Proceedings of the 5th International Conference on Signal Processing and Machine Learning
DOI: 10.54254/2755-2721/121/2025.19559

A Comparative Analysis of YOLOv5nu, YOLOv8n, and YOLOv11n Models for Blood Cell Detection and Classification

Tingting Li[1,a,*]

[1]*College of Engineering, University of Wisconsin-Madison, Wisconsin, United States of America*
a. tli442@wisc.edu
**corresponding author*

Abstract: Object detection, which detects and classifies objects, is widely used in different fields. One such field is in the medical field, and more specifically it can be used to detect molecules in microscopy images to improve blood test efficiency. Nowadays, plentiful hospital labs are still manually counting blood cells which is time-consuming, and it is likely to make manual errors. A solution to this problem is to enable accurate object detection specifically designed for identifying molecules. To achieve this goal, the comparison analysis of different YOLO models, one of the most popular models in object detection, on a blood cell dataset is valuable as it evaluates the models' accuracy, which provides insight into their strengths and weaknesses. The paper considers three types of YOLO models, YOLOv5nu, YOLOv8n, and YOLO11n models, on detecting and classifying red blood cells (RBC), white blood cells (WBCs), and Platelets. The experiment results show that all three models have high precision and recall rates, which means that they can identify most of the molecules accurately. This indicates a promising future of integrating object detection on blood cell count to speed up medical analysis.

Keywords: Computer Vision, Deep Learning, Object Detection.

1. Introduction

Object detection is widely used for detecting different kinds of objects from one or various types of classes, such as blood molecules [1]. There are multiple ways of presenting the detected object information. One of the ways is simply enclosing the object's location in the image with a bounding box and placing the object class near the bounding box [1, 2]. One famous and widely used object detection algorithm that uses a bounding box to present its information is called You Only Look Once (YOLO) [2, 3]. Redmon et al. presented the first YOLO in 2015 [3]. Since then, YOLO has significantly improved and several iterations of the model have been developed, including YOLOv5nu, YOLOv8n, and YOLO11n [2, 3].

Object detection is extensively applied across different fields, including medical imaging, robotics, and security [1]. Particularly in the medical field, its ability to enable automatic detection and localization of abnormalities in medical images has made it a crucial role [4]. It has done excellent jobs in medical image identification, such as body organ identification [5]. There has been a growing interest in object detection technologies in assisting patient health evaluation [6]. One common and crucial area is object detection for blood count [7]. Enabling the identification of the number of blood cells is one of the most vital and essential tools in the first stage of analyzing a patient's health [8, 9]. The abnormality in the concentration of blood molecules, particularly WBCs and RBCs, is closely related to common blood diseases, such as anemia and leukemia [9].

Even though blood count is a common practice, many hospital laboratories manually count the blood cells using devices that require an extended period, which are prone to inaccurate results [4, 8, 9]. In other cases, laboratories are limited by the number of expensive equipment to perform a large number of blood counts [7, 9]. Therefore, enabling object detection that uses deep learning models that are specifically designed for identifying molecules in images could be meaningful in assisting laboratory and medical professionals [6]. However, a model that can produce high accuracy in detecting blood count is needed [4]. YOLO's ability to detect abnormality in medical images provides it potential to enhance the precision and efficiency of medical diagnoses [4, 10].

Therefore, to achieve this goal this study aims to compare the performance of different YOLO families, including YOLOv5nu, YOLOv8n, and YOLO11n models, in detecting and classifying RBCs, WBCs, and Platelets. By training three different YOLO models on the same blood cell database, the primary objective of the report is to evaluate each model's accuracy in detecting blood cells to determine their strengths and weaknesses.

For the structure of the paper, the paper begins with an overview of the main architectures and key innovations made in YOLOv5nu, YOLOv8n, and YOLO11n models. Followed by a description of the dataset, experiment setup, and training procedure. Then, a presentation and analysis of the experiment's results are shown. Lastly, a conclusion summarizes the overall findings and provides suggestions for future research.

2. Data and Method

2.1. Data Description and Preprocessing

The Blood Cell Count and Detection (BCCD) dataset is used for running blood cell detection tasks using different YOLO models. The dataset contains images of three categories of blood cells: RBCs, WBCs, and Platelets, and their annotation. Specifically, it has 364 JPEG images with a resolution of 640x480 pixels, a sample image in Figure 1, and 364 annotations for each image in PASCAL Visual Object Classes (VOC) format for object detection. In short, in the dataset, each JPEG image has a corresponding .xml file that contains information about the labeled objects (RBCs, WBCs, Platelets).

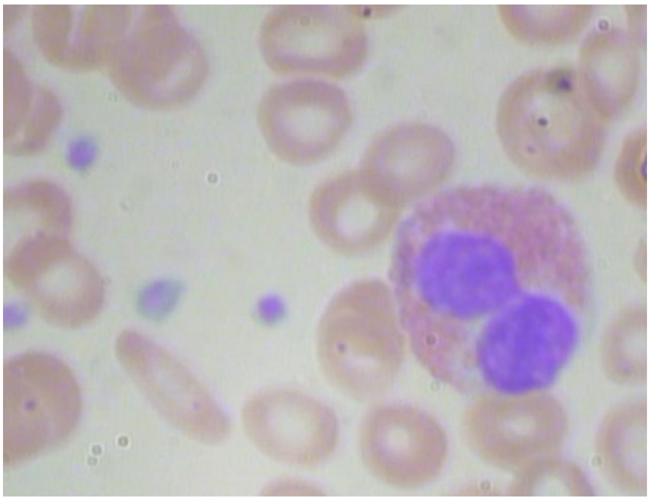

Figure 1: Sample Image_410 of BCCD Dataset.

The annotation files in the dataset are in VOC format, which means a data preprocessing step is necessary for converting each file from VOC format to a YOLO format annotation file, which is a more compact format because YOLO model training for object detection requires a different annotation structure to process the data efficiently. YOLO format annotation files are in text file format (.txt) where each line in the file corresponds to three of the objects-RBCs, WBCs, and Platelets. This preprocessing step is accomplished by creating a Python script that automatically converts all the feed in XML files into corresponding YOLO format annotation txt files with the same name. On top of completing the data preprocessing step, a data configuration file (.yaml) file is created to include paths to the dataset and class names.

2.2. Overview of Models

For this experiment, three different models are being used to run object detection on the BCCD dataset, which are YOLOv5nu, YOLOv8n, and the most recently released YOLO 11n. The experiment will mostly focus on analyzing the blood cell detection results to determine the strengths and weaknesses of the three models due to their differences in model architecture. The main architectures of the YOLO families contain three part-backbone, neck, and head. The backbone uses input images to extract features, the neck is for aggregating features and passing them to the head for prediction, and the head is for categorizing the objects and generating the bounding boxes inside those boxes for object detection.

YOLOv5nu, made for object detection tasks, has the fastest CPU speech in the entire YOLOv5n family. One of the key features of YOLOv5u is it challenges the tradeoff between speed and accuracy by enabling accurate real-time detection. It uses Cross Stage Partial Network-Darknet (CSP-Darknet) structure for the backbone, and Cross Stage Partial Path Aggregation Network (CSP-PAN) and Spatial Pyramid Pooling-Fast (SPPF) structures for the neck. Moreover, the YOLOv8n model for object detection tasks incorporates an advanced backbone and neck architecture compared to previous YOLO versions. It uses a Crossover and Fusion (C2f) module for backbone architecture, and an anchor-free split Ultralytics head to improve object detection and feature extraction accuracy and speed. Lastly, YOLOv11 is the most recent version of the YOLO family. Similarly, YOLOv11n is designed for object detection tasks. It uses a newly innovative Cross Stage Partial with Kernak size 2(C3k2) for the backbone, Spatial Pyramid Pooling-Fast (SPPF), and Convolutional block with Parallel Spatial Attention (C2PSA) structures for the neck to enhance feature extraction and models performance.

3. Experiment Results and Discussion

3.1. Experimental Setup and Procedure

For the experiment configuration, the dataset, which contains 364 images and their corresponding annotation files, is divided into a training set, an evaluation set, and a testing set, with a split ratio of 78:11:11, respectively. Specifically, the training set comprised 285 images, the evaluation set contained 40 images, and the testing set contained 39 images.

After setting up the experiment configuration, the training of three different types of models (YOLOv5nu, YOLOv8n, and YOLO11n) is accomplished by simply running the training command on the BCCD dataset's command-line interface (CLI), where configuration can be directly set inside the command line. During the training process of each model, the learning rate and batch size are adjusted to ensure the model achieves optimal performance and converges efficiently.

The specific configurations used for training the YOLOv5nu, YOLOv8n, and YOLO11n models are shown in Table 1. This experiment trains all three models over 30 epochs with a batch size of 8, with an input image of size 640 × 640, a learning rate of 0.001, and using Adam Optimizer.

Table 1: YOLOv5un, YOLOv8n, and YOLO11n training configurations.

Configuration	Value
Learning Rate	0.001
Batch Size	8
Number of Epoch	30
Optimizer	Adam
Image Size	640

3.2. Experimental Results and Analysis

Analyzing the experiment results of all three models, their metrics indicate that the models have converged, as their performance metrics, such as precision, recall, and mAP, have become stable over the last 10 epochs.

Precision is the ratio of the true positive to the total predicted positive (Total of the true positive and false positive). Calculating the average of the metrics across the last 10 epochs, all three models have a precision rate above 80%, which means all models are unlikely to predict molecules when there are none. Among the three modules, the YOLOv8n model has the highest precision rate at predicting molecules of nearly 89% as shown in the second row of Table 2.

The recall rate is the ratio of true positives to the total actual positive. All three models have a recall rate of 90% and above, indicating that they are all good at detecting molecules when there are. Among them, YOLOv5nu models have the highest recall rate of 95% presented in the third row of Table 2.

Intersection over Union (IoU) measures the overlap between the predicted bounding box and the true bounding box. ‘Mean Average Precision at IoU = 50’ (mAP@50) only calculates the precision for predictions that have IoU greater than or equal to 0.5. As shown in the fourth row of Table 2, all models have an average mAP@50 at around 94% percent indicating a high model’s detection performance with a less strike tolerance for bounding box overlap. On the other hand, presented in the fifth row of Table 2, for mAP50-90, calculated using various IoU thresholds from 0.50 to 0.95, all models have a much lower metric of around 67%. This indicates that the models’ performance across multiple IoU thresholds is unsatisfactory for molecular detection.

Table 2: Average performance metrics of the final 10 epochs for various models.

Metrics	YOLOv5nu	YOLOv8n	YOLO11n
Precision	0.830968	0.88999	0.874569
Recall	0.951767	0.91197	0.908924
mAP50	0.939269	0.942764	0.938761
mAP50-95	0.672152	0.676476	0.669005

Figures 2, 3, and 4 are examples of predicted results on the test set using corresponding models. As shown in all three figures, the square bounding boxes are slightly bigger than the molecules, which might be due to most molecules being irregularly circular. To ensure that the bounding box encloses the entire cell, the bounding boxes are not tightly surrounding the edges of each cell, which might be the cause of a low mAP50-90 rate.

For the completely visible cells, as shown in the figures, all three models can detect and classify the cell type with high accuracy. However, models missed detect some cells that are only partially visible, with a portion of them cut off at the edge of the image. Comparing all three figures, YOLOv5nu has the least miss detections, and YOLO11n has the most miss detections as it is misidentifying some cells with only a small portion cut off.

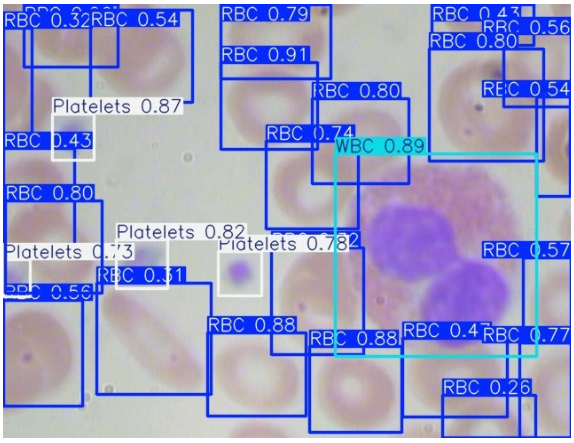

Figure 2: Model Prediction Result on Image_410 Using YOLOv5nu.

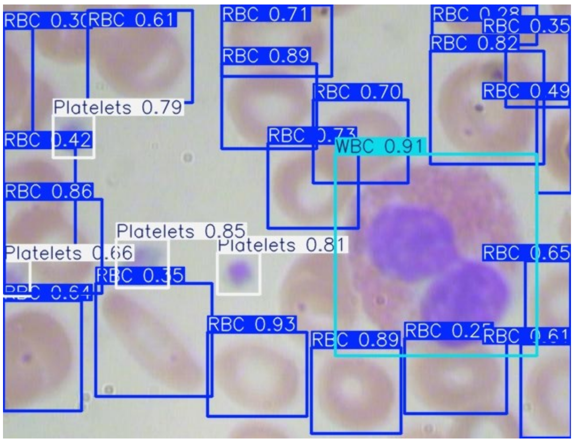

Figure 3: Model Prediction Result on Image_410 Using YOLOv8n.

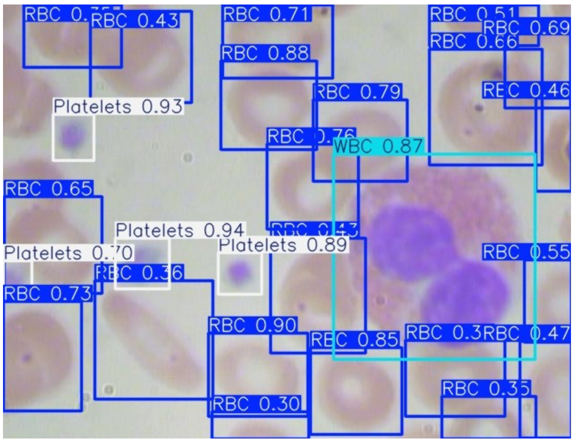

Figure 4: Model Prediction Result on Image_410 Using YOLO11n.

4. Conclusion

The main objective of the study is to train three different YOLO models (YOLOv5nu, YOLOv8n, and YOLO11n) on the same blood cell dataset and analyze their accuracy and effectiveness to determine their strengths and weaknesses in detecting RBCs, WBCs, and Platelets in medical images. Overall, both three models have relatively high accuracy and recall rates. However, all of them have relatively low mAP50-95 rates, which could be caused by several factors. One factor might be the bounding boxes not tightly enclosing the molecules due to irregular circular shapes. Another factor might be the dataset contains images that have heavily overlapping molecules, which might prevent it from distinguishing them and ultimately lead to poor localization and low mAP50-95 rate. In addition to the low mAP50-95 rate, there is also a limitation to this experiment. Due to the computational constraints, a relatively small dataset size was chosen for the experiment. However, an effort is made to reduce the negative effect, such as by carefully choosing the configurations of each model. Moreover, the experiment only contains three classes that are relatively distinct from one another. Overall, even with a relatively low mAP50-95 rate, YOLO models already demonstrate a good precision and recall rate in identifying RBCs, WBCs, and Platelets. This indicates a high possibility that its application in the medical field will significantly increase clinical diagnosis efficiency and accuracy.

References

[1] *Amit, Y., Felzenszwalb, P., & Girshick, R. (2021). Object detection. In Computer Vision: A Reference Guide (pp. 875-883). Cham: Springer International Publishing.*

[2] *Baburaj, B., Roy, A. J., & Brite Sun George, D. T. Y. M. A COMPARATIVE STUDY ON DIFFERENT YOLO VERSIONS.*

[3] *Jiang, P., Ergu, D., Liu, F., Cai, Y., & Ma, B. (2022). A review of YOLO algorithm developments. Procedia Computer Science, 199, 1066-1073.*

[4] *Qureshi, R., RAGAB, M. G., ABDULKADER, S. J., ALQUSHAIB, A., SUMIEA, E. H., & Alhussian, H. (2023). A comprehensive systematic review of yolo for medical object detection (2018 to 2023).*

[5] *Kaur, A., Singh, Y., Neeru, N., & others. (2022). A survey on deep learning approaches to medical images and a systematic look up into real-time object detection. Archives of Computational Methods in Engineering, 29, 2071–2111.*

[6] *Shahzad, M., Ali, F., Shirazi, S. H., Rasheed, A., Ahmad, A., Shah, B., & Kwak, D. (2024). Blood cell image segmentation and classification: A systematic review. PeerJ Computer Science, 10, e1813.*

[7] *Alhazmi, L. (2022). [Retracted] Detection of WBC, RBC, and Platelets in Blood Samples Using Deep Learning. BioMed research international, 2022(1), 1499546.*

[8] *Shinde, S., Oak, J., Shrawagi, K., & Mukherji, P. (2021, December). Analysis of WBC, RBC, platelets using deep learning. In 2021 IEEE Pune Section International Conference (PuneCon) (pp. 1-6). IEEE.*

[9] *Cruz, D., Jennifer, C., Castor, L. C., Mendoza, C. M. T., Jay, B. A., Jane, L. S. C., & Brian, P. T. B. (2017, December). Determination of blood components (WBCs, RBCs, and Platelets) count in microscopic images using image processing and analysis. In 2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM) (pp. 1-7). IEEE.*

[10] *Yeerjiang, A., Wang, Z., Huang, X., Zhang, J., Chen, Q., Qin, Y., & He, J. (2024). YOLOv1 to YOLOv10: A Comprehensive Review of YOLO Variants and Their Application in Medical Image Detection. Journal of Artificial Intelligence Practice, 7, 112-122.*