

# Bone Marrow Cell Classification: Performance Analysis of Custom and Prebuilt CNN Models

---

## Introduction

This document summarizes the performance analysis of two Convolutional Neural Network (CNN) models applied to the Bone Marrow Cell Classification Dataset.

The task at hand involves classifying bone marrow cells into various categories based on image data. The analysis compares a pre-built CNN model and a custom-designed CNN model to determine their classification abilities.

The following metrics were considered for evaluation: Accuracy, Precision, Recall, F1-Score, AUC-ROC, and Confusion Matrix.

## Dataset Description

The dataset used in this project consists of images of bone marrow cells categorized into 7 classes.

The images are organized into subfolders, each corresponding to a different category. In addition to the images, a CSV file provides the mapping of images to their respective labels.

The dataset was accessed through Kaggle, and the images were pre-processed and resized for compatibility with the CNN models.

## Data Preprocessing

To prepare the dataset for training, the images were resized to 224x224 pixels, which is a standard input size for CNN models.

The images were also normalized to a range of 0-1 to improve the model's convergence during training.

Additionally, labels were one-hot encoded to match the output format

expected by the CNN models.

## Model Architecture

Two types of CNN models were trained: a pre-built CNN model (such as VGG16 or ResNet) and a custom-designed CNN model.

The pre-built model was fine-tuned on the dataset by adding a few dense layers on top of the pre-trained base to tailor it to the specific classification task.

The custom CNN model was designed from scratch, featuring multiple convolutional layers, max-pooling layers, and dropout for regularization.

Both models were trained using categorical cross-entropy as the loss function, with accuracy as the evaluation metric.

## Model Training and Evaluation

Both models were trained on a subset of the dataset due to time constraints.

The training was done for a small number of epochs

to reduce training time while still providing sufficient data for comparison.

The models were evaluated using a validation set that was set aside during the training process.

The following metrics were calculated for both models:

- Accuracy: Measures the proportion of correct predictions.
- Precision: Indicates how many of the predicted positive cases were actually correct.
- Recall: Measures how many of the actual positive cases were correctly identified.
- F1-Score: The harmonic mean of precision and recall.
- AUC-ROC: The area under the receiver operating characteristic curve, which indicates the model's ability to discriminate between classes.
- Confusion Matrix: Provides a visual representation of the model's performance across different classes.

## Results and Comparison

The results from the evaluation of both models were compared across the following metrics:

- Accuracy: Both models performed similarly in terms of accuracy, with the prebuilt model showing a slight edge.
- Precision: The precision was comparable between the models, with slight variations based on the class distribution.
- Recall: The recall values indicated how well each model was able to identify positive cases.
- F1-Score: Both models demonstrated good balance between precision and recall, with the custom model showing a marginally higher F1-score in some classes.
- AUC-ROC: Both models achieved high AUC-ROC scores, indicating their strong discriminatory ability.
- Confusion Matrix: The confusion matrices highlighted areas where both models confused certain classes, showing opportunities for further improvement.

## Conclusion

Both the prebuilt and custom CNN models demonstrated strong performance in the Bone Marrow Cell Classification task.

The prebuilt model, leveraging transfer learning, showed a slight advantage in terms of accuracy and AUC-ROC, likely due to the benefit of pre-trained weights.

The custom CNN model, while performing similarly, showcased its potential for adaptation to specific tasks through its tailored architecture.

Further improvements can be made by fine-tuning the models, increasing the dataset size, or exploring more advanced model architectures.

# RESULTS:



