AUTOMATED DETECTION OF WHITE BLOOD CELLS AND PLATELETS FROM MICROSCOPIC IMAGES USING DEEP LEARNING MODELS

Student: Thao Nguyen Xuan Mai[1], Quang Van Nguyen[2]

[1]Class 21AD; Faculty of Computer Science; Email: thaomnx.21ad@vku.udn.vn

[2]Class 20AD; Faculty of Computer Science; Email: nvquang.20it12@vku.udn.vn

Instructor: Phap Do Cong Nguyen

Unit: Faculty of Computer Science; Email: ndcphap@vku.udn.vn

Abstract. This study focuses on the detection of white blood cells (WBCs) and platelets using deep learning(DL) models on microscopic image data. The accurate identification and categorization of these blood cell types are essential for various medical diagnostic applications. The proposed methodology includes data preprocessing and inputting the preprocessed data into the model. The experimental results demonstrate that the proposed method employed in this study has the potential to achieve mAP@0.5 of 0.989 and mAP@0.5:0.95 of 0.909, indicating highly accurate detection of white blood cells and platelets. Experimental results show the approach's efficacy in accurately detecting WBCs and platelets. The findings highlight the potential of our approach in automating the analysis of microscopic images and contribute to the advancement of medical diagnostics in this domain.

Keywords: White Blood Cells Detection, Platelets Detection.

1. Introduction

The classification and localization of platelets [1] and white blood cells [2] (WBCs), including lymphocytes, monocytes, eosinophils, basophils, and neutrophils in medical practice contribute to the diagnosis and monitoring of immune and infections [3]. WBCs play a crucial role in protecting the body against bacteria [4], viruses, and cancer, while platelets are essential for the clotting [5] process. The classification and localization provide vital information for assessing the immune and blood system status, aiding in the diagnosis and monitoring of corresponding diseases.

The dataset "A dataset for microscopic peripheral blood cell images for development of automatic recognition systems" [6] has proven valuable for developing classification models that categorize white blood cells (WBCs) and platelets. As seen in [7], the authors utilized this dataset for deep learning models, achieving accuracy results of over 99%. Similarly, in the article [8], outstanding results were obtained using this diverse dataset. However, the potential of this dataset for more practical object detection has yet to be explored.

In this work, we demonstrate the feasibility of adapting an existing dataset "A dataset for microscopic peripheral blood cell images for development of automatic recognition systems", originally intended for classification, to instead enable precise object detection using state-of-the-art models. Our key contribution is manually annotating tight bounding boxes around all cell instances in the dataset, transforming the labels from image-level to instance-level. This updated dataset provides a new benchmark for detecting individual WBC and platelet cells.

Leveraging these updated annotations, we train the state-of-the-art YOLOv8 [9] model for accurate detection of WBCs and platelets, unlocking new benchmark potential beyond classification. Our experiments highlight the feasibility of this dataset transformation, with strong detection performance obtained. Practical clinical functionality demands precise cell instance detection rather than coarse image-level labels alone. In summary, our key contribution is refining an established medical imaging dataset to unlock new potential. We demonstrate the value of adapting resources to overcome limitations of previous classification-only approaches. The outcomes of our study - a wholly re-annotated dataset and strong detection model performance - further validate the broad applicability of the blood cell dataset for augmenting dataset value.

The remainder of this paper is structured as follows. Section 2 reviews related works. Section 3 presents our proposed approach. Section 4 details experiments. Finally, Section 5 concludes and discusses future directions.

2. Related works

The research on improving the detection and recognition of blood cell types through microscopic image analysis has always attracted attention from the scientific community and the medical research community.

There have been notable studies in blood cell detection. For example, the research conducted by Hợp et al. (2022) [10] provides an overview of blood and various diagnostic methods for blood disorders. They analyze the application of deep learning techniques for detection tasks and construct models for blood cell detection. Park et al. (2019) [11] proposed a deep learning-based approach specifically for the detection of red and WBCs, achieving high accuracy using a convolutional neural network (CNN) architecture. Kutlu et al. (2020) [3] focused on detecting and classifying different types of leukocytes, contributing to the development of automated systems for blood cell

analysis. Zhang et al. (2020) [12] developed a deep learning framework to detect abnormal blood cells, aiding in disease diagnosis. Rahadi et al. (2020) [13] explored image processing techniques to automatically identify and distinguish between red and white blood cells. Lastly, Koyuncu et al. (2020) [14] proposed a multi-task deep learning model combining CNNs and RNNs for the simultaneous detection and classification of various blood cell types. These studies collectively highlight the potential of deep learning models in improving automated blood cell analysis and detection. This research shows promise in the field of medical image analysis and disease diagnosis. However, it is worth noting that while there has been a greater emphasis on large datasets for classification-related studies, research related to detection, which holds higher practical applications, has been conducted on relatively fewer datasets.

Our work is most similar to that of [7], [8] in aiming to detect individual WBCs. However, we utilize the public dataset from [6] and transform the labels from image-level to instance-level to enable object detection. To our knowledge, no existing work has adapted this established medical imaging dataset for detection tasks. We demonstrate strong performance detecting platelets and white blood cells with state-of-the-art YOLOv8 models by leveraging our dataset advancement. Overall, our work builds upon a proven public resource and unlocks new research directions through refining dataset labeling.

Overall, the research community is actively exploring and advancing the field of blood cell detection and recognition, utilizing deep learning techniques and addressing challenges specific to medical image analysis.

3. Proposed approach

In this research, we collected data from [9]. After data collection, we performed data selection, preprocessing, and utilized YOLOv8 model. This model was used for the detection of blood cells in microscopic images.

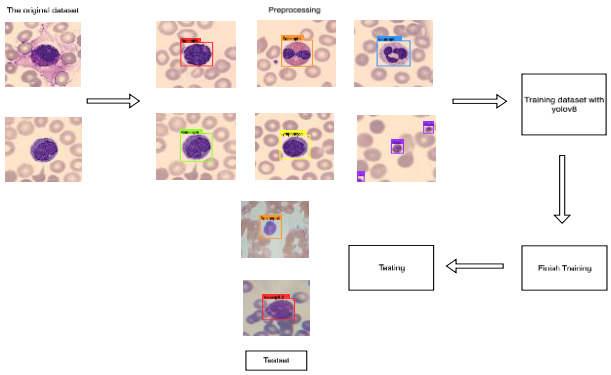

Fig. 1. Proposal Aproach

3.1. Data preprocessing

Preprocessing image data plays a crucial role in object detection, requiring several important steps to prepare the data for the detection model. Initially, the images are sourced and loaded, followed by re-sizing to ensure uniformity in size for seamless model processing. Normalization of pixel values is then performed, typically involving scaling them within a desired range, often achieved by dividing by 255. The subsequent task involves creating ground truth bounding boxes by accurately labeling objects and assigning corresponding class labels. To accomplish this, we utilized Roboflow, a platform that facilitates the process of drawing bounding boxes around the objects of interest. With Roboflow, we were able to annotate the images in the dataset by manually outlining the bounding boxes around the desired objects, such as white blood cells and platelets. Subsequently, we partitioned the dataset into three distinct subsets: the training set, the testing set, and the validation set. The training set consisted of 7,232 images, while the testing set and the validation set comprised 1,000 images each. This division allowed us to train the model on a substantial portion of the data, evaluate its performance on unseen data during testing, and further assess its generalization capabilities using the validation set. Lastly, additional preprocessing tasks, such as converting image formats and generating annotation files or metadata, may be undertaken to meet specific requirements. Overall, these preprocessing steps serve as a fundamental groundwork for effective object detection.

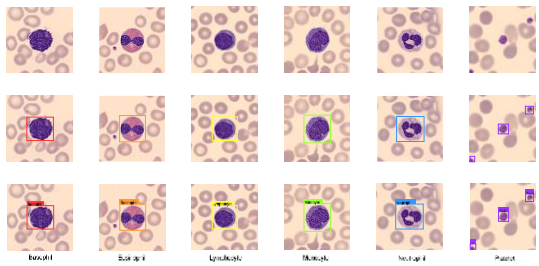

Fig. 2. Annotating the dataset by outlining the bounding boxes around the objects

3.2. Deep Learning models

Detection Transformers(DeTR)

DeTR [15] is an end-to-end model that can be trained directly on input images and object labels, without the need for intermediate steps like region proposal or anchor box computation. On standard benchmarks like COCO, DeTR has demonstrated performance on par with or exceeding traditional object detectors.

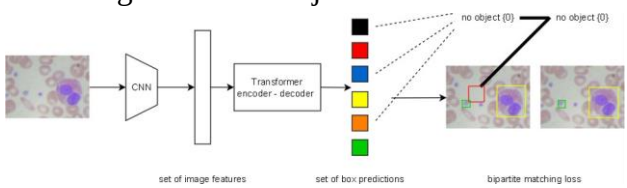

Fig. 3. DeTR model architecture

Yolov8

YOLOv8 [16] is an advanced real-time object detection model. It builds upon the YOLO series and introduces a powerful backbone network, anchor boxes, and optimization techniques. YOLOv8 can detect objects of different scales and aspect ratios using multi-scale feature fusion. It incorporates focal loss and IoU loss for improved accuracy in challenging scenarios. YOLOv8 is

widely used in surveillance, autonomous driving, and medical imaging applications, including the detection of blood cells. Its real-time processing capabilities and high accuracy make it a popular choice for object detection tasks.

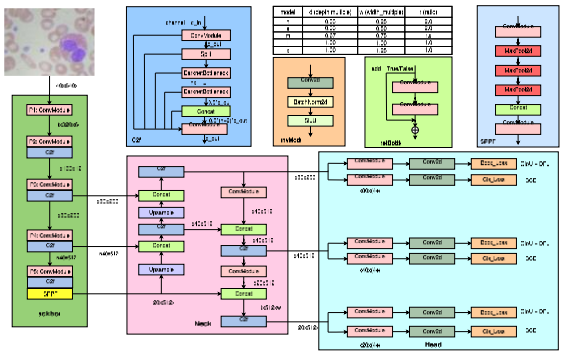

Fig. 4. YOLOv8 model architecture

Faster RCNN

Faster R-CNN [17] is a two-stage object detection model that introduced the Region Proposal Network (RPN) to efficiently generate region proposals. This eliminated the need for the slower selective search used in earlier R-CNN models, significantly improving detection speed while maintaining accuracy.The RPN shares convolutional features with the classification and bounding box regression network, making the overall pipeline much more efficient. Faster R-CNN has become a widely adopted framework due to its effectiveness and efficiency.

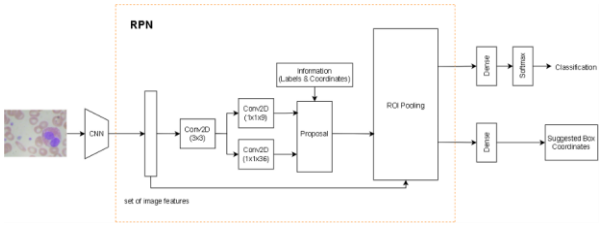

Fig. 5. Faster RCNN model architecture

4. Experiments

4.1. Dataset

The research project utilizes the dataset [6]. This dataset comprises 17,092 individual images of blood cells captured under a microscope. The images are divided into six main categories: neutrophils, eosinophils, basophils, lymphocytes, monocytes, and platelets. Each image has a size of 360x360 pixels and is in JPG format.

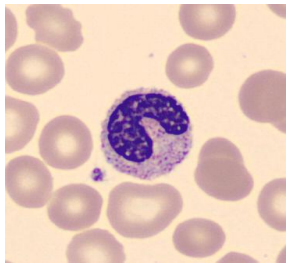

Fig. 6. Image from dataset

After data collection, we conducted a thorough data cleaning process where we removed non-standard data such as noisy and blurry images. This resulted in a refined dataset of 9,293 images deemed suitable for model training. However, since the original dataset was primarily designed for the classification task, an additional step was necessary to transform it into an object detection dataset, that is drawing bounding boxes around the objects.

Table 1. Statistics for each type of cell in the dataset

Type of Blood Cells	Quantity
Neutrophil	2,072
Basophil	1,424
Eosinophil	1,872
Monocyte	1,657
Lymphocyte	1,470
Platelets	1,967

In summary, the provided data indicates the quantities of different types of blood cells. Neutrophils are the most abundant type with 2,072 cells, while basophils are relatively less abundant with 1,424 cells. Eosinophils have a slightly higher quantity of 1,872 cells, and monocytes have a similar quantity of 1,657 cells. Lymphocytes are relatively less abundant with 1,470 cells. Platelets, not being white blood cells, have the highest quantity of 1,967 cells among all the mentioned cell types. Overall, the data shows variations in the quantities of different blood cell types in the sample.

4.2. Evaluation metrics

For the object detection task of WBCs and platelets, several commonly used experiment metrics can be used to evaluate the detection model's performance.

IOU

In our experiments, we utilize Intersection over Union (IOU) as a metric to measure the overlap between the predicted bounding box and the ground truth bounding box. IOU is commonly used as a threshold to distinguish between true positives and false positives in object detection.

$$IOU = Area\frac{of \cap}{Areaof \cup}$$

Precision

Precision is computed as the ratio of true positive detections (TP) to the total number of positive detections, encompassing both true positive and false positive (FP) instances.

$$Precision = \frac{TP}{TP + FP}$$

Here, TP refers to scenarios where the IOU between a predicted bounding box and its corresponding ground truth bounding box equals or exceeds the IOU threshold, while FP denotes cases where the IOU falls below the IOU threshold

Recall

We use recall as a metric to evaluate the model's capability to identify all relevant objects in a dataset. It is computed by dividing the number of true positive

detections by the total number of ground truth objects, which include TP and False Negative (FN).

$$Recall = \frac{TP}{TP + FN}$$

In this formula, TP refers to the case when the IOU between a predicted bounding box and its corresponding ground truth bounding box is greater than or equal to the IOU threshold. Conversely, FN situations arise when an object's bounding box remains undetected, resulting in missed detection by the model.

mAP 0.5

The mean average precision (mAP) at 0.5 is a metric used to evaluate the performance of an object detection model in detecting blood cells. It measures the model's ability to accurately detect blood cells with a confidence threshold of 0.5. A higher mAP at 0.5 indicates better performance in accurately detecting blood cells while minimizing false positives and false negatives.

mAP 0.5:0.95

The mAP at 0.5:0.95 is a metric that measures the performance of an object detection model in accurately detecting objects with high confidence. It calculates the average precision across a range of confidence thresholds from 0.5 to 0.95, considering true and false positive detections. A higher mAP at 0.5:0.95 indicates better performance in achieving a balance between precision and recall, resulting in accurate detections with minimal false positives.

4.3. Experiment settings

In our study, we employed the pretrained DeTR, Faster RCNN and YOLOv8 models for the object detection task, which had been trained on a large dataset prior to our experiments. The models was pretrained to learn general object detection features and patterns. We fine-tuned the pretrained models by training them for 50 epochs with a learning rate of 0.001, a decay rate of 1e-5, a momentum of 0.9, and a batch size of 16. During training, we gradually reduced the learning rate using a decay value of 1e-5 after each epoch to fine-tune the models's performance.

4.4. Experiment results

The following are the experimental results of the deep learning models on the prepared dataset:

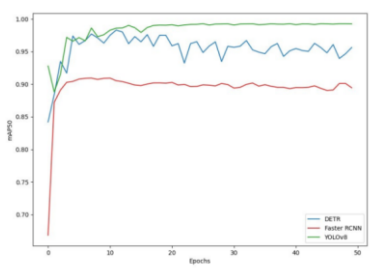

Fig. 7. The mAP 0.5 of each model

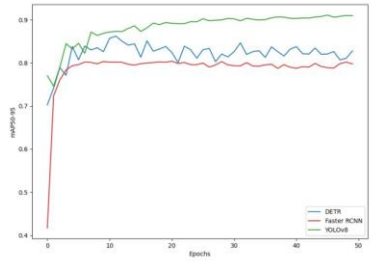

Fig. 8. The mAP 0.5:0.95 of each model

Based on the figure above, we can observe the following: The Yolo model initially achieved an approximate mAP of 0.93, which rapidly increased and reached a peak of 0.989. The mAP increased significantly during the first 10 epochs and then gradually plateaued. The DeTR model started with a mAP of around 0.84 and exhibited a sharp increase during the first 20 epochs. However, it later experienced fluctuations and maintained a relatively stable mAP value, with the highest recorded at 0.979. The Faster RCNN model had an initial mAP of approximately 0.65. It showed a significant improvement over the next two epochs, reaching a peak mAP of 0.911. Subsequently, the mAP showed minimal fluctuations and remained at a relatively high value.

Simultaneously, the Yolo model demonstrates an initial mAP at 0.5:0.95 of approximately 0.76, which quickly increases. It reaches its highest mAP at 0.5:0.95 of 0.909. The mAP at 0.5 shows rapid improvement within the first 20 epochs, followed by a gradual increase and stabilization. On the other hand, the DeTR model starts with an mAP at 0.5:0.95 of around 0.71. It exhibits significant improvement in the first 10 epochs, followed by considerable fluctuations. The highest recorded mAP at 0.5 is 0.850. Lastly, the Faster RCNN model initially has an mAP at 0.5:0.95 of approximately 0.42. It shows substantial improvement in the next two epochs, followed by minimal fluctuations. The model achieves its highest mAP at 0.5 of 0.812.

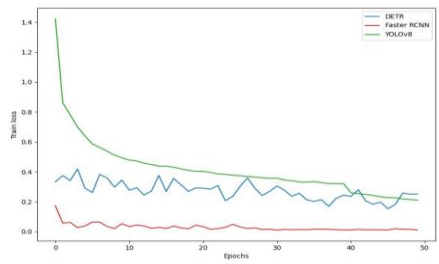

Fig. 9. The train loss of each model

The Yolo model starts with a train loss value of over 1.4 and consistently decreases throughout the training process, reaching its lowest value of 0.448. The DeTR model has an initial train loss value of approximately 0.37 and exhibits significant fluctuations during training, eventually achieving its lowest value of 0.154. The Faster RCNN model begins with a train loss value around 0.2, experiences slight decreases, and stabilizes during training, reaching its lowest loss value of 0.011.

Table 2. Best Performances of Three models

Model	mAP@0.5	mAP@0.5:0.95	Loss	Precision	Recall
DeTR	0.979	0.850	0.154	0.956	0.901
Faster R-CNN	0.911	0.812	**0.011**	0.899	0.92
YOLOv8	**0.989**	**0.909**	0.448	**0.986**	**0.983**

In summary, the performance of the three models can be summarized as follows: DeTR achieves high mAP@0.5 and mAP@0.5:0.95 scores, indicating accurate object detection and localization. YOLOv8 demonstrates excellent precision and recall, making it proficient in minimizing false positives and detecting a high proportion of true positives. Faster R-CNN performs competitively in terms of precision and recall, with a relatively lower loss value.

Overall, YOLOv8 stands out for its strong performance across multiple metrics, while DeTR and Faster R-CNN also showcase commendable results in specific areas. The choice of the best model depends on the specific requirements and priorities of the application.

5. Conclusions

Our significant achievement lies in the successful adaptation of a previously processed dataset from the classification task to the object detection task for white blood cells and platelets. Through our experiments, we have demonstrated that this dataset yields high-quality results, with the YOLOv8 model achieving mAP@0.5 of 0.989 and mAP@0.5:0.95 of 0.909, highlighting its value for object detection purposes in the medical domain. The reliable capabilities of YOLOv8 offer extensive prospects for its application in healthcare and disease diagnosis, potentially enhancing accuracy and efficiency in blood cell image analysis and improving the medical diagnostic process.

References

[1] Michelson, A.D., Cattaneo, M., Frelinger, A. and Newman, P. eds., 2019. Platelets. Academic press.

[2] Bøyum, A., 1964. Separation of white blood cells. Nature, 204(4960), pp.793-794.

[3] Kutlu, H., Avci, E. and Özyurt, F., 2020. White blood cells detection and classification based on regional convolutional neural networks. Medical hypotheses, 135, p.109472.

[4] Stewart, E.J., 2012. Growing unculturable bacteria. Journal of bacteriology, 194(16), pp.4151-4160.

[5] Kalafatis, M., Egan, J.O., van't Veer, C., Cawthern, K.M. and Mann, K.G., 1997. The regulation of clotting factors. Critical Reviews™ in eukaryotic gene expression, 7(3).

[6] Acevedo A, Merino A, Alférez S, Molina Á, Boldú L, Rodellar J. A dataset of microscopic peripheral blood cell images for development of automatic recognition systems. Data in brief. 2020 Jun 1;30.

[7] Yang, J., Shi, R., Wei, D., Liu, Z., Zhao, L., Ke, B., Pfister, H. and Ni, B., 2023. MedMNIST v2-A large-scale lightweight benchmark for 2D and 3D biomedical image classification. Scientific Data, 10(1), p.41.

[8] Sadafi, A., Salehi, R., Gruber, A., Boushehri, S.S., Giehr, P., Navab, N. and Marr, C., 2023, October. A continual learning approach for cross-domain white blood cell classification. In MICCAI Workshop on Domain Adaptation and Representation Transfer (pp. 136-146). Cham: Springer Nature Switzerland.

[9] Glenn, J. Ultralytics yolov8. https://github.com/ultralytics/ultralytics (2023).

[10] Hợp, Đ.X., 2022. Ứng dụng kỹ thuật học sâu để phát hiện, phân loại tế bào máu trên ảnh hiển vi chụp tiêu bản máu ngoại vi. Digital library - Ha Noi university of science and technology (hust.edu.vn)

[11] Kim, G., Jo, Y., Cho, H., Min, H.S. and Park, Y., 2019. Learning-based screening of hematologic disorders using quantitative phase imaging of individual red blood cells. Biosensors and Bioelectronics, 123, pp.69-76.

[12] Qiu, W., Guo, J., Li, X., Xu, M., Zhang, M., Guo, N. and Li, Q., 2020, December. Multi-label detection and classification of red blood cells in microscopic images. In 2020 IEEE International Conference on Big Data (Big Data) (pp. 4257-4263). IEEE.

[13] Rahadi, I., Choodoung, M. and Choodoung, A., 2020, May. Red blood cells and white blood cells detection by image processing. In Journal of Physics: Conference Series (Vol. 1539, No. 1, p. 012025). IOP Publishing.

[14] Koyuncu, C.F., Gunesli, G.N., Cetin-Atalay, R. and Gunduz-Demir, C., 2020. DeepDistance: a multi-task deep regression model for cell detection in inverted microscopy images. Medical Image Analysis, 63, p.101720.

[15] Carion, N., Massa, F., Synnaeve, G., Usunier, N., Kirillov, A. and Zagoruyko, S., 2020, August. End-to-end object detection with transformers. In European conference on computer vision (pp. 213-229). Cham: Springer International Publishing.

[16] Ju, R.Y. and Cai, W., 2023. Fracture Detection in Pediatric Wrist Trauma X-ray Images Using YOLOv8 Algorithm. arXiv preprint arXiv:2304.05071.

Girshick, R., 2015. Fast r-cnn. In Proceedings of the IEEE international conference on computer vision (pp. 1440-1448)