West Visayas State University
COLLEGE OF INFORMATION AND COMMUNICATIONS TECHNOLOGY
Luna St., La Paz, Iloilo City 5000
Iloilo, Philippines
* Trunkline: (063) (033) 320-0870 loc 1403 * Telefax No.: (033) 320-0879
* Website: www.wvsu.edu.ph * Email Address: cict@wvsu.edu.ph

Automated Multi-Class White Blood Cell Classification for Early Neonatal Sepsis Using Custom WBCNet CNN

A Final Project Documentation
Presented to the Faculty of the
College of Information and Communications Technology
West Visayas State University
La Paz, Iloilo City

In Partial Fulfillment of the Requirements for the Subject
CCS 248 - Artificial Neural Networks

By
Esponilla, Wilfame
Gegawin, Ariane Pearl
Libuna, Donjie
BSCS 3-B AI

December 2025

I. Problem Statement

Neonatal sepsis, a life-threatening bloodstream infection affecting newborns, requires rapid and accurate diagnosis for timely intervention, yet traditional microscopic examination methods can delay treatment decisions.

The manual analysis of white blood cell (WBC) types from peripheral blood smears is a time-consuming, labor-intensive process prone to human error and inter-observer variability, particularly in critical care settings such as neonatal intensive care units.

Our project aims to develop an automated multi-class WBC classification system using Convolutional Neural Networks (CNN) to accurately identify eight distinct blood cell types from microscopic images, thereby supporting early risk stratification of neonatal sepsis through rapid identification of abnormal cell distributions and the presence of immature granulocytes.

II. Current Solutions and Limitations

Currently, clinical laboratories rely primarily on manual microscopic examination performed by trained hematologists and medical technologists to classify WBCs and calculate critical diagnostic ratios such as the immature-to-total neutrophil (I:T) ratio. This conventional approach, while considered the gold standard, suffers from significant drawbacks including prolonged turnaround times (often 2-4 hours), subjective interpretation variations between different observers, and the requirement for highly skilled personnel who may not be available 24/7 in resource-limited settings.

Automated hematology analyzers exist in the market, such as the Sysmex XN-series and Beckman Coulter DxH systems, which can perform complete blood counts with five-part differential; however, these systems have limitations in accurately identifying immature

granulocytes and abnormal cell morphologies that are crucial for sepsis diagnosis. Recent research has explored machine learning approaches for blood cell classification, with studies demonstrating the feasibility of deep learning models achieving 85-95% accuracy on binary classification tasks, though most existing solutions focus on limited cell types or require extensive preprocessing that limits real-world clinical applicability.

Moreover, current automated solutions often lack the granularity needed for neonatal sepsis assessment, as they typically do not provide detailed classifications of all eight clinically relevant cell types (neutrophils, eosinophils, basophils, lymphocytes, monocytes, immature granulocytes, erythroblasts, and platelets) necessary for comprehensive hematological analysis.

The integration of artificial intelligence in clinical hematology is still in its nascent stages, with most AI-based systems remaining in research environments rather than being deployed in routine clinical practice. Furthermore, existing deep learning models often rely on proprietary datasets that are not publicly available, making it difficult for researchers and clinicians to validate or adapt these systems to their specific clinical contexts.

Our proposed solution addresses these gaps by developing an open, transparent CNN-based classification system trained on a comprehensive, expert-annotated dataset that covers all eight critical cell types, with specific focus on supporting the automated calculation of sepsis indicators such as the I:T ratio and absolute neutrophil count (ANC).

III. Dataset Selection

Primary Dataset: Barcelona Hospital Blood Cell Images

For this project, we have selected the Barcelona Hospital Blood Cell Images dataset available on Kaggle as our primary data source. The original dataset contains 17,092 high-quality images of individual blood cells captured from peripheral blood smears, covering eight clinically relevant cell classes: neutrophils, eosinophils, basophils, lymphocytes, monocytes, immature granulocytes, erythroblasts, and platelets. Each image measures 360×363 pixels in RGB format, providing sufficient resolution to capture morphological features such as nuclear shape, cytoplasmic characteristics, and granularity patterns. All images have been carefully annotated by expert pathologists from the Hospital Clinic of Barcelona, ensuring accurate ground truth labels suitable for clinical and research applications.

For this project, we have reduced the dataset to approximately 3,000 images while maintaining representation across all cell classes. This smaller subset provides a manageable dataset size for training a deep learning model while still capturing the diversity of blood cell types. The model is trained using a CNN architecture, allowing us to develop a custom WBCNet model from scratch tailored to this dataset.

The dataset is particularly relevant for neonatal sepsis applications because it includes immature granulocytes as a distinct class. Elevated levels of immature granulocytes are a key hematological indicator for assessing sepsis risk in newborns. By including this class, the dataset enables models trained on it to potentially support early detection and clinical decision-making in neonatal care.

IV. Solution Overview

The proposed solution is an AI-driven white blood cell (WBC) classification system based on a WBCNet architecture, a custom built Convolutional Neural Network (CNN) designed for deep feature extraction. The primary objective is to provide fast and accurate identification of eight WBC types—neutrophils, eosinophils, basophils, lymphocytes, monocytes, immature granulocytes, erythroblasts, and platelets—reducing manual workload and minimizing human error in laboratory analysis. The system also supports neonatal sepsis screening by enabling automated computation of hematological ratios, such as the immature-to-total (I:T) ratio, which is critical for early detection and timely intervention.

The workflow begins with dataset loading and preprocessing, which standardizes image size, format, and pixel intensity to ensure consistent input for the WBCNet model. Data augmentation techniques, including rotations, flips, brightness adjustments, and zoom variations, are applied to increase dataset diversity and improve model generalization, particularly for underrepresented classes. Images are then fed into the custom CNN model, which learns hierarchical features directly from the data during training. The model is trained through a structured pipeline incorporating classification loss functions, learning-rate scheduling, and mini-batch optimization. Real-time performance tracking is facilitated through Matplotlib-generated accuracy and loss plots, providing transparency and interpretability throughout the learning process. The approach prioritizes clarity and modularity, ensuring that the system can easily be extended to future medical imaging tasks or incorporated into real-time NICU screening workflows.

V. Network Structure

The custom WBCNet CNN architecture was designed to balance computational efficiency with high discriminative capability. The network consists of multiple convolutional blocks, each containing convolutional layers, ReLU activation functions, batch normalization layers, and max-pooling operations. These components work together to capture morphological cues such as nuclear shape, cytoplasmic granularity, contour irregularities, and staining variations.

Feature extraction proceeds from fine-grained low-level features toward deeper abstract representations that describe cell type–specific patterns. The extracted feature maps are flattened and passed into fully connected layers optimized for eight-class classification. The final output layer applies softmax activation to estimate class probabilities. This architecture remains lightweight enough for deployment on standard clinical or embedded systems yet powerful enough to model subtle hematological differences central to diagnostic accuracy.

VI. Tools and Development Environment

The project was developed using the PyTorch deep learning framework. Core libraries include Torchvision for transforms and dataset handling, NumPy for numerical operations, and Matplotlib for graph generation and performance visualization. The project structure follows reproducible best practices, with separate sections for preprocessing, model definition, training utilities, and evaluation.The development environment was Google Colab, which provided cloud-based GPU acceleration, interactive notebooks, and model-tracking capabilities. Colab allowed for efficient experimentation, iterative testing, and reproducibility, eliminating the need for high-performance local hardware while ensuring smooth integration with deep learning workflows.

VII. Overall Accuracy

After completing the training process, the custom WBCNet model achieved a validation accuracy of 98.32%, reflecting strong discriminative capability across all eight white blood cell categories. This level of performance shows that the model effectively learned the morphological features required for reliable multi-class classification directly from the dataset. Although certain pretrained CNN architectures may reach slightly higher accuracy levels, the customized WBCNet design offers the advantage of feature representations specifically tuned to the characteristics of blood smear images, making it well suited for clinical applications.

VIII. Hyperparameter Tuning

Key hyperparameters were systematically optimized to achieve the best performance. Parameters such as learning rate, batch size, number of epochs, and optimizer type were iteratively tested. The final configuration was determined to be a learning rate of 0.001, batch size of 32, 20 epochs, and the Adam optimizer. Because the WBCNet architecture does not include dropout, the model relies on data augmentation and batch normalization as the primary regularization mechanisms. Augmentation includes random flips, random rotations, color transformations, and normalization, producing varied samples that reduce overfitting risks. These combined strategies support generalization across diverse cell morphologies while preserving the integrity of clinically relevant features.

IX. Evaluation Scores

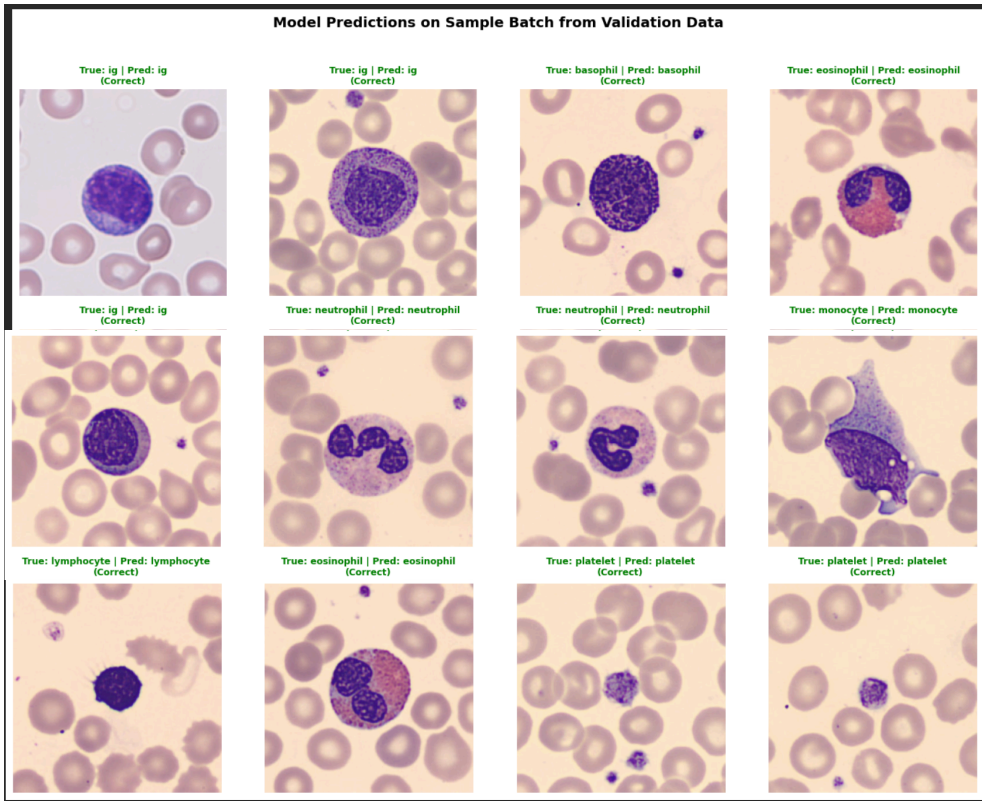
The model's performance was evaluated using precision, recall, F1-score, and overall accuracy metrics. Macro-averaged results reflected uniformly strong performance, with precision, recall, and F1-score all recorded at 98%. Class-wise evaluation showed that

well-represented categories such as eosinophils, neutrophils, and platelets achieved exceptionally high precision and recall, with eosinophils and platelets reaching perfect scores. Basophils, erythroblasts, lymphocytes, and immature granulocytes likewise demonstrated high discriminative performance, maintaining recall values between 96-100%. The only comparatively lower recall was observed in monocytes with 96%, likely due to greater morphological overlap with neighboring leukocyte types. Despite this, the overall validation accuracy of 98.32% indicates strong separation across all classes and suggests that misclassifications were minimal and largely confined to cells with highly similar visual features. Collectively, these evaluation results demonstrate that the WBCNet model is robust, reliable, and well suited for automated hematological analysis, supporting its potential integration into clinical workflows for neonatal diagnostic assessment.

X. Documentation

```
...
    Classification Report:

                  precision    recall  f1-score   support

       basophil       0.99      0.98      0.99       180
     eosinophil       1.00      1.00      1.00       467
    erythroblast      0.96      1.00      0.98       218
              ig      0.96      0.97      0.97       450
     lymphocyte       1.00      0.96      0.98       196
       monocyte       0.98      0.96      0.97       238
     neutrophil       0.97      0.99      0.98       480
       platelet       1.00      1.00      1.00       334

       accuracy                           0.98      2563
      macro avg       0.98      0.98      0.98      2563
   weighted avg       0.98      0.98      0.98      2563


    Overall Validation Accuracy: 0.9832
```

```
[OK] Batch Accuracy: 100.00% (32/32 correct)
  Displayed 12 predictions with True vs Predicted labels

Per-class performance across the validation subset:
  - basophil: 176/180 (97.78%)
  - eosinophil: 465/467 (99.57%)
  - erythroblast: 218/218 (100.00%)
  - ig: 435/450 (96.67%)
  - lymphocyte: 189/196 (96.43%)
  - monocyte: 228/238 (95.80%)
  - neutrophil: 475/480 (98.96%)
  - platelet: 334/334 (100.00%)
```



References

Kaggle. (n.d.). Blood Cells Image Dataset. Retrieved from https://www.kaggle.com/datasets/unclesamulus/blood-cells-image-dataset

ScienceDirect. (2024). Machine learning approaches for white blood cell analysis in sepsis detection. Retrieved from https://www.sciencedirect.com

Hospital Clinic of Barcelona. (2023). Barcelona Hospital Blood Cell Images Dataset Documentation.

Manroe, B. L., et al. (1979). The neonatal blood count in health and disease. I. Reference values for neutrophilic cells. Journal of Pediatrics, 95(1), 89-98.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 770-778.

Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

Tan, M., & Le, Q. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. International Conference on Machine Learning, 6105-6114.