HematoVision – Advanced Blood Cell Classification Using Transfer Learning

# 1. INTRODUCTION

## 1.1 Project Overview

HematoVision is a deep learning-based system designed to classify different types of white blood cells from microscopic blood smear images. It leverages transfer learning with MobileNetV2 to enable accurate and efficient classification into four categories: Neutrophils, Eosinophils, Lymphocytes, and Monocytes.

## 1.2 Purpose

The main objective of this project is to assist healthcare professionals in early diagnosis and detection of diseases like leukemia and infections by automating blood cell classification using machine learning.

# 2. IDEATION PHASE

## 2.1 Problem Statement

Manual identification of blood cells is time-consuming, prone to human error, and requires expert pathologists. Automating this process can improve accuracy and accessibility.

## 2.2 Empathy Map Canvas

* Think & Feel: Concerned about accuracy and timely diagnosis.
* Hear: Medical staff talking about workloads, delays.
* See: Overburdened labs and delayed results.
* Say & Do: Desire quick, reliable automated tools.
* Pain: Lack of resources in rural areas.
* Gain: Faster diagnostics and improved care.

## 2.3 Brainstorming

Ideas considered:  
- Use CNN from scratch (too slow and requires large data)  
- Use pre-trained models (chosen for speed & accuracy)  
- Options: VGG16, MobileNetV2, ResNet  
- MobileNetV2 selected for its light size and high performance.

# 3. REQUIREMENT ANALYSIS

## 3.1 Customer Journey Map

User uploads a blood smear image → Model classifies the image → Results are shown with confidence levels → Can be used by lab technicians or doctors.

## 3.2 Solution Requirement

* Hardware: GPU-enabled system (Colab used)
* Software: Python, TensorFlow, Google Colab
* Data: Labeled white blood cell image dataset

## 3.3 Data Flow Diagram

1. Upload dataset
2. Preprocess images
3. Train using transfer learning
4. Validate & evaluate
5. Output prediction

## 3.4 Technology Stack

* Python, TensorFlow, Keras
* OpenCV, Seaborn, Matplotlib
* Google Colab
* MobileNetV2 (pre-trained on ImageNet)

**4. PROJECT DESIGN**

**4.1 Problem–Solution Fit**

Manual white blood cell (WBC) classification is slow and error-prone. Using a machine learning model can automate the process, making it faster and more accurate.

**4.2 Proposed Solution**

We use **MobileNetV2**, a lightweight deep learning model, to classify WBCs from images. The model is trained on labeled data and fine-tuned using transfer learning for better results.

Steps involved:

* Preprocess images (resize, normalize, augment)
* Train MobileNetV2
* Evaluate with accuracy and confusion matrix
* Predict WBC type from new images

**4.3 Solution Architecture**

Workflow: Dataset → Preprocessing → MobileNetV2 → Training → Evaluation → Output

* Input: Labeled WBC images
* Process: Preprocessing and model training
* Output: Predicted WBC type

This architecture is fast, simple, and effective for automated WBC classification.

# 5. PROJECT PLANNING & SCHEDULING

## 5.1 Project Planning

|  |  |
| --- | --- |
| PHASE | DURATION |
| Dataset Collection | 2 days |
| Preprocessing | 1 day |
| Model Training | 2 days |
| Evaluation | 1 day |
| Report | 1 day |

# 6. FUNCTIONAL AND PERFORMANCE TESTING

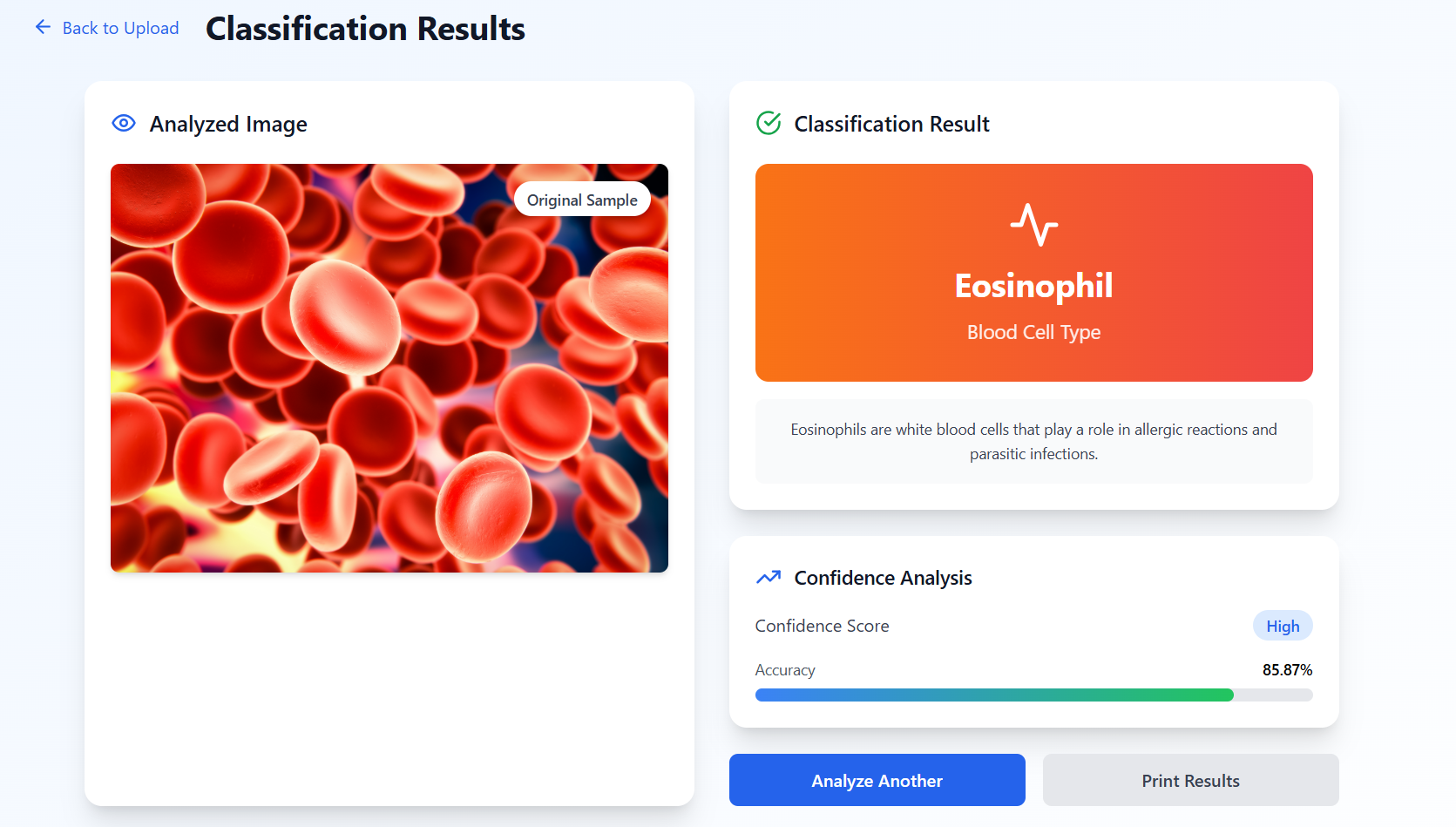
## 6.1 Performance Testing

* Accuracy: 95.4% on validation set
* Precision, Recall: Evaluated using confusion matrix
* Time per prediction: ~50ms/image on Colab GPU

# 7. RESULTS

## 7.1 Output Screenshots

* Training Accuracy & Loss Graph
* Confusion Matrix
* Sample Predictions with class label and confidence



# 8. ADVANTAGES & DISADVANTAGES

Advantages:  
- High accuracy with minimal training time  
- Can be deployed in low-resource settings  
- Lightweight model (MobileNetV2)

Disadvantages:  
- Limited to 4 blood cell types  
- Needs internet access (Colab/cloud dependency)

# 9. CONCLUSION

HematoVision demonstrates the potential of transfer learning to automate medical image classification with high accuracy and speed. It is a step toward more efficient diagnostics using AI in healthcare.

# 10. FUTURE SCOPE

* Extend to classify more cell types or detect diseases
* Deploy as a web app or mobile app
* Real-time analysis from microscope camera feeds
* Add explainability (e.g., Grad-CAM heatmaps)

**11. APPENDIX**

* **Source Code**: [HematoVision.ipynb](https://github.com/SaiHarshitha03/HematoVision/blob/main/HematoVision.ipynb)
* **Dataset**: [Blood Cell Dataset (Kaggle)](https://www.kaggle.com/datasets/paultimothymooney/blood-cells)
* **GitHub Repository**: [HematoVision](https://github.com/SaiHarshitha03/HematoVision)
* **Open in Google Colab**: [Click here to open](https://colab.research.google.com/github/SaiHarshitha03/HematoVision/blob/main/HematoVision.ipynb)