Artificial Intelligence and Machine Learning-STB3

HematoVision: Advanced Blood Cell Classification Using Transfer Learning

Team ID: LTVIP2025TMID46585

Team Members:

Poralla Venkata Phani Kumar

M Yagna Sree

Kolagana Saikumar

Meka Keerthi

S Bhargav

Introduction

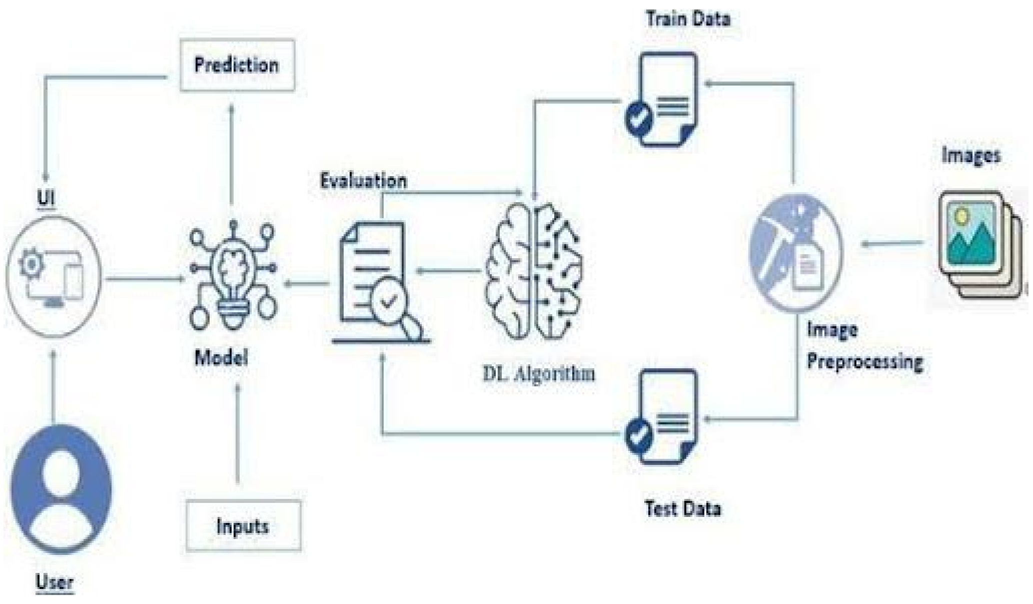
HematoVision aims to develop an accurate and efficient model for classifying blood cells by employing transfer learning techniques. Utilizing a dataset of 12,000 annotated blood cell images, categorized into distinct classes such as eosinophils, lymphocytes, monocytes, and neutrophils, the project leverages pre-trained convolutional neural networks (CNNs) to expedite training and improve classification accuracy. Transfer learning allows the model to benefit from pre-existing knowledge of image features, significantly enhancing its performance and reducing computational costs. This approach provides a reliable and scalable tool for pathologists and healthcare professionals, ensuring precise and efficient blood cell classification

Objectives

By the end of this project, you will:

- Know fundamental concepts and techniques used for Deep Learning.
- Gain a broad understanding of data.
- Have knowledge of pre-processing the data/transformation techniques on outliers and some visualization concepts.

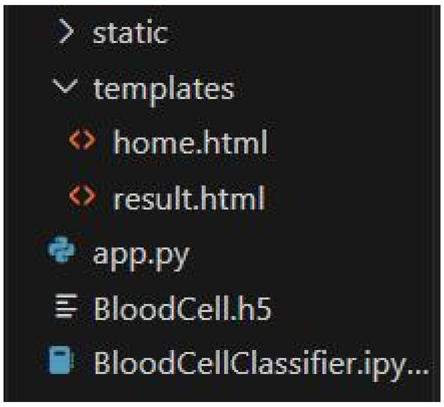
Project Architecture



Project Flow

1. Data Collection: Collect or download the dataset that you want to train.
2. Data pre-processing: Data Visualization, Data Augmentation
3. Model building
4. Application Building

Project Structure



```
> static
v templates
    home.html
    result.html
  app.py
  BloodCell.h5
  BloodCellClassifier.ipy...
```

1. **Data Collection:**

1.1. Collect the data

There are many popular open sources for collecting the data. Eg: kaggle.com, UCI repository, etc. This dataset contains 12,500 augmented images of blood cells (JPEG) with accompanying cell type labels (CSV). There are approximately 3,000 images for each of 4 different cell types grouped into 4 different folders (according to cell type). The cell types are Eosinophil, Lymphocyte, Monocyte, and Neutrophil.

Link: https://www.kaggle.com/datasets/paultimothymooney/blood-cells/data

2. Pre-processing:

2.1. Data Visualization

The provided Python code imports necessary libraries and modules for image manipulation. It selects a random image file from a specified folder path. Then, it displays the randomly selected image using IPython's Image module. This code is useful for showcasing random images from a directory for various purposes like data exploration or testing image processing algorithms.

2.2. Data Augmentation

Data augmentation is a technique commonly employed in machine learning, particularly in computer vision tasks such as image classification, including projects like the BloodCells Classification . The primary objective of data augmentation is to artificially expand the size of the training dataset by applying various transformations to the existing images, thereby increasing the diversity and robustness of the data available for model training. This approach is particularly beneficial when working with limited labeled data.

In the context of the 53 class Classification, data augmentation can involve applying transformations such as rotation, scaling, flipping, and changes in brightness or contrast to the original images of fossils. These transformations help the model generalize better to variations and potential distortions present in real-world images, enhancing its ability to accurately classify unseen data.

This is a crucial step but this data is already cropped from the augmented data so. this time it is skipped accuracy is not much affected but the training time increased.

3. Model Building:

Mobilenet V2 Transfer-Learning Model:

The MobileNetV2-based neural network is created using a pre-trained MobileNetV2 architecture with frozen weights. The model is built sequentially, incorporating the MobileNetV2 base, a flattening layer, dropout for regularization, and a dense layer with SoftMax activation for classification into four categories of blood cells. The model is compiled using the Adam optimizer and categorical cross-entropy loss. During training, which spans 5 epochs, a generator is employed for the training data, and validation is conducted with callbacks such as Model Checkpoint and Early Stopping. The best-performing model is saved as "blood_cell.h5" for future use. The model summary provides an overview of the architecture, showcasing the layers and parameters involved.

4. Application Building:

In this section, we will be building a web application that is integrated into the model we built. A UI is provided for the uses where he has to enter the values for predictions. The enter values are given to the saved model and prediction is showcased on the UI.

This section has the following tasks

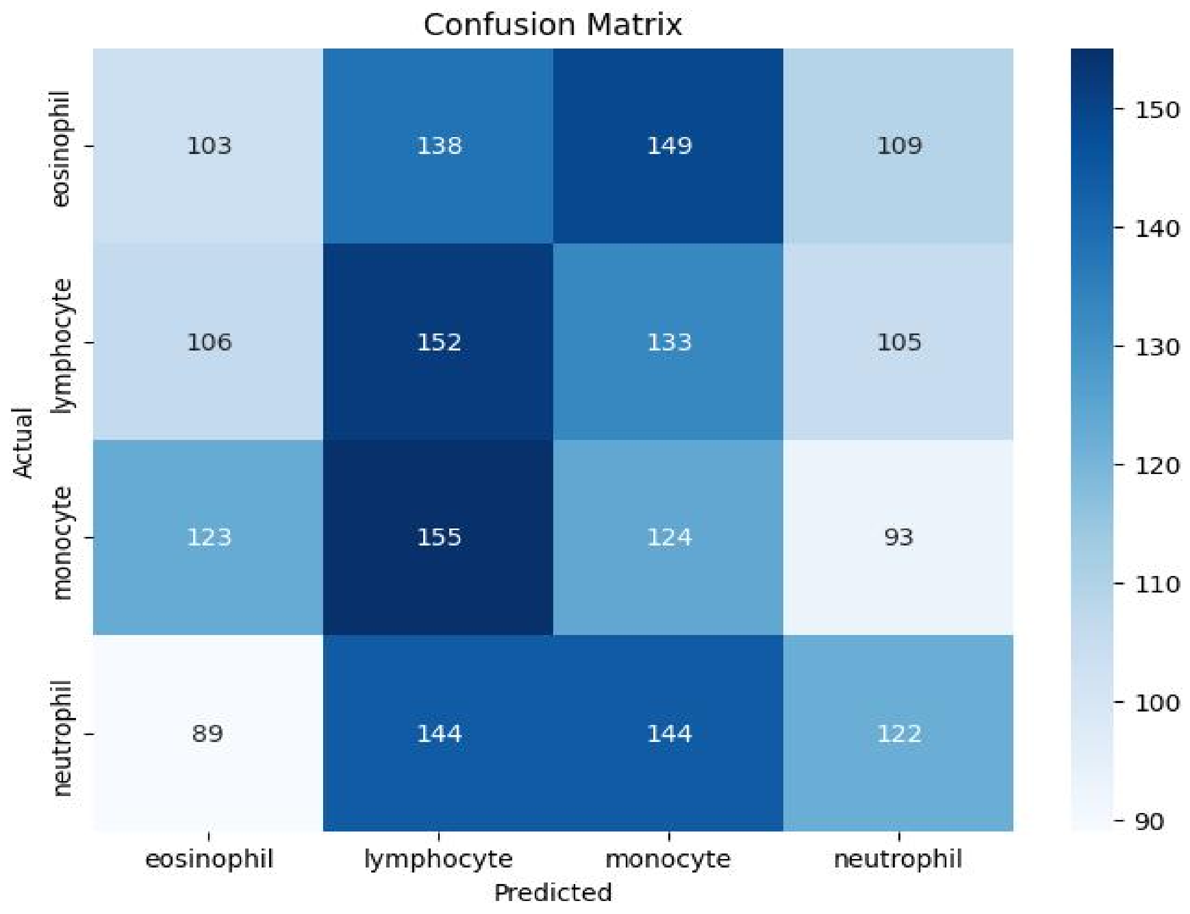
- Building HTML Pages
- Building server-side script

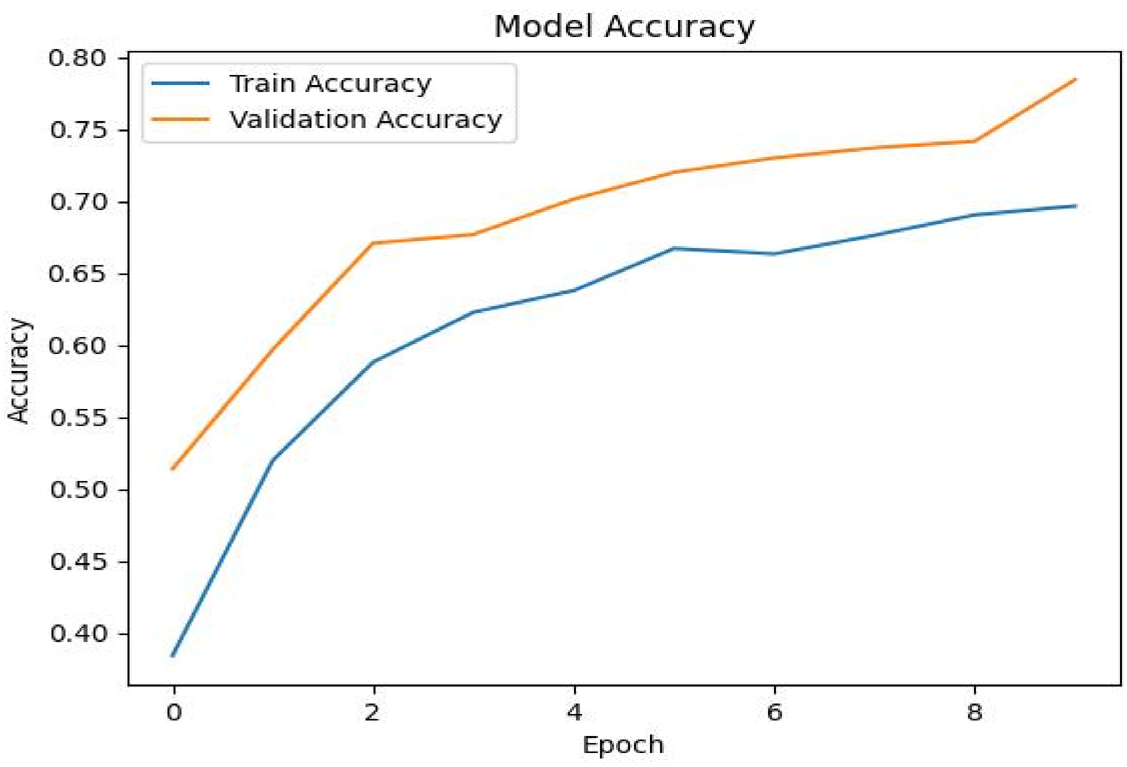
For this project create two HTML files namely

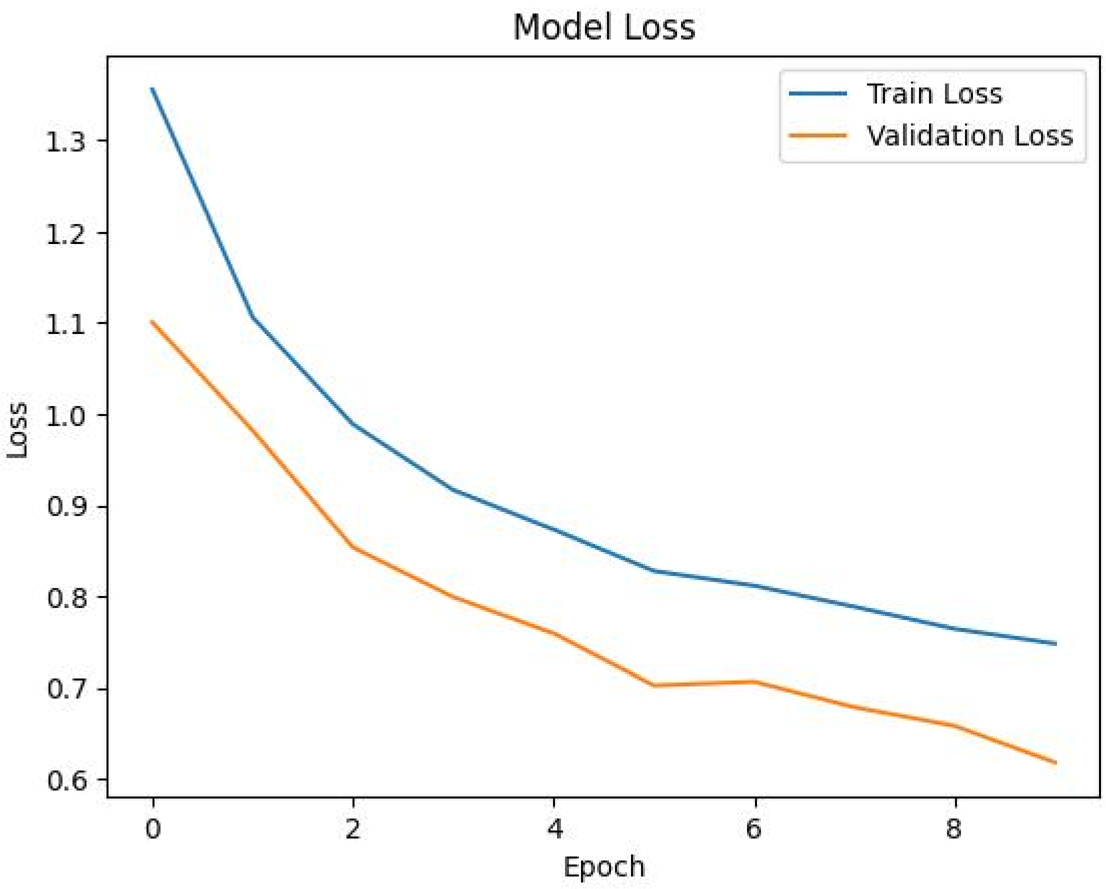
- home.html

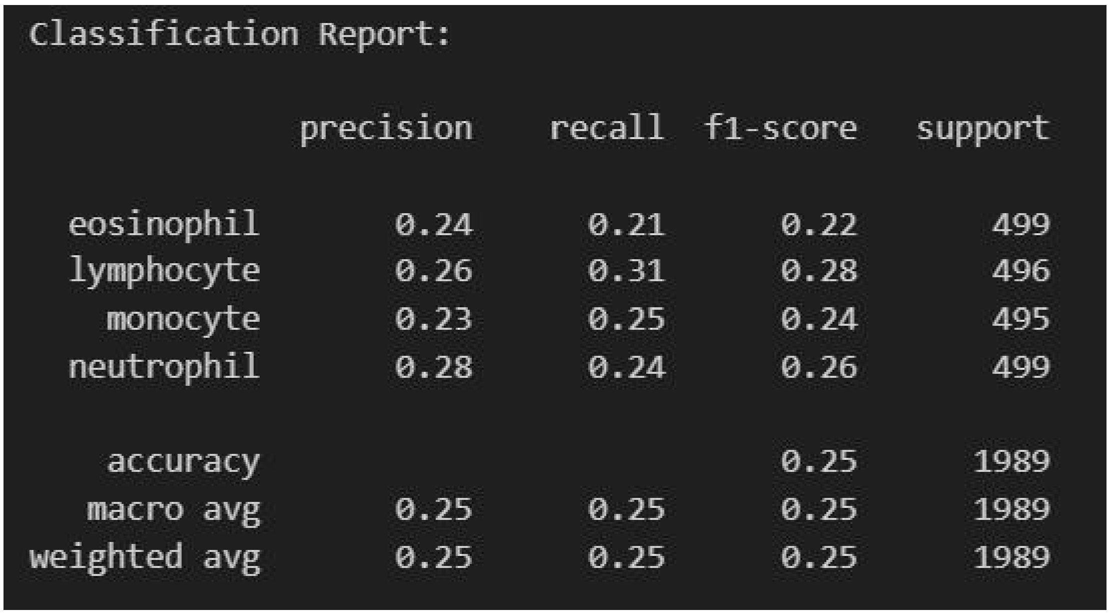
- result.html

Output:




Model Accuracy
Train Accuracy
Validation Accuracy
Accuracy
0.80
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0.40
0
2
4
6
8
Epoch


Model Loss
Train Loss
Validation Loss
Loss
1.3
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0
2
4
6
8
Epoch

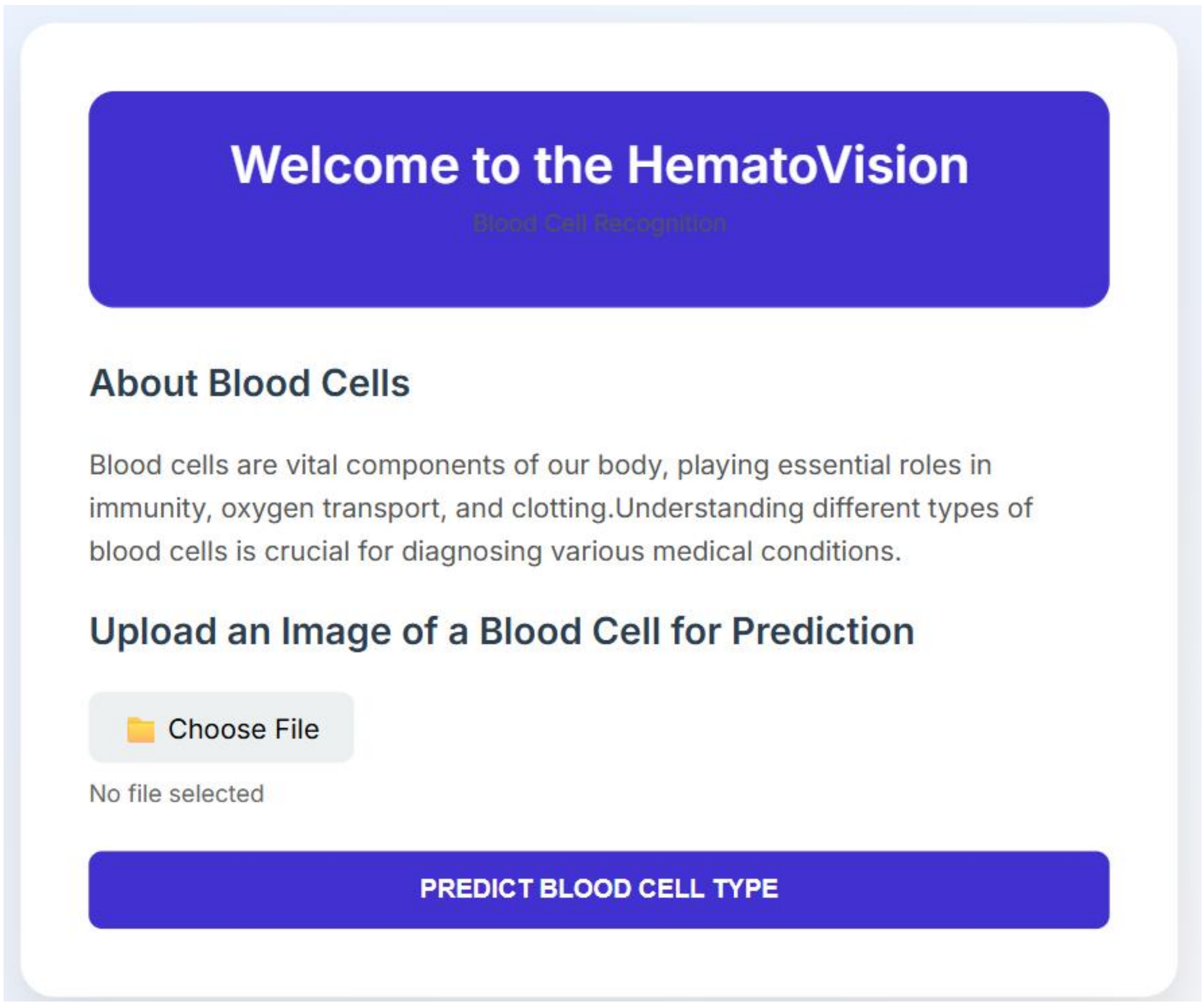


```
Classification Report:

              precision    recall  f1-score   support

  eosinophil       0.24      0.21      0.22       499
  lymphocyte       0.26      0.31      0.28       496
    monocyte       0.23      0.25      0.24       495
  neutrophil       0.28      0.24      0.26       499

    accuracy                           0.25      1989
   macro avg       0.25      0.25      0.25      1989
weighted avg       0.25      0.25      0.25      1989
```



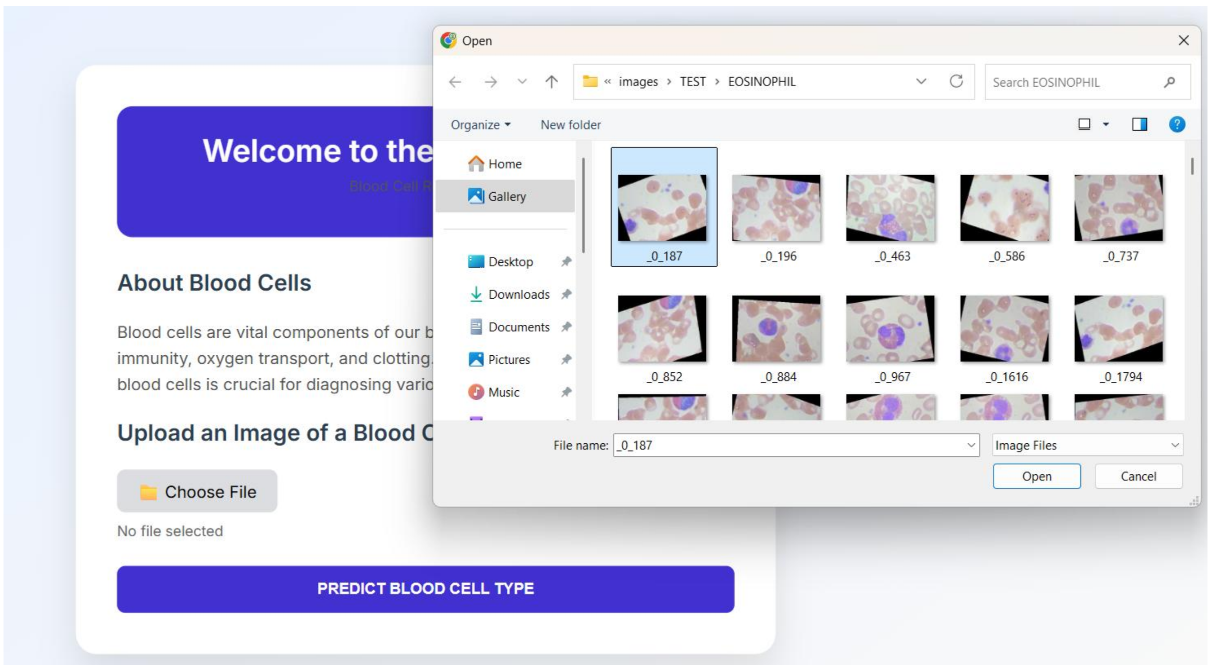
Welcome to the HematoVision

Blood Cell Recognition

About Blood Cells

Blood cells are vital components of our body, playing essential roles in immunity, oxygen transport, and clotting.Understanding different types of blood cells is crucial for diagnosing various medical conditions.

Upload an Image of a Blood Cell for Prediction


Welcome to the

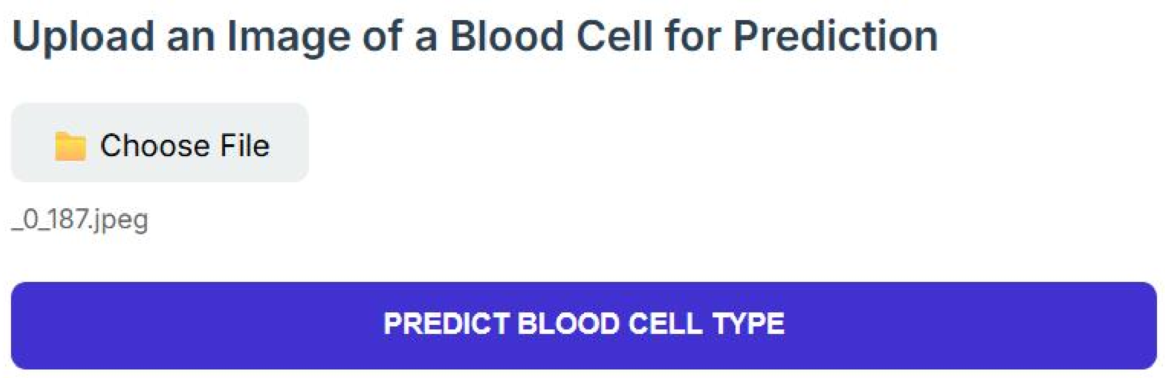
About Blood Cells

Blood cells are vital components of our b
immunity, oxygen transport, and clotting.
blood cells is crucial for diagnosing vario

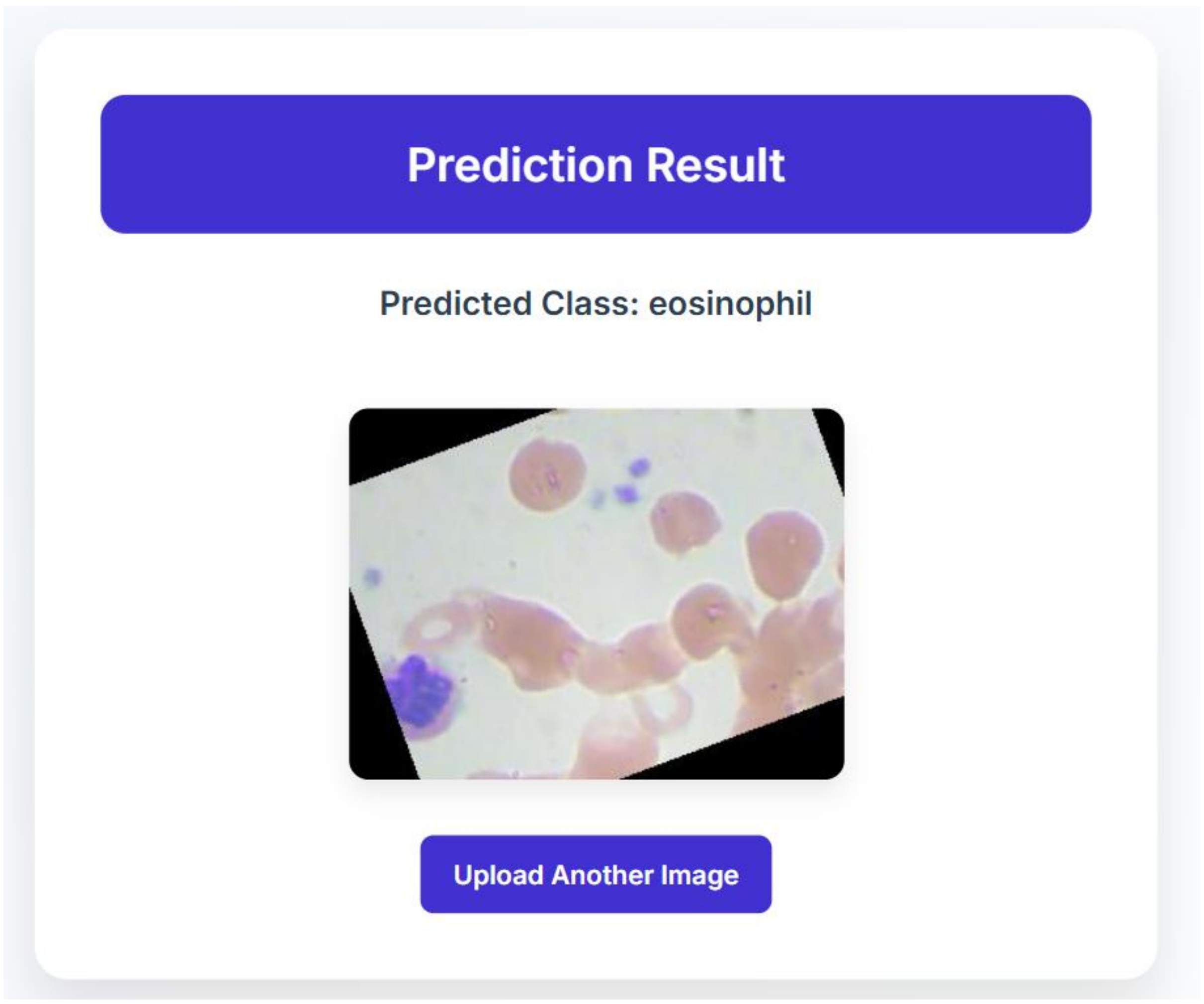
Upload an Image of a Blood C

No file selected

File name: _0_187


Upload an Image of a Blood Cell for Prediction

_0_187.jpeg



Prediction Result

Predicted Class: eosinophil