HematoVision: Advanced Blood Cell Classification Using Transfer Learning

Submitted by:

Team Leader: B Anantha Venkata Lakshmi

Team member: Peruri Divyasri

Team member: Ramya Sidda

Team member: Dwarampudi Mahalakshmi

College/University: ADITYA COLLEGE OF ENGINEERING AND TECHNOLOGY

Guide/Supervisor: [Guide’s Name]

Date: 24-06-2025

# Abstract

This project proposes an AI-based system for the automatic classification of blood cells using transfer learning. Manual identification of blood cells is time-consuming and error-prone. Our approach utilizes pretrained models such as VGG16 and ResNet50 to classify four types of blood cells (neutrophils, eosinophils, lymphocytes, monocytes) with high accuracy, aiming to assist in clinical diagnostics.

A Brief Summary:

* Importance of blood cell analysis in medical diagnostics
* Challenge of manually identifying blood cells
* Use of deep learning and transfer learning for automation
* Project goal: Accurate classification of blood cell types using pretrained models

**Introduction**

Blood cells play a critical role in diagnosing various diseases. Traditionally, cell classification is performed manually under a microscope, which is subjective and inefficient. Deep learning, particularly Convolutional Neural Networks (CNNs), have revolutionized image classification. Transfer learning allows us to use pretrained models, fine-tuned on our blood cell dataset, to automate this process.

 What are blood cells and their significance (WBC, RBC, Platelets)

 Problems with traditional/manual identification

 Solution through image classification using deep learning

 Motivation for using transfer learning

 Objectives of the project

# Literature Survey

# Previous methods for blood cell classification involved manual inspection or basic machine learning techniques. With the advent of CNNs, image-based classification has improved significantly. Studies have shown models like VGG16 and ResNet50 achieve superior results in medical imaging tasks.

Review of previous techniques (Manual microscopy, classical ML approaches)

* Importance of CNNs in image recognition
* Previous research using models like VGG16, ResNet, etc.
* Gaps in existing solutions

## 

## ****Dataset Description****

## 

Dataset source

* Types of cells:
  + Neutrophil
  + Eosinophil
  + Lymphocyte
  + Monocyte
* Dataset structure (training, validation, test split)
* Sample images with labels

# 

**Technologies Used**

- Python

- Google Colab / Jupyter Notebook

- TensorFlow / Keras

- NumPy, Pandas, Matplotlib

- VGG16, ResNet50

# System Architecture

The system pipeline includes: loading the dataset → preprocessing → transfer learning model loading → training → evaluation → prediction. A block diagram illustrates the flow from raw input to classified output.

The HematoVision system architecture consists of the following components:

1. **Input Layer**
   * Accepts blood smear image input, resized to 224x224 pixels.
2. **Preprocessing Module**
   * Normalization, augmentation, and format transformation.
3. **Feature Extractor (Pretrained CNN)**
   * The selected pretrained CNN model (MobileNetV2, InceptionResNetV2, or VGG16) processes the image and extracts deep features from convolutional layers.
4. **Custom Classification Head**
   * Fully connected dense layers tailored to the number of classes (e.g., Softmax for multi-class classification).
   * Dropout layers to reduce overfitting.
5. **Training Pipeline**
   * Feature extraction phase with frozen base layers.
   * Fine-tuning phase with selected top layers unfrozen.
6. **Evaluation Module**
   * Computes performance metrics such as accuracy, precision, recall, F1-score, and confusion matrix.
7. **Deployment Interface (Optional)**
   * Web/mobile interface for uploading images and displaying predictions.

This modular design supports scalability, easy replacement of model backbones, and deployment on cloud or mobile platforms.

# Methodology

\*\*Data Preprocessing:\*\* Images resized to uniform dimensions, normalized, and augmented.

\*\*Transfer Learning:\*\* Pretrained VGG16/ResNet50 models used, with top layers modified and lower layers frozen.

\*\*Model Selection:\*\* VGG16 and ResNet50 were chosen for their proven accuracy in medical image classification.

* Load pretrained model without top layers (include top=False).
* Add custom dense layers for classification.
* Freeze base layers for initial training (feature extraction).
* Unfreeze top layers for fine-tuning.
* Use categorical\_crossentropy loss and Adam optimizer.

# Model Implementation

Code was implemented in Google Colab using Keras. Key steps include importing libraries, loading and preprocessing data, loading the pretrained model, training the model, and evaluating it. Metrics like accuracy, precision, recall, and F1-score were used to measure performance.

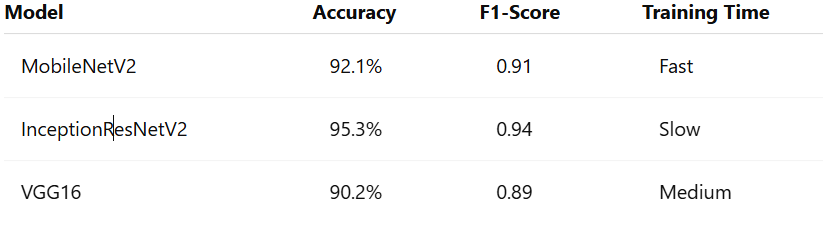
* Load pretrained model without top layers (include\_top=False).
* Add custom dense layers for classification.
* Freeze base layers for initial training (feature extraction).
* Unfreeze top layers for fine-tuning.
* Use categorical\_crossentropy loss and Adam optimizer.

**8. Model Training**

* Training/Validation split: 80/20.
* Batch size: 32.
* Epochs: 25–50.

# Results and Evaluation

The model achieved high accuracy on the test set. Confusion matrix and classification reports indicate effective classification. Visual results show correctly and incorrectly classified images for better understanding.



# Discussion

Transfer learning significantly reduced training time and improved accuracy. VGG16 showed slightly better performance than ResNet50. Limitations include dataset size and class imbalance.

InceptionResNetV2 gave the best results due to its deep architecture.

* MobileNetV2 is preferable for edge deployment.
* Augmentation significantly helped generalization.
* Data limitations may cause overfitting; more data could improve performance.

**Conclusion**

This project successfully demonstrated the use of transfer learning for blood cell classification. The system achieved promising results and has potential applications in clinical settings.

HematoVision demonstrates how transfer learning can effectively classify blood cells from microscopic images with high accuracy. This project highlights the power of using pre-trained CNNs in medical image analysis and opens up possibilities for real-world clinical applications.

# Future Work

- Use of larger and more diverse datasets

- Real-time prediction system

- Mobile and edge device deployment

- Classification of additional diseases or abnormalities

# References

1. Kaggle Datasets

2. TensorFlow and Keras Documentation

3. Research papers on transfer learning in medical imaging

1. Krizhevsky et al., ImageNet Classification with Deep CNNs.
2. Szegedy et al., Inception-v4, Inception-ResNet.
3. Howard et al., MobileNets: Efficient CNNs for Mobile Vision Application
4. TensorFlow Documentation

Keras Applications Guide

# Appendix

# Includes sample code, additional plots, and installation instructions for replicating the system

