Project 6
Few-Shot White Blood Cell Image Classification

Project Overview

White blood cell (WBC) classification from microscopic blood smear images is a fundamental task in hematology, supporting disease screening, diagnosis of blood disorders, and clinical decision-making. Deep learning–based image classification models have demonstrated strong performance in automated WBC recognition; however, these models typically require large, well-annotated datasets curated by trained experts, which are costly and time-consuming to obtain.

This project investigates how white blood cell classification performance degrades under limited annotation availability, and how different convolutional neural network (CNN) and transformer-based architectures cope with varying proportions of labeled training data. By systematically reducing the percentage of labeled samples in a WBC classification dataset, this study aims to benchmark classical CNN models and modern Vision Transformer (ViT) models, establishing a clear performance–annotation trade-off. The project will first target an international conference and then be extended toward a Q2 journal publication, incorporating more advanced learning strategies.

Research Objectives

- Quantify WBC classification performance under different proportions of labeled data (e.g., 1%, 5%, 10%, 25%, 50%, 100%).
- Benchmark popular CNN-based and transformer-based classification models.
- Analyze robustness and generalization under annotation scarcity.
- Provide qualitative and quantitative insights into misclassification patterns across WBC subtypes.
- Produce a conference paper and an extended Q2 journal submission.

Methodology

1. Dataset Preparation:
- Use a white blood cell image dataset with multi-class labels (e.g., neutrophils, lymphocytes, monocytes, eosinophils, basophils).
- Apply standardized preprocessing (color normalization, resizing, data augmentation).
- Construct multiple training subsets by sampling different proportions of labeled images.
- Keep validation and test sets fixed to ensure fair comparison.

2. Limited Annotation Simulation:
- Systematically vary the proportion of labeled data used for training.
- Maintain consistent data splits and training protocols across experiments.

3. Model Benchmarking:
- Train and evaluate CNN-based models (e.g., ResNet, DenseNet, EfficientNet).
- Train and evaluate transformer-based models (e.g., ViT, hybrid CNN–ViT models).
- Use consistent optimization, augmentation, and evaluation settings for fair comparison.

4. Evaluation Metrics:
- Accuracy
- Precision, Recall, F1-score (macro and per-class)
- ROC-AUC (where applicable)
- Confusion matrices

5. Visualization and Analysis:
- Performance–annotation curves
- Per-class performance breakdown
- Failure case and misclassification analysis

Timeline (6 Months)

Phase 1 (Months 1–3): Conference Target
- 1 white blood cell classification dataset
- Multi-class classification under limited annotation
- Benchmark CNN and ViT-based models
- Conference submission (e.g., ICISN, CITA, MAPR)
- Compute: Google Colab, Kaggle

Phase 2 (Months 4–6): Journal Target
- Extend to additional datasets or dataset splits
- Explore advanced paradigms (semi-supervised, self-supervised, few-shot learning)
- Optional extension to joint detection or segmentation tasks
- Q2 journal submission
- Compute: AIVN GPUs, Private servers