Early Detection of White Blood Cell Abnormalities Using Attentive CNN with Class Weighting and Focal Loss

Muhamad Indra[1*] and Hilman Ferdinandus Pardede[2†]

[1*]Faculty of Information Technology, Universitas Nusa Mandiri, Jakarta, 12540, Indonesia.
[2]Research Center for Artificial Intelligence and Cybersecurity, National Research and Innovation Agency (BRIN), Jakarta, 12710 , Indonesia.

*Corresponding author(s). E-mail(s): 14230008@nusamandiri.ac.id;
Contributing authors: hilm001@lipi.go.id;
†These authors contributed equally to this work.

Abstract

Early detection of abnormal white blood cells (WBCs), such as in leukemia, is crucial for supporting accurate hematological diagnosis. However, manual identification through microscopic images is relatively time-consuming and constrained by the high morphological similarity between cell, making it difficult to identify consistently. This study develops a deep learning-based WBC classification model that addresses data imbalance and incorporates an attention mechanism. The model utilizes ResNet50 and EfficientNetB0 backbones, employing Class Weighting and Focal Loss for balancing, as well as the CBAM attention block to emphasize relevant cellular morphological features. The dataset was obtained from Munich University Hospital and underwent ROI segmentation, data augmentation, and evaluation using accuracy, F1-score, and confusion matrix. Results show that the ResNet50 model achieved an accuracy of 95.86% with a macro F1-score 75.15%, demonstrating improved detection performance for abnormal classes. This approach contributes to the development of reliable and efficient image-based automatic classification systems for hematological diagnostics.

Keywords: White Blood Cell Abnormalities, Image Classification, Deep Learning, Data Balancing, Attention Mechanism

1 Introduction

White blood cells (WBC) play an important role in the human immune system, namely in the fight against infection and the maintenance of physiological stability[1]. Abnormalities in the WBC, such as those occurring in leukemia, severe infection, or immunological disorders, can be an early indicator of serious hematologic disease. Therefore, early detection of WBC abnormalities is crucial in supporting the diagnosis, treatment and follow-up of patients[2].

Conventionally, WBC identification is performed manually through microscopic image observation by laboratory analysts. Although this method has been used for a long time, the process tends to be time consuming, requires specialized skills, and has the potential to produce varying results due to the morphological similarities between cell types that are difficult to distinguish visually[3]. This challenge becomes even more complex when dealing with large volumes of data and the need for accurate and rapid diagnostic results[4].

Advances in artificial intelligence technology, particularly Deep Learning, have opened up new opportunities in the development of white blood cell classification systems based on digital images. Deep learning models such as Convolutional Neural Networks (CNN) have been proven to be capable of recognizing visual patterns and performing automatic classification with high accuracy [5]. However, the main challenge in implementing these models is the presence of imbalanced data, where the number of samples in the abnormal class is significantly smaller than in the normal class, which can lead to bias in the model training process[6].

A study by Dasariju et al.[7] highlighted similar challenges when using the Random Forest algorithm for automatic leukocyte classification, especially immature cells such as monoblasts, promyelocytes, and myeloblasts. Although the multi-Otsu segmentation approach and colour-morphology features in the LAB colour space successfully achieved high accuracy of around 93% and an AUC-ROC of 0.98, the model's performance was still affected by the imbalance of minority classes, such as promyelocytes, which were far fewer in number. The classification results became less sensitive to rare cell types, which ultimately reduced the reliability of the model in real clinical scenarios.

Similarly, Elhassan et al[8]developed a deep learning-based classification method with automatic feature extraction techniques (CMYK-Moment Localisation) and a special CNN architecture that can recognise various types of WBC, including difficult types such as MOB and NGB. This model achieved an accuracy of up to 97% on internal and external datasets. However, extreme imbalance in class distribution remains a challenge, as high accuracy does not guarantee fairness in classification for minority classes. This indicates that without an appropriate imbalance handling strategy, model performance can be misleading and less representative.

This study aims to develop a white blood cell (WBC) classification model based on Transfer Learning with the integration of Attention Mechanism and data balancing techniques, such as Class Weighting and Focal Loss, to improve the model's performance in identifying abnormal white blood cells, especially in minority classes that are difficult to recognise. Data imbalance is a common issue in medical data, where

positive cases (minority) are far fewer than negative cases. This can cause the learning model to be biased towards the majority class and ignore the important class that is actually the main focus[9]. This approach is based on findings from previous studies showing that high accuracy does not always reflect fair and representative model performance, especially when dealing with imbalanced class distributions. The model was tested using the AML_Cytomorphology_LMU dataset, which consists of single-cell morphology images from peripheral blood smears of 100 Acute Myeloid Leukemia (AML) patients and 100 non-malignant controls, collected at the University Hospital of Munich between 2014 and 2017. This dataset includes fifteen different cell categories, including several immature types such as monoblasts, promyelocytes, and myeloblasts, whose limited presence poses a unique challenge in automatic classification.

2 Methods

This research was conducted through several main stages, starting from the collection of white blood cell image datasets, data pre-processing, dataset division for training and testing, model training with deep learning algorithms, to model performance evaluation.

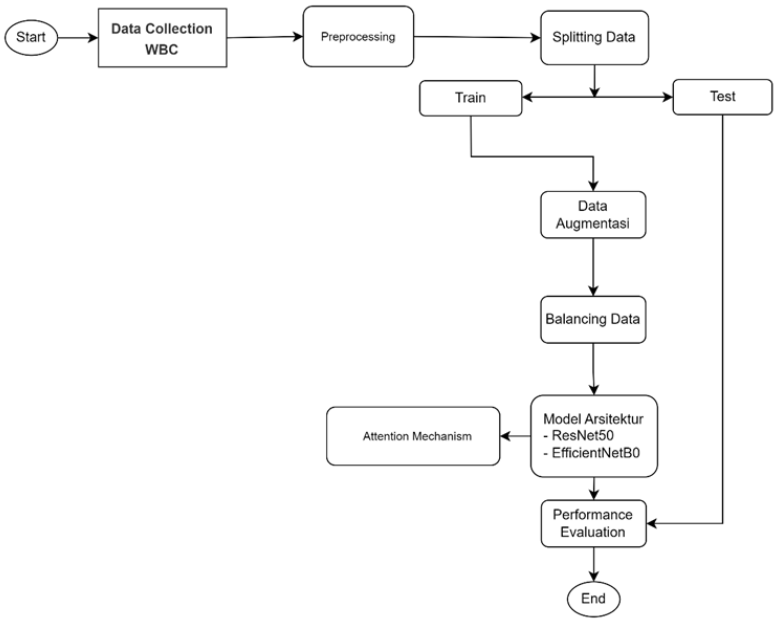

Fig. 1 Research flowchart block diagram

2.1 Dataset Collection

The dataset used in this study consists of microscopic images of white blood cells obtained from the public repository AML Cytomorphology LMU. This dataset

includes various types of white blood cell disorders, such as leukaemia, as well as normal and abnormal cell categories.

This dataset contains various types of cells, both normal and abnormal. Normal cells include basophils, eosinophils, band neutrophils, segmented neutrophils, monocytes, and typical lymphocytes. Abnormal cells include erythroblasts, smudge cells, atypical lymphocytes, metamyelocytes, monoblasts, myeloblasts, myelocytes, promyelocytes, and bilobed promyelocytes—many of which are associated with acute leukaemia.

The number of samples in each class varies significantly, leading to a highly imbalanced class distribution. This imbalance poses a major challenge in the training process of the classification model.

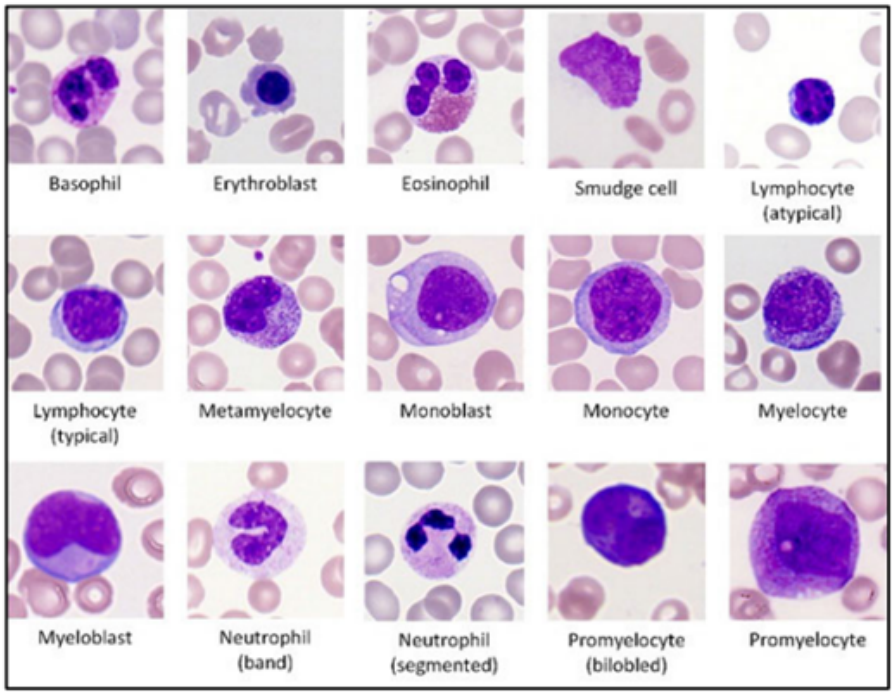

Fig. 2 Dataset WBC Normal and Abnormal

The number of samples across the white blood cell (WBC) classes exhibits substantial variation, as shown in Table 1. This results in a highly imbalanced class distribution, where certain classes such as Segmented_Neutrophil and Typical_Lymphocyte dominate the dataset, while others like Atypical_Lymphocyte or Metamyelocyte are severely underrepresented. Such imbalance presents a significant challenge during model training, as it can lead to biased predictions favoring majority classes and reduced sensitivity to minority classes.

Table 1 Number of Samples per White Blood Cell Class

Class Name	Number of Samples
Atypical_Lymphocyte	11
Band_Neutrophil	109
Basophil	79
Eosinophil	424
Erythroblast	78
Metamyelocyte	15
Monoblast	26
Monocyte	1789
Myeloblast	3268
Myelocyte	42
Promyelocyte	70
Promyelocyte_Bilobled	18
Segmented_Neutrophil	8484
Smudge_cell	18
Typical_Lymphocyte	3937

2.2 Pre Processing and Segmentation

In the preprocessing stage, white blood cell images are processed to ensure optimal input quality for the classification model. The process begins by loading the image and converting its colour space to CMYK format, then only the Cyan (C) component is taken because it is considered the most informative for highlighting the core details of the cell[8]. This component is then segmented using the Otsu thresholding method to separate the cell nucleus from the background. Thresholding itself serves to separate the pixels of the cell nucleus object from the background based on lighting or brightness levels, thereby helping to clarify the boundaries of the cell nucleus area to be analysed further[9]. After the initial segmentation, a series of morphological operations were performed, such as opening to remove fine noise points, closing to close narrow gaps between nucleus areas, and hole filling to fill small holes within the nucleus object[10]. This process helps to produce a cleaner and more solid binary mask to enhance the quality of the segmentation. Additionally, a denoising process using the Non-Local Means method is applied to the cropped image to reduce residual noise, ensuring that important details are preserved[11].

After the preprocessing stage is complete, the centroid of the nucleus is identified using spatial moments. The maximum distance from the centroid to the mask boundary is calculated as the base radius, then expanded by approximately 30% to ensure that the cytoplasm area around the nucleus remains covered[12]. Thus, the Region of Interest (ROI) is defined as a square area formed based on this adaptive radius. The ROI is directly cropped from the original image, then adjusted to a standard size through padding to ensure uniformity, and normalised in quality to be ready for use in subsequent processing stages[13].

This step produces an ROI that not only accurately represents the cell nucleus but also preserves the biological context surrounding the nucleus that is relevant for

white blood cell analysis or classification. With this strategy, the preprocessing and ROI extraction processes complement each other to ensure consistent, structured, and optimal image input quality in the haematology medical image classification pipeline.

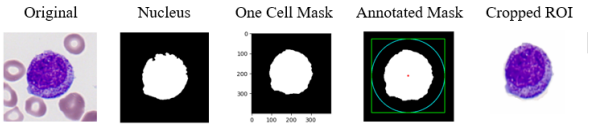

Fig. 3 Region of Interest (ROI) resulting from pre-processing and segmentation

2.3 Data Splitting

The data distribution was set at 80% for training data and 20% for test data[14]. This proportion was chosen so that the model would have sufficient data to learn representative variations in white blood cell morphology, ranging from cell nucleus shape and size to cytoplasm texture[15]. Meanwhile, the test data played an important role in objectively evaluating the model's performance under real conditions.

2.4 Data Augmentation

To improve the model's generalisation ability and enrich the diversity of white blood cell images, data augmentation was performed using ImageGenerator from Keras. The augmentation techniques included random rotation up to 30°, horizontal and vertical shifts of 10% each, zoom in/out in the range of 0.8 - 1.2, and a slight shear of 5%[16]. Flipping transformations were performed horizontally and vertically, provided they remained biologically relevant. Lighting variations were also added within a brightness range of 0.8 - 1.2 to mimic different imaging conditions. All tranformation used the "reflect" fill mode to avoid artefacts at the image edges. This approach ensures that the model can recognise cell morphology in various spatial and photometric conditons without losing its biological validity[17].

2.5 Data Balancing Techniques

In addition to augmentation, this study also applied data balancing techniques to address the imbalance in the amount of data between white blood cell classes. This imbalance can affect model performance because classes with fewer data tend to be ignored. Therefore, the class weighting method was used to give greater weight to minority classes, so that the contribution of prediction errors in those classes became more significant during the learning process. Additionally, the focal loss loss function was used, which is designed to focus the model's learning on samples that are difficult to classify.

2.6 Modeling

In this study, the model selection was carried out by considering the ability of deep learning architecture to recognise visual patterns in white blood cell images. The

models used included two architectures, namely ResNet50 and EfficientNetB0. These two models were selected because they have characteristics that are suitable for image classification tasks, especially in detecting variations in white blood cell morphology, such as cell nucleus shape and cytoplasm texture[18],[19].

ResNet50 was chosen because it has a residual block architecture with shortcut connections that are effective in addressing degradation issues in very deep networks, enabling it to extract features deeply without losing important information. Meanwhile, EfficientNetB0 was used because it is designed with the principle of compound scaling, which optimises depth, width, and image resolution in a balanced manner, resulting in high classification performance with a relatively efficient number of parameters.

2.6.1 ResNet-50

ResNet-50 is a CNN architecture from the ResNet (Residual Networks) family developed by Microsoft Research Asia, knows for its efficiency in image categorisation and optimal depth compared to other variants such as ResNet-18 and ResNet-101[20]. ResNet is designed to overcome the vanishing gradient problem that often arises in deep neural networks by applying residual blocks through skip connection [21]. According to Pardede et al.[22], this approach allows the model to learn the difference (residual) between input and output, rather than the entire function, thereby improving training stability and enabling the construction of deeper networks without performance degradation.

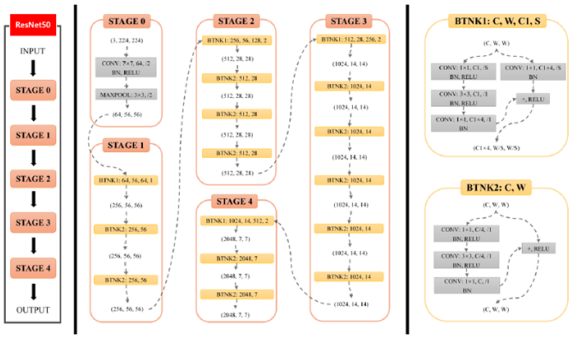

Fig. 4 ResNet-50 Architecture

2.6.2 EfficientNetB0

EfficientNet is a convolutional neural network design and scaling technique that uses compound coefficients to scale depth, width, and resolution parameters evenly[19]. The EfficientNet series is built on an optimised base architecture and successful scaling strategy. EfficientNet-B0 to EfficientNet-B7 are the eight EfficientNet architectures.

The more blocks used, the more parameters generated, but also the higher the accuracy. EfficientNet-B0 was trained using over one million images from the ImageNet database.

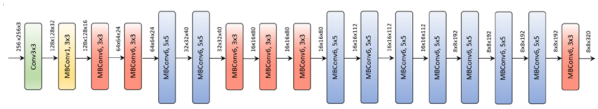

Fig. 5 EfficientNet-B0 Architecture

2.7 Convolutional Block Attention Module (CBAM)

This study applies the Convolutional Block Attention Module, a lightweight attention module embedded in the CNN architecture to strengthen feature representations by emphasising relevant information and suppressing less important information[23][24]. CBAM consists of two attention stages: Channel Attention uses a squeeze-and-excitation, and Spatial Attention leverages spatial pooling and 2D convolution [23][24]. this module has proven effective in various computer vision tasks such as medical imaging classification, object detection, and segmentation without significantly increasing computational load [23]-[25].

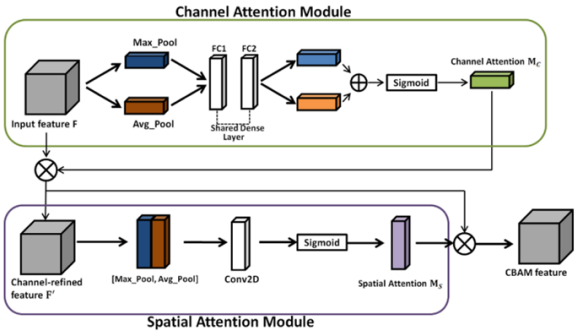

Fig. 6 CBAM Architecture

2.8 Evaluation

The proposed models were evaluated with quantitative evaluation such as such as accuracy, macro precision, recall, F1 score and Matthews Correlation Coefficienct (MCC) . These parameters are used to indicate how effectively the classifier worked with the test data. The values of TP (true positive), TN (true negative), FP (false positive), and FN (false negative) are used to calculate accuracy, precision, recall, and F1 score as follows [26][27][28][29][30].

$$\text{Accuracy} = \frac{\text{True Positives (TP)} + \text{True Negatives (TN)}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100 \tag{1}$$

$$\text{Macro Precision} = \frac{1}{N}\sum_{i=1}^{N}\text{Precision}_i \tag{2}$$

$$\text{Macro Recall} = \frac{1}{N}\sum_{i=1}^{N}\frac{TP_i}{TP_i + FN_i} \tag{3}$$

$$\text{Macro F1-score} = \frac{1}{N}\sum_{i=1}^{N}\frac{2 \cdot \text{Precision}_i \cdot \text{Recall}_i}{\text{Precision}_i + \text{Recall}_i} \tag{4}$$

$$\text{MCC} = \frac{TP \cdot TN - FP \cdot FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{5}$$

3 Results

Figure 7 presents the Grad-CAM visualization comparing the baseline and CBAM-enhanced models. The baseline model shows dispersed activations across the nucleus and background, indicating a lack of focus on diagnostically relevant regions. After integrating CBAM, the activations become more conenctrated around the nucleus, where key morphological features such as color and texture are prominent. The *Difference* map further demonstrates increased attention (red/orange) in relevant regions and reduced focus (blue) on background areas, confirming that CBAM effectively guides the model toward meaningful morphological structures and improves interpretabilltiy.

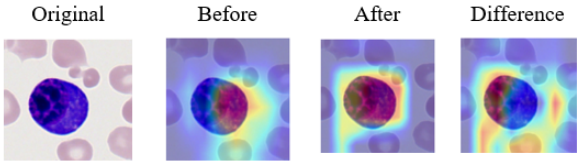

Fig. 7 Comparison before and after the implementation of the Attention mechanism

These qualitative observations are further supported by quantitative performance evalutations. To evaluate the effectiveness of different strategies in addressing class imbalance in white blood cell (WBC) image classification, we experimented with two baseline architecture; ResNet50 and EfficientNetB0. Both models were tested with enhancements such as Convolutional Block Attention Module (CBAM), Class Weighting (CW), Focal Loss (FL), and combinations of these techniques. The results are detailed below and summarized in Tables 2 and 3 .

ResNet50 was selected as the primary baseline due to its performance in medical image classification tasks. The baseline model achieved an accuracy of 95.81%, MCC of 0.9400, and macro F1-score of 68.29%. Although the overall accuracy and MCC were high, the macro precision 68.82% and macro recall 68.78% revealed limitations in recognizing minority WBC classes, indicating the influence of class imbalance on model sensitivity. Adding CBAM slightly reduced accuracy by 0.38% to 95.43% and MCC by 0.0052, but substantially improved macro metrics, with macro precision increasing by

13.10%, macro recall by 5.16%, and macro F1 - score by 5.95% in relative terms. These results demonstrate that CBAM effectively enhanced the model's ability to attend to discriminative spatial - channel features of minority classes, improving class balance despite a marginal drop in overall accuracy.

Applying Class Weighting yielded an accuracy of 95.34% and MCC of 0.9336. Compared to the baseline, macro precision increased by 4.97%, macro recall by 9.99% and macro F1 - score by 6.48% in relative terms. This suggests that class weighting improved the model's responsiveness to underrepresented samples while maintaining global prediction stabillity.

When Focal Loss was applied independently, the model achieved 95..56% accuracy and 0.9378 MCC, accompanied by signinficant macro metric improvements - macro precision up by 14.37%, macro recall up by 8.30% and macro F1 - score up by 9.49% relative to baseline. This consistent enhancement across all macro measures indicates that Focal Loss succesfully reduced majority-class dominance and increased attention to difficult or minority samples without notable accuracy degradation.

Combining Focal Loss and Class Weighting did not yield further gains. Accuracy dropped by 1.83% to 93.98% and MCC decreased to 0.9150. Nonetheless, macro recall and macro F1 - score improved by 8.43% and 4.60%, respectively, relative to baseline, indicating increased sensitivity but reduced classification stability - likely due to over-compensation fo minority classes.

The combination of CBAM and Class Weighting showed minimal performance change, with 95.48% accuracy and 0.9365 MCC. Macro precision, recall, and F1 - score improved only slightly by 0.28%, 1.91% and 1.10%, respectively. This indicates limited synergy between attention mechanism and class reweighting. The CBAM with Focal Loss configuration produced comparable results, with 95.62% accuracy, 0,9373 MCC and modes macro gains precision +4.60%, recall +1.18% and F1 - score +2.02%, showing that combined effect was beneficial but not optimal.

The best-performing configuration was the combination of CBAM, CW, and FL. This setup achieved an accuracy of 95.86% (a. o.05% increase) and MCC of 0.9408, accompanied by substantial improvements in macro metrics - macro precision up 22.49%, macro recall up 8.87%, and macro F1 - score up 10.04% relative to baseline. These results confirm that combining attention, adaptive loss, and reweighting mechanisms yields the most balance model, effectively improving sensitivity toward minority classes while maintaining high generalization and predictive consistensy. Detailed results for all ResNet50 configurations are provided in Tables 2.

Table 2 Performance Comparison of ResNet50 Variants for WBC Classification

Model Experiment	Accuracy	Macro			MCC
		Precision	Recall	F1-Score	
ResNet50	95.81%	68.81%	68.78%	68.29%	0.9400
ResNet50 + CBAM	95.43%	77.81%	72.34%	72.35%	0.9348
ResNet50 + CW	95.34%	72.23%	75.65%	72.96%	0.9336
ResNet50 + FL	95.56%	78.71%	74.49%	74.78%	0.9378
ResNet50 + FL + CW	93.98%	70.60%	74.58%	71.44%	0.9150
ResNet50 + CBAM + CW	95.48%	69.00%	70.10%	69.04%	0.9365
ResNet50 + CBAM + FL	95.62%	71.98%	65.59%	69.67%	0.9373
ResNet50 + CBAM + FL + CW	**95.86%**	**84.26%**	**74.88%**	**75.15%**	**0.9408**

Note: CBAM = *Convolutional Block Attention Module*, CW = *Class Weighting*, FL = *Focal Loss*

EfficientNetB0, a lightweight yet powerful model, was also evaluated as a second baseline. The initial baseline performance achieved 95.37% accuracy, MCC of 0.9338, macro precision of 70.99%, macro recall of 62.64%, and macro F1-score of 65.12%. These results indicate solid overall performance, but the relatively lower macro recall and F1-score suggest limited sensitivity to minority WBC classes.

Introducing CBAM improved the accuracy by 0.46% and increased MCC by 0.0066 in absolute terms. However, this enhancement was accompanied by a decrease in macro precision by 5.75%, macro recall by 1.53%, and macro F1-score by 3.33%, all in relative terms. This pattern reveals that while CBAM improved global performance metrics, it offered limited gains in class-level balance and even reduced minority class recognition.

Using class weighting independently caused a slight accuracy drop of 0.36% and a sharp decline in MCC by 0.0632 in absolute terms. However, it resulted in a significant increase in macro precision by 14.06%, macro recall by 6.69%, and macro F1-score by 9.31%, all measured relatively. These results suggest that class weighting strongly improved the model's sensitivity to minority classes but introduced higher variance and inconsistency in overall predictions.

Applying focal loss alone yielded an accuracy of 95.43%, reflecting a minor decline of 0.06%, while MCC increased slightly by 0.0007. Macro precision rose by 0.27% relative to baseline, but macro recall dropped by 3.78%, and macro F1-score declined by 1.64%. These mixed changes indicate that focal loss alone was insufficient to address imbalance and only modestly altered the model's performance.

Surprisingly, the combination of focal loss and class weighting did not outperform the baseline. The accuracy dropped by 0.08% and MCC by 0.0011. Meanwhile, macro precision declined by 4.90%, macro recall by 2.52%, and macro F1-score by 3.16%, all in relative terms. This suggests a potential incompatibility between the focal loss and class weighting when applied together in EfficientNetB0.

The combination of CBAM and class weighting produced the worst performance among all configurations. While macro recall increased significantly by 19.36%, macro precision dropped sharply by 28.93% and macro F1-score fell by 12.44%, with the

lowest accuracy recorded at 90.63%. This severe imbalance implies the model overcompensated toward minority classes, resulting in an overall loss of precision and reliability.

In contrast, the CBAM and focal loss combination delivered better results. Accuracy increased by 0.90%, and MCC improved by 0.0127 in absolute terms. Macro recall rose by 4.58%, macro F1-score improved by 2.87%, and macro precision decreased slightly by 1.04%, all in relative terms. While bias toward majority classes remained, this configuration offered more balanced improvements than most alternatives.

Finally, the best configuration for EfficientNetB0 was achieved by combining CBAM, focal loss, and class weighting. Although accuracy dropped by 0.27% and MCC by 0.0036, the model achieved a macro recall increase of 18.08% and a macro F1-score increase of 8.86%, with macro precision decreasing only slightly by 0.44%, all relative to the baseline. This outcome indicates that the three-way combination most effectively balanced sensitivity to minority classes with overall performance. A detailed summary of all configurations is presented in Tables 3 .

Table 3 Performance Comparison of EfficientNetB0 Variants for WBC Classification

Model Experiment	**Accuracy**	**Macro**			**MCC**
		Precision	**Recall**	**F1-Score**	
EfficientNetB0	95.37%	70.99%	62.64%	65.12%	0.9338
EfficientNetB0 + CBAM	95.83%	66.91%	61.68%	62.95%	0.9404
EfficientNetB0 + CW	95.73%	80.97%	66.83%	71.18%	0.8706
EfficientNetB0 + FL	95.43%	71.18%	60.27%	64.05%	0.9345
EfficientNetB0 + FL + CW	95.29%	67.51%	61.06%	63.06%	0.9327
EfficientNetB0 + CBAM + CW	90.63%	50.45%	74.77%	57.02%	0.8706
EfficientNetB0 + CBAM + FL	96.27%	71.85%	65.51%	66.99%	0.9465
EfficientNetB0 + CBAM + FL + CW	95.10%	70.68%	73.95%	70.89%	0.9302

Note: CBAM = *Convolutional Block Attention Module*, CW = *Class Weighting*, FL = *Focal Loss*

4 Discussion

The experiment demonstrate the complexity of handling class imbalance in WBC image classification using deep learning models. While both ResNet50 and EfficientNetB0 achieved high balance accuracy and MCC, their macro-level metric - particularly recall and F1 - score highlighted deficiencies in detecting minority classes.

For ResNet50, neither CBAM nor class weigthing alone sufficiently improved minority class performance. CBAM slightly increased accuracy but reduced macro precision and recall, suggesting an overemphasis on majority features. Class weighting alone raised recall but reduced precision. Focal Loss alone degraded all macro metrics, but its combination with class weighting improved recall and F1 - score with minimal impact on overall accuracy. The most balanced result came from integrating CBAM, Focal Loss and Class Weighting enhancing minority class sensitivity without sacrificing overall performance.

In contrast, EfficientNetB0 responded differently. Class weighting improved macro precision but significantly lowered MCC, indicating instability. Focal Loss and Class weighting combined were ineffective. However, the integration of all three methods – CBAM, Focal Loss and Class weighting yielded the best balance, improving recall and F1 - score while maintaining acceptable accuracy and MCC.

There results suggest that imbalance strategies shoule align with model architecture. ResNet50 benefited from Focal Loss and Class weighting, while EfficientNetB0 required a combination of all three techniques. Overall, a multi-strategy approach appears necessary to address imbalance in biomedical imaging. Future work should explore adaptive or meta-learning methods to dynamically tune class sensitivity during training and evaluate these strategies across broader datasets.

5 Conclusion

This study investigated the effectiveness of multiple strategies to address class imbalance in white blood cell (WBC) image classification, focusing on two deep learning architectures: ResNet50 and EfficientNetB0. Various configurations involving Convolutional Block Attention Module (CBAM), Class Weighting (CW), and Focal Loss (FL) were evaluated in isolation and combination.

The results showed that while baseline models achieved high overall accuracy and MCC, they suffered from reduced sensitivity to minority classes, as reflected in macro-level metrics. Individually, CBAM and class weighting offered partial improvements but often introduced trade-offs between precision and recall. Focal Loss alone was inadequate in addressing the imbalance.

The most effective configuration for both architectures was the integration of CBAM, Focal Loss, and Class Weighting. In ResNet50, this setup yielded balanced improvements in macro metrics with minimal loss in accuracy. In EfficientNetB0, the same combination achieved the best recall and F1-score across all experiments, demonstrating its robustness across model types.

These findings highlight the importance of combining complementary techniques to improve class-level generalization in imbalanced biomedical image datasets. Future research may explore dynamic or adaptive reweighting strategies, integration with transformer-based architectures, or generalization across multi-center WBC datasets to enhance clinical applicability.

Acknowledgements. Not applicable.

Declarations

Funding. Not applicable.

Conflict of interest. The authors declare that they have no confict of interest. .

Ethics approval and consent to participate. Not applicable.

Consent for publication. Not applicable.

Data availability. The dataset used in this study is publicly available and can be accessed at: https://www.cancerimagingarchive.net/collection/aml-cytomorphology_lmu/.

Materials availability. All materials used in this study are publicly available through standard libraries or repositories.

Code availability. The code used in this study is not publicly available.

Author contribution.

Author contributions. All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by [Author A], [Author B], and [Author C]. The first draft of the manuscript was written by [Author A], and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript. .

References

[1] Ferda Ernawan, M.F.Y.A. Kartika Handayani: Light gradient boosting with hyper parameter tuning optimization for covid-19 prediction. (IJACSA) International Journal of Advanced Computer Science and Applications, 514–523 (2022) https://doi.org/10.14569/IJACSA.2022.0130859

[2] Oh Joo KweonID, M.-K.L.H.R.K. Yong Kwan Lim: Red and white blood cell morphology characterization and hands-on time analysis by the digital cell imaging analyzer di-60. PLoS ONE **17**, 1–13 (2022) https://doi.org/10.1371/journal.pone.0267638

[3] O. J. Kweon, M.K.L. Y. K. Lim, Kim, H.R.: The impact of pathologist review on peripheral blood smears: A college of american pathologists q-probes study of 22 laboratories. Arch. Pathol. Lab. Med **149**(7), 618–626 (2025) https://doi.org/10.5858/arpa.2024-0117-cp

[4] Yawo M. Kobara, A.D.N.F.H.A.M.P. Ikpe Justice Akpan: Artifcial Intelligence and Data Science Methods for Automatic Detection of White Blood Cells in Images, (2025). https://doi.org/10.1007/s10278-025-01538-y

[5] Marcio Dorn, P.H.N.B.C.F.E.A.A.K.C.S.A. Bruno Iochins Grisci: Comparison of machine learning techniques to handle imbalanced covid-19 cbc datasets. vol. 7, pp. 1–34. PeerJ Computer Science (2021). https://doi.org/10.7717/peerj-cs.670

[6] Katz, B.Z., Benisty, D., Sayegh, Y., Lamm, I., Avivi, I.: Remote Digital Microscopy Improves Hematology Laboratory Workflow by Reducing Peripheral Blood Smear Analysis Turnaround Time vol. 13, pp. 1108–1115 (2022). https://doi.org/10.1055/a-1957-6219

[7] Dasariraju, S., Huo, M., McCalla, S.: Detection and classification of immature leukocytes for diagnosis of acute myeloid leukemia using random forest algorithm, vol. 7, pp. 1–12 (2020). https://doi.org/10.3390/bioengineering7040120

[8] Elhassan, T.A.M., Rahim, M.S.M., Swee, T.T., Hashim, S.Z.M., Aljurf, M.: Feature Extraction of White Blood Cells Using CMYK-Moment Localization and Deep Learning in Acute Myeloid Leukemia Blood Smear Microscopic Images (2022). https://doi.org/10.1109/ACCESS.2022.3149637

[9] Merlina, N., Noersasongko, E., Andono, P.N., Soeleman, M.A., Riana, D.: Optimization of the Preprocessing Method for Edge Detection on Overlapping Cells at PAP Smear Images (2023). https://doi.org/10.30630/joiv.7.2.1329

[10] Fatonah, N.S., Tjandrasa, H., Fatichah, C.: Identification of acute lymphoblastic leukemia subtypes in touching cells based on enhanced edge detection (2020). https://doi.org/10.22266/IJIES2020.0831.18

[11] Lepcha, D.C., Dogra, A., Goyal, B., Goyal, V., Kukreja, V., Bavirisetti, D.P.: A constructive non-local means algorithm for low-dose computed tomography denoising with morphological residual processing (2023). https://doi.org/10.1371/journal.pone.0291911

[12] Muthumanjula, M., Bhoopalan, R.: Detection of White Blood Cell Cancer using Deep Learning using Cmyk-Moment Localisation for Information Retrieval (2022). https://doi.org/10.36548/jismac.2022.1.006

[13] Khan, R.U., Almakdi, S., Alshehri, M., Haq, A.U., Ullah, A., Kumar, R.: An intelligent neural network model to detect red blood cells for various blood structure classification in microscopic medical images (2024). https://doi.org/10.1016/j.heliyon.2024.e26149

[14] Nguyen, Q.H., Dao, D.T., Le, T.A., Ngo, D.T., Le, M., Le, T.T.: Influence of data splitting on performance of machine learning models in prediction of shear strength of soil (2021). https://doi.org/10.1155/2021/4832864

[15] Kouzehkanan, Z.M., Tavakoli, M., Ghassemi, N., Shakeri, M.Z., Rezvanian, A.T., Dehkordi, R.H., Dinarvand, R., Helfroush, M.S., Asadi, M.R., Gharehbaghi, S., Ghaffari, M., Boostani, R., Amini, A.M., Marateb, H.R.: A large dataset of white blood cells containing cell locations and types, along with segmented nuclei and cytoplasm. Nature Publishing Group (2022). https://doi.org/10.1038/s41598-021-04426-x . https://doi.org/10.1038/s41598-021-04426-x

[16] Gao, X., Saha, R.K., Prasad, M.R., Roychoudhury, A.: Fuzz Testing Based Data Augmentation to Improve Robustness of Deep Neural Networks. ACM (2020). https://doi.org/10.1145/3377811.3380415 . https://doi.org/10.1145/3377811.3380415

[17] Kandel, I., Castelli, M., Manzoni, L.: Brightness as an Augmentation Technique for Image Classification (2022). https://doi.org/10.28991/ESJ-2022-06-04-015 . https://doi.org/10.28991/ESJ-2022-06-04-015

[18] Toptaş, M., Toptaş, B., Hanbay, D.: Classifying white blood cells using combining different convolutional neural networks. Springer (2025). https://doi.org/10.1007/s11042-025-20879-y . https://doi.org/10.1007/s11042-025-20879-y

[19] Aksoy, A.: An Innovative Hybrid Model for Automatic Detection of White Blood Cells in Clinical Laboratories (2024). https://doi.org/10.3390/diagnostics14182093 . https://doi.org/10.3390/diagnostics14182093

[20] Sarker, I.H.: Machine Learning: Algorithms, Real-World Applications and Research Directions (2021). https://doi.org/10.1007/s42979-021-00592-x

[21] Vujović, : Classification model evaluation metrics. International Journal of Advanced Computer Science and Applications (IJACSA) **12**(6), 599–606 (2021) https://doi.org/10.14569/IJACSA.2021.0120670

[22] Pardede, H.F., et al.: Plant Diseases Detection with Low Resolution Data Using Nested Skip Connections, (2020). https://doi.org/10.1186/s40537-020-00332-7

[23] Wang, Y., Chen, X., Li, J., Lu, Z.: Convolutional block attention module–multimodal feature-fusion action recognition: Enabling miner unsafe action recognition. Sensors **24**(14) (2024) https://doi.org/10.3390/s24144557

[24] Zakariah, M., Alnuaim, A.: Recognizing human activities with the use of convolutional block attention module. Egyptian Informatics Journal **27**(September), 100536 (2024) https://doi.org/10.1016/j.eij.2024.100536

[25] Wu, S., Yamauchi, H.: A lightweight binarized convolutional block attention module: B-cbam. Journal of Advanced Information Technology **16**(5), 751–759 (2025) https://doi.org/10.12720/jait.16.5.751-759

[26] Maza, D., Ojo, J.O., Akinlade, G.O.: A predictive machine learning framework for diabetes. Turkish Journal of Engineering **8**(3), 583–592 (2024) https://doi.org/10.31127/tuje.1434305

[27] Aguilar-Ruiz, J.S., Michalak, M.: Classification performance assessment for imbalanced multiclass data. Scientific Reports **14**(1), 1–10 (2024) https://doi.org/10.1038/s41598-024-61365-z

[28] Uddin, S., Haque, I., Lu, H., Moni, M.A., Gide, E.: Comparative performance analysis of k-nearest neighbour (knn) algorithm and its different variants for disease prediction. Scientific Reports **12**(1), 1–11 (2022) https://doi.org/10.1038/s41598-022-10358-x

[29] Ivanov, I.G., Kumchev, Y., Hooper, V.J.: An optimization precise model of stroke data to improve stroke prediction. Algorithms **16**(9), 1–16 (2023) https://doi.org/10.3390/a16090417

[30] Zhang, J., Chen, L., Abid, F.: Prediction of breast cancer from imbalance respect using cluster-based undersampling method. Journal of Healthcare Engineering **2019**, 7294582 (2019) https://doi.org/10.1155/2019/7294582