Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12811
e-ISSN : 2541-2019
p-ISSN : 2541-044X

White Blood Cell Detection Using Yolov8 Integration with DETR to Improve Accuracy

Shinta Jitny Ayu Nugraha[1)*], Bayu Erfianto[2)]
[1,2)] Universitas Telkom, Bandung, Jawa Barat, Indonesia
[1)]shintajitny@student.telkomuniversity.ac.id, [2)] erfianto@ telkomuniversity.ac.id

Submitted : Jul 30, 2023 | **Accepted** : Jul 30, 2023 | **Published** : Jul 31, 2023

Abstract: One of the body's most crucial blood cell kinds is the white blood cell. White blood cells, called leukocytes, are crucial for the body's defence mechanism and against hazardous foreign substances, tumour cells, and infectious bacteria. This paper suggests a computer-based automated system for detecting white blood cells using the YOLOV8 transformer and white blood cell analysis in digital images of blood cells. The Generate process uses Yolov8. In Generate, this will produce image processing in the form of annotation results on each type of white blood cell and dataset with COCO format. The DETR Model training conducted in this study is to increase the accuracy value of the white output of the blood cell picture formation. Test results using recall, precision, f1 score and object detection values. In the lymphocyte and basophil datasets, the number of white blood cell images used is only 10 images. Following the results of training from yolov8 using Roboflow, the results were increased relatively high, with an average increase of 0.68 in all five images of white blood cells. This test also gets an average improvement in detection results from Yolo to DETR, getting a fairly significant result of 68%, which is because YOLO cannot handle undetected objects (which are not in the training dataset; furthermore, DETR can handle multiple objects in a single image. Typically, detecting traditional objects such as YOLO requires repeatedly multiple object detection with a fixed batch size.

Keywords: Coco; DETR; Object Detection; Wbc; Yolo;

INTRODUCTION

The human body needs blood to transport oxygen and other nutrients. Blood, a red body fluid, is formed by respiratory proteins containing iron and hemoglobin. In addition, blood is divided into different types, each performing certain bodily functions. In addition, blood is grouped into different types, each performing a specific task for the body. White blood cells are one of the body's most important types of blood cells(Reynaldo & Lina, 2019). White blood cells, also known as leukocytes, are important to the body's defense system and fight infectious microorganisms, tumor cells, and dangerous foreign substances. Leukocytes consist of monocytes, lymphocytes, basophils, eosinophils, and neutrophils(Ak et al., 2018).

This paper suggests a computer-based automated system for detecting white blood cells using transformer analysis and white blood cell YOLOv8 in digital images of blood cells. Analyzing white blood cells using digital images requires only peripheral blood sample images, so they are not subjective and, therefore, faster and more accurat(Sri Indrawanti & Prakarsa Mandyartha, 2018). The machine learning component known as deep learning has developed into a powerful tool for classifying images(Minarno et al., n.d.). Convolutional neural networks (CNNs) and transformer-based object identification models may be divided into categories based on the model's structure. Transformer employs self-attentive methods, whereas CNN extracts and processes information via convolutional operations. There are two types of object detection networks: one-stage and two-stage, with benefits

*name of corresponding author

and drawbacks.(Leng et al., 2023). In stage one, use You Only Look on v8 (YOLOv8). They use the sigmoid function as an activation function for object scores, which indicates the possibility that the bounding box contains objects. They also use the SoftMax function for class probabilities, which indicates the possibility that objects belong to every possible class(Terven & Cordova-Esparza, 2023). In natural language processing, transformer-based models have been widely used.

Vision transformation mode (ViT)(Dosovitskiy et al., 2020). In natural language processing, transformer-based models have been widely used. Google's proposed vision transformer (ViT) model demonstrates the prospect of transformer applications in the field and begins in-depth research on transformers. For accurate and promising detection,transformer-based network backbones such as Swin transformers can be combined with classic CNN detection. Transformers submitted by Facebook based on end-to-end object detection network(Carion Nicolas et al., 2020; Liu et al., 2021). In this study, the authors performed image processing of white blood cells using a roboflow platform using the YOLOv8 model. Roboflow is a platform for annotating and labeling and can produce results using the same format as COCO(Park & Jun, 2022). A large-scale image dataset with high complexity, Common Objects in Text Dataset (COCO) displays complex scene images with many small objects and footnotes with very detailed outlines. COCO was also designed to enable the study of interaction matters(Caesar et al., n.d.). There are research questions in this study that are relevant to the research.
Why should DETR use the COCO dataset?(RQ1)
What factors can affect the detection results of yolov8 with roboflow tools?
The research questions will be addressed later in the discussion section.

LITERATURE REVIEW

Based on a study completed by Fathi E. etc. current study used neuro-fuzzy and group data processing techniques to determine the presence of acute leukemia in children using a complete blood count test. Acute lymphoblastic and acute myeloid leukaemia can be diagnosed using the suggested method. Multi-layer neural networks are used in adaptive neuro-fuzzy inference (ANFIS), which employs fuzzy logic and neural network learning algorithms to create nonlinear mappings between the input and output regions. AN FIS and the group method of data handling (GMDH) will be employed to identify the kind of childhood leukaemia. This study analyzed the application of neuro-fuzzy and group data processing methods for identifying acute leukemia in children based on a full blood count test. It was based on a study carried out by Fathi Ehsan and colleagues. The suggested method can assist in differentiating between acute lymphoblastic and acute myeloid leukemia, two different kinds of leukemia. Multi-layer neural networks are used in adaptive neuro-fuzzy inference (ANFIS), which employs fuzzy logic and neural network learning algorithms to create nonlinear mappings between input and output regions. The kind of childhood leukemia will be identified using ANFIS and the group method of data handling (GMDH). (Fathi et al., 2020)

Cinthia Espinoza-Del Angel and Aurora Feemat-Diaz researched the Comparison of the Accuracy of Color Spaces in Cell Features Classification in Images of Leukemia types ALL and MM. This suggests using analysis of major components with statistical descriptors as input variables to ensure that color space represents a high degree of color space. Comparison of the Color Space Accuracy in Cell Features Leukemia classification kinds ALL and MM. In order to guarantee that colour space accurately reflects a large portion of colour space, this advises employing analysis of key components using statistical descriptors as input variables. The kNN process, as well as the confidential matrix, ensures predictive model accuracy. Photo segmentation is attempted using HSV space, considering that it separates image intensity from color data, and the information base image does not have uniform brightness. The water-limiting algorithm is used for the resulting photographs to separate overlapping cores (Espinoza Cinthia & Femat Aurora, 2022).

Research conducted by P. Abhishek, P. Vinod, and P. Rajeshwari on Object Detection is a computer vision and autonomous driving system; object recognition and detection are essential. We want to create a system that delivers real-time solutions without compromising performance or accuracy. The models deliver high-performance outcomes are becoming more and more crucial as the importance of computer vision increases every day. High-performance algorithms that tackle real-world issues have significantly improved due to the exponential growth in computing power and the rising

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12811
e-ISSN : 2541-2019
p-ISSN : 2541-044X

popularity of deep learning. Although it has been trained on a bigger dataset, our model may be improved by enabling users to identify the necessary objects. (Rajeshwari et al., 2019).

Research results conducted by Depu Meng, Xiaokang Chen and friends of DETR Techniques recently developed to detect objects using transformer encoders and decoder architectures and produce promising results. In this work, we address crucial difficulties with delayed training convergence and offer conditional cross-attention techniques for quick DETR training. The cross-attention of DETR drives our method. It depends on content embedding to localize four extremities and forecast boxes, increasing the demand for high-quality content embeddings and training complexity. Our DETR approach learns conditional spatial queries from decoder embeddings for multi-head cross-attention. Each cross-attention head may attend a band of distinct regions, such as a single item's extremities or a region inside the object, which is an advantage. (Meng et al., n.d.).

METHOD

The system to be constructed in this study is a white blood cell image detection system that integrates yolov8 and DETR. The flow of the system is illustrated in Figure 1:

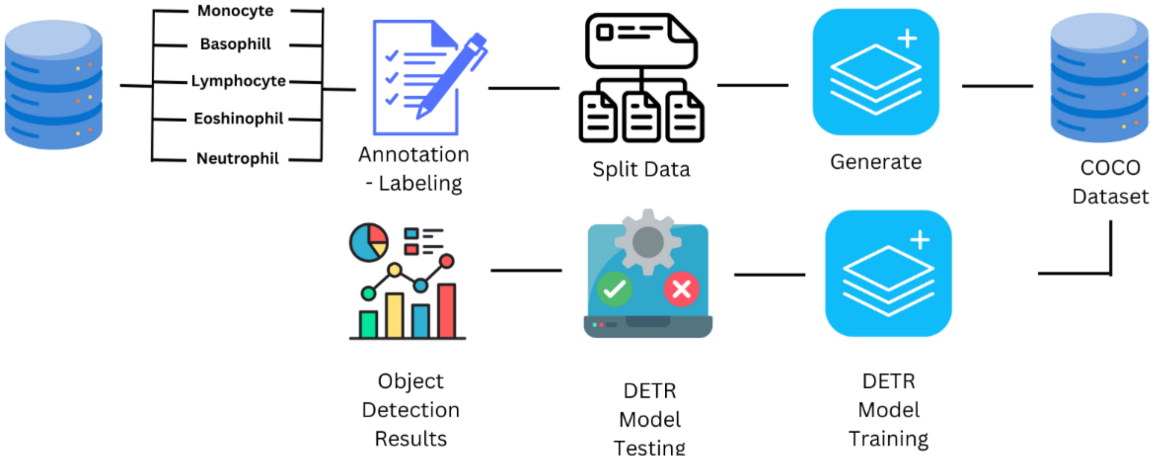
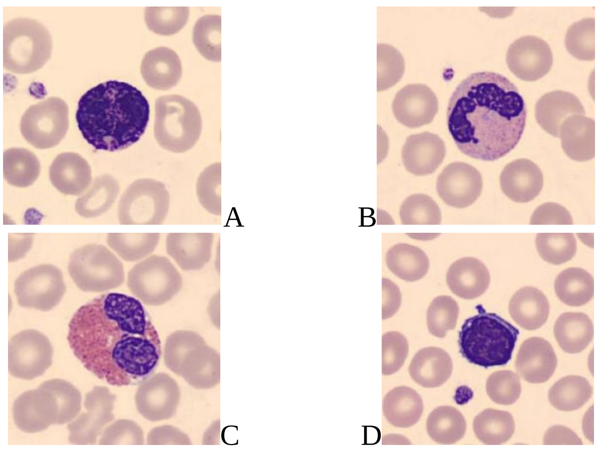


Figure 1 System Design
source: property researcher

Dataset

The dataset is from the Kagle website, a large online data repository. The study used 104 white blood cell images of 10 basophils, 10 lymphocytes, 34 eosinophils, 25 monocytes, and 28 neutrophils. Figure 2 shows an example of a white blood cell. This dataset has not been processed or annotated.



*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12811
e-ISSN : 2541-2019
p-ISSN : 2541-044X

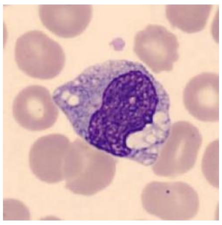
E 

Figure 2 White Blood Cell (A-Basofil, B-Neutrophil,C-Eosinopil, D-Limposit, E- Monocytet)
source: https://www.kaggle.com/datasets/bzhbzh35/peripheral-blood-cell

Annotation-labeling

Roboflow's label and annotation used a bounding box. In the labelling and annotation process, there are two labels in each image: red blood cell (RBC) and the type of white blood cell corresponding to the group. An example of labelling can be seen in Figure 3:

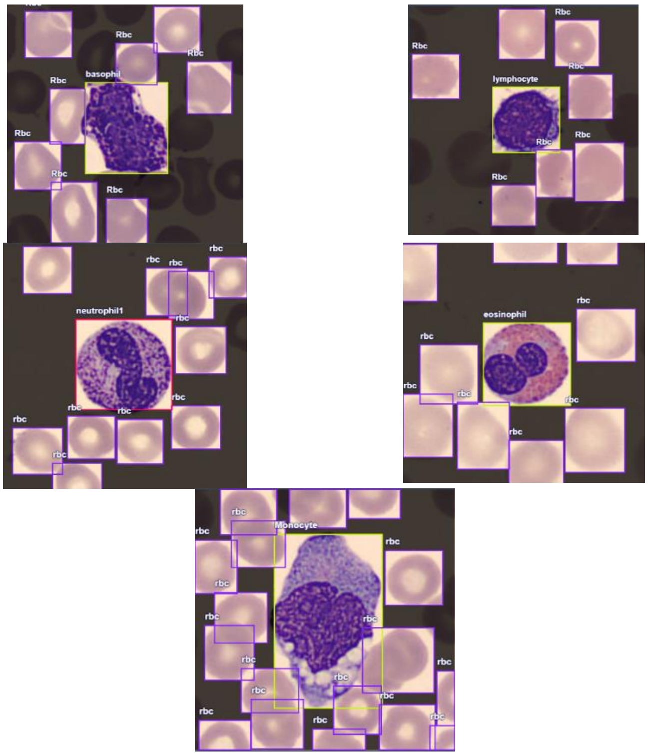


Figure 3 White blood cell labeling
source: property researcher

Split Data

After annotating and labelling, the data splits. Because in this study, using roboflow, the data split automatically. The results of the data split can be seen in Table 1:

Table 1 Split Data Result

White Blood Cell	Data Train	Data Valid	Data Test
Basofil	7	2	1
Limposit	7	2	1
Neutropil	18	6	4
Eosinopil	24	7	3
Monosit	17	5	2

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

Generate Dataset

In the next stage, Roboflow will systematically change the image size to 640x640. The Generate process uses Yolov8. In Generate, this will produce image processing in the form of annotations on each type of white blood cell and dataset with COCO format.

DETR Generate Model

The DETR Model training conducted in this study was to increase the accuracy of the results of generating white blood cell image images performed with the yolov8 model via the roboflow platform. DETR uses a Roboflow processing result dataset with COCO format during this training process.

DETR Testing Model

The DETR testing model is to test the image of white blood cells using the DETR model. For data sharing, this training is the same as in Table 1. Detr retrieves or pulls data via API KEY from roboflow during the training process.

Evaluation

DETR (Detection Transformer) is an object detection model that processes image data using transformer techniques. Various sizes, including the following, can be used to evaluate the DETR model:

1. The F1 score evaluates the model's effectiveness in recognizing objects while considering precision and recall. The harmonic average accuracy and recall are used to determine the F1 score. The excellent F1 score DETR model performs effectively on object detection in various scenarios(Suherman et al., 2023). The score of F1 is shown in the following equation

$$F1\ Score = \frac{2\ x\ Precision\ x\ recall}{Precision + recall} \quad (2)$$

2. Precision and recall measure how well a model recognizes objects by computing true positive, false positive, and false negative results. Precision is assessed by comparing the total of true positives, false positives, and true positives instead of recall, which is obtained by comparing true positives, erroneous positives, and true positives. The model works effectively at object identification in various contexts, as seen by excellent recall and DETR model accuracy(Aldi et al., 2023).

$$Recall = \frac{(TP)}{(TP + FN)} \quad (3)$$

$$Precision = \frac{(TP)}{(TP + FP)} \quad (4)$$

RESULT

In this section, Table 2 shows the test results using the results of recall, precision, and f1 score. In the lymphocyte and basophil datasets, the number of white blood cell images used is only 10 images. As for the product of neutrophil eosinophils and monocyte, it has a high yield.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12811
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Table 2 Evaluation Result

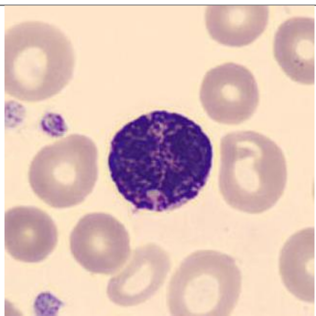
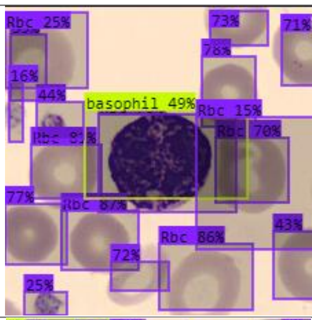
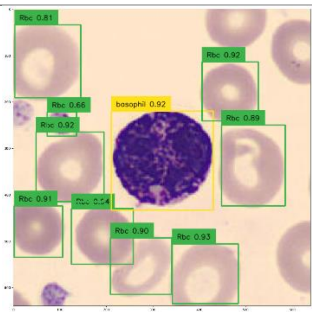
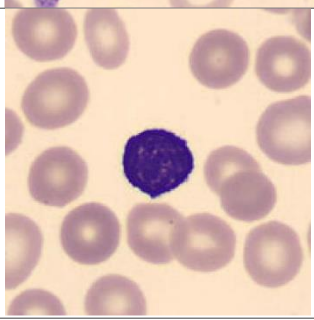
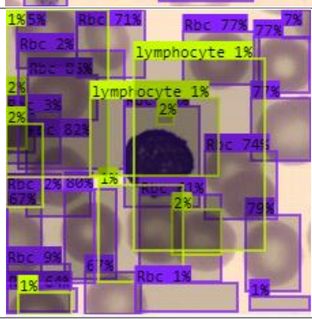
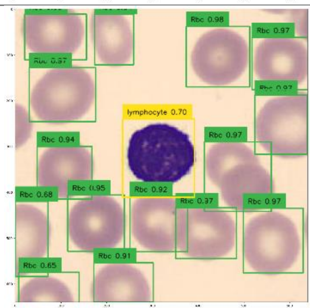
White blood cell type	Precision	Recall	F1 Score
Basophil	60,9%	77,5%	68,2%
Lymphocyte	87,4%	33,3%	48,22%
Neutrophil	84,4%	93,2%	90,73%
Eosinophil	91,5%	87,5%	89,45%
Monocyte	92.0%	97.5%	94,67%

Furthermore, the results of improved detection of yolov8 with DETR are shown in Table 3. Following the results of Table 3, training of yolov8 using Roboflow received a relatively high increase in results, with an average increase of 0.68 in all five images of white blood cells. The highest detection results are found in lymphocytes 0,88

Table 3 table of enhanced detection results

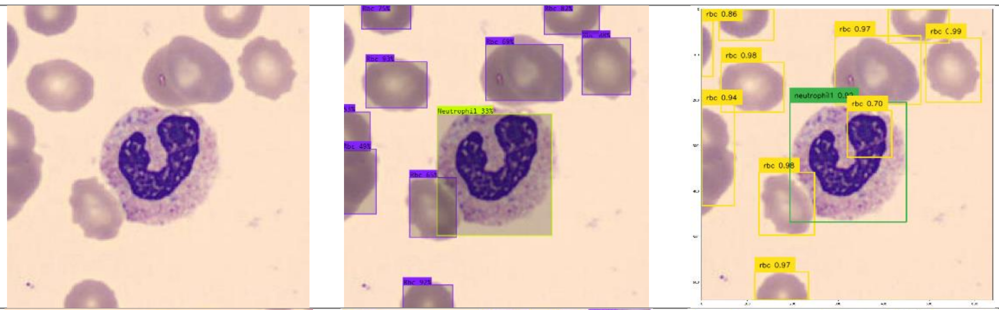
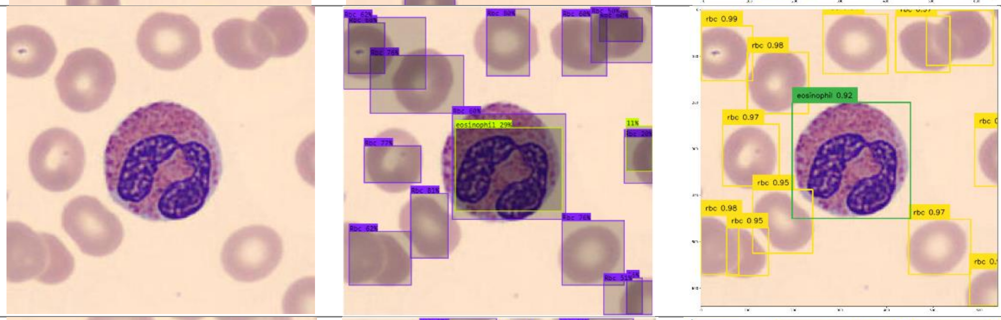
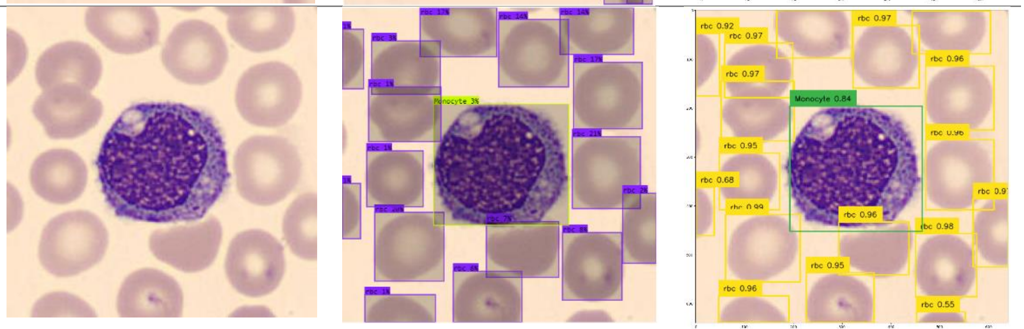
White Blood Cell Type	YOLOv8	DETR	Increase Model Detetion
Basophil	49%	92%	49%
Lymphocyte	2%	70%	88%
Neutrophil	29%	92%	63%
Eosinophil	33%	92%	59%
Monocyte	3%	84%	81%

Table 4 object detection image result

White blood cell type	Dataset	YOLOv8	DETR
Basophil			
Lymphocyte			

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12811
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Neutrophil	
Eosinophil	
Monocyte	

DISCUSSIONS

The discussion section discusses research-related topics to provide a more comprehensive explanation of the responses. There are two questions in this study.

Why should DETR use the COCO dataset? (RQ1)

An object detection method using a transformer model, DETR (Detection Transformer), is designed specifically for object segmentation and detection tasks. There are several reasons why people choose to train DETR with a dataset from COCO (Common Objects in Context):

1. **A wide and diverse COCO dataset**: with over 330,000 images covering 80 object categories, allows DETR to view various objects, backgrounds, and poses. By training DETR on the COCO dataset, models can better understand objects in various situations.
2. **Rich Annotations:** The COCO dataset has rich annotations that help DETRs better understand the shape and position of objects. This annotation includes information about the existing object's frame (bounding box) and proper pixel segmentation for each object.
3. **Major Benchmarks:** COCO has developed into one of the main task and object segmentation detection standards. Training DETR on COCO allows models to be tested and compared to other work in the object detection community. This facilitates the evaluation of the performance and progress of the suggested techniques.
4. **Engineering Progress:** Many new object detection models and segmentation techniques have been developed and tested on the COCO dataset. This model can leverage its knowledge of this method and compare the results with other models tested on the same dataset by training DETR on the same dataset.

*name of corresponding author

e-ISSN : 2541-2019
p-ISSN : 2541-044X

What factors can affect the detection results of yolov8 with the roboflow? (RQ2)
Some factors can affect the YOLOV8 Roboflow's siltation. Here are some of these:

1. **Dataset Quality:** A complete, diverse and clean dataset will improve the performance of the model and its ability to detect more accurately, as the quality of the dataset used to train the model is an important factor in the outcome
2. **The size of the dataset:** The model's generalization is influenced by the amount of training data that includes a variety of object variations, lighting conditions, and background. The more datasets used, the better the model detects objects.
3. **Preprocessing:** Detection results can be affected by initial data processing before training. Model performance can be improved by proper normalization, augmentation, and data processing
4. **Model Architecture:** YOLOV8 is a version of the YOLO model, and its structure affects its performance. The hyperparameter setting of the model can influence the detection results.
5. **Training Process:** The inference speed and accuracy of the model can be affected by the size of the network, including the number of layers and the filter's dimensions.
6. **Detection Object Selection:** Results are affected by the object the model wants to detect. Some objects may be more difficult to detect because of their complexity or size.
7. **Hardware and Infrastructure:** The hardware and infrastructure used to run the YOLOV8 model will also affect its performance. The inference speed will be increased using a GPU or TPU compared to the standard CPU.
8. **Framework and Library version:** To get the best results, use the latest and compatible version of the deep learning framework and the library used to train and run models.
9. **Object Difficulty Level:** Simple objects may be easier to identify than objects with complex shapes, sizes, or patterns. The accuracy of detection findings is also affected by the number of difficulties.

Improving the accuracy and performance of the YOLOV8 model with Roboflow will be helped by considering all of the above elements and optimizing each step of the object detection process.

CONCLUSION

Based on test results and analysis on white blood cell detection using yolov8 integration with DETR to impact detection. In this study, we tested a dataset of 5 types of white blood cells using yolov8 and DETR models. This test also gets an average improvement in detection results from Yolo to DETR, getting a fairly significant result of 68%, which is because YOLO cannot handle undetected objects (which are not in the training dataset; furthermore, DETR can handle multiple objects in a single image. Typically, detecting traditional objects such as YOLO requires repeatedly multiple object detection with a fixed batch size. The transformer architecture used in DETR allows the model to perform global mapping of the entire image while considering each object's global context. It helps to structure the context for each object better and allows DETR to address problems related to overlapping or adjacent objects.

REFERENCES

Ak, S. B., Analis, J., Poltekkes, K., & Makassar, K. (2018). ANALISIS JUMLAH LEUKOSIT DAN JENIS LEUKOSIT PADA INDIVIDU YANG TIDUR DENGAN LAMPU MENYALA DAN YANG DIPADAMKAN. *Jurnal Media Analis Kesehatan*, *1*(1).

Aldi, F., Nozomi, I., Sentosa, R. B., & Junaidi, A. (2023). Machine Learning to Identify Monkey Pox Disease. *Sinkron*, *8*(3), 1335–1347. https://doi.org/10.33395/sinkron.v8i3.12524

Caesar, H., Uijlings, J., & Ferrari, V. (n.d.). *COCO-Stuff: Thing and Stuff Classes in Context*. http://calvin.inf.ed.ac.uk/datasets/coco-stuff

Carion Nicolas, Massa Francisco, Synnaeve Gabriel, Usunier Nicolas, Kirillov Alexander, & Zagoruyko Sergey. (2020). *End-to-End Object Detection with Transformers* (A. Vedaldi, H. Bischof, T. Brox, & J.-M. Frahm, Eds.). Springer International Publishing. https://doi.org/10.1007/978-3-030-58452-8

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 7, Number 3, July 2023
DOI : https://doi.org/10.33395/sinkron.v8i3.12811
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). *An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale*. http://arxiv.org/abs/2010.11929

Espinoza Cinthia, & Femat Aurora. (2022). *Comparison of Accuracy of Color Spaces in Cell Features Classification in Images of Leukemia types ALL and MM* [Revista Mexicana de Ingeniería Biomédica]. https://doi.org/10.17488/RMIB.43.2.3

Fathi, E., Rezaee, M. J., Tavakkoli-Moghaddam, R., Alizadeh, A., & Montazer, A. (2020). Design of an integrated model for diagnosis and classification of pediatric acute leukemia using machine learning. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, *234*(10), 1051–1069. https://doi.org/10.1177/0954411920938567

Leng, B., Wang, C., Leng, M., Ge, M., & Dong, W. (2023). Deep learning detection network for peripheral blood leukocytes based on improved detection transformer. *Biomedical Signal Processing and Control*, *82*. https://doi.org/10.1016/j.bspc.2022.104518

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., & Guo, B. (2021). *Swin Transformer: Hierarchical Vision Transformer using Shifted Windows*. http://arxiv.org/abs/2103.14030

Meng, D., Chen, X., Fan, Z., Zeng, G., Li, H., Yuan, Y., Sun, L., & Wang, J. (n.d.). *Conditional DETR for Fast Training Convergence*.

Minarno, A. E., Aripa, L., Azhar, Y., & Munarko, Y. (n.d.). *Classification of Malaria Cell Image Using Inception-V3 Architecture*. www.joiv.org/index.php/joiv

Park, Y. H., & Jun, T. H. (2022). Classification of Soybean [Glycine max (L.) Merr.] Seed Based on Deep Learning Using the YOLOv5 Model. *Plant Breeding and Biotechnology*, *10*(1), 75–80. https://doi.org/10.9787/PBB.2022.10.1.75

Rajeshwari, P., Abhishek, P., Srikanth, P., & Vinod, T. (2019). Object Detection: An Overview. In *Published in International Journal of Trend in Scientific Research and Development (ijtsrd)* (Issue 3). http://creativecommons.org/licenses/by/4.0

Reynaldo, D., & Lina,). (2019). *KAJIAN TENTANG PENDETEKSIAN SEL DARAH PUTIH DENGAN TEKNIK SEGMENTASI WATERSHED*.

Sri Indrawanti, A., & Prakarsa Mandyartha, E. (2018). Deteksi Limfoblas pada Citra Sel Darah Menggunakan Fitur Geometri dan Local Binary Pattern. In *JNTETI* (Vol. 7, Issue 4).

Suherman, E., Rahman, B., Hindarto, D., & Santoso, H. (2023). Implementation of ResNet-50 on End-to-End Object Detection (DETR) on Objects. *SinkrOn*, *8*(2), 1085–1096. https://doi.org/10.33395/sinkron.v8i2.12378

Terven, J., & Cordova-Esparza, D. (2023). *A Comprehensive Review of YOLO: From YOLOv1 and Beyond*. http://arxiv.org/abs/2304.00501

*name of corresponding author