Blood Cell Classification Using Deep Learning

Abstract:

This project aims to classify blood cells using the YOLO (You Only Look Once) object detection framework, an advanced and efficient method for object detection in real-time scenarios. Blood cell classification is critical in medical diagnostics, enabling the detection of conditions like anemia, infections, and leukemia by analyzing microscopic images.

The project focuses on the classification of 8 specific blood cell types: Basophil, Erythroblast, Immature Granulocyte (IG), Lymphocyte, Monocyte, Neutrophil, Platelet, and Erythrocyte. The dataset, sourced from Kaggle, is preprocessed by splitting it into training, validation, and test sets, ensuring proper distribution for accurate model evaluation.

The YOLO model is trained using PyTorch and the ultralytics library, employing customized parameters to optimize performance. Key techniques, such as data augmentation and hyperparameter tuning, enhance the model's accuracy. After training, the model is tested on unseen images to evaluate its prediction accuracy and class visualization capabilities.

The results demonstrate the potential of YOLO-based systems for accurate and efficient blood cell classification, which could significantly benefit medical practitioners by providing automated insights from microscopic images.

Dataset:

Download Link for the Data:
https://www.kaggle.com/datasets/unclesamulus/blood-cells-image-dataset?resource=download

Code and Functions:

- Test Folder Creation(Line7): Randomly moves files from a source to a target. This is useful for organizing the dataset.
- 10% File Selection for Validation(Line63): Moves 10% of the files from the training set to the validation set.
- Library Installation(Line101): Installation code for YOLO and PyTorch.
- Model Training(Line107): Code for training the YOLO model with the specified data.
- Test and Prediction(Line136): Runs the trained model on a test image.
- Class Prediction Visualization(Line145): Visualizes the model's predictions and adds prediction results to the image.
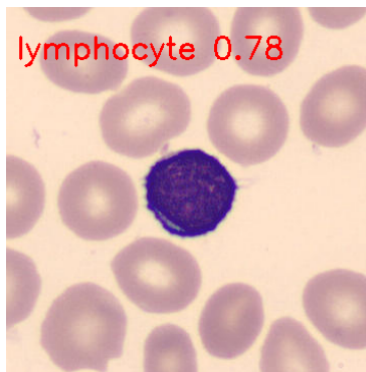
Results:

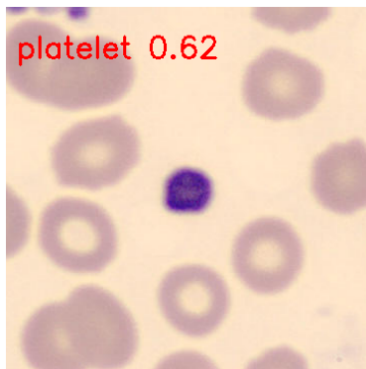## Random Test Results for Blood Cell Classification:

Below are randomly selected test results from the model's predictions. Each example shows the predicted blood cell type alongside the actual type for comparison purposes. This section is intended to demonstrate the model's performance visually.
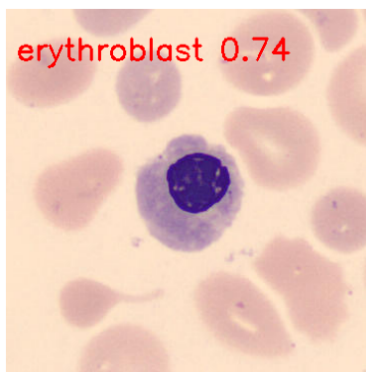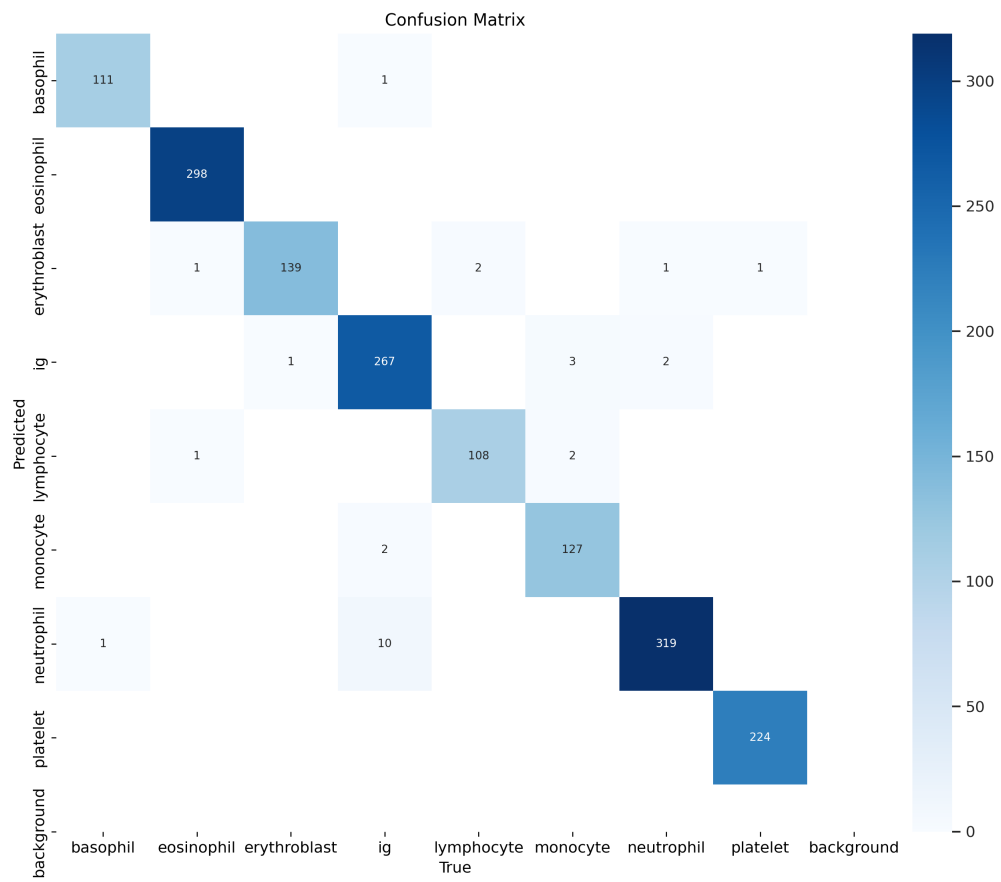

Actual Type: Erythroblast

Actual Type: Lymphocyte

Actual Type: Platelet

Confusion Matrix

Confusion Matrix

Training Results Table

Epoch	Time (s)	Train/Loss	Metric/Accuracy Train	Metric/Accuracy Val	Val/Loss	Lr/p0	Lr/p1	Lr/p2
1	685.238	0.37652	0.86737	0.99815	0.41188	0.0670414	0.000332915	0.000332915
2	1370.25	0.16494	0.9636	1	0.11684	0.0339095	0.000534331	0.000534331
3	2046.21	0.1285	0.96484	0.99938	0.11584	0.000645571	0.000603747	0.000603747
4	2724.33	0.0985	0.97717	1	0.07199	0.000406	0.000406	0.000406
5	3398.7	0.05733	0.98273	1	0.06069	0.000208	0.000208	0.000208

References and Acknowledgments:

- Kaggle Blood Cell Dataset: https://www.kaggle.com/datasets/unclesamulus/blood-cells-image-dataset?resource=download
- Thanks to the Ultralytics team for developing the YOLO framework.

Author:

Batuhan Kılıçarslan