Erythroblast Cells: ML Models for Multiclass Classification in Single Image and Mixed Magnification.

Afnan Abdul Gafoor

Project Guide: Nirmal Punjabi, IIT Bombay

DH 307: R & D Project
Week 5 & 6

March 11, 2025

Overview

1. Task Overview
2. Findings and Results
3. Conclusion
4. Task Overview
5. Findings and Results
6. Conclusion

Task Overview

- **Individual Images:** Analyze each image separately and record performance metrics per class.
- **Combined Images:** Use stitched images to assess overall detection performance.

Findings and Results for Individual Images

Class	Images	Precision (P)	Recall (R)
All	3535	0.954	0.955
Basophil	259	0.981	0.988
Eosinophil	630	0.973	0.987
Erythroblast	320	0.972	0.968
Immunoglobin	574	0.949	0.960
Lymphocyte	234	0.996	0.990
Monocyte	281	0.961	0.989
Neutrophil	664	0.959	0.978
Platelets	397	0.803	0.738
Lymphoblast	176	0.996	1.000

Table: Performance metrics for individual blood cell classes

Findings and Results for Combined Images

Metric	Value
True Positives	2750
False Positives	559
False Negatives	778
Precision	0.831
Recall	0.779
F1 Score	0.804
Total Predictions	3309
Total Ground Truth	3528
Avg. Predictions per Image	0.936
Accuracy	0.673

Table: Overall Performance Metrics for Combined Images (3x3)

Conclusion

- The model performs well on individual images with high precision and recall.
- Performance on combined images is lower due to false positives and false negatives.
- **Augmentation and label correction** are required to improve predictions on combined images.

Task Overview

- Use **MakeSense.ai** to accurately label the **Platelet** class.
- Train the model on the newly labeled dataset.
- Experiment with augmentation techniques such as mixup and mosaic.

Findings and Results for Individual Images

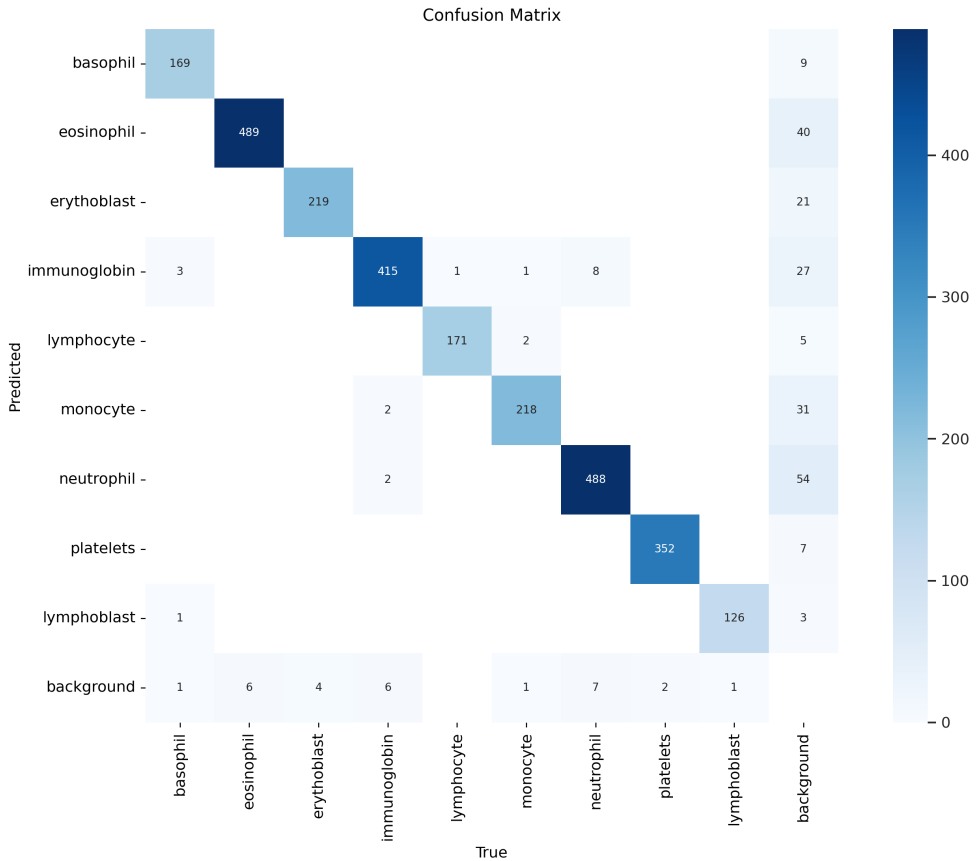

Figure: Confusion Matrix for Individual Images

The Platelet class is fully relabeled with no incorrect detections.

Findings and Results for Individual Images

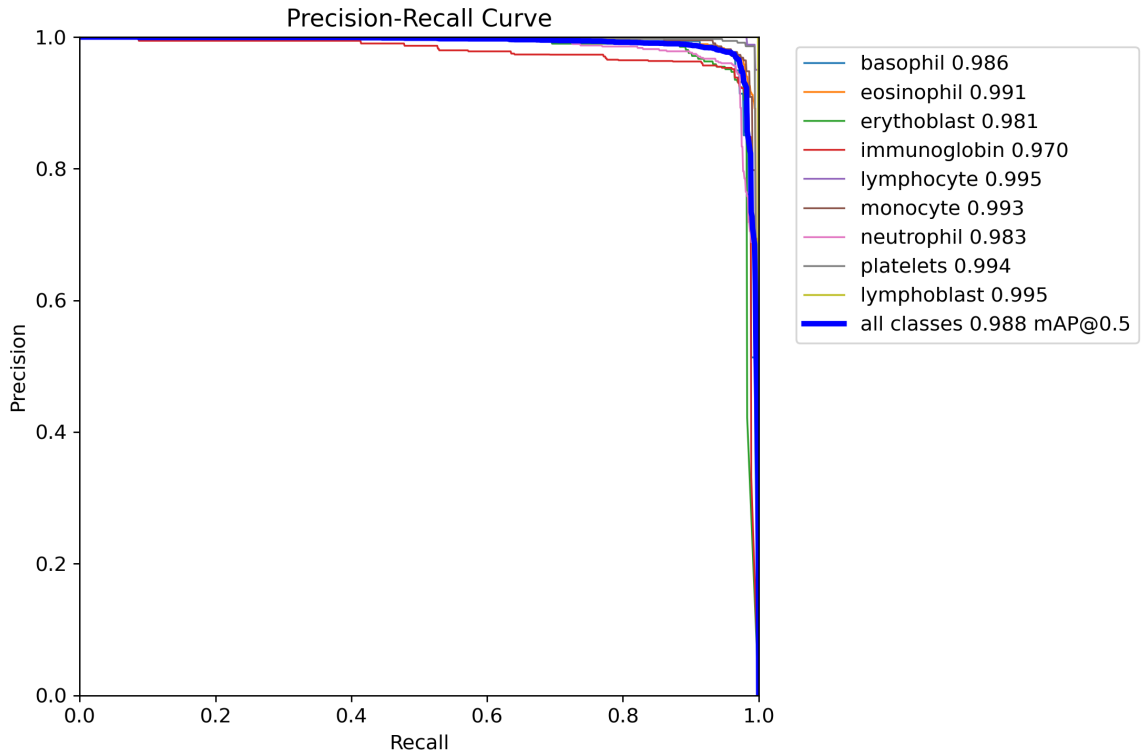

Figure: Precision-Recall Curve for Individual Images

Findings and Results for Individual Images

Class	Images	Precision (P)	Recall (R)
All	2694	0.978	0.983
Basophil	174	0.988	0.989
Eosinophil	495	0.972	0.977
Erythroblast	223	0.955	0.969
Immunoglobin	425	0.953	0.981
Lymphocyte	172	0.987	1.000
Monocyte	222	0.995	0.962
Neutrophil	503	0.978	0.980
Platelets	354	0.990	0.992
Lymphoblast	127	0.983	1.000

Table: Performance metrics for individual blood cell classes

Conclusion

- The Platelet class now performs very well, leading to a significant improvement in overall precision and recall.
- The model is trained with the following parameters:
- Mixup: Blends two images and their labels to create a composite image.
- Mosaic: Combines four training images into one.

Findings and Results for 2x2 Combined Images

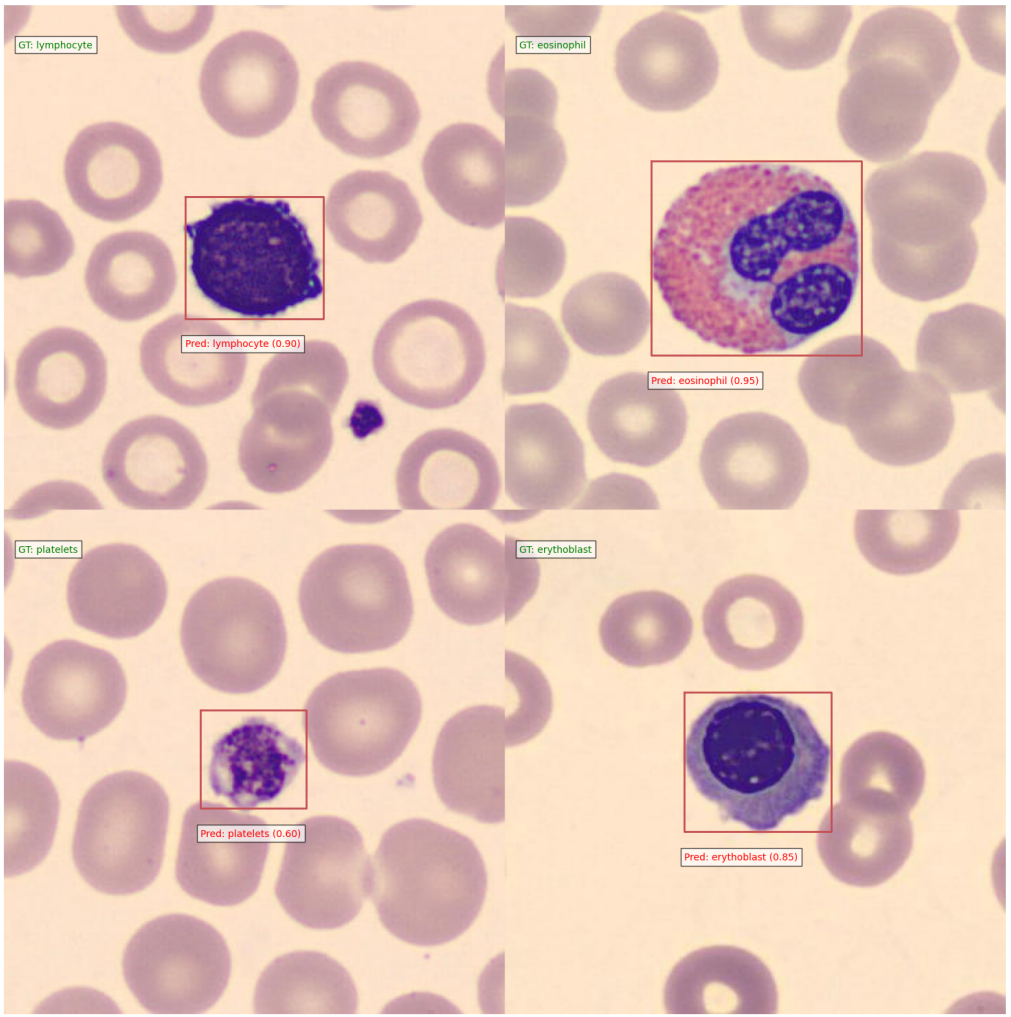

Figure: Detection on 2x2 image (Mosaic: 1.0, Mixup: 0.5)

Findings and Results for 2x2 Combined Images

Metric	Value
True Positives	2634
False Positives	189
False Negatives	59
Precision	0.933
Recall	0.978
F1 Score	0.955
Total Predictions	2823
Total Ground Truth	2693
Avg. Predictions per Image	1.048
Accuracy	0.914

Table: Overall Performance Metrics for Combined Images (2x2 Images with Mosaic: 1.0, Mixup:0.5)

Findings and Results for 3x3 Combined Images

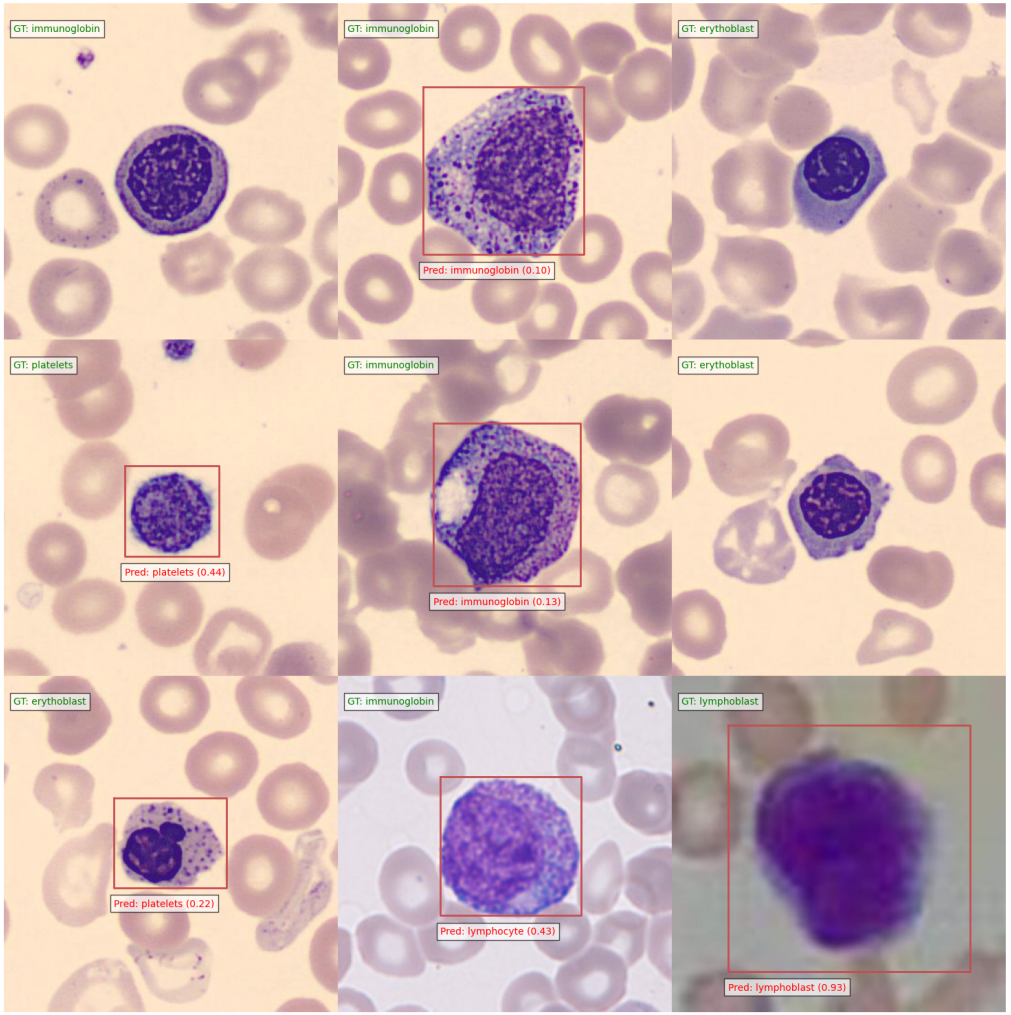

Figure: Detection on 3x3 image (Mosaic: 1.0, Mixup: 0.5)

Findings and Results for 3x3 Combined Images

Metric	Value
True Positives	270
False Positives	17
False Negatives	2422
Precision	0.941
Recall	0.100
F1 Score	0.181
Total Predictions	287
Total Ground Truth	2692
Avg. Predictions per Image	0.106
Accuracy	0.100

Table: Overall Performance Metrics for Combined Images (3x3 Images with Mosaic: 1.0, Mixup:0.5)

Findings and Results for 3x3 Combined Images

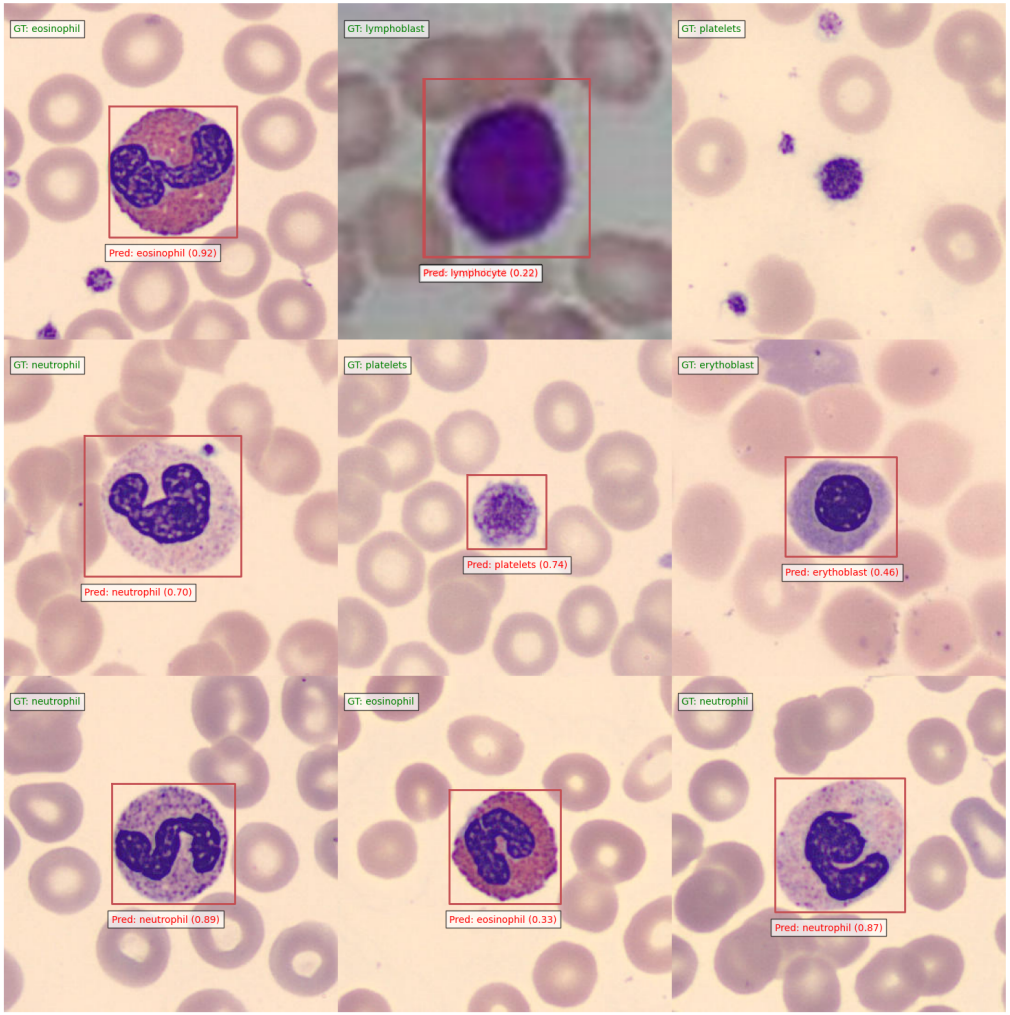

Figure: Detection on 3x3 image (Mosaic: 0.5, Mixup: 0.5)

Findings and Results for 3x3 Combined Images

Metric	Value
True Positives	1748
False Positives	451
False Negatives	944
Precision	0.795
Recall	0.649
F1 Score	0.715
Total Predictions	2199
Total Ground Truth	2692
Avg. Predictions per Image	0.816
Accuracy	0.556

Table: Overall Performance Metrics for Combined Images (3x3 Images with Mosaic: 0.5, Mixup:0.5)

Findings and Results for 3x3 Combined Images

Metric	Value
True Positives	1882
False Positives	714
False Negatives	810
Precision	0.725
Recall	0.699
F1 Score	0.712
Total Predictions	2596
Total Ground Truth	2692
Avg. Predictions per Image	0.964
Accuracy	0.553

Table: Overall Performance Metrics for Combined Images (3x3 Images with Mosaic: 0.1, Mixup: 1.0)

Conclusion

- The models performed well on 1×1 and 2×2 images, likely due to the mosaic augmentation, which trains on 2×2 images.
- Performance on 3×3 images was not satisfactory. However, reducing the mosaic value led to improvements.
- To enhance generalization, the training set should include a balanced mix of 1×1, 2×2, and 3×3 images.

Plan for this Week

- Train with a mix of 1×1, 2×2, and 3×3 images in the dataset.
- Experiment with different augmentation techniques to improve generalization.