International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Blood Cancer Cell Detection Using YOLO With 3D Depth Imaging for Automated Hematological Diagnosis

Darshan M[1], Nayim A Dalawai[2], Prathamraj P B[3], Aditya C. R[4]

[1]Student, AI ML, Vidyavardhaka College of Engineering
[2,3]Assistant Engineer, Robotics Engineer, Stem Avishkar Pvt Ltd.
[4]Associate Professor, AI ML, Vidyavardhaka College of Engineering

Abstract
Early detection of blood cancers remains a critical challenge due to the limitations of conventional manual blood smear examinations, which are slow and dependent on expert interpretation. In response, this study presents an AI-driven approach that leverages the YOLO object detection framework paired with a standard webcam for real-time analysis of blood samples. Trained on a dataset of major blood cell types, the system demonstrates promising capability in distinguishing abnormal cells. By offering a cost-effective and accessible diagnostic aid, the method shows potential to support pathologists in the early screening of leukemia, lymphoma, myeloma, and related disorders.

Keywords: Blood Cancer Detection, YOLO Object Detection, 3D Depth Imaging, Automated Hematological Diagnosis, Deep Learning in Healthcare, Medical Image Analysis, Computer-Aided Diagnosis, Artificial Intelligence in Oncology

1. INTRODUCTION

Blood cancers such as leukemia, lymphoma, and myeloma originate in the bone marrow or lymphatic system and disrupt the normal production and function of blood cells. These malignancies often cause structural and morphological changes in white blood cells, red blood cells, and platelets, which can be observed through microscopic analysis of blood smears. Identifying such abnormalities at an early stage plays a decisive role in guiding diagnosis and treatment. However, traditional examination methods remain labor-intensive, subjective, and highly dependent on expert evaluation.

Advances in artificial intelligence have opened new possibilities for automated screening, particularly through deep learning–based object detection models. Among these, the YOLO (You Only Look Once) framework has gained recognition for its ability to perform rapid and accurate localization and classification of objects within images. When applied to hematological samples, YOLO can detect and differentiate cell types in real time, offering a pathway to faster and more consistent identification of irregular cells linked to blood cancers.

By combining accessible imaging tools with advanced detection algorithms, this approach introduces a direction where diagnostic support systems become faster, more reliable, and potentially more available in diverse clinical settings. The promise of such methods invites deeper exploration into how technology can complement medical expertise in the early detection of life-threatening diseases.

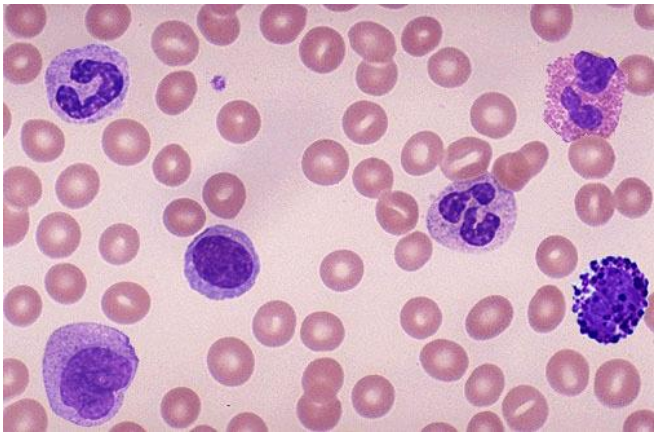


Figure 1. Representative microscopic image of a peripheral blood smear.

The image displays a mixture of red blood cells (pink circular structures) and various types of white blood cells (purple-stained cells with distinct nuclei). Normal red blood cells appear as uniform discs, while the white blood cells exhibit different shapes and granule patterns depending on their type (such as neutrophils, lymphocytes, and monocytes). Structural or morphological abnormalities in these cells often serve as important indicators of blood cancers like leukemia, lymphoma, and myeloma. Detecting such variations is essential for early diagnosis, and serves as the foundation for applying deep learning–based models such as YOLO for automated recognition and classification.

2. Background Study

In India, blood cancers have emerged as a significant public health challenge, with both new cases and deaths steadily increasing over time. Projections from national cancer registries indicate that the incidence of leukemia, lymphoma, and myeloma is likely to climb further by 2025, with the highest impact seen among older age groups. Studies also show that the risk rises sharply with advancing age, and clear differences between men and women provide important clues about how these diseases affect the population.

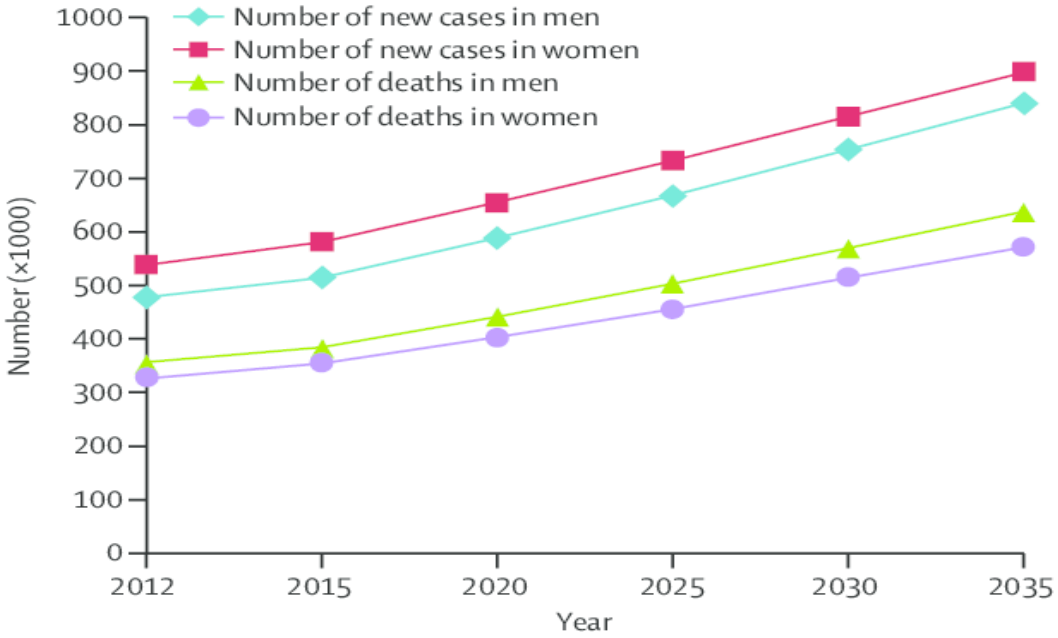


Figure 2. Projected trends of overall cancer incidence and mortality in India (up to 2025) based on national registry estimates and global projections.

Cancer is emerging as an increasingly serious health issue in India. The National Cancer Registry Programme (NCRP) estimates **1.46 million new cancer cases in 2022**, and projects this to rise to around **1.57 million by 2025**, marking a roughly **12.8% increase** over the period Indian Journal of Medical Research+2Indian Journal of Medical Research+2. In tandem, the country's overall disease burden—measured in **disability-adjusted life years (DALYs)**—is expected to climb from **26.7 million in 2021** to **29.8 million by 2025** BioMed Central.

Among all cancers, blood-related types—including **leukemias, lymphomas, and multiple myeloma**—play a significant role. According to GLOBOCAN 2022 data, India saw over **120,000 new blood cancer cases** that year; **leukemia** alone accounted for nearly **49,883 diagnoses**, followed by **non-Hodgkin lymphoma (~39,736 cases)** and **Hodgkin lymphoma (~9,611)** Indian Journal of Medical Research.

Alarmingly, recent regional observations point to rising trends among younger populations. In **New Delhi (2025)**, there are at least **3,000 annual cases of acute myeloid leukemia (AML)**, with growing incidence among adults aged **30–40** The Times of India+1. Similarly, in **Nagpur (2024–25)**, cases of pediatric acute lymphoblastic leukemia (ALL) rose from **22 in 2021–22** to **35**, reflecting an **82% surge in one year** The Times of India.

3. Technical Background

The YOLO model is a widely recognized deep learning framework for object detection, designed to achieve both high speed and accuracy in real-time applications. Unlike traditional methods that rely on sliding windows or region proposal networks, YOLO treats detection as a single regression task, directly predicting bounding box coordinates, objectness scores, and class probabilities from the entire image in a single pass. The input image is divided into a grid, and each grid cell predicts whether an object is present, along with its location and class label. The model architecture typically consists of three main components: a backbone network for feature extraction, a neck for combining multi-scale features, and a detection head that produces the final predictions. Over successive versions, YOLO has evolved from basic convolutional backbones to more advanced structures such as CSPDarknet and anchor-free detection heads, as seen in YOLOv8. This continuous development has enabled YOLO to detect objects of varying sizes and complexities effectively. In the context of hematology, this capability is highly beneficial since blood smear images contain numerous small and overlapping cells. YOLO can rapidly identify and classify cell types such as neutrophils, lymphocytes, eosinophils, and platelets, while also highlighting abnormal structures that may indicate early signs of leukemia, lymphoma, or myeloma. Its ability to process multiple frames per second using even standard imaging devices like webcams makes YOLO an ideal solution for cost-effective and real-time medical image analysis, providing valuable assistance to pathologists in early screening and diagnosis.

4. YOLOv8 Network Architecture (Ultralytics, 2023)

The figure illustrates the internal design of the YOLOv8 object detection network, which is systematically divided into three primary components: the **Backbone**, the **Neck**, and the **Head**. The **Backbone** is the feature extraction stage, where the input image (commonly 640 × 640 pixels) is processed through successive convolutional (Conv) layers and advanced C2f blocks. These operations progressively downsample the spatial resolution of the image while simultaneously enriching it with deeper semantic representations. As the feature maps pass through the backbone, the model learns to capture low-level visual cues such as edges and textures, as well as higher-level object structures. At the end of this stage,

the Spatial Pyramid Pooling – Fast (SPPF) module is introduced, which is specifically designed to expand the receptive field efficiently. By applying multiple pooling operations in parallel, the SPPF enables the network to better capture multi-scale contextual information, ensuring robust recognition of objects that vary in size and shape.

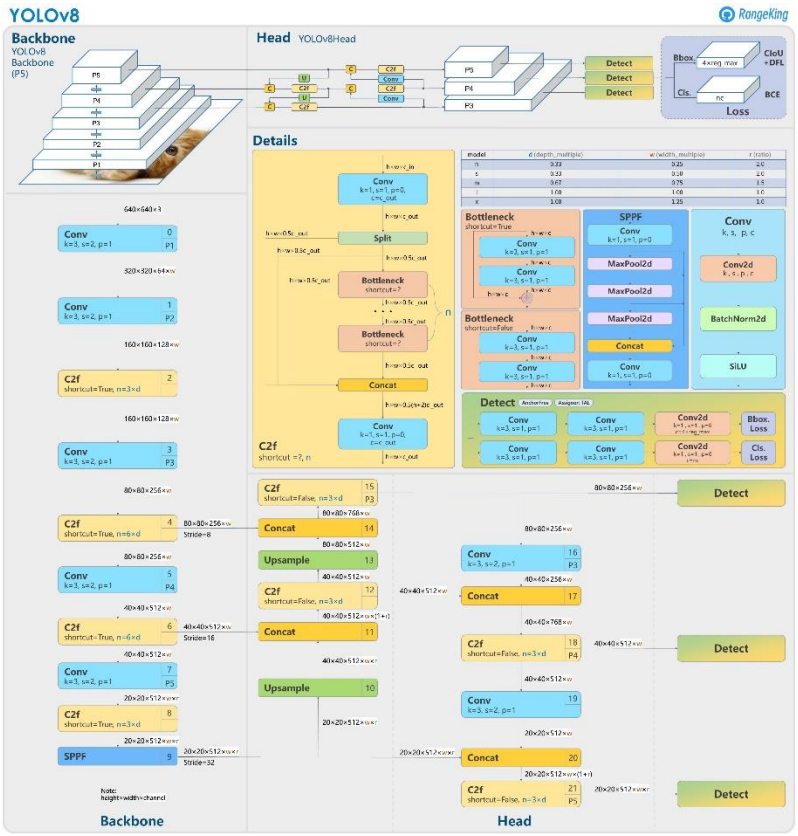


Fig 3., Ultralytics (2023). YOLOv8 Architecture Discussion

Available at: https://github.com/orgs/ultralytics/discussions/3610.

The **Neck** serves as a bridge between the Backbone and the Head, playing a critical role in fusing information from different stages of the network. YOLOv8 employs a combination of Feature Pyramid Networks (FPN) and Path Aggregation Networks (PAN) to achieve multi-scale feature integration. Within this section, features from deeper layers are **unsampled** and then **concatenated** with features from earlier layers. This process merges high-resolution spatial detail with rich semantic information, significantly improving the model's ability to detect both small and large objects. Such capability is particularly crucial for medical applications like blood smear analysis, where cells differ widely in size and often appear in dense or overlapping clusters.

Finally, the **Head** is responsible for generating the actual predictions. Unlike earlier YOLO versions that relied on predefined anchor boxes, YOLOv8 adopts an **anchor-free detection strategy**. For each scale of the feature maps, the Head directly predicts the bounding box coordinates, the objectness score (probability that an object exists in that location), and the class probabilities. This streamlined design not only reduces computational overhead but also improves the model's flexibility when applied to datasets with objects of irregular shapes or varying aspect ratios, such as abnormal blood cells in pathological images.

The **Outputs and Loss** function complete the detection process. The outputs consist of bounding boxes

tightly surrounding each detected object, along with class probabilities that indicate the type of object, such as a neutrophil, lymphocyte, or platelet. To ensure that the model learns effectively, YOLOv8 optimizes its parameters using a compound loss function. This function integrates three penalties: localization loss, which measures how accurately predicted bounding boxes align with ground truth annotations; classification loss, which evaluates the correctness of the predicted class labels; and objectness loss, which assesses the confidence that a box contains a valid object. By balancing these three objectives, the network achieves both high accuracy and real-time inference capability, making it well-suited for critical applications such as early blood cancer screening.

Blood Cell Types Considered for Detection

In this study, the YOLOv8 architecture was employed to detect and classify seven distinct categories of blood cells. The selected cell types represent the most relevant populations for both normal hematology and the identification of abnormalities that may indicate early stages of blood-related disorders. The categories include **basophils, eosinophils, monocytes, neutrophils, platelets, and lymphocytes**. Each of these cells plays a critical role in maintaining immune defense and overall blood physiology, and their abnormal structure or count often serves as an early marker of leukemia, lymphoma, or other hematological malignancies.

The dataset used for training and evaluation was initially sourced from publicly available repositories on **Kaggle**, where blood smear images were annotated according to cell type. To ensure clinical reliability, the dataset was subsequently **validated at Janani Hospital, Gulbarga**, under the guidance of **Dr. Rohan K**, who confirmed the accuracy of cell-type labelling. This process provided a dual advantage: the broad accessibility of online datasets combined with the assurance of expert medical verification. By focusing on these seven blood cell categories, the model was designed not only to distinguish normal haematological features but also to highlight subtle morphological variations that may support the early screening of blood cancers.

Experimental Setup and Model Training

The training of the YOLOv8 model was carried out on a system equipped with an **Intel Core i7-10700F processor (2.90 GHz, 16 cores)**, **16 GB RAM**, and support for **DirectX 12**. The system configuration ensured that computationally demanding tasks such as dataset preprocessing, annotation, and model training could be performed efficiently.

For dataset preparation, the **LabelMe** tool was used to annotate blood smear images, creating bounding boxes for each of the seven target cell types. Since YOLOv8 requires training data in a specific format, the annotated images in **JPG** format were converted into corresponding **TXT files** containing bounding box coordinates and class labels. This conversion ensured compatibility with the YOLO training pipeline.

The training process was executed through the **Anaconda Prompt** environment, where all required dependencies were installed prior to execution. Essential libraries such as **Ultralytics** (for YOLOv8 implementation) and **TensorFlow** (for additional deep learning utilities) were installed and configured. Following successful setup, the training was initiated, and the model weights were iteratively updated. Upon completion, the model generated a final checkpoint file, **best.pt**, representing the best-performing trained weights based on validation accuracy.

For **real-time detection**, the system was integrated with an **Intel 3D Depth Camera**, which provided high-resolution live video streams. The trained YOLOv8 model processed these streams to detect and

classify blood cells in real time. The depth camera enhanced detection accuracy by offering robust image quality and improved spatial information, making the system more effective for practical screening tasks.

End-to-end runtime algorithm

Loop over frames from the webcam

1. **Read frame** $I \in R^{H \times W \times 3}$
2. **Letterbox** to 640×640 while keeping aspect ratio (details below).
3. **Forward pass** through YOLOv8 → multi-scale predictions.
4. **Decode** raw outputs into boxes + class probs.
5. **Compute scores** and **filter** by conf=0.75.
6. **Non-Maximum Suppression (NMS)** to remove duplicates.
7. **Map boxes back** to original frame coordinates.
8. **Draw** boxes/labels; **show** and **save** frame.

Letterbox resize

Let the original frame be H×WH\times WH×W. To fit 640×640 without distortion:

Scale $s = min\{\frac{640}{H}, \frac{640}{W}\}$

Resized size $H' = [sH]$, $W' = [sW]$.

Pad amounts $p_x = \frac{640-W'}{2}$, $p_y = \frac{640-H'}{2}$.

Forward map (orig → net):

x′=s x + p_x , y′= s y+p_y.

Inverse map (net → orig):

x=$\frac{x'-p_x}{s}$, y=$\frac{y'-p_y}{s}$.

Prediction heads and decoding

YOLOv8 is **anchor-free** and predicts at three strides S∈{8,16,32}. At each location p=(u , v) on a stride-SSS feature map, the model outputs:

- **Bounding box** via distance distribution (DFL) for each side:

$$p_{L,}p_T, p_R, p_B \in \Delta\text{^}K$$

Decode each side as an **expected value**:

$$d_L = \sum_{i=0}^{K} i(p_L\ )i, d_T = \sum_{i=0}^{K} i(p_T)i, d_R = \sum_{i=0}^{K} i(p_R)i, d_B = \sum_{i=0}^{K} i(p_B)i.$$

Convert to **pixel distances** by multiplying the stride : $d' = S * d$.

The grid center in input space is

$$x_c = S(u + 0.5), y_c = S(v + 0.5).$$

The **box corners** are

$$x_1 = x_c - d'_L, y_1 = y_c - d'_T, x_2 = x_c - d'_R, y_2 = y_c - d'_B.$$

Objectness logit o → probability via sigmoid σ(o) ∈ (0,1).

Classes logits $z \in R^c$ → probabilities via softmax

$$p_c = \frac{e^z}{\sum_j e^z}, c \in \{1, \ldots\ldots, C\}.$$

For each class ccc, the **final score** at that location is

$$score_c = \sigma(o).p_c.$$

Your conf=0.75 keeps detections with scorec≥0.75.

Non-Maximum Suppression (NMS)

To remove overlapping duplicates per class:

- IoU between boxes A, B:

$$\text{IoU(A,B)} = \frac{|A \cap B|}{|A \cup B|} = \frac{max\,(0, min(x_2^A, x_2^B) - max(x_1^A, x_1^B)) * max\,(0, min(y_2^A, y_2^B) - max(y_1^A, y_1^B))}{(x_2^A - x_1^A)(y_2^A - y_1^A) + (x_2^B - x_1^B)(y_2^B - y_1^B) - |A \cap B|}$$

- Greedy NMS algorithm: g ↓.

Pick top box B1B_1B1; suppress any remaining box Bj with IoU(B1,Bj)>(Bτ(e.g.,0.5). Repeat with the next remaining box.

This yields a set of non-overlapping high-confidence detections.

Mathematical proof" style clarifications

While object detection is empirical, there are a few **provable properties** in your pipeline:

1. **Probability bounds**

Because σ(·) ∈ (0,1) and$\sum_c p = 1(softmax)$,

0<$score_c = \sigma(o)p_c \leq 1$.

Hence your threshold 0.750.750.75 is a **well-defined probability lower bound** for kept detections.

2. **Geometry-consistent decoding** .

The DFL decoding produces **non-negative distances** from the cell center to each side; converting to corners via

(x1, y1, x2, y2)=$(x_c - d'_{L,} y_c - d'_{T,} x_c + d'_R, y_c + d'_B)$

guarantees x2≥x1,y2≥y1, i.e., valid boxes.

3. **Coordinate consistency through letterbox.**

Because the inverse map (x,y)=$\left(\frac{x'-p_x}{s}, \right), ((y' - p_y)/s)$ is **affine**, it preserves straight edges and IoU rankings; therefore, NMS decisions taken in network space remain consistent after mapping back to the original frame.

5. Results

Training Performance Analysis

The training curves illustrate the convergence behavior of the YOLO model over 100 epochs. The loss functions, including box loss, classification loss, and distribution focal loss, consistently decreased during training and validation, indicating effective learning and improved bounding box regression. The evaluation metrics show stable precision, while recall fluctuated during the initial epochs before gradually improving. Both mAP@50 and mAP@50–95 demonstrated an upward trend, reflecting enhanced detection accuracy as training progressed. These results confirm that the model was able to generalize well to unseen validation data while maintaining robustness in cell detection tasks.

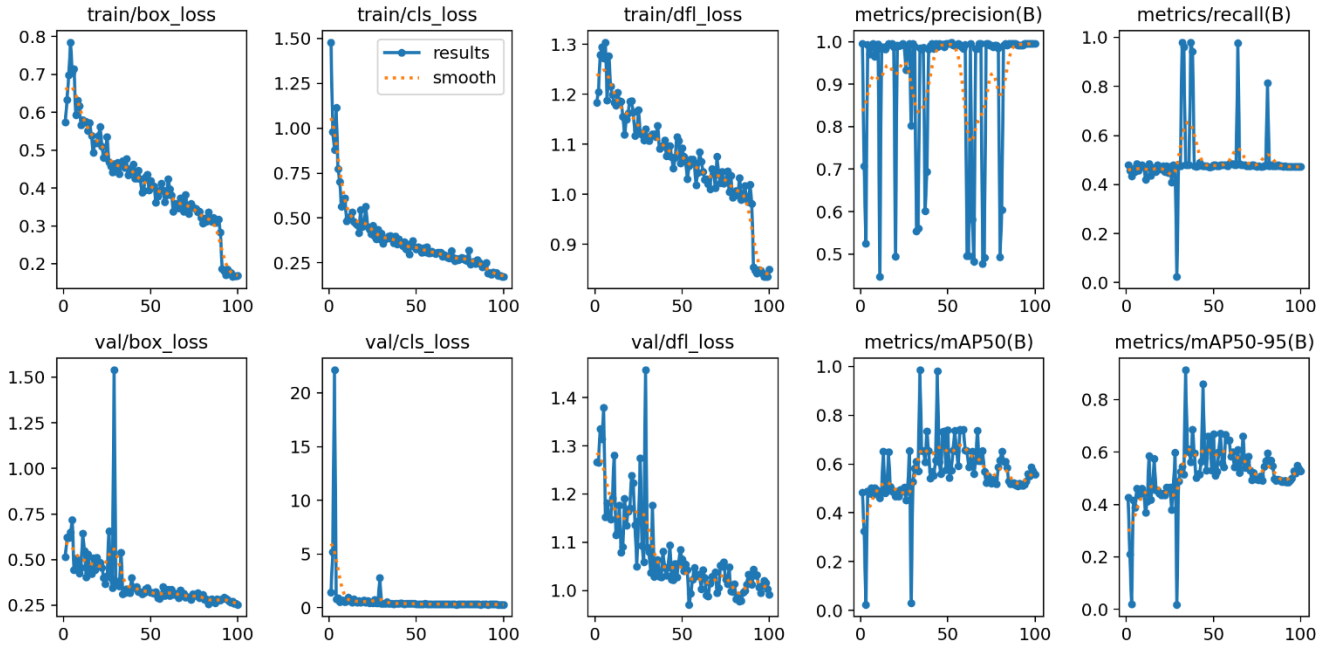
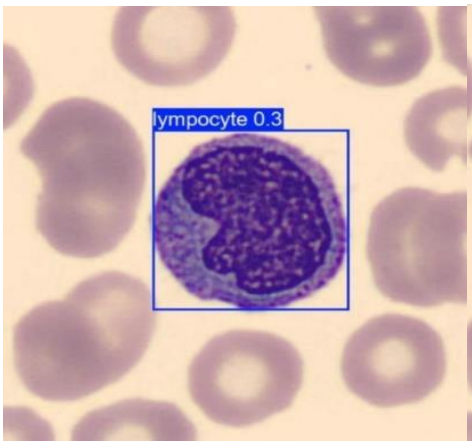


Fig 4. Training and validation loss and accuracy curves of YOLO model

Blood Cell and Cancer Cell Detection with 3D Depth Imaging

In this study, a 3D depth camera was integrated with YOLO to detect and classify blood cells. The depth information provided by the camera enhanced the quality of cell visualization, allowing clearer differentiation between normal and abnormal structures. The system was able to identify monocytes, lymphocytes, platelets, and basophils with high confidence, as illustrated in Figures X–X. Extending this approach towards malignant cell recognition offers a promising pathway for early blood cancer detection. The combination of real-time object detection and depth imaging reduces reliance on manual microscopy, providing faster and more reliable support for clinical diagnostics.



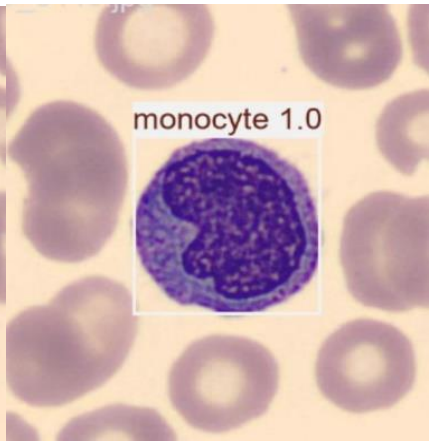


Fig 5. Lympocyte **Fig 6. Monocyte**

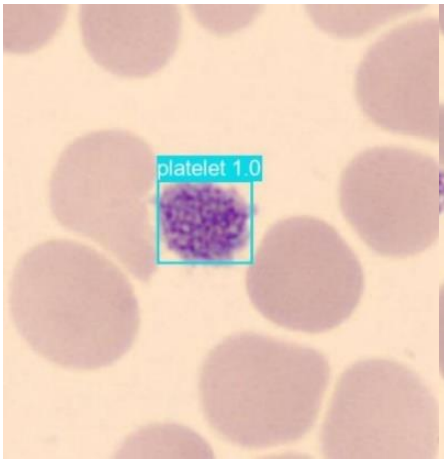


Fig 7. Platelet

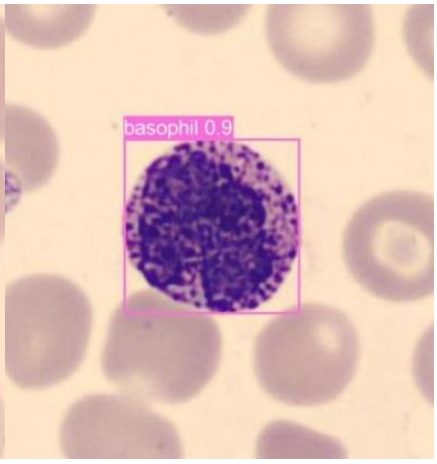


Fig 8. Basophil

6. Conclusion

This work demonstrated the use of a 3D depth camera combined with the YOLO detection framework for identifying blood cells and extending the approach towards cancer cell recognition. The system was able to detect and classify different blood cell types with promising accuracy, supported by strong training performance and evaluation metrics. By integrating depth information with real-time object detection, the method reduces dependency on manual microscopic analysis and opens the door for faster, more consistent, and automated hematological diagnostics. While further refinement and larger datasets are needed to achieve clinical deployment, the results highlight the potential of AI-driven imaging systems as supportive tools for early detection of blood cancers.

7. Acknowledgement

I would like to express my sincere gratitude to Vidyavardhaka College of Engineering, Mysuru, Department of Artificial Intelligence and Machine Learning (AIML) for their continuous support and encouragement throughout this project. I also extend my heartfelt thanks to STEM Avishkar Pvt. Ltd., Mysuru for providing valuable guidance and technical assistance, which played a crucial role in the successful completion of this work.

References

1. Indian Journal of Medical Research (IJMR), 2022. Cancer incidence estimates for 2022 & projection for 2025: Results from the National Cancer Registry Programme – India. Indian J Med Res. Available at: https://ijmr.org.in/cancer-incidence-estimates-for-2022-projection-for-2025-result-from-national-cancer-registry-programme-india/ [Accessed 22 August 2025].
2. BMC Cancer, 2022. Cancer burden and disability-adjusted life years (DALYs) projections in India. BMC Cancer, 22(10). Available at: https://bmccancer.biomedcentral.com/articles/10.1186/s12885-022-09578-1 [Accessed 22 August 2025].
3. GLOBOCAN/IARC, 2022. Global Cancer Observatory: Cancer Today – India fact sheet. Lyon, France: International Agency for Research on Cancer. Available at: https://gco.iarc.fr/today/ [Accessed 22 August 2025].
4. Times of India, 2025. Blood cancer spikes among young adults in Delhi. Times of India, 15 March. Available at:

https://timesofindia.indiatimes.com/city/delhi/blood-cancer-spikes-among-young-adults/articleshow/121464178.cms [Accessed 22 August 2025].

5. Times of India, 2025. Rise in blood cancer cases among children in central India a cause for concern: Doctors. Times of India, 18 March. Available at: https://timesofindia.indiatimes.com/city/nagpur/rise-in-blood-cancer-cases-among-children-in-central-india-a-cause-for-concern-docs/articleshow/121473879.cms [Accessed 22 August 2025].
6. Nature Reviews Clinical Oncology, 2023. Haematological malignancies in India: epidemiology and burden. Nature Reviews Clin Oncol. Available at: https://www.nature.com/articles/s41408-023-00853-3 [Accessed 22 August 2025].
7. N. A. Dalawai, P. B. Prathamraj, B. S. Dayananda, A. P. Jyothi, and R. Suresh, "Real-Time Skin Disease Detection and Classification Using YOLOv8 Object Detection for Healthcare Diagnosis," International Journal of Multidisciplinary Research and Analysis, vol. 7, no. 12, pp. 5429–5443, Dec. 2024, doi: 10.47191/ijmra/v7-i12-09. (Print ISSN: 2643-9840, Online ISSN: 2643-9875, Impact Factor: 8.22).
8. F. Sultana, A. Sufian, and P. Dutta, "Advancements in Image Classification using Convolutional Neural Network," presented at the Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kolkata, India, 2018, pp. 122–129.
9. Abas, S.M., Abdulazeez, A.M. and Zeebaree, D.Q., 2022. A YOLO and convolutional neural network for the detection and classification of leukocytes in leukemia. Indonesian Journal of Electrical Engineering and Computer Science, 25(1), pp.200-213.
10. Mehta, P., Vaghela, R., Pansuriya, N., Sarda, J., Bhatt, N., Bhoi, A.K. and Srinivasu, P.N., 2025. Benchmarking yolo variants for enhanced blood cell detection. International Journal of Imaging Systems and Technology, 35(1), p.e70037.
11. Naing, K.M., Kittichai, V., Tongloy, T., Chuwongin, S. and Boonsang, S., 2021. The evaluation of acute myeloid leukaemia (AML) blood cell detection models using different YOLO approaches. bioRxiv, pp.2021-08.
12. Naing, K.M., Kittichai, V., Tongloy, T., Chuwongin, S. and Boonsang, S., 2024. The detection and classification of acute myeloid leukaemia blood cell images based on different YOLO approaches. Bulletin of Electrical Engineering and Informatics, 13(2), pp.1147-1158.
13. Zhang, D., Bu, Y., Chen, Q., Cai, S. and Zhang, Y., 2024. TW-YOLO: An innovative blood cell detection model based on multi-scale feature fusion. Sensors, 24(19), p.6168.
14. Liu, C., Li, D. and Huang, P., 2021, December. ISE-YOLO: Improved squeeze-and-excitation attention module based YOLO for blood cells detection. In 2021 IEEE International Conference on Big Data (Big Data) (pp. 3911-3916). IEEE.