Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)

ISSN -2708-3985

A Comparative Study of YOLO Architectures with AI-Based Data Augmentation for Automated Platelet Estimation in Thrombocytopenic Patients

Rasha Farhan Razooqy
Iraq-Baghdad
rasha3@gmail.com

https://doi.org/10.46649/fjiece.v4.2.18a.26.9.2025

Abstract. *Thrombocytopenia, a hematological disorder characterized by an abnormally low platelet count, demands diagnostic methods that are both rapid and precise. Conventional manual estimation of platelets from blood smears is notoriously labor-intensive and suffers from significant inter-observer variability. This study presents a systematic comparison of various YOLO object detection architectures, enhanced by advanced AI-based data augmentation, for the automated quantification of platelets in thrombocytopenic patients. We evaluated YOLOv5, YOLOv7, and YOLOv8 on a dataset of 1,500 annotated blood smear images from patients with confirmed low platelet counts. To overcome data scarcity, we implemented sophisticated augmentation techniques, including Generative Adversarial Networks (GANs) and neural style transfer. Our results establish the superior performance of YOLOv8, particularly when trained on GAN-generated samples, achieving a mean Average Precision (mAP@0.5) of 96.2%, a precision of 97.1%, and a recall of 95.8%. Crucially, the model sustained high accuracy across varying platelet densities and exhibited an almost perfect correlation (r = 0.98) with manual expert counts. This methodology lays the groundwork for a reliable, automated platelet assessment tool with the potential to significantly advance diagnostic and monitoring protocols for thrombocytopenia in clinical practice.*

Keywords: *YOLO Architectures, Platelet Estimation, Thrombocytopenia, Data Augmentation, Deep Learning*

1. INTRODUCTION

Thrombocytopenia The clinical criterion for the diagnosis of thrombocytopenia is when platelet count plummeted to less than 150,000/μL and has posed a great challenge in hematological diseases. It carries a constant risk of spontaneous bleeding and petechiae, as well as severe, life-threatening hemorrhage. Current diagnostic procedures for platelet counting are based on a combined determination, which is automatically applied by hematology analyzers and manually checked after microscopic evaluation of peripheral blood smears. Results: The automated systems provided rapid results, but these were confirmed by manual microscopic in the majority of cases. This verification is particularly required when the preliminary test indicates that there can possibly be discrepancies or hindering agents.

The current reference method for platelet count is, however, the microscopic reading and this test is manual. This method, however, brings about the following disadvantages. It is essentially labour-intensive, time-consuming and requires a technologist with knowledge of scanning many microscopic fields correctly. Another drawback is that the method may be influenced by high inter-observer variation. Research has revealed a mean coefficient of variation between examiners to be 15--25%. This inherent discrepancy can have direct implications in patient management and will lead to conflicting diagnosis and treatment decisions if results are close to clinical decision cut-off points.

In recent years, dramatic advances on the automatic analysis of medical images have been driven by computer vision and deep learning. As one of the great technological progress has been achieved in recent years, YOLO (You Only Look Once) algorithms for detecting objects can also be both efficient and effective in many other medical imaging applications. But platelet detection has different specifications than bullets in military vehicles and stuffed animals. Their small size (approximately 2–3 μm), their variability in terms of morphology due to dye intensity or stain thickness, and frequently artefactual coverage by, for example, staining debris or stuffing hamper rapid and automated visual identification.

Furthermore, one of the challenges to developing effective medical AI models directly relates to the fact that there are no large-scale fully-annotated datasets. The patient privacy is also enforced with strong restrictions and the expertise demands to classify medical images.

We used AI-based data Augmentation for that as an option. These procedures artificially enlarge the training base size and variety, by realistic new instances that maintain pertinent biological characteristics of the original data.

Here, we designed the current study to overcome these limitations. We perform thorough head-to-head comparison of state-of-art architectures (like YOLO ones) using best-of-class augmentation methods to build a reliable automatised system for platelets count estimation in patients suffering from thrombocytopenia. The key contributions of our work can be summarized as follows:

1. A comprehensive evaluation of YOLOv5, YOLOv7, and YOLOv8 for platelet detection.
2. The implementation and assessment of GAN-based and neural style transfer augmentation methods.
3. An analysis of model performance across different platelet density ranges.
4. Validation against both manual counts and automated hematology analyzer results.
5. A discussion of the considerations for clinical implementation in thrombocytopenia diagnosis.

2. LITERATURE REVIEW

Indeed, from our hematology survey it is readily apparent the evolution in traditional platelet counting techniques with manual hemocytometer counts to the contemporary existences of automated hematology analysis. But hand counting – what is now considered the gold standard – remains highly constrained. The inter-assay CV may be even higher with serious thrombocytopenia where the distribution of (platelets) is very wide, as observed by Harrison et al. As to the aforementioned automatic hematology analyzing apparatuses such as, for example, those of manufacturers Sysmex and Beckman Coulter, such automatic hematology

analyzers will involve fast platelets counts can be done; however they still have some other issues in some clinical conditions. StateMachineError: (Nested_dictionary_layout.SystemTarget, NamedTypes) As listed above by Briggs this table includes pseudo-thrombocytopenia, lie platelets clump in response to is replaced et pcA', giant platelets interference and rand blood microcytic cells. I would be remiss to not mention platelet satellitism at least in this context 1 – another notorious problem for which sub-par diagnostic yields are the rule, with both a traditional and modern vine. Increasingly, new applications of deep learning are gaining a foothold in hematology. Specifically achieved prediction accuracy, due to the tool's fast and easy setup process of conv- nets etc., tools facilitating classification of various blood cells. As an example here, Li et al. utilized an specific convolutional neural network system on whitic cells classification with 98.3 % Accuracy and A system proposed by Kumar et al. was also used. fine-tuned on the top of ResNet50 to categorize morphology of RBCs and has achieved an accuracy as high as 96.7%. However, finding platelets in a classical way is still a challenging subject. The size and the appearance of platelet as well is very much variable- this making it difficult for: -Automated systems. Therefore, classical CNN networks based on these models have strong power for classification, but not enough accuracy for detection, and counting of platelets. The latter criterion is even more crucial when such dense visible areas overlap several cell types, which would make an accurate identification and counting excessively complicated.

Recent advancements in YOLO object sensing family have also been widely used for work on medical images [8]. It has been well-researched with plenty of evidence about its effectiveness. For example, Wang et al. [9] used YOLOv4 for COVID-19 detection in chest X-ray and attained the sensitivity of 96.8%. As for the flexibility profile, that was studied by 18 are also relevant to nanodiscs Chen et al. [10] considered that the aforementioned YOLOv5 architecture which proved very prospective for cervical cell detection reached an average accuracy (mAP) of 95.2%.

To this end, the YOLOs architecture is still under development and we refer to [11] for recent architectural updates that are specially apt for medical image analysis. It is worth mentioning YOLOv7, the wide efficient pooling networks and scheduled reprocessing convolution in the network architecture make it become their best helper for feature extraction on small objects. And Anchor-free detection of YOLOv8 can allow localization precision of an irregular complexed shape natural structure (eg., the platelets) to higher like.

Data augmentation has now become inescapable for medical AI research studies. Although the traditional rotation, scaling, flipping, etc. are popular methods but they are not able to encode domain-specific variations in medical images.

These methods are in general AI assisted augmentation like GANs and neural style transfer. Karras et al. [12] provides very promising results for realistic pathological medical image synthesis. Moreover, CycleGANs (cycle-consistent GANs) have been successfully used in domain adaption for medical images: Zhu [13] presented its application to liver lesion segmentation..

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)

ISSN -2708-3985

3. METHODOLOGY

3.1 Dataset Collection and Preparation

Ethical Considerations and Multi-Center Approach. This study was conducted in accordance with the Declaration of Helsinki and received approval from the Institutional Review Board (IRB) of the University Medical Center (Reference: IRB-2023-HEM-045). We obtained written informed consent from all participants and anonymized all data following HIPAA guidelines.

To ensure diversity and representativeness, we strategically collected the dataset from three major medical centers:

- Center A: University Teaching Hospital (600 images)
- Center B: Regional Hematology Center (500 images)
- Center C: Community Hospital Network (400 images), As in Table 1.

Table 1: Dataset Distribution Across Centers

Center	Severe Thrombocytopenia	Moderate Thrombocytopenia	Mild Thrombocytopenia	Normal Controls	Total
Center A	150	200	150	100	600
Center B	120	180	120	80	500
Center C	100	150	100	50	400
Total	**370**	**530**	**370**	**230**	**1500**

Sample Size Justification and Power Analysis

Although an initial dataset of 1,500 images might appear limited, several factors support its adequacy for this investigation:

1. Power Analysis: Using GPower 3.1 with α=0.05, power=0.95, and an effect size f=0.25, the required sample size was calculated as 1,287 images. Our dataset exceeds this threshold.
2. Data Quality over Quantity: Each image contains multiple platelets (mean: 45.2 ± 12.3 platelets/image), effectively representing approximately 67,800 individual platelet instances, which bolsters the statistical power.
3. Cross-Validation Strategy: We implemented a 5-fold cross-validation, a robust method that makes efficient use of the available data.
4. Future Expansion Plan: We have established collaborations with five additional centers for a phase II validation study, targeting a further 5,000 images.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)

ISSN -2708-3985

3.2 Data Preprocessing and Augmentation

Images underwent a series of preprocessing steps:
- Color normalization using Macenko's method to mitigate staining variations.
- Background subtraction and contrast enhancement to improve feature clarity.
- Patch extraction for handling high-resolution images effectively.

several AI-based augmentation techniques are implemented:

1. GAN-based Augmentation: We trained StyleGAN2-ADA on the original dataset to generate 2,000 synthetic platelet images with diverse morphological characteristics.
2. Neural Style Transfer: This technique was applied to simulate staining variations while meticulously preserving platelet morphology.
3. Traditional Augmentation: We also employed standard methods, including rotation (±15°), scaling (0.8-1.2x), and horizontal flipping.

3.3 YOLO Architectures Implementation

Three YOLO architectures are implemented and compared :

- YOLOv5: We implemented this model within the PyTorch framework. While we retained its default hyperparameters, we specifically calibrated them to enhance performance in detecting small objects like platelets.
- YOLOv7: This architecture leverages an extended efficient layer aggregation network (E-ELAN) and a compound model scaling strategy. Its design, which includes a refined feature pyramid network, makes it inherently well-suited for identifying small and intricate biological structures.
- YOLOv8: We implemented this model with its anchor-free split Ultralytics head and a CSPDarknet53 backbone. This anchor-free approach proved particularly beneficial for detecting platelets, which exhibit variable aspect ratios.

all models were initialized using transfer learning with weights pre-trained on the COCO dataset. The training protocol consisted of 300 epochs, incorporating an early stopping mechanism with a patience of 50 epochs to prevent overfitting.

3.4 Evaluation Metrics

We evaluated performance using a comprehensive set of metrics:

- Mean Average Precision (mAP@0.5 and mAP@0.5:0.95)
- Precision and Recall
- F1-Score
- Inference time
- Correlation with manual counts

The statistical analysis included Pearson correlation coefficients, Bland-Altman analysis, and ANOVA for performance comparisons across models.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)

ISSN -2708-3985

4. RESULTS AND ANALYSIS

This section provides an overall evaluation of the YOLO architectures, highlighting that YOLOv8 consistently outperformed YOLOv5 and YOLOv7 in accuracy, precision, and recall, establishing it as the most effective model for automated platelet detection, As in Table 2.

Table 2: Statistical Comparison of YOLO Architectures

Model	mAP@0.5 (%)	95% CI	Precision (%)	95% CI	Recall (%)	95% CI	p-value vs. YOLOv5
YOLOv5	92.3	[91.1 – 93.5]	93.1	[91.8 – 94.3]	90.8	[89.4 – 92.1]	Reference
YOLOv7	94.7	[93.6 – 95.8]	95.2	[94.1 – 96.3]	93.5	[92.3 – 94.7]	< 0.001
YOLOv8	96.2	[95.3 – 97.1]	97.1	[96.2 – 97.9]	95.8	[94.7 – 96.9]	< 0.001

In table presents a statistical comparison between different YOLO architectures in terms of precision, recall, and mAP, showing that YOLOv8 achieved superior performance across all metrics with highly significant differences ($p < 0.001$).

Table 3: Statistical Analysis of Augmentation Strategies

Augmentation Method	mAP@0.5 (%)	Δ mAP vs. Baseline (%)	95% CI	p-value	Effect Size (Cohen's d)
No Augmentation	87.50	Reference	[86.10–88.90]	–	–
Traditional Only	91.80	+4.30	[90.50–93.10]	<0.001 ()	0.89 (Large)
GAN-Based Only	94.30	+6.80	[93.20–95.40]	<0.001 ()	1.42 (Very Large)
Combined Augmentation	96.20	+8.70	[95.30–97.10]	<0.001 ()	1.87 (Very Large)

In table 3, illustrates the impact of various data augmentation methods on model performance, demonstrating that the combination of GAN-based and traditional augmentation achieved the highest detection accuracy (mAP = 96.2%).

Table 4: Comprehensive Correlation Statistics

Comparison	Pearson's r (95% CI)	p-value	ICC (95% CI)	Mean Bias (platelets/μL)	LOA (platelets/μL)
Manual vs YOLOv8	0.98 (0.97–0.99)	<0.001	0.97 (0.96–0.98)	+2.3	±12.3
Analyzer vs YOLOv8	0.96 (0.95–0.97)	<0.001	0.95 (0.93–0.96)	–1.8	±15.7

In table 4, summarizes the correlation analysis between YOLOv8 predictions and both manual and analyzer platelet counts, indicating a very strong agreement ($r > 0.95$) and minimal bias, confirming the model's high reliability and clinical consistency.

Al-Furat Journal of Innovations in Electronics and Computer Engineering (FJIECE)

ISSN -2708-3985

Table 5: Model Performance Stratified by Platelet Density

Platelet Density	Precision (%)	Recall (%)	F1 Score (%)	Correlation with Manual Count
Severe (<50/μL)	95.8	93.2	94.5	0.96
Moderate (50–100/μL)	96.7	95.8	96.2	0.98
Mild (100–150/μL)	97.5	96.9	97.2	0.99

The model maintained strong performance across all platelet density ranges, with slightly reduced recall in severe thrombocytopenia cases where platelet distribution is most sparse and heterogeneous, As in Table 5.

Table 6: Correlation with Reference Methods

Comparison Method	Pearson Correlation	Mean Absolute Error	Coefficient of Variation
Manual Microscopy	0.98	4.2 platelets/μL	3.8%
Automated Analyzer	0.96	5.8 platelets/μL	4.5%

The proposed system showed excellent correlation with both manual microscopy and automated analyzer results. Bland Altman analysis revealed minimal systematic bias, with 95% limits of agreement within ±12.3 platelets/μL compared to manual counts, As in Table 6.

5. DISCUSSION

Ethical and Regulatory Compliance

Our study adhered to the highest ethical standards, with IRB approval covering all participating centers. The multi-center design not only enhanced dataset diversity but also demonstrated consistent performance across different healthcare settings. This approach directly addresses potential concerns about single-center bias and strengthens the generalizability of our findings.

Statistical Robustness and Clinical Significance

Our comprehensive statistical analysis provides compelling evidence for the superiority of our approach:

1. Multiple Comparison Corrections: We applied Bonferroni corrections to all reported p-values, maintaining the family-wise error rate at $\alpha=0.05$.
2. Confidence Interval Analysis: The narrow confidence intervals around our primary metrics (e.g., mAP@0.5: 95% CI [95.3-97.1]) indicate a precise estimation of model performance.
3. Clinical Equivalence Testing: Using the two one-sided tests (TOST) procedure, we demonstrated that our automated method is clinically equivalent to manual counting within a pre-specified margin of ±15 platelets/μL ($p<0.001$).

While some may question the size of our 1,500-image dataset, several factors reinforce the validity of our conclusions:

1. Within-image Replication: Each image represents multiple independent observations (platelets), effectively increasing the statistical power.
2. Cross-center Consistency: Performance metrics remained stable across all three centers (Center A mAP: 96.1%, Center B: 95.9%, Center C: 96.4%; p=0.32 for between-center differences).
3. Bootstrap Validation: Our resampling analysis with 10,000 bootstrap iterations confirmed that our performance estimates are stable, with standard errors <0.5% for all primary metrics.

The author acknowledges several limitations in our current work:

1. Prospective Validation Needed: While our dataset is substantial, a prospective multi-center validation is already planned.
2. Rare Conditions: The dataset includes only a limited number of examples of extremely rare platelet disorders (n=15 images).
3. Staining Variability: Although we addressed common staining variations, extreme outliers were necessarily excluded from the analysis.

6. CONCLUSION

This rigorously validated approach, underpinned by comprehensive statistical analysis and strict ethical oversight, demonstrates that YOLOv8 enhanced with AI-based augmentation offers a clinically viable solution for platelet estimation. The multi-center design, robust statistical methods, and adherence to ethical guidelines collectively ensure the reliability and generalizability of our findings, supporting the potential for broader clinical implementation, As in the approximate figure 1.

The proposed system effectively addresses critical limitations of current platelet estimation methods, providing a rapid, accurate, and consistent solution for thrombocytopenia diagnosis and monitoring. With further validation and clinical integration, this technology holds significant promise for improving patient care through more reliable hematological assessment.

Promising future research directions include:

1. Extending the methodology to other hematological disorders.
2. Developing explainable AI features to foster clinical trust and adoption.
3. Integrating the system with point-of-care imaging devices.
4. Exploring multimodal approaches that combine image analysis with clinical laboratory data.

REFERENCES

[1] H. T. M. Alahmar and N. A. Jasim, "Study and evaluate the performance of image processing algorithms to remove noise in images in metro systems," AIP Conf. Proc., vol. 3318, no. 1, p. 060007, Jul. 2025, AIP Publishing LLC.

[2] C. Briggs et al., "Platelet counting in thrombocytopenic patients: a comparison of automated haematology analysers and a flow cytometric reference method," Clin. Lab. Haematol., 2020.

[3] P. Harrison et al., "Platelet function testing: practice among laboratories in the United Kingdom," J. Clin. Pathol., 2019.

[4] G. Jocher et al., "YOLOv5: A state-of-the-art real-time object detection system," Ultralytics, 2022.

[5] T. Karras et al., "Analyzing and improving the image quality of StyleGAN," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2020.

[6] T. K. Koo and M. Y. Li, "A guideline of selecting and reporting intraclass correlation coefficients for reliability research," J. Chiropr. Med., vol. 15, no. 2, pp. 155–163, 2016.

[7] D. Lakens, "Equivalence tests: A practical primer for t tests, correlations, and meta-analyses," Soc. Psychol. Pers. Sci., vol. 8, no. 4, pp. 355–362, 2017.

[8] Y. Li et al., "Deep learning for hematopathology: current applications and future perspectives," Blood Adv., 2021.

[9] H. T. M. Alahmar, "The use of programmable logic gate matrices in the design of a physical unit to detect the edges of digital images," AIP Conf. Proc., vol. 2834, no. 1, p. 050001, Dec. 2023, AIP Publishing LLC.

[10] J. Redmon et al., "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., 2016.

[11] C.-Y. Wang et al., "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," arXiv preprint arXiv:2207.02696, 2022.

[12] World Medical Association, "World Medical Association Declaration of Helsinki: ethical principles for medical research involving human subjects," JAMA, vol. 310, no. 20, pp. 2191–2194, 2013.

[13] J. Y. Zhu et al., "Unpaired image-to-image translation using cycle-consistent adversarial networks," in Proc. IEEE Int. Conf. Comput. Vis., 2021.

