Blood Cell Detection

Nguyen Anh Quan, Pham Xuan Trung, Hoang Ha Dang, Hua Hai Minh, Phung Duc Thai
University of Science and Technology of Hanoi

***Abstract*—Bioengineering applies engineering principles to biology and medicine, facilitating healthcare applications. Traditional blood cell classifying techniques are time-consuming and prone to inaccuracies, necessitating more efficient and automated approaches. In response to this, we carried out this study to experiment with various deep learning algorithms, aiming to streamline the process and improve accuracy. By eliminating human intervention, our system offers a more efficient and reliable solution for analyzing human biological samples in healthcare settings.**

KEYWORDS

—Blood Cell Detection, Deep Learning, YOLO, Detectron2, Retinanet

I. INTRODUCTION

Blood Cell Detection is a crucial medical diagnostic test and is used to evaluate an individual's well-being. Blood comprises three main types of cells: Red Blood Cells (RBCs), White Blood Cells (WBCs), and Platelets. RBCs, or erythrocytes, are a type of blood cell that is made in the bone marrow and found in the blood. Red blood cells contain a protein called hemoglobin, which carries oxygen from the lungs to all parts of the body. WBCs, also known as leukocytes, are a type of blood cell that is made in the bone marrow and found in the blood and lymph tissue. White blood cells are part of the body's immune system. They help the body fight infection and other diseases. Types of white blood cells are granulocytes (neutrophils, eosinophils, and basophils), monocytes, and lymphocytes (T cells and B cells). Platelets are tiny, disc-shaped pieces of cells that are found in the blood and spleen. Platelets are pieces of very large cells in the bone marrow called megakaryocytes. They help form blood clots to slow or stop bleeding and to help wounds heal. Having too many or too few platelets or having platelets that don't work as they should can cause problems. RBCs are the most common type of blood cell, and they make up 40-45% of the blood cells. Platelets are also found in huge numbers in the blood cells. WBCs make up 1% of the blood cells. Most of the time, blood cell detection is done manually using a hematology analyzer along with other laboratory equipment which can be time-consuming.

In recent years, Computer-Aided Diagnosis (CAD) models have been used for automatic medical diagnosis purposes. Deep learning, image processing, and object detection can be used for automatic blood cell detection. In object detection, ground truth annotations, which include bounding boxes around the objects, are required for model training. Deep learning in object detection employs neural networks to recognize and pinpoint objects within images and videos, essential for applications like autonomous vehicles and surveillance systems. Convolutional Neural Networks (CNNs) are commonly utilized for various visual recognition tasks. Over the years, CNNs have heavily been used in medical image analysis because of their excellent capacity to learn and extract features. CNNs can capture more useful features and can better detect objects on input images.

In this study, we propose a deep-learning-based blood cell detection method. We implemented three types of deep learning models for detecting and counting three types of blood cells: RBCs, WBCs, and Platelets – "You Only Look Once" (YOLO), EfficientDet, and Faster R-CNN. YOLO is a state-of-the-art object detection model that is one of the fastest object detection models and is known for its efficiency, accuracy, and versatility. EfficientDet is a family of single-stage object detection models based on the EfficientNet backbone. It uses a number of optimizations to achieve high performance while maintaining low latency. After the training of these models, various comparisons were made to assess their efficiency, accuracy, and performance.

The primary contribution of this work is to demonstrate the possibility of deploying a deep learning-driven methodology for blood cell detec-

tion, thereby expanding the scope of potential applications in Computer-Aided Diagnosis (CAD).

II. Related Works

In the field of Blood Cell Detection, a multitude of methodologies has been explored to enhance accuracy and efficiency. For example, Rabia et al. [1] utilize the support vector machine (SVM) and various machine learning models to achieve an automatic blood cell classification from blood cell microscopic images and Lee et al. [2] have leveraged convolutional neural networks (CNNs) to automatically extract discriminative features from raw images, showcasing promising results in identifying various blood cell types. In addition, Kevin Barrera [3] and his partners have further pushed the boundaries by utilizing architectures tailored specifically for blood cell detection, often integrating techniques such as data augmentation and generative adversarial networks (GANs). This exploration of diverse methodologies, from traditional image processing to deep learning approaches, alongside the integration of multiple modalities, reflects a continuous pursuit within the field of blood cell detection to enhance accuracy and efficiency for various applications.

III. The Benchmark Blood Cell Detection Dataset

Blood cell detection relies heavily on high-quality datasets for training and evaluation. In this study, we utilize a benchmark blood cell detection dataset that comprises a diverse range of blood smear images obtained from various sources. The dataset consists of annotated images, where ground truth annotations are provided for each blood cell type: Red Blood Cells (RBCs), White Blood Cells (WBCs), and Platelets. This dataset is split into train, testing, and validation datasets(87% for training, 9% for testing, and 4% for validation). Two sample images from the dataset are shown in Figure 1 and Figure 2, illustrating the diversity and complexity of the images included. These images serve as examples of the types of data used in our study and highlight the challenges associated with blood cell detection tasks.

IV. Methods

In this section, we demonstrate the three deep learning-based methods employed for blood cell detection: YOLOv8 (You Only Look Once), Detectron 2 and RetinaNet.

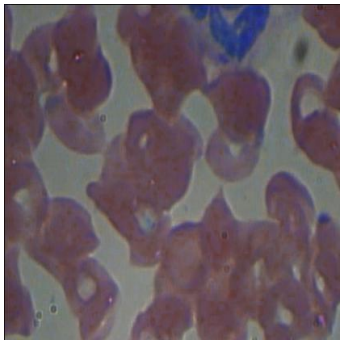

Fig. 1: Blood cell 1

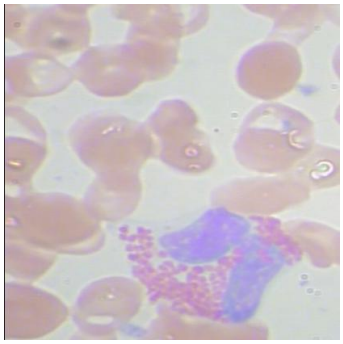

Fig. 2: Blood cell 2

A. YOLOv8 (You Only Look Once)

The first model that is used to perform the study is the well-known YOLOv8 with backbone P5, this refers to the final output stage of the model's backbone network. P5 captures high-level features from the input image with a large receptive field. This means it can identify objects even at lower resolutions. P5 consists of a Spatial Pyramid Pooling (SPP) layer. The SPP layer can process features at various scals, enhancing the model's ability to detect objects of different sizes. The Neck takes the feature maps from the backbone and combines information from different resolutions to create a richer representation for the head. The Head is responsible for making the final predictions. The head uses these rich feature maps to predict bounding boxes, object confidence scores, class probabilities. YOLOv8 contains 365 layers, 68155497 parameters. The loss functions used in YOLOv8 are classification loss, distributed focal loss, box loss.

B. Detectron 2

Detectron 2 is a next-generation open-source object detection system from Facebook AI Research. It is the successor of Detectron and maskrcnn-benchmark and supports many computer vision research projects and production applications in Facebook. It can be used for detection tasks such as bounding-box detection, instance and semantic segmentation, and person keypoint detection.

For this project, we choose the Base (Faster) R-CNN with Feature Pyramid Network (Base-RCNN-FPN), which is the basic bounding box detector extendable to Mask R-CNN. Faster R-CNN detector with FPN backbone is a multi-scale detector that realizes high accuracy for detecting small to big objects, making itself the de-facto standard detector.

There are three main parts in the Detectron2's architecture. Firstly, the Backbone Network extracts feature maps from the input image at different scales. Then it uses Region Proposal Network to detect object regions from multi-scale features with confidence scores. Finally, the Box Head will crop and warp feature maps using proposal boxes into multiple fixed-size features, and obtains fine-tuned box locations and classification results via fully-connected layers with maximum 100 boxes. The model has 83 layers and 41,699,936 parameters with the input size of the image is 256x256.

C. Retinanet

Another popular network designed for object detection tasks is RetinaNet. The network is introduced by Tsung-Yi Lin et al. in "Focal Loss for Dense Object Detection" [4]. One of the main innovations of RetinaNet is the focal loss function, handling the problem of class imbalance. The focal loss function down-weights the loss assigned to well-classified examples, on the other hand, focuses more on the hard examples that often be misclassified during the training process.

The architecture of RetinaNet is simple, it uses a Feature Pyramid Network (FPN) for backbone, efficiently extracting feature maps at multiple scales from the input image. The detection model has two main subnetwoks, one is for classification and the other is for regression.

In the project, the model uses ResNet50 as backbone, and is pre-trained with the COCO object detection dataset. The model has 160 layers and 36,394,120 parameters in total. Images are preprocessing and resize to 640x640 before passed into the model.

D. Evaluation metrics

We evaluate the performance of our blood cell detection models using several key metrics, including Recall, Precision, and Average Precision

1) Recall: Recall is the fraction of instances in a class that the model correctly classified out of all instances in that class. Recall is a good measure to determine when the cost of False Negative is high.

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

2) Precision: Precision is the measure of instances correctly classified as belonging to a specific class out of all instances the model predicted to belong to that class. Precision is a good measure to determine when the cost of a False Positive is high.

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

3) Average Precision(AP): The average precision score (AP) measures the balance between precision and recall. The score is the area under the precision-recall curve, ranging from 0 to 1. When both precision and recall are high, the AP score is high

V. Results

A. YOLOv8 Loss

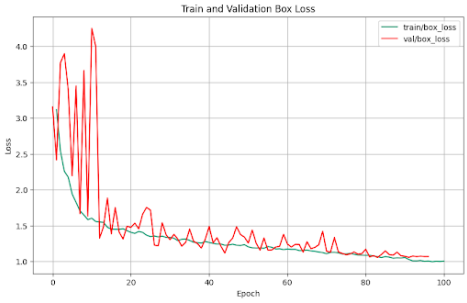

Fig. 3: YOLOv8 Box Loss

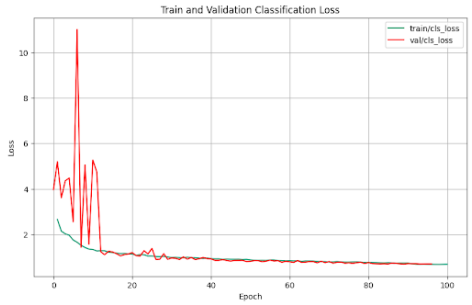

Fig. 4: YOLOv8 Classification Loss

The figures illustrate the classification and box loss of the YOLOv8 model during the training process after 100 epochs. The validation loss appears to be significantly high (over 4. for box loss and 10. for classification loss) compared with the training loss in about the first 25 epochs but after this period, this loss seems to become stable, and the distance between training and validating curved is narrowed.

B. Detectron 2 Loss

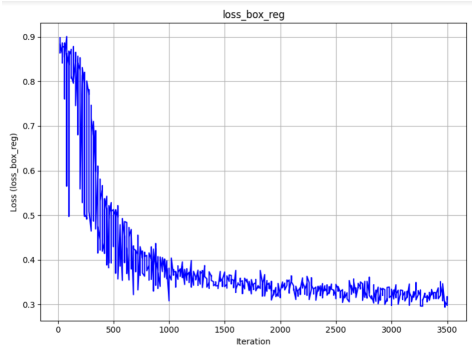

Fig. 5: Detectron 2 Box Loss

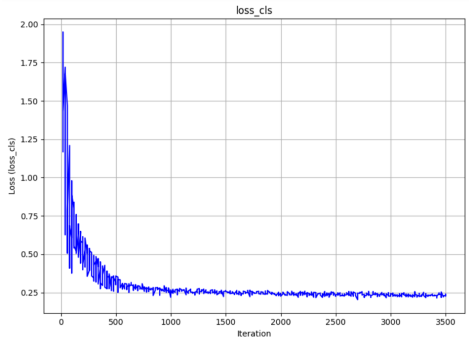

Fig. 6: Detectron 2 Classification Loss

The positive results are also depicted in the visualization of the classification and the box loss of the Detectron 2 model after 3500 iterations of training (each iteration fits the model with a batch of images from a dataset, in this scenario, the batch size is 4). The losses started from a large number and were reduced remarkably in the first 1000 iterations and converged in the rest of the training process.

C. RetinaNet Loss

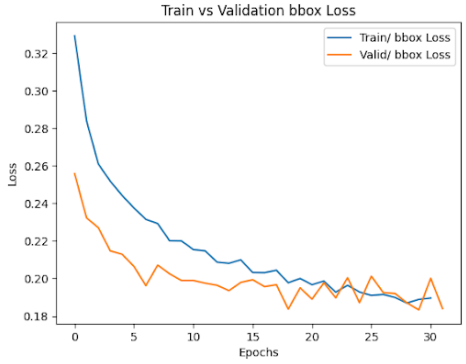

Fig. 7: Retinanet Box Loss

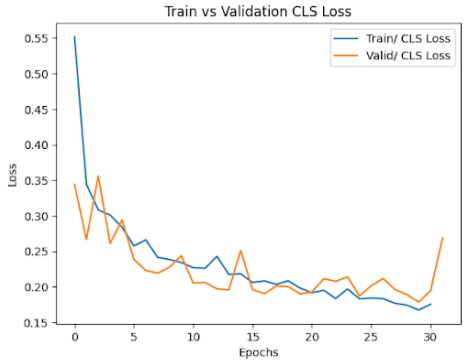

Fig. 8: Retinanet Classification Loss

Despite the limited computational resources restricting the number of epochs to 30 during training, RetinaNet has demonstrated remarkable performance optimization due to its complex architecture. The loss between training and validating seems to be close to each other and tends to reduce if the training period continues.

Overall, the experimented models showcase proficiency in minimizing loss throughout the learning process, indicating effective training. Furthermore, they demonstrate the ability to generalize well to unseen data, thereby contributing to their applicability in real-world scenarios with practical examples.

Models	mAP	Recall	Precision
Yolov8	0.631	0.879	0.85
Detectron2	0.555	0.641	0.899
Retinanet	0.83	0.941	0.953

TABLE I: Evaluation Metrics

D. Evaluation Metrics

Table 1 provides a comparative analysis of the performance results for three object detection models: Yolo V8, Detectron 2, and Retinanet with the proposed metrics: mAP, Recall and Precision. RetinaNet consistently achieves the highest scores across all metrics, indicating its superior performance in accurately detecting objects in images. Detectron 2 follows with intermediate scores, while Yolo V8 demonstrates the lowest performance among the evaluated models. Therefore, we can consider using RetinaNet as a preferable choice for deploying real-world healthcare applications aimed at detecting and analyzing blood cells accurately and efficiently.

VI. Conclusion

In conclusion, the study successfully examines blood cell detection using YOLO v8, Detectron2, and RetinaNet models and returns promising outcomes, highlighting their effectiveness in accurately detecting blood cells within images. These findings suggest significant advancements in blood cell analysis, offering improved efficiency and accuracy, with potential applications in medical diagnostics and biomedical research.

However, several challenges appeared during the training process of these models, including limitations in computational resources, availability of comprehensive datasets, and validation complexities. Additionally, the high training time required poses practical constraints on implementation. Future work should focus on addressing these challenges through model optimization, leveraging transfer learning techniques, conducting clinical validation studies, and optimizing the calculational process to enhance efficiency and scalability.

References

[1] R. Asghar *et al.*, "Classification of white blood cells using machine and deep learning models: A systematic review," 2023. [Online]. Available: https://arxiv.org/ftp/arxiv/papers/2308/2308.06296.pdf

[2] J.-W. L. Shin-Jye Lee, Pei-Yun Chen, "Complete blood cell detection and counting based on deepneural networks," 2022. [Online]. Available: https://www.researchgate.net/publication/362697125_Complete_Blood_Cell_Detection_and_Counting_Based_on_Deep_Neural_Networks

[3] K. Barrera *et al.*, "Automatic generation of artificial images of leukocytes and leukemic cells using generative adversarial networks (syntheticcellgan)," 2023. [Online]. Available: https://doi.org/10.1016/j.cmpb.2022.107314

[4] R. G.-K. H. P. D. Tsung-Yi Lin, Priya Goyal, "Focal loss for dense object detection," 2017. [Online]. Available: https://arxiv.org/abs/1708.02002