Erythroblast Cells: ML Models for Multiclass Classification in Single Image and Mixed Magnification.

Afnan Abdul Gafoor

Project Guide: Nirmal Punjabi, IIT Bombay

DH 307: R & D Project

Week 1

January 20, 2025

Overview

Problem Statement

- Develop a Machine Learning model for classifying different types of erythroblast cells with mixed magnifications from a single image.

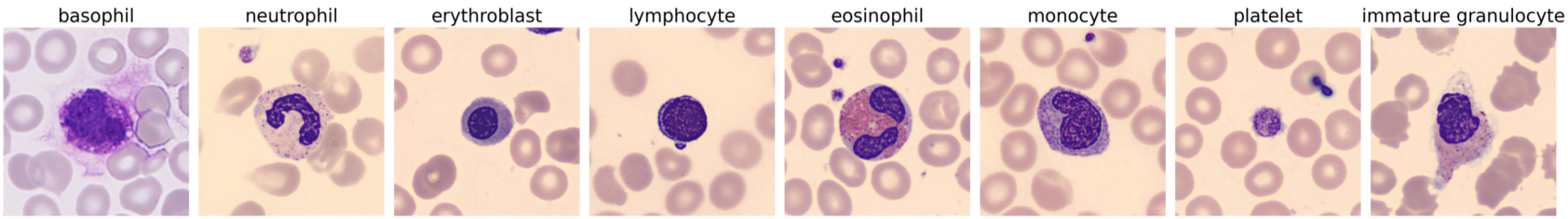

Figure: Sample images from each class of the dataset.

Proposed Approach

Dataset Creation:

- Use the base dataset containing various blood cell types.
- **Magnify and combine** images with **OpenCV** to create composite images.

Model Building:

- Build a convolutional neural network (CNN) to classify erythroblast cells.
- Use pre-trained models like **ResNet50** for transfer learning.
- Train the model with the augmented composite dataset.

Evaluation:

- Evaluate model performance using **accuracy, precision, recall**, and **F1-score**.
- Perform a **confusion matrix** analysis for multi-class classification.

Dataset Creation

Objective:

- Create a composite image by combining multiple cell images with varying magnifications.
- Use random positioning and magnifications to simulate real-world data variability.

Approach:

- **Image Magnification:** Apply random scaling (0.5x, 1x, 1.5x) using **OpenCV** `cv2.resize()`.
- **Image Combination:** Use random placement of resized images on a blank canvas with **OpenCV** array slicing.

Metadata:

- Store information on image class, magnification, and bounding box coordinates.
- Metadata is saved in a structured format (e.g., JSON or CSV).

B. Goswami et al.

Backbone Model:

- ResNet-50 chosen for feature extraction due to superior accuracy (98.72%) and computational efficiency.
- Pre-trained on ImageNet, fine-tuned for blood cell classification.

Classifiers:

- Evaluated traditional ML classifiers: SVM, XGB, KNN, RF.
- ResNet-50 used as a feature extractor; classifiers trained on extracted features.

Methodology:

- Normalized images (224x224 pixels) with mean and standard deviation from ImageNet.
- Employed cross-entropy loss function and Adam optimizer for ResNet-50 fine-tuning.
- 5-fold cross-validation to assess model performance.

Architecture

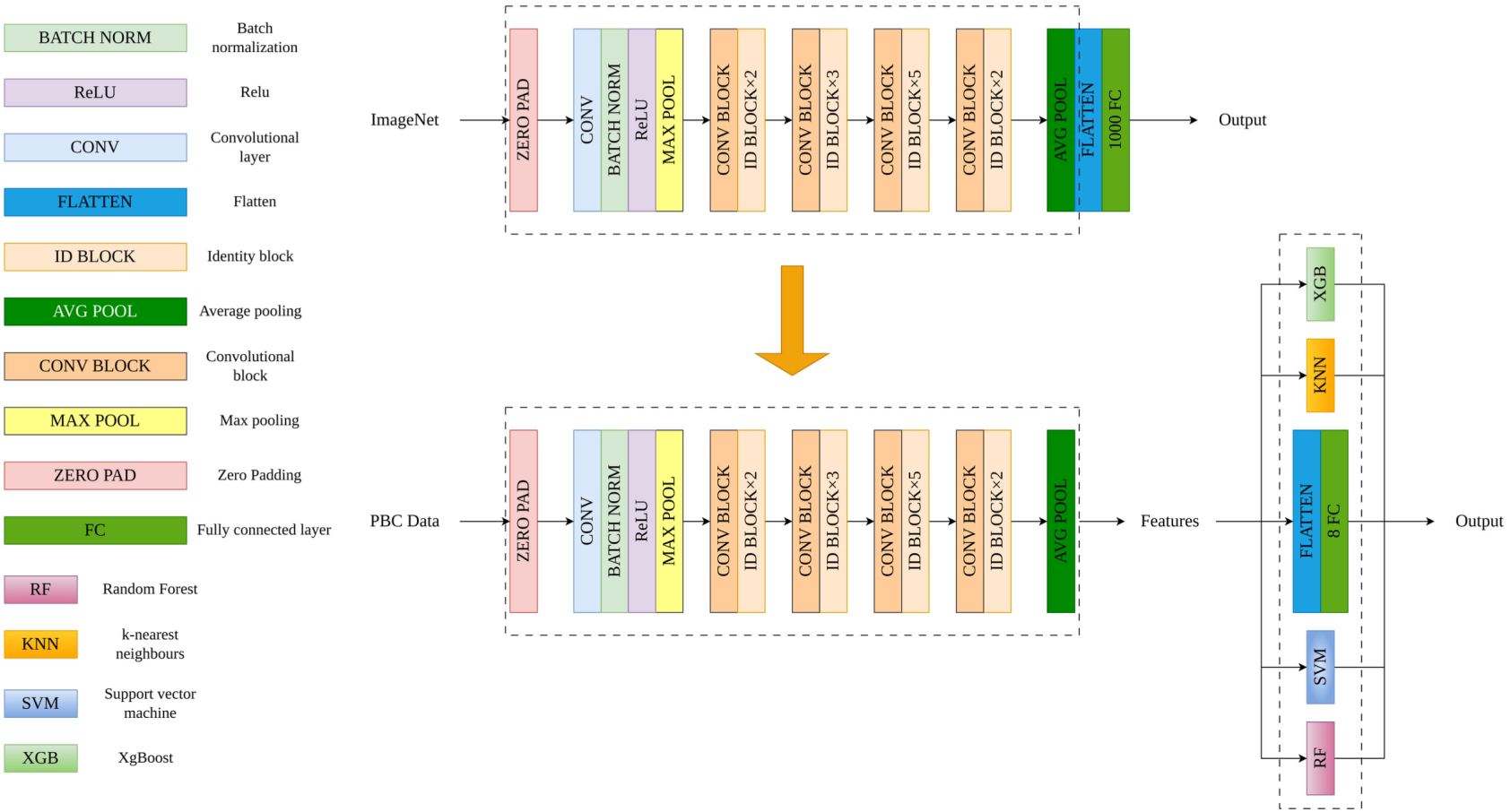

Figure: Classifier Enhanced ResNet-50 Model Architecture

Microcell-Net: A Deep Neural Network

Backbone Model:

- Microcell-Net, a deep neural network designed specifically for multi-class classification of microscopic blood cell images.
- Utilized convolutional layers to extract spatial and hierarchical features from the images.

Optimization and Training:

- Adaptive moment estimation (Adam) optimizer employed for efficient parameter updates.
- Cross-entropy loss function used to handle multi-class classification tasks.
- Data augmentation techniques applied to mitigate overfitting and enhance model robustness.

Architecture

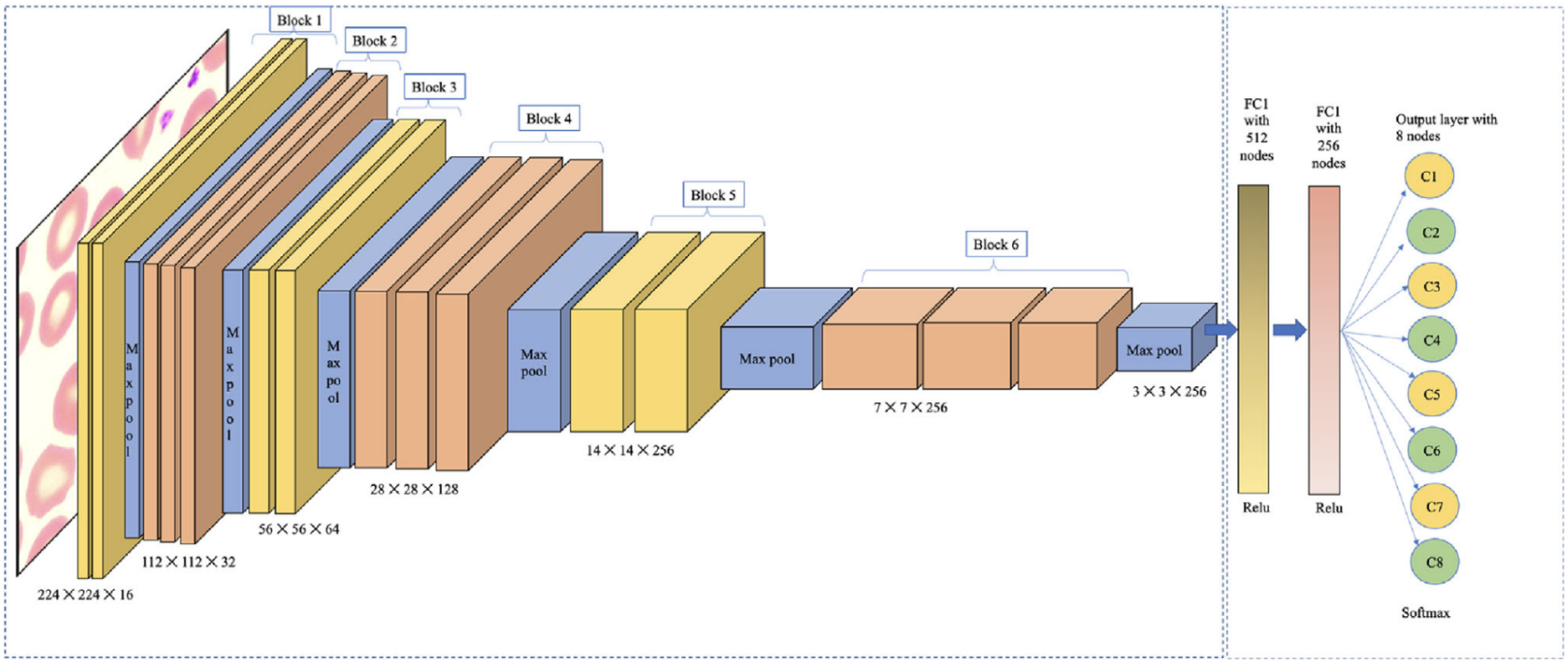

Figure: CNN Architecture of Microcell-Net

Plan for This Week

Focus: Dataset Creation with OpenCV

- Generate composite images by combining 8 classes randomly.
- Leverage OpenCV for image processing and manipulation.
- Ensure balanced representation of all classes in the generated images.
- Validate the created dataset to ensure consistency and correctness.

Goal:

- Create a diverse and well-structured dataset ready for training and analysis.

References

B. Goswami, A. B. Somaraj, P. Chakrabarti, R. Gudi, and N. Punjabi. (2024)

Classifier Enhanced Deep Learning Model for Erythroblast Differentiation with Limited Data

arXiv preprint arXiv:2411.15592.

https://arxiv.org/abs/2411.15592.

K. Dwivedi, M. K. Dutta. (2023)

Microcell-Net: A Deep Neural Network for Multi-class Classification of Microscopic Blood Cell Images

Expert Systems, 40(7), e13295.

https://doi.org/10.1111/exsy.13295.