# Bone Marrow Cell Recognition: Training Deep Object Detection with A New Loss Function

Dehao Huang<sup>1</sup>, Jintao Cheng<sup>3</sup>, Rui Fan<sup>4,5</sup>, Zhihao Su<sup>1</sup>, Qiongxiong Ma<sup>1,\*</sup>, Jie Li<sup>2,\*</sup>,

<sup>1</sup>Guangdong Provincial Key Laboratory of Nanophotonic Functional Materials and Devices,

School of Information and Optoelectronic Science and Engineering, South China Normal University,

Guangzhou 510006, China.

<sup>2</sup>Department of Hematology, Nanfang Hospital, Southern Medical University, Guangzhou 510515, China.

<sup>3</sup>School of Physics and Telecommunication Engineering, South China Normal University,

Guangzhou 510006, China.

<sup>4</sup> College of Electronic and Information Engineering, Tongji University, Shanghai 201804, P. R. China.
 <sup>5</sup> Shanghai Research Institute for Intelligent Autonomous Systems, Shanghai 201210, P. R. China.

Email: huangdehao919@gmail.com, chengjintao@unity-drive.com, rui.fan@ieee.org, suuuzh@sina.com, maqx@m.scnu.edu.cn, ljjean@163.com

Abstract—For a long time, bone marrow cell morphology examination has been an essential tool for diagnosing blood diseases. However, it is still mainly dependent on the subjective diagnosis of experienced doctors, and there is no objective quantitative standard. Therefore, it is crucial to study a robust bone marrow cell detection algorithm for a quantitative automatic analysis system. Currently, due to the dense distribution of cells in the bone marrow smear and the diverse cell classes, the detection of bone marrow cells is difficult. The existing bone marrow cell detection algorithms are still insufficient for the automatic analysis system of bone marrow smears. This paper proposes a bone marrow cell detection algorithm based on the YOLOv5 network, trained by minimizing a novel loss function. The classification method of bone marrow cell detection tasks is the basis of the proposed novel loss function. Since bone marrow cells are classified according to series and stages, part of the classes in adjacent stages are similar. The proposed novel loss function considers the similarity between bone marrow cell classes, increases the penalty for prediction errors between dissimilar classes, and reduces the penalty for prediction errors between similar classes. The results show that the proposed loss function effectively improves the algorithm's performance, and the proposed bone marrow cell detection algorithm has achieved better performance than other cell detection algorithms.

*Index Terms*—Bone marrow smear, Deep learning, Loss function, Cell detection algorithm, YOLOv5.

## I. INTRODUCTION

The morphological examination of bone marrow cells plays a vital role in the diagnosis of blood diseases. It can identify the morphological characteristics of blood diseases such as acute leukemia, hemolymphoid tumors, and dysplasia [1], [2]. However, the analysis of cell morphology

\* Corresponding author.

is mainly done manually by experienced hematologists now. This process is not only cumbersome and labor-intensive but also subjective factors that will bring negative effects [3]. Therefore, a computer-assisted quantitative automatic analysis system is essential, and the cell detection algorithm as its core has aroused the interest of many researchers.

In recent years, the research on cell detection algorithms can be divided into traditional methods based on image processing and shallow machine learning, and methods based on deep learning. The general steps of traditional methods include cell segmentation, feature extraction, and cell classification. In some previous studies, image processing algorithms such as intensity clustering, watershed transformation, and adaptive thresholding were used as segmentation methods [4], traditional classifying algorithms such as Support Vector Machine (SVM) and Naive Bayes Classifier were used as cell classifiers [5], [6]. Although they have achieved excellent results, they are still not enough to apply automatic analysis of bone marrow smear. Traditional methods have the limitations of low learning ability and the need for artificial features [4]. Algorithms based on deep learning do not need to manually design features but learn features from a large amount of data and have powerful feature extraction capabilities. At the same time, deep learning methods can also be mixed and manually adjusted to maintain strong feature extraction capabilities while maintaining strong feature extraction capabilities. Presently, in some studies on cell classification and detection algorithms based on deep learning, they have achieved better results than traditional methods [4], [7]–[10]. Therefore, this paper focuses on a bone marrow cell detection algorithm based on deep learning.

Currently, cell detection algorithms based on deep learning develop steadily. Kutlu et al. used Faster-RCNN for end-toend cell detection in 5 classes and achieved excellent results [7]. Wang et al. used YOLOv3 and Single Shot MultiBox Detector(SSD) for end-to-end cell detection in 11 classes and achieved outstanding performance [10]. In the existing re-

This work was supported by Special Funds for the Cultivation of Guangdong College Students' Scientific and Technological Innovation ("Climbing Program" Special Funds) (No. pdjh2021b0134); Key-Area Research and Development Program of Guangdong Province (No. 2020B090922006); Featured Innovation Project of Guangdong Education Department (No. 2019KTSCX034); Science and Technology Program of Guangzhou (No. 202002030165, 2019050001).



Fig. 1. Including the following content: the structure of the proposed bone marrow cell detection algorithm, sub-modules, and the process of using samples to train the model. CSP module means Cross Stage Partial module, SPP module means Spatial Pyramid Pooling module, FPN means module of Feature Pyramid Networks, PAN means module of Pixel Aggregation Network, NMS means Non-maximum suppression, and CIoU means Complete IoU [11]. Part (a) is the image of the local area of the bone marrow smear, Part (b) is the sample label corresponding to the image (including cell location, class), and Part (c) is the structure diagram of the proposed method. The sub-modules of the proposed algorithm can be divided into the following three categories: Normal core module of YOLOv5, Effective core module of YOLOv5, and Proposed loss function, corresponding to the three color labels at the bottom of the figure. Both the Normal core module of YOLOv5 and the Effective core module of YOLOv5 has outstanding advantages in detecting bone marrow cells. The following is the algorithm training process: 1. the image input algorithm is forwarded to the Prediction part to obtain the detection result of the algorithm, 2. the loss part inputs the detection result to the loss function and the sample label corresponding to the image, 3. finally backpropagates and optimizes the model parameters.

search, the accuracy of the cell detection algorithm is sufficient for the commercial automatic analysis system of peripheral blood smears. However, the cell detection algorithm is still not effective enough to be used in an automatic bone marrow smear analysis system. (The bone marrow cell detection algorithm in the paper refers to the cell detection algorithm used to detect bone marrow cells.) The cell detection algorithm based on deep learning for bone marrow cell detection has the following three difficulties:

 Bone marrow cells classes are divided according to the series and stages of the cells, and the number of classes exceeds 30. The classes of adjacent stages belonging to the same series have a high degree of similarity so that it is difficult to distinguish. In contrast, the classes belonging to different series have apparent differences so that it is easier to distinguish. The above-mentioned inter-class relationship contains information that helps the algorithm classify cells, but it is still difficult for the algorithm to make full use of this information.

- In bone marrow smears, the cells are densely distributed, and there are many adherent cells. The algorithm should have sufficient cell segmentation capabilities.
- 3) The classification of part of cell classes needs to be combined with the context information of the surrounding cells in the image for classification judgment. The algorithm should have the ability to extract global context information.

In view of the above three problems, the main contributions of this paper include the following two points:

 This paper proposes a novel loss function. The finegrained classes are defined as finely divided classes according to the principle of different series and different stages. The coarse-grained classes are defined as roughly divided classes according to the principle of merging fine-grained classes of the same series and adjacent stages. A coarse-grained class often contains more than one fine-grained class, and the algorithm directly predicts the fine-grained classes. The proposed loss function will increase the penalty for the model's prediction of coarse-grained classification errors during algorithm training, try to use the similarity between finegrained classes belonging to the same coarse-grained class to strengthen the supervision of model feature extraction.

2) This paper proposes a bone marrow cell detection algorithm based on YOLOv5 network [11], trained with the novel loss function. The proposed algorithm can segment cells effectively because of powerful feature extraction capabilities, and also can extract context information because of a sub-module that can expand the receptive field. Compared with the existing cell detection algorithms, the proposed bone marrow cell detection algorithm has superior performance.

#### II. RELATED WORKS

Cell detection algorithms based on deep learning can be divided into two types. One type of algorithm includes two steps of segmentation and classification: the two-step algorithm. The other type of algorithm completes the end-to-end detection: the end-to-end algorithm. Then the following are analyzed separately.

The two-step cell detection algorithms are relatively common, which first uses the cell segmentation algorithm to segment the image of a single cell from the image and then uses the cell classification algorithm to classify each single cell image. For example, the method in [12] uses a threshold segmentation algorithm to segment the single-cell image and then uses the Alexnet image classification model to classify the single-cell image. The method in [13] uses Faster-RCNN network to segment the single-cell image and then uses Alexnet to perform fine-grained classification. For cell detection tasks, the segmentation and classification algorithms included in the two-step algorithm are relatively independent and clear to optimize separately. The excellent segmentation algorithm can be combined with the excellent classification algorithm [4], [8] for cell detection. However, it will have the following limitations when the two-step algorithm is used to detect bone marrow cells: 1. the classification of single cells is challenging to combine context information, 2. the training process is cumbersome, 3. the detection speed is low.

With the rapid improvement of end-to-end object detection network performance, many researchers have begun to use end-to-end cell detection algorithms that are mainly based on end-to-end target detection networks. For example, the method in [7] uses the Faster-RCNN network to implement an endto-end cell detection algorithm and trains them used transfer learning methods. The method in [10] is based on YOLOv3 and SSD to achieve end-to-end cell detection algorithms. Compared with the two-step cell detection algorithm, the end-to-end cell detection algorithm has the advantages of a simple training process and fast running speed. What's more, it is relatively easy to combine the contextual information. At present, the main limitation of the end-to-end cell detection algorithm is that it cannot make full use of the characteristics of the bone marrow cell detection task, such as the similarities between the types of bone marrow cells mentioned in Chapter I.

#### **III. METHODS**

The YOLOv5 network is a classic one-stage detection network, and it is the latest work of the YOLO architecture [14] series. Compared with the previous work YOLOv4 network, YOLOv5 network has achieved better accuracy and faster inference speed, and nearly 90% has reduced the volume of model parameters.

This paper proposes a bone marrow cell detection algorithm based on the YOLOv5 network. Figure 1 shows the following information: the structure of the algorithm model, essential sub-modules, and training method. The following content of this chapter introduces two types of essential sub-modules, corresponding to the Effective core modules and Proposed loss function in the Figure 1.

## A. Effective Core Module

The following four sub-modules in the proposed method have four advantages in bone marrow cell detection tasks:

- 1) SPP module [15]. The algorithm adds the SPP module to the Neck part, which helps the feature map merge local and global features. SPP can extract context-related information of bone marrow cells.
- 2) Mosaic data enhancement. Lack of data is a major challenge in the bone marrow cell detection task. The algorithm adds Mosaic data enhancement at the input end. It uses four pictures to be spliced by random scaling, random cropping, and random arrangement to enrich the data set.
- 3) Adaptive anchor boxes calculation. Bone marrow cells are mainly circular, and most of the cells are similar in shape and size. The algorithm adds an adaptive anchor boxes calculation at the input end, making it easy to find the best anchor boxes to improve the model's ability to locate cells.
- 4) CIoU loss [16]. In the labeled samples of the bone marrow cell detection task, the rectangular boxes of the labeled cells are often unable to frame the cells due to manual labeling accurately. However, it can be ensured that the center of labeled boxes and the aspect ratio of human-labeled data is approximately the same as ground truth. The proposed algorithm uses CIoU loss, which gains the advantage in considering the overlap area, center point distance, and aspect ratio, making the penalty for the model more effective.

## B. Proposed Loss Function

The YOLOv5 network originally calculates the loss of each labeled ground-truth bounding box using Eq.1, including three parts, namely positioning loss, confidence loss, and classification loss. The way to calculate classification loss is as shown in Eq.2, where  $S^2$  is the number of grid cells, B is the number of anchor boxes that each grid cell shall predict,  $I_{ij}^{obj}$  represents whether the labeled ground-truth bounding box of the current calculation loss belongs to the j anchor box of the i grid cell, and classes represent all the categories of the sample,  $\delta$  means the function of Sigmoid, $\hat{p}_i(c)$  is the true probability of class c. In contrast,  $p_i(c)$  is the predicted probability of class c. The classification loss part calculates the sum of the classification loss of the labeled ground-truth bounding box and all its corresponding anchor boxes.

$$Loss = Loss_{box} + Loss_{obj} + Loss_{cls} \tag{1}$$

$$Loss_{cls} = \sum_{i=0}^{S \times S} \sum_{j=0}^{B} I_{ij}^{obj} \times \sum_{c \in classes} [\widehat{p}_i(c) \log(\delta(p_i(c))) + (1 - \widehat{p}_i(c)) \log(1 - \delta(p_i(c)))]$$
(2)

The experiment found that if we only train the ability of the proposed algorithm to detect the position of bone marrow cells, the algorithm performs exceptionally well, with accuracy and recall reaching over 99%. Therefore, the limitation of the proposed bone marrow cell detection algorithm lies in the classification of bone marrow cells instead of segmentation.

Based on the above judgment, the weights of each part of the loss function of the YOLOv5 network were adjusted first. A basic weights parameter  $\alpha$  is set to satisfy  $\alpha > 1$ , and the three-part weights of the loss function of the YOLOv5 network is adjusted according to Eq.3 to make it more sensitive to classification loss, which make algorithm focus more on learning classification problems.

$$Loss = Loss_{box} + Loss_{obj} + \alpha Loss_{cls}$$
(3)

Next, try to optimize the loss function by using the similarity between the fine-grained classes of bone marrow cells. In the fine-grained classes, the cell classes of the same series at different stages have common characteristics. However, when calculating the classification loss value, the prediction errors of different series of cells and the prediction errors of the same series of cells are equivalent to the penalty effect of the model, which is not conducive to the supervised model to learn more useful features. The proposed loss tries to use the following methods to solve the above problems: Using the proposed loss function Eq.4 to replace the classification part of the loss function Eq.3.  $\gamma$  in Eq.4 is the same as in Eq.5. When the class of the labeled ground-truth bounding box is inconsistent with the coarse classification class predicted by the j anchor box of i grid cell, the classification loss of the anchor box will be multiplied by the weights  $(1 + \beta)$  to satisfy  $\beta > 0$ .

$$Loss_{cls} = \sum_{i=0}^{S \times S} \sum_{j=0}^{B} I_{ij}^{obj} (1+\gamma) \times \sum_{c \in classes} [\widehat{p}_i(c) \log(\delta(p_i(c))) + (1-\widehat{p}_i(c)) \log(1-\delta(p_i(c)))]$$
(4)

$$\gamma = \begin{cases} \beta & predict_{coarse} \neq target_{coarse} \\ 0 & predict_{coarse} = target_{coarse} \end{cases}$$
(5)

In Eq.3 and Eq.4,  $\alpha$  and  $\beta$  are two hyperparameters. Increasing  $\alpha$  can increase the model's attention to classification, and increasing  $\beta$  can increase the penalty weight for prediction errors of coarse-grained classes.

#### IV. EXPERIMENTAL RESULTS AND DISCUSSION

The bone marrow cell labeling data set we used for training and testing came from the Southern Hospital of Southern Medical University and was collected by optical microscope and camera. The data set included 2,549 labeled images from 144 patients, 15,642 labeled individual cells. The size of each image is 1024 x 684, and the total number of types of bone marrow cells is 36. (An example is shown in Figure 2.) To ensure a low correlation between the training set and the test set, we divide the data set with patients as the division unit, the training set contains data of 117 patients, and the test set contains data of 27 patients.



Fig. 2. Example of data set

The data set included 36 fine-grained classes, including lymphocyte, Promyelocyte, megaloblasts, Metamyelocyte, megaloblasts, late megaloblasts, etc. Based on the principle of the same series, adjacent stages, the existence of common features, and the need for cell counting when diagnosing blood diseases, we merged 36 fine-grained classes and obtained 14 coarsegrained classes(An example is shown in Figure 3).



Fig. 3. Example of fine-grained classes merged into coarse-grained classes. Polychromatic normoblast and Orthochromatic normoblast belonging to the coarse-grained classes include Polychromatic normoblast and Orthochromatic normoblast belonging to the fine-grained classes, and Myelocyte and Metamyelocyte belonging to the coarse-grained classes include Myelocyte and Metamyelocyte belonging to the fine-grained classes.

Method and Hyperparameters		mAP@0.5 fine-grained classification	mAP@0.5 coarsed-grained classification
SSD [10]		21.18%	37.38%
Faster-RCNN [7]		34.97%	56.83%
YOLOv5 + Normal loss function <sup>1</sup>		44.65%	66.00%
YOLOv5 + Class weighted loss function <sup>2</sup>	$\alpha = 2.50$	48.93%	66.68%
	$\alpha = 2.75$	48.48%	66.77%
	$\alpha = 3.00$	48.76%	66.28%
	$\alpha = 3.25$	48.50%	67.00%
	$\alpha = 3.50$	49.68%	66.14%
YOLOv5 + Proposed loss function <sup>3</sup>	$\alpha = 2, \ \beta = 1$	51.14%	69.71%

TABLE I Experimental results

<sup>1</sup> The normal loss is the original loss function of YOLOv5 network.

<sup>2</sup> The class weighted loss is the loss function that increases the weight of the classification part.

<sup>3</sup> The proposed loss is the loss function proposed in the paper

We use the Pytorch deep learning framework and NVIDIA RTX 2070 GPU for training and testing in the following experiments. When experimenting with the methods of other papers, we use the hyperparameters proposed in those papers. When experimenting with the proposed bone marrow cell detection algorithm in this paper, the following hyperparameters are set for training: batch size is 18, total training epochs is 80, and the image size after resize is 640x640.

In order to fully demonstrate the advantages of the algorithm proposed in this paper, we have selected the following two indicators to analyze the performance of the cell detection algorithm quantitatively:

- 1) The algorithm predicts the mAP(mean average precision) indicator of fine-grained classes [17], IOU threshold is set to 0.5, after this referred to as mAP@0.5.
- 2) The algorithm predicts the mAP@0.5 indicator of the coarse-grained class. It should be added that both the classes of algorithm prediction and sample annotation are fine-grained classes. The result of the algorithm predicting the coarse-grained classes is obtained from the mapping rule of the fine-grained classes to the coarse-grained classes.

Using the data set of this paper, the cell detection algorithms which have excellent performance in the current field [7], [10] and the bone marrow cell detection algorithm proposed in this paper are trained, measured, and evaluated. Each experiment was repeated three times to take the average value to avoid the contingency, the indexes obtained in Table I.

Since the weight of the classification part of the proposed loss function is unevenly distributed between  $\alpha$  and  $\alpha + \beta$ , we selected five  $\alpha$  which are equidistant distributed between  $\alpha$  and  $\alpha(1 + \beta)$  in the class weighted experiment, verifying that proposed loss is not the result of adjusting the weight of the classification section.

From the experimental results in Table I, the following conclusions can be analyzed:

- Proposed bone marrow cell detection algorithm performs well. The mAP@0.5 in the case of fine-grained classification reaches 51.14%. The mAP@0.5 in the case of coarse-grained classification reaches 69.71%, which is significantly better than the other two methods [7], [10]. Compared with the Faster-RCNN combined with transfer learning method [7], the proposed algorithm improves the mAP@0.5 of fine-grained classification by 16.17% and the mAP@0.5 of coarse-grained classification by 12.88%.
- 2) Increase the weight of the classification part in the loss function, the model can focus more on the classification of cells, which can effectively improve the performance of the bone marrow cell detection algorithm. Compared with the model with normal loss function, the model with class weighted loss function improves the mAP@0.5 of fine-grained classification by 5.03%.
- 3) Use the loss proposed in this paper to adjust the penalty for model prediction of coarse-grained class errors and fine-grained class errors, which can effectively improve the performance of the YOLOv5 network for bone marrow cell detection algorithms. Compared to model with class weighted loss function, the model with proposed loss function improves the mAP@0.5 of fine-grained classification by 1.46%, and the mAP@0.5 of coarsegrained classification by 3.57%. The results show that the use of the relationship between the coarse-grained classes and the fine-grained classes to strengthen the supervised training of the model is likely to help the algorithm learn the characteristics of the cell classes more effectively.

It should be noticed that when using cell morphology to diagnose patients, some common blood diseases can determine the diagnosis results by coarse-grained classification and counting of cells. The proposed algorithm effectively improves the model's performance in predicting the coarse classes. , Which improves the feasibility of the bone marrow cell detection algorithm in the automatic analysis system.

In the end, we evaluated the running speed of the proposed bone marrow cell detection algorithm: The model deployed on openvino uses Intel i7-10700 CPU to detect cells in a single image. It takes 73ms, which can reach the requirements of most automatic analysis systems for the efficiency of cell detection algorithms.

#### V. CONCLUSION AND FUTURE WORK

The bone marrow cell detection algorithm is the core of the meaningful bone marrow cell morphology automatic analysis system. This paper proposes a bone marrow cell detection algorithm based on the novel loss function and YOLOv5 network. The experimental results show that the proposed bone marrow cell detection algorithm has a better performance than other cell detection algorithms. The final mAP@0.5 of the fine-grained classification reaches 51.14%, and the mAP@0.5 of the coarse-grained classification reaches 69.71%. In future research, we will continuously try to combine model optimization and increase the sample size of the data set to improve the algorithm's accuracy in classifying cells.

#### REFERENCES

- S. Malempati, S. Joshi, S. Lai, D. A. Braner, and K. Tegtmeyer, "Videos in clinical medicine. bone marrow aspiration and biopsy," *N Engl J Med*, vol. 361, no. 15, pp. e28, 2009.
- [2] M. Kaur, A. P. Singh Rana, S. Kapoor, and A. Puri, "Diagnostic value of bone marrow aspiration and biopsy in routine hematology practice," *J Clin Diagn Res*, vol. 8, no. 8, pp. FC13–6, 2014.
- [3] Y. Y. Wu, T. C. Huang, R. H. Ye, W. H. Fang, S. W. Lai, P. Y. Chang, W. N. Liu, T. Y. Kuo, C. H. Lee, W. C. Tsai, and C. Lin, "A hematologist-level deep learning algorithm (bmsnet) for assessing the morphologies of single nuclear balls in bone marrow smears: Algorithm development," *Jmir Medical Informatics*, vol. 8, no. 4, pp. 17, 2020.
- [4] 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*, vol. 12, no. 12, pp. e0189259, 2017.
- [5] J. Prinyakupt and C. Pluempitiwiriyawej, "Segmentation of white blood cells and comparison of cell morphology by linear and naive bayes classifiers," *Biomed Eng Online*, vol. 14, pp. 63, 2015.
- [6] S. H. Rezatofighi and H. Soltanian-Zadeh, "Automatic recognition of five types of white blood cells in peripheral blood," *Comput Med Imaging Graph*, vol. 35, no. 4, pp. 333–43, 2011.
- [7] H. Kutlu, E. Avci, and F. Ozyurt, "White blood cells detection and classification based on regional convolutional neural networks," *Med Hypotheses*, vol. 135, pp. 109472, 2020.
- [8] F. Qin, N. Gao, Y. Peng, Z. Wu, S. Shen, and A. Grudtsin, "Fine-grained leukocyte classification with deep residual learning for microscopic images," *Comput Methods Programs Biomed*, vol. 162, pp. 243–252, 2018.
- [9] A. I. Shahin, Y. Guo, K. M. Amin, and A. A. Sharawi, "White blood cells identification system based on convolutional deep neural learning networks," *Comput Methods Programs Biomed*, vol. 168, pp. 69–80, 2019.
- [10] Q. Wang, S. Bi, M. Sun, Y. Wang, D. Wang, and S. Yang, "Deep learning approach to peripheral leukocyte recognition," *PLoS One*, vol. 14, no. 6, pp. e0218808, 2019.
- [11] Glenn Jocher, Alex Stoken, Jirka Borovec, NanoCode012, Christopher-STAN, Liu Changyu, Laughing, tkianai, Adam Hogan, lorenzomammana, yxNONG, AlexWang, Laurentiu Diaconu, Marc, wanghaoyang, ml5ah, Doug, Francisco Ingham, Frederik, Guilhen, Hatovix, Jake Poznanski, Jiacong Fang, Lijun Yu, changyu98, Mingyu Wang, Naman Gupta, Osama Akhtar, PetrDvoracek, and Prashant Rai, "ultralytics/yolov5: v3.1 - bug fixes and performance improvements," 2020.

- [12] Amjad Rehman, Naveed Abbas, Tanzila Saba, Syed Ijaz ur Rahman, Zahid Mehmood, and Hoshang Kolivand, "Classification of acute lymphoblastic leukemia using deep learning," *Microscopy Research and Technique*, vol. 81, no. 11, pp. 1310–1317, 2018.
- [13] Jane Hung and Anne Carpenter, "Applying faster r-cnn for object detection on malaria images," in *Proceedings of the IEEE conference* on computer vision and pattern recognition workshops, pp. 56–61.
- [14] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 779–788.
- [15] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, pp. 1904–1916, 2015.
- [16] Zhaohui Zheng, Ping Wang, Dongwei Ren, Wei Liu, Rongguang Ye, Qinghua Hu, and Wangmeng Zuo, "Enhancing geometric factors in model learning and inference for object detection and instance segmentation," arXiv preprint arXiv:2005.03572, 2020.
- [17] Mark Everingham, Luc Van Gool, Christopher K. I. Williams, John Winn, and Andrew Zisserman, "The pascal visual object classes (voc) challenge," *International Journal of Computer Vision*, vol. 88, no. 2, pp. 303–338, 2009.