) = \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases} \quad [11]$$

The ranking of general scores is as follows:

$$L_{RS} = loss_{RS}(i) - loss_{RS}^*(i) \quad [12]$$

It is divided into two parts: the sorting loss of the original sequence and the sorting loss within the positive sample.

For the original sorting sequence, the sorting loss is:

$$loss_{RS}(i) = \frac{\sum_{j \in N} H(x_{ij})}{r(i)} + \lambda \frac{\sum_{j \in P} H(x_{ij})(1 - y_j)}{r_p(i)} \quad [13]$$

Where r is the sorting position, λ is the weighting coefficient, H is the step function, and y is the IOU value of the prediction box and GTbox.

For the internal loss of positive samples (sorted loss), since there are no negative samples, the score loss is 0. The sort loss is that:

$$loss_{RS}^*(i) = \lambda \frac{\sum_{j \in P} (H(x_{ij}) \& y_j \geq y_i)(1 - y_j)}{\sum_{j \in P} (H(x_{ij}) \& y_j \geq y_i)} \quad [14]$$

Where H is the step function and y is the IOU value of the prediction box and GTBox. It can be seen from the above formula that for positive samples with no change in relative ranking position, the original ranking loss is the same as that sorted by ranking, which will offset each other in the final loss calculation. It will weaken the impact on the correctly ranked positive samples. However, the positive samples whose original ranking is after the negative samples have a large score loss, which promotes the model to enhance the detection ability of these positive samples. The rank sort loss (22) proposed here is a special case of the generalized loss where $\lambda=1$, and in our model we set $\lambda=0.5$ through hyperparameter search.

Training details

During training, the momentum algorithm is selected as the gradient descent method, where momentum = 0.9, the learning rate is initialized by 0.001, the weight attenuation is 0.0001, and batch size = 1 for each iteration. The learning rate attenuation adopts linear attenuation, and the learning rate attenuation every 4 rounds is 0.3 times that of the original, with model iterative training of 20 rounds.

After training, the test set is preprocessed with the same augmentation method. The predicted target category and the coordinates of the upper left and lower right corners of the rectangular box of the target will be obtained and displayed in the image sample in a visual manner.

Statistical analysis

We compute the metrics over all types of cells and take their average as the final performance. Recall, Precision, F1-score is adopted as the clinical evaluation metrics for cell classification. AP@50 is adopted as the computer vision evaluation metric for cell detection.

Hardware devices

The hardware devices include an optical microscope (collecting the field of view under the microscope), digital equipment (camera, transforming optical images into electronic signals, i.e., digital images), and host (used to install and execute programs and models for cell recognition).

Microscope

Bone marrow smear specimens can be directly observed by the eyepiece after optical magnification of the microscope.

Image digitizing equipment

This equipment adopts a high-definition digital camera and is connected to the computer through USB interface. The digital microscope integrates the optical microscope and the image digitizing device to directly output digital microscope images.

Mainframe

The computer completes the functions of digital image processing, storage, diagnosis report generation, analysis, and statistics in the medical micro image processing system. It is the core of the medical micro image processing system.

Peripherals

The image of the system and the results of processing and analysis can be displayed on the display, or printed on the report form through a color image printer. The process is described below:

- (I) Place the bone marrow specimen on the stage of the optical microscope, and adjust the focal length to achieve a clear image effect;
- (II) The optical signal is transformed into a color digital image through the camera for further

processing. The white balance, exposure, and other parameters of the camera can be adjusted to make the digital image clear and recognizable;

- (III) The mathematical model of bone marrow morphological pattern classification is established based on deep learning technology. Through the constructed recognition model, the cells entering the field of vision are classified and detected;
- (IV) Combined with the expert diagnosis indexes of clinical bone marrow morphology, the system test is carried out, and the digital bone marrow image database based on deep learning is established to realize computer-aided diagnosis;
- (V) Through the above hardware device, combined with neural network and bone marrow cell morphology, a bone marrow morphology detection system based on deep learning can be constructed to realize the automatic detection and recognition of bone marrow cells.

Results

Workflow

The bone marrow specimens of 70 patients with leukemia from the Hematology Department of The Seventh Medical Center of Chinese PLA General Hospital after examination were used for digital scanning of whole slides, a database of bone marrow morphological characteristics was established, and data collection of the image database was realized. We randomly selected 80% of the images as the training set and the remaining 20% of images as the testing set. We filtered the categories of cells and kept those cell categories with number of cell samples greater than 100. A total of 20 classification categories were obtained. To handle the diversity of feature distributions, we enriched the data augmentation process by combining 5 data enhancement methods: image horizontal flip, vertical flip, image rotation, image translation, and image adding Gaussian noise. As shown in *Figure 1*, we developed our model based on the Faster R-CNN. In view of the decline of recognition accuracy caused by many types of bone marrow cell images and unbalanced sample distribution, a generalized average precision loss (G-AP loss) was designed so that the model could decrease negative samples effectively in training. Finally, combined with the diagnosis indicators of clinical bone marrow morphology experts, the results were systematically verified and tested, and a new type of

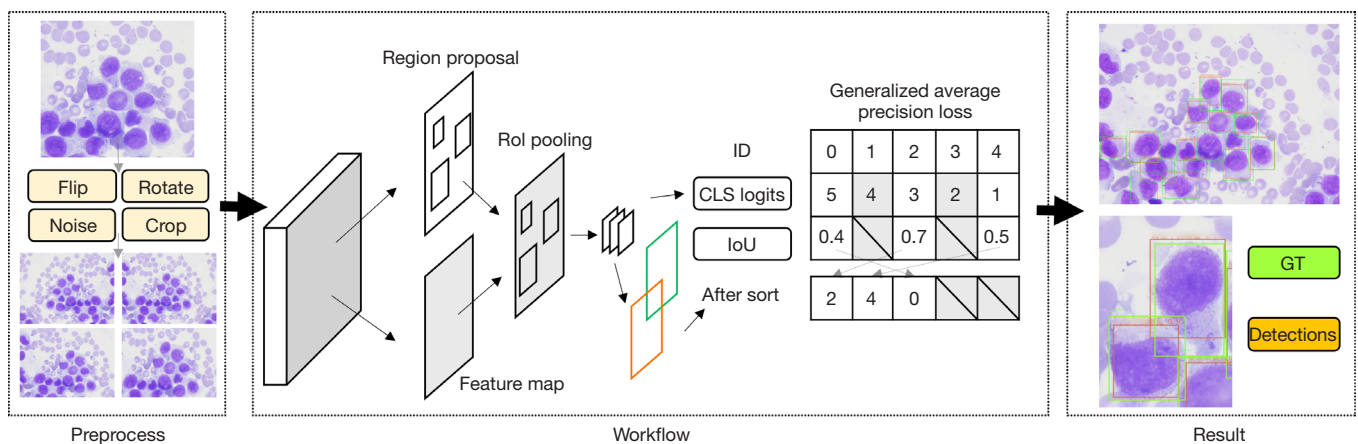


Figure 1 Workflow of the system. The feature maps are extracted by the backbone of ResNet-101. Based on Faster R-CNN, the model detects the bounding boxes of the objects and the corresponding confidence. The results are firstly sorted by their classification logits, followed by the generalized average precision loss to relieve the effect caused by class imbalance (Wright's stain, $\times 1,000$). R-CNN, Region-Convolutional Neural Network; CLS, classification; IoU, Intersection of Union; GT, ground truth.

hematological disease diagnosis model was established.

Model performance

The testing set contained 798 images with 6,847 labeled cells. We performed ablation studies and reported the results of 4 models: baseline (the original Faster R-CNN model), ranksort [Faster R-CNN with original ranksort ($\lambda=0.5$), and our methods (rank sort $\lambda=0.5$ + enriched data augmentation). We summarized the recall and precision values in *Table 2*, and the AP@50 and F1-score values in *Table 3*. The results showed that our method consistently improved performance when adding ranksort, data augmentation, and generalized ranksort to the baseline. Compared to the baseline, although our method resulted in a slight drop of recall ($\sim 4\%$), it improved the precision by 26.4%, the F1-score by 12.1%, and the AP@50 by 3%. Our model achieves a recall of 0.710, precision of 0.496, AP@50 of 0.533, F1-score of 0.575.

We also drew the average precision (AP) curve to demonstrate the superiority of our methods. As shown in *Figure 2*, our approach demonstrates better AP values among all thresholds compared to the baseline and baseline with original rank sort loss. We also present the visualization results in *Figure 3*. Our method effectively reduces the missed detected targets and is thus more precise compared to the baseline.

Augmented reality microscope

Finally, we integrated the software into the microscope system to build an augmented reality system. As shown in *Figure 4*, the system includes an optical microscope (collecting the field of view under the microscope), digital equipment (camera, transforming optical images into electronic signals, i.e., digital images), and host (used to install and execute programs and models for cell recognition). The deep learning model is established on the host for computer-aided diagnosis. Clinical tests showed that with assistance from the newly developed diagnosis system, we could finish 200 fields of analysis within 16 minutes, 4.8 s for each image. As a reference, a well-trained expert needs around 40–50 s to perform a diagnosis for each field.

Discussion

We developed a new morphological diagnosis system for bone marrow cells. The model is based on the Faster R-CNN. To deal with the diversity of feature distributions, we enriched the data augmentation process by combining 5 data enhancement methods: image horizontal flip, vertical flip, image rotation, image translation, and image adding Gaussian noise. Due to the highly imbalanced data distribution, we designed the generalized rank loss, which provided extra optimization to make positive samples with larger prediction score. Our model has successfully

Table 2 Precision and recall of different types of bone marrow cells

Cell type	Recall			Precision		
	Baseline	Ranksort	Our method	Baseline	Ranksort	Our method
Eosinophilic lobulated nuclei	0.673	0.764	0.818	0.319	0.477	0.523
Basophilic normoblast	0.615	0.558	0.596	0.215	0.349	0.425
Protoplasma cell	0.604	0.505	0.535	0.293	0.389	0.47
Promyelocyte	0.71	0.96	0.935	0.247	0.358	0.497
Megalocytocytes	0.863	0.678	0.664	0.448	0.553	0.522
Lymphocyte atypia	0.913	0.738	0.738	0.4675	0.567	0.633
Neutrophils	0.745	0.61	0.67	0.423	0.502	0.51
Polychromatic erythroblast	0.662	0.577	0.606	0.399	0.438	0.473
Neutrophilic metamyelocyte	0.643	0.534	0.562	0.301	0.385	0.375
Neutrophilic myelocyte	0.699	0.573	0.649	0.327	0.429	0.464
Naive lymphocyte	0.851	0.766	0.793	0.543	0.526	0.649
Naive plasma cell	0.838	0.731	0.769	0.456	0.451	0.479
Neutrophilic granulocyte band form	0.758	0.686	0.761	0.468	0.476	0.521
Naive monocyte	0.624	0.538	0.524	0.265	0.337	0.366
Metarubricyte	0.778	0.74	0.803	0.543	0.57	0.579
Protolymphocyte	0.8	0.822	0.833	0.468	0.53	0.554
Mature lymphocyte	0.68	0.712	0.753	0.398	0.397	0.395
Myeloblast	0.668	0.611	0.66	0.376	0.453	0.43
Abnormal promyelocytic granulocytes	0.841	0.849	0.877	0.512	0.41	0.489
Protomonocyte	0.81	0.711	0.725	0.462	0.502	0.493
Mean	0.740	0.680	0.710	0.393	0.457	0.496

integrate the model into the software and hardware system of the augmented reality microscope, which has a significant potential values in clinical usage.

Although experiments demonstrate the significant performance improvement of our proposed methods over most of the metrics, there are still some limitations of our current system. Firstly, our model is trained on a dataset from only 70 patients with 4,451 fields. Thus, a variety of bone marrow cell types are not included in the training set. Secondly, these cell types are only from a single center, meaning that our model may be biased towards the particular population and machines from the center dataset. Thirdly, our model cannot effectively learn rare samples, therefore we had to eliminate cell types with numbers of samples less than 100 before training.

Our hardware devices also need to be improved. Firstly, due to manual operation issues, different areas of the glass slide have different thicknesses, so it is necessary to constantly adjust the focal length of the microscope in the process of observation to keep the field of vision clear, and focusing will take some time. It also takes time to find the right view during initial adjustment. In addition, as some cells are not correctly detected, it requires technicians to spend some time to correct the wrong predictions to record the results.

In the future, we will continue to improve the model by involving more specimens and multi-center datasets for training and design few-shot learning methods to learn rare samples. We will also introduce an autofocus microscope to improve the system's efficiency further.

Table 3 AP@50 and F1-scores of different types of bone marrow cells

Cell type	AP@50			F1-score		
	Baseline	Ranksort	Our method	Baseline	Ranksort	Our method
Eosinophilic lobulated nuclei	0.498	0.644	0.692	0.4330	0.587	0.638
Basophilic normoblast	0.358	0.308	0.468	0.3180	0.43	0.496
Protoplasma cell	0.381	0.334	0.367	0.3950	0.44	0.5
Promyelocyte	0.431	0.431	0.458	0.3670	0.445	0.535
Megalocytocytes	0.626	0.526	0.504	0.5900	0.609	0.584
Lymphocyte atypia	0.775	0.643	0.653	0.6180	0.641	0.681
Neutrophils	0.496	0.445	0.51	0.5400	0.551	0.579
Polychromatic erythroblast	0.442	0.374	0.409	0.4980	0.498	0.531
Neutrophilic metamyelocyte	0.356	0.33	0.394	0.4100	0.447	0.45
Neutrophilic myelocyte	0.416	0.398	0.478	0.4450	0.491	0.541
Naive lymphocyte	0.730	0.625	0.691	0.6630	0.624	0.714
Naive plasma cell	0.604	0.525	0.571	0.5910	0.558	0.59
Neutrophilic granulocyte band form	0.508	0.482	0.586	0.5790	0.562	0.618
Naive monocyte	0.323	0.323	0.336	0.3720	0.414	0.431
Metarubricyte	0.591	0.582	0.68	0.6400	0.644	0.673
Protolymphocyte	0.628	0.657	0.669	0.5900	0.644	0.665
Mature lymphocyte	0.437	0.468	0.488	0.5020	0.51	0.518
Myeloblast	0.435	0.438	0.489	0.4810	0.521	0.521
Abnormal promyelocytic granulocytes	0.709	0.685	0.699	0.6370	0.553	0.628
Protomonocyte	0.527	0.483	0.528	0.5880	0.589	0.587
Mean	0.517	0.485	0.533	0.513	0.539	0.575

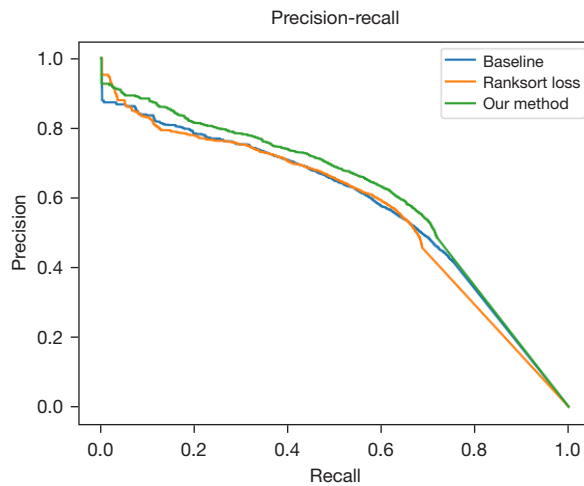


Figure 2 AP curves of different methods. X-axis is the recall value and y-axis is the precision value of cell detections under develop thresholds. AP, average prevision.

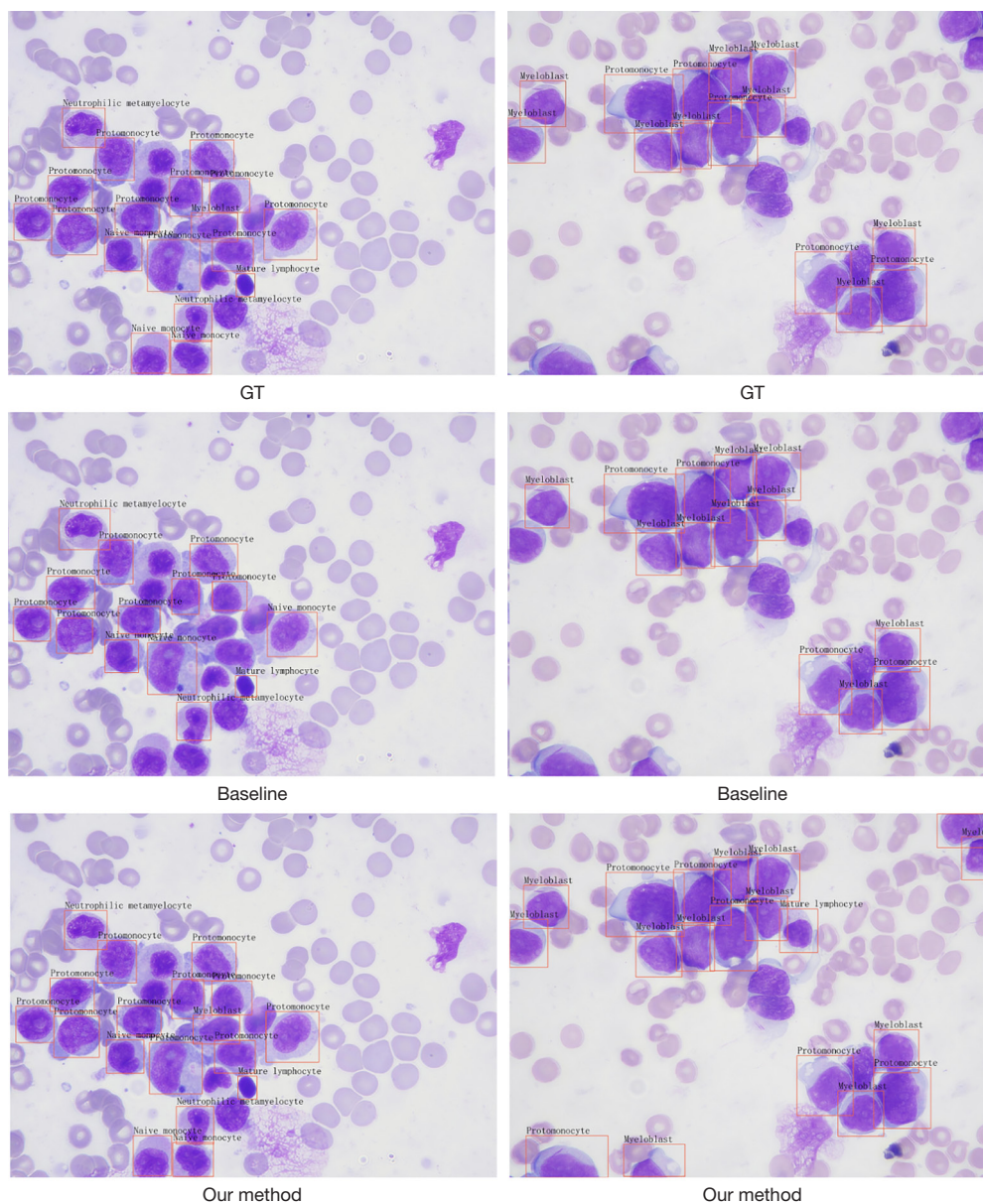


Figure 3 Comparison of visualization results for different methods. Both type and prediction scores are shown (Wright's stain, $\times 1,000$). GT, ground truth.

Conclusions

In this article, we establish a new morphological diagnosis system of bone marrow cells based on the deep learning object detection framework. The model is based on the Faster R-CNN (23), and we proposed a general score ranking loss to solve the problem of long tail data distribution. We also verified this system with 70 bone marrow specimens of leukemia patients, containing 4,451

bone marrow fields, which proved that it can realize intelligent recognition with high efficiency. The software is finally integrated into the microscope system to build an augmented reality system, with clinical test show that the newly developed diagnosis system can respond more rapidly comparing with a well-trained expert for diagnosis. Thus, establishing an efficient algorithm and training through a large number of images can break the dependence of bone marrow cell morphology analysis on human recognition and

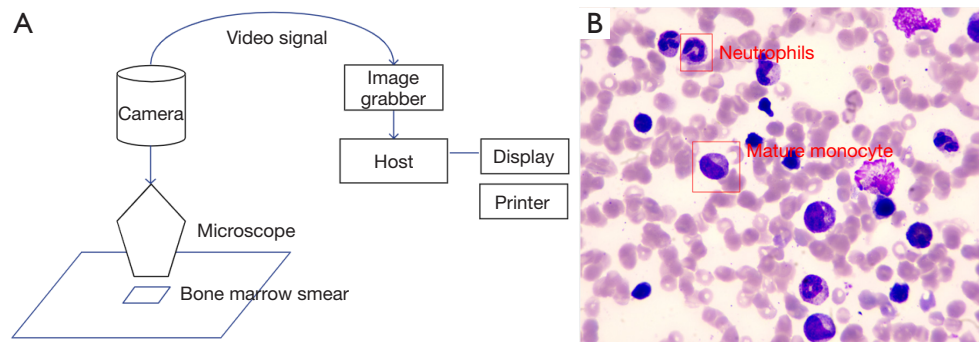


Figure 4 Framework of the whole system of augmented reality microscope. (A) Structure of the augmented reality microscope system; (B) a sample image displayed in the field of view (Wright's stain, $\times 1,000$).

realize intelligent recognition.

Acknowledgments

We thank Lingxia Han, Jianxia Xu, Yulan Wang, and Xin Yin for helping collect the data.

Funding: This work was supported by the National Science Foundation of China (grant No. 31900979 to LX), CCF-Tencent Open Fund (to LX), and 2019 Medical Big Data and Artificial Intelligence R & D Project (2019MBD-048 to JL).

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at <https://atm.amegroups.com/article/view/10.21037/atm-22-486/rc>

Data Sharing Statement: Available at <https://atm.amegroups.com/article/view/10.21037/atm-22-486/dss>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://atm.amegroups.com/article/view/10.21037/atm-22-486/coif>). ZZ and JY are from Hanyuan Pharmaceutical Co., Ltd. The other authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by Ethics Committee of Seventh Medical Center of Chinese PLA

General Hospital (No. 2022-25) and informed consent was taken from all individual participants.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Pelcovits A, Niroula R. Acute Myeloid Leukemia: A Review. *R I Med J* (2013) 2020;103:38-40.
2. Thomopoulos TP, Bouhla A, Papageorgiou SG, et al. Chronic myelomonocytic leukemia - a review. *Expert Rev Hematol* 2021;14:59-77.
3. Masilamani V, Devanesan S, AlSalhi MS, et al. Fluorescence spectral detection of acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML): A novel photodiagnosis strategy. *Photodiagnosis Photodyn Ther* 2020;29:101634.
4. Geng X, Zhang M, Wang X, et al. Selective and sensitive detection of chronic myeloid leukemia using fluorogenic DNAzyme probes. *Anal Chim Acta* 2020;1123:28-35.
5. Palmieri R, Buccisano F, Maurillo L, et al. Current strategies for detection and approach to measurable residual disease in acute myeloid leukemia. *Minerva Med* 2020;111:386-94.
6. Asnafi AA, DerisZayeri Z, Shahrabi S, et al. Chronic

- myeloid leukemia with complex karyotypes: Prognosis and therapeutic approaches. *J Cell Physiol*2019;234:5798-806.
7. Koczkodaj D, Popek-Marciniak S, Zmorzyński S, et al. Examination of clonal evolution in chronic lymphocytic leukemia. *Med Oncol*2019;36:79.
 8. Warnat-Herresthal S, Perrakis K, Taschler B, et al. Scalable Prediction of Acute Myeloid Leukemia Using High-Dimensional Machine Learning and Blood Transcriptomics. *iScience*2020;23:100780.
 9. Eckardt JN, Bornhäuser M, Wendt K, et al. Application of machine learning in the management of acute myeloid leukemia: current practice and future prospects. *Blood Adv*2020;4:6077-85.
 10. Wang J, Dao FT, Yang L, et al. Characterization of somatic mutation-associated microenvironment signatures in acute myeloid leukemia patients based on TCGA analysis. *Sci Rep* 2020;10:19037.
 11. Coombes CE, Abrams ZB, Li S, et al. Unsupervised machine learning and prognostic factors of survival in chronic lymphocytic leukemia. *J Am Med Inform Assoc*2020;27:1019-27.
 12. Li Z, Lam YW, Liu Q, et al. Machine Learning-Driven Drug Discovery: Prediction of Structure-Cytotoxicity Correlation Leads to Identification of Potential Anti-Leukemia Compounds. *AnnuIntConf IEEE Eng Med BiolSoc*2020;2020:5464-7.
 13. Niazi MKK, Parwani AV, Gurcan MN. Digital pathology and artificial intelligence. *Lancet Oncol* 2019;20:e253-61.
 14. Suman G, Panda A, Korfiatis P, et al. Convolutional neural network for the detection of pancreatic cancer on CT scans. *Lancet Digit Health* 2020;2:e453.
 15. Roy D, Panda P, Roy K. Tree-CNN: A hierarchical Deep Convolutional Neural Network for incremental learning. *Neural Netw*2020;121:148-60.
 16. Zhang Y, Lin H, Yang Z, et al. Neural network-based approaches for biomedical relation classification: A review. *J Biomed Inform* 2019;99:103294.
 17. Kermany DS, Goldbaum M, Cai W, et al. Identifying Medical Diagnoses and Treatable Diseases by Image-Based Deep Learning. *Cell* 2018;172:1122-1131.e9.
 18. Ouyang N, Wang W, Ma L, et al. Diagnosing acute promyelocytic leukemia by using convolutional neural network. *ClinChimActa*2021;512:1-6.
 19. Bibi N, Sikandar M, Ud Din I, et al. IoMT-Based Automated Detection and Classification of Leukemia Using Deep Learning. *J HealthcEng*2020;2020:6648574.
 20. Zhang C, Wu S, Lu Z, et al. Hybrid adversarial-discriminative network for leukocyte classification in leukemia. *Med Phys* 2020;47:3732-44.
 21. Wang CW, Huang SC, Lee YC, et al. Deep learning for bone marrow cell detection and classification on whole-slide images. *Med Image Anal* 2022;75:102270.
 22. K OksuzK, Cam BC, Akbas E, et al. Rank & Sort Loss for Object Detection and Instance Segmentation. *ICCV* 2021;pp. 3009-3018
 23. Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Trans Pattern Anal Mach Intell*2017;39:1137-49.
- (English Language Editor: C. Betlazar-Maseh)

Cite this article as: Liu J, Yuan R, Li Y, Zhou L, Zhang Z, Yang J, Xiao L. A deep learning method and device for bone marrow imaging cell detection. *Ann Transl Med* 2022;10(4):208. doi: 10.21037/atm-22-486