

AI Powered Leukocyte Classification, A Machine Learning Approach To Blood Diagnostics

Janelli M. Mendez, DIT¹, Geoffrey B. Layco, MIS², Jopher F. Reyes, MIT³, Mark Ericson B. Baladad, MMPHA, RMT⁴, Josephine C. Milan, MSMT, RMT⁵

^{1,2,3} College of Computer Studies and Engineering, LORMA Colleges, CLI Campus, San Juan La Union, Philippines

^{4,5} College of Medical Laboratory Studies. LORMA Colleges, Carlatan Campus, San Fernando, La Union, Philippines

³department, College or University/Industry, City, Country.

Email: ¹janelli.mendez@lorma.edu, ²jlayco@lorma.edu, ³Jopher.reyes@lorma.edu, ⁴

Markericson.baladad@lorma.edu, ⁵Josephine.milan@lorma.edu

Orchid Id number: ¹0009-0001-9127-4896, ²0009-0007-2112-3877,

Abstract:

Diagnosis of various diseases including infections, immune disorders, and leukemia requires white blood cell (WBC) classification. The traditional methods take long hours for WBC classification through manual examination by the hematologists, which may lead to errors by human judgment, as it is difficult and requires much expertise. But with recent developments in deep learning, automated WBC classification is gaining interest due to its accuracy and speed. This study investigates a deep learning approach to WBC classification using the VGG16 convolutional neural network (CNN) architecture.

Transfer learning is used in our model in which a pre-trained VGG16 network has been fine-tuned using a dataset containing labeled WBC images. The dataset contains images of neutrophils, lymphocytes, monocytes, and eosinophils for a truly comprehensive classification task. The model's WBC type differentiation ability is evaluated based on the accuracy, precision, re-call, and F1-score for the different classes. The model generalization is further enhanced using data augmentation, thus preventing overfitting.

The experimental results have shown that the VGG16-based model achieves a very high classification accuracy when compared with other traditional machine learning-based approaches and even some custom deep learning architectures. It demonstrates that deep learning, and especially VGG16, could serve as an important tool to automate WBC classification and thereby cut down on the time for diagnostic purposes while assisting hematological analyses by medical professionals. Future work may involve the integration of the model into real-time diagnostic systems for rare WBC abnormalities in the dataset and the evaluation of more sophisticated architectures for maximized performance. The study showcases how deep learning can change the paradigm of medical image analysis and take diagnostics to another level of accuracy.

Keywords: Deep Learning, White Blood Cell Classification, VGG16, Convolutional Neural Networks (CNNs), Medical Image Analysis..

INTRODUCTION:

White blood cells (WBCs), or leukocytes, play a crucial role in the immune system by defending the body against infections and foreign invaders. Abnormalities in WBC count and morphology are indicative of various medical conditions, such as bacterial and viral infections, autoimmune diseases, and hematological malignancies like leukemia. Accurate and timely classification of WBCs is vital for early diagnosis and effective treatment planning. Traditionally, WBC classification is performed manually by hematologists through microscopic examination of stained blood smear slides. While effective, this method is time-consuming, prone to human error, and requires significant expertise [1].

With advancements in artificial intelligence (AI) and deep learning, automated WBC classification has gained increasing attention. Deep learning, particularly convolutional neural networks (CNNs), has demonstrated exceptional performance in medical image analysis, offering higher accuracy and efficiency compared to traditional machine learning approaches [2]. Among the various CNN architectures, VGG16 has been widely adopted for image classification tasks due to its deep feature extraction capabilities and simple yet effective architecture [1,2]. VGG16 consists of 16 layers with small convolutional filters, allowing it to capture intricate patterns in images while maintaining computational efficiency.

Several studies have explored deep learning approaches for WBC classification. For instance, [3] implemented a CNN-based model for leukocyte classification and reported high accuracy compared to conventional machine learning techniques. Similarly, [3] investigated transfer learning with pre-trained models like VGG16 and ResNet, demonstrating their effectiveness in medical image analysis. These studies highlight the potential of deep learning in automating WBC classification and improving diagnostic efficiency.

According to [3,4] developed a CNN-based model for leukocyte classification, achieving superior accuracy compared to traditional machine learning methods. This study emphasized the ability of CNNs to learn complex features from microscopic blood smear images. A deep CNN model specifically designed for WBC classification. The model was trained on a large dataset of stained peripheral blood smear images, containing different types of leukocytes, including neutrophils, lymphocytes, monocytes, and eosinophils. The study implemented data preprocessing techniques, such as contrast enhancement and image augmentation, to improve model robustness and generalization.

One of the key aspects of their approach was feature extraction using CNNs, which eliminated the need for manual feature engineering. Traditional machine learning techniques, such as support vector machines (SVMs) and k-nearest neighbors (KNN), require handcrafted feature extraction, which can be time-consuming and less effective in capturing intricate cellular structures. In contrast, CNNs automatically learn relevant spatial features, making them highly effective for WBC classification.

Investigated transfer learning using pre-trained models like VGG16 and ResNet for WBC classification, demonstrating their effectiveness in medical image analysis. Transfer learning, which leverages pre-trained models on large datasets, can significantly improve performance, particularly when dealing with limited medical image data. [5,6]

ResNet (Residual Network) introduces skip connections that help mitigate the vanishing gradient problem in deep networks [7]. This allows for more efficient training and better feature extraction, particularly in complex datasets like medical images.

Developed an advanced system for white blood cell (WBC) classification, integrating image processing, segmentation, and transfer learning to create a highly efficient and accurate classification pipeline. Their work addressed key challenges in automated hematological analysis, such as dealing with noise in microscopic images, accurately segmenting WBCs from surrounding components (e.g., red blood cells and platelets), and leveraging deep learning models for high-precision classification. By combining traditional image processing techniques with state-of-the-art deep learning approaches, [7] provided a robust framework for automated WBC analysis. [8]

Microscopic blood smear images often contain artifacts, uneven illumination, and variations in staining intensity, which can hinder accurate classification. [10] implemented several image preprocessing techniques to enhance image quality before classification, including: Contrast enhancement: To improve visibility of WBC features. Noise reduction: Applying median filtering to remove unwanted noise. Color normalization: Standardizing stain variations across different samples.

This study aims to leverage the VGG16 architecture for automated WBC classification by fine-tuning the pre-trained model on a dataset of labeled WBC images. The key objectives of this research include developing a deep learning-based model for accurate classification of different WBC types, including neutrophils, lymphocytes, monocytes, and eosinophils. Evaluating the model's performance using key metrics such as accuracy, precision, recall, and F1-score. Comparing the effectiveness of VGG16-based classification with traditional machine learning approaches and other deep learning architectures. Exploring the impact of data augmentation techniques on improving model generalization and robustness.

The impact of this study is addressing the challenges in manual white blood cell classification. Time-consuming in reviewing and classifying WBCs in blood smears requires significant expertise and effort. Prone to Human Error due to factors like fatigue and inter-observer variability can lead to inconsistencies in diagnosis. Labor-Intensive in resource-limited settings, the availability of trained hematologists is often insufficient.

This study addresses these challenges by automating the WBC classification process using deep learning, image processing, and segmentation techniques, significantly improving diagnostic accuracy and efficiency.

2) METHODS AND METHODOLOGY:

This study adopted Mixed Method of design, which quantitative and applied approach, focusing on the development and evaluation of deep learning models for automated WBC classification. The methodology will be structured into the following phases: Dataset Acquisition and Preparation using the Image Collection using the microscopic blood smear images will be collected with the help of College of Medical Laboratory Science of Lorma Colleges.

The dataset will encompass images of various WBC types (neutrophils, lymphocytes, monocytes, eosinophils, basophils) and potentially include images with pathological abnormalities with the help of the College of Medical Laboratory Science. Image acquisition parameters (e.g., magnification, staining protocol) will be documented to ensure consistency. Image Preprocessing using the images will be preprocessed to enhance quality and reduce noise.

Preprocessing steps may include the color normalization to standardize staining variations. Contrast enhancement to improve cell visibility that will be annotated by the Expert Medical Laboratory Professionals. Noise reduction using filtering techniques. Image resizing and cropping to standardize image dimensions. Image Annotation where in the Expert Medical Laboratory Professionals will annotate the WBC images, labeling each cell with its corresponding type. Annotations will be validated by Medical Laboratory Experts to ensure accuracy and minimize inter-observer variability. Annotations will be saved in a format compatible with the deep learning framework.

During the Data augmentation techniques will be employed to increase the size and diversity of the training dataset, improving model robustness and generalization.

Augmentation techniques may include different phases such as Rotation, flipping, and scaling. Brightness and contrast adjustments. Elastic deformations. Adding small amounts of noise.

The dataset will be partitioned into three subsets: Training set will be used to train the deep learning models. Validation set will be used to tune hyperparameters and monitor model performance during training. Test set will be used to evaluate the final performance of the trained models on unseen data.

The researchers adopted the CNN architecture which has 6 layers. The architecture is designed to learn spatial hierarchies of features automatically and adaptively, from low to high-level patterns.

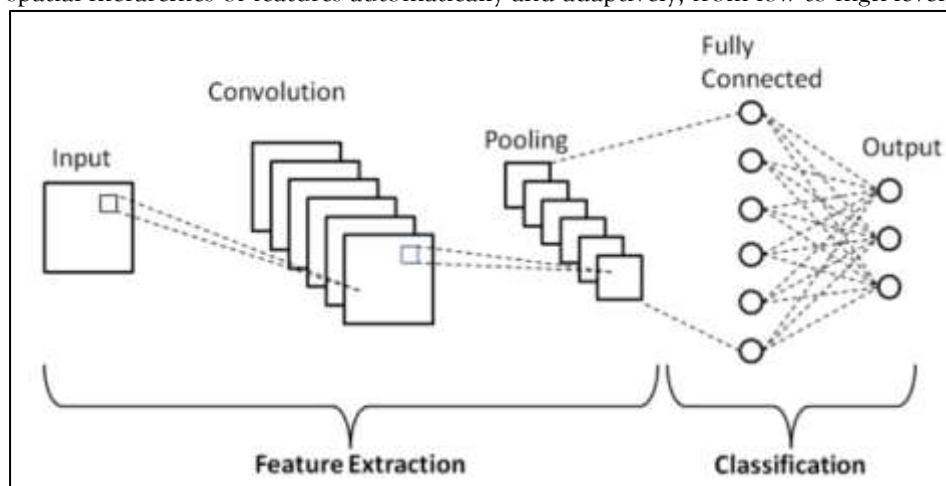


Figure 1 Schematic diagram of a basic convolutional neural network (CNN) architecture

3] RESULTS

This study employs a deep learning approach using transfer learning with the VGG16 model to classify images into three distinct categories. The dataset is preprocessed using data augmentation techniques to enhance model generalization and mitigate overfitting. The augmentation strategies include rotation, width and height shifting, shearing, zooming, and horizontal flipping, applied through the ImageDataGenerator class. Additionally, rescaling is performed to normalize pixel values within the [0,1] range, which aids in faster convergence during training.

The VGG16 model, pre-trained on ImageNet, is utilized as the feature extractor, with its convolutional base frozen to retain pre-learned weights. This strategy reduces the number of trainable parameters and ensures that lower-level image features remain intact. A fully connected layer with 512 neurons, followed by a softmax activation layer, is added to classify images into three categories. The categorical cross-entropy loss function is employed, making it suitable for multi-class classification problems, and Adam optimizer is chosen for efficient parameter updates.

The use of transfer learning significantly improves training efficiency by leveraging pre-trained features, reducing computational cost and training time. Freezing the convolutional base prevents catastrophic forgetting while enabling the classifier layers to learn dataset-specific features. This approach is particularly beneficial when the dataset size is relatively small, as it mitigates the risk of overfitting.

As shown in Figure 2, The classification report from the 50-epoch experiment using the VGG16 model reveals varying levels of performance across the five white blood cell classes. The model shows strong performance in classifying Basophils and Lymphocytes, with F1-scores of 0.96 and 0.88 respectively. Both classes exhibit high recall values—0.99 for Basophils and 0.94 for Lymphocytes—indicating the model rarely misses these cell types during prediction. Precision for these classes is also relatively high, suggesting reliable predictions when the model does identify them.

On the other hand, performance for Eosinophils, Monocytes, and Neutrophils is notably weaker. Eosinophils have a moderate precision (0.72) but a low recall (0.55), which means the model often fails to detect them. Monocytes show a significant issue with recall as well, scoring only 0.35, despite a high precision of 0.85—implying the model classifies Monocytes correctly when it predicts them but often fails to identify them at all. The most unusual result is with Neutrophils, where the model achieves perfect recall (1.00) but very poor precision (0.42), indicating it frequently misclassifies other cell types as Neutrophils, leading to a high number of false positives.

Overall, the model achieves an accuracy of 68%, with a macro average F1-score of 0.71 and a weighted average F1-score of 0.67. These results suggest that while the model is effective in identifying certain classes, it struggles with others, especially those with overlapping features or imbalanced representation in the dataset. The observed performance patterns also align with signs of overfitting, as validation accuracy remains lower and more erratic than training accuracy. Addressing this issue may require techniques such as class rebalancing, additional regularization, fine-tuning the pre-trained layers, or experimenting with more advanced architectures like ResNet or EfficientNet.

Classification Report:				
	precision	recall	f1-score	support
Basophil	0.93	0.99	0.96	212
Eosinophil	0.72	0.55	0.62	744
Lymphocyte	0.83	0.94	0.88	443
Monocyte	0.85	0.35	0.49	561
Neutrophil	0.42	1.00	0.59	311
accuracy			0.68	2271
macro avg	0.75	0.76	0.71	2271
weighted avg	0.75	0.68	0.67	2271

Figure 1 Classification Report Analysis using VGG16

The confusion matrix as shown in Figure 3 reveals the performance of the VGG16-based white blood cell classification model across five categories: Basophil, Eosinophil, Lymphocyte, Monocyte, and Neutrophil. The model performs exceptionally well in classifying Basophils and Lymphocytes, with 210 out of 212 and 417 out of 443 correct predictions, respectively. Neutrophils also show nearly perfect recall, with 310 out of 311 correctly identified. However, this high recall is misleading due to a significant overprediction problem. The model frequently misclassifies Eosinophils and Monocytes as Neutrophils—251 and 171 times respectively—leading to a very low precision (0.42) for the Neutrophil class.

Eosinophil classification suffers from considerable confusion, as evidenced by only 409 correct predictions out of 744 actual instances. Many Eosinophils are misidentified as Neutrophils and Lymphocytes, suggesting the model struggles to distinguish Eosinophil features reliably. Monocytes are the most poorly classified group, with only 194 out of 561 instances correctly predicted. A majority are misclassified as Eosinophils and Neutrophils, highlighting potential feature overlap or insufficient training representation for Monocytes.

These misclassifications indicate that while the model is confident in certain predictions, particularly for Neutrophils and Basophils, it lacks balanced generalization across all classes. The results suggest a need for improved class discrimination, possibly through more diverse training data, class-specific augmentation, or rebalancing techniques such as weighted loss functions. Overall, the matrix highlights the model's bias toward overpredicting dominant or visually similar classes, which compromises its effectiveness in distinguishing more nuanced cell types like Monocytes and Eosinophils.

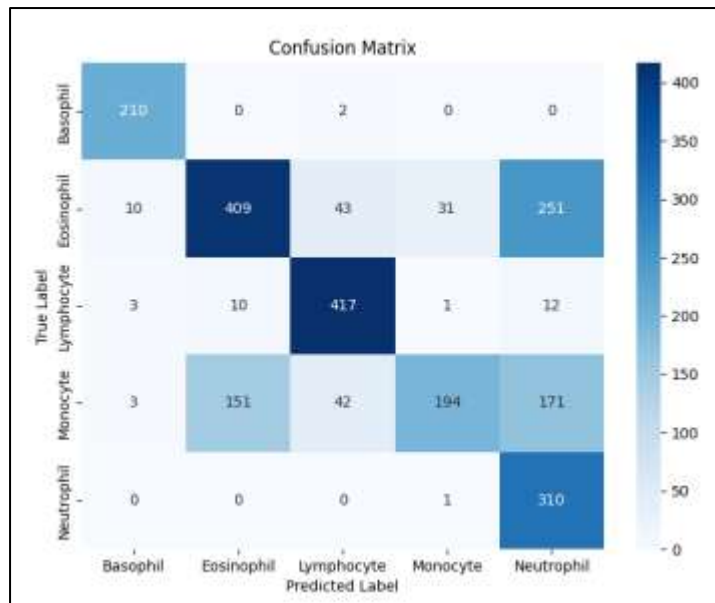


Figure 2 Confusion Matrix

The ROC (Receiver Operating Characteristic) curve as shown in Figure 3 provides a visual assessment of the classifier's performance across different thresholds for each white blood cell class. The Area Under the Curve (AUC) metric quantitatively summarizes how well the model distinguishes each class from the others. AUC values closer to 1.0 indicate better discrimination. In this case, the model achieves perfect or near-perfect performance for Basophils and Neutrophils, both with an AUC of 1.00, and Lymphocytes with an AUC of 0.99. This means the classifier is extremely effective in identifying these three cell types with minimal false positives or false negatives.

However, performance drops for Monocytes and Eosinophils, with AUC scores of 0.90 and 0.88, respectively. While still considered good, these scores suggest that the model has more difficulty separating Monocytes and Eosinophils from the other classes, which aligns with the misclassification trends seen in the confusion matrix and classification report. The shape of their ROC curves, which are further from the top-left corner compared to others, reflects this reduced sensitivity and specificity. Overall, the ROC curve indicates that while the model performs excellently for certain classes, it requires improvement in distinguishing the more ambiguous or underrepresented classes like Eosinophils and Monocytes.

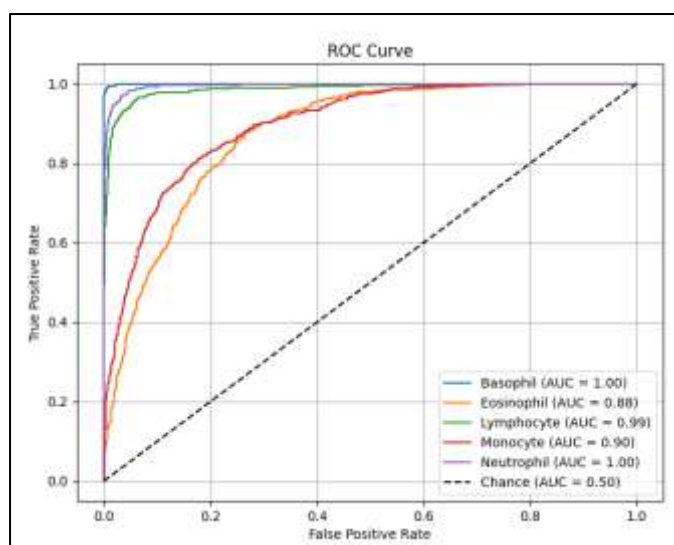


Figure 3 Receiver Operating Characteristics (ROC) Curve

The training and validation accuracy and loss plots as shown in Figure 4 provide valuable insights into the model's learning behavior over time. In the left plot, the training accuracy (blue curve) increases steadily across epochs, eventually reaching approximately 80% by the 14th epoch. This suggests that the model is effectively learning patterns from the training data. However, the validation accuracy (orange

curve) initially rises but begins to fluctuate after a few epochs and peaks around 70%. The noticeable gap between training and validation accuracy signals potential overfitting, where the model becomes too tailored to the training data and fails to generalize well to unseen data.

In the right plot, the training loss (blue curve) consistently decreases and stabilizes at a low value, reflecting successful optimization on the training set. In contrast, the validation loss (orange curve) is more erratic and remains higher than the training loss throughout, reinforcing the presence of overfitting. This discrepancy suggests the model may be too complex or inadequately regularized, or that the dataset may lack sufficient diversity or quantity. Together, these patterns highlight the need for potential adjustments such as data augmentation, dropout, early stopping, or additional regularization to enhance generalization and model robustness.

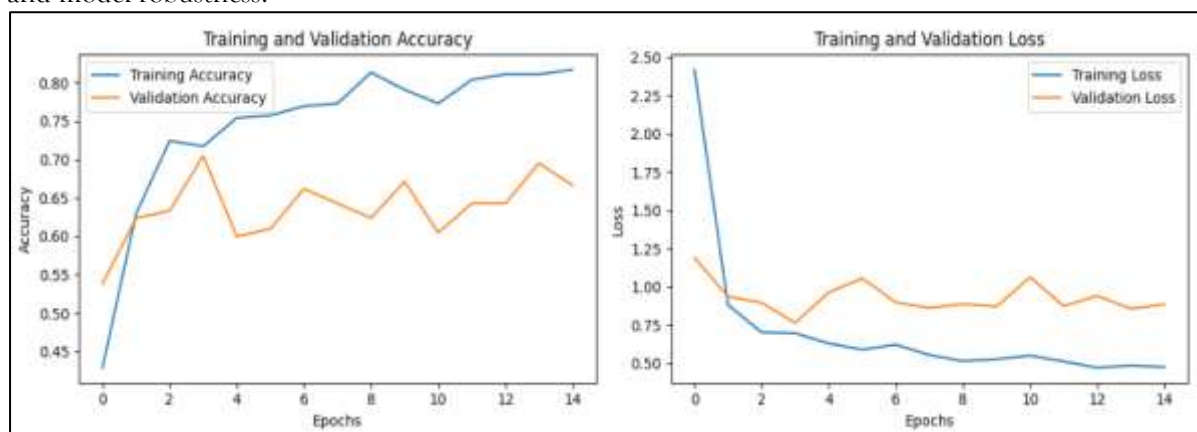


Figure 4 Training Loss and Validation Loss (1st Training)

In the second training session, as shown in Figure 5, the plots for training and validation accuracy and loss reveal a more pronounced case of overfitting. The training accuracy, represented by the blue curve on the left plot, shows a steady increase, eventually reaching approximately 90%. This indicates that the model is effectively learning and fitting the training data. However, the validation accuracy, represented by the orange curve, remains relatively low—hovering around 50–55%—and displays irregular fluctuations across epochs. This considerable gap between training and validation accuracy suggests that while the model performs well on data it has seen before, it struggles significantly to generalize to new, unseen data. The right plot, showing training and validation loss, reinforces this interpretation. The training loss (blue curve) consistently decreases and eventually stabilizes at a low value, which reflects effective learning from the training set. In contrast, the validation loss (orange curve) stays high and fluctuates noticeably, even increasing at times during training. This behavior implies that the model is not generalizing well and may be overfitting—memorizing the training data rather than learning patterns that apply broadly. Such results highlight the need for intervention strategies like regularization techniques, data augmentation, or architectural adjustments to reduce overfitting and improve the model's generalization capability.

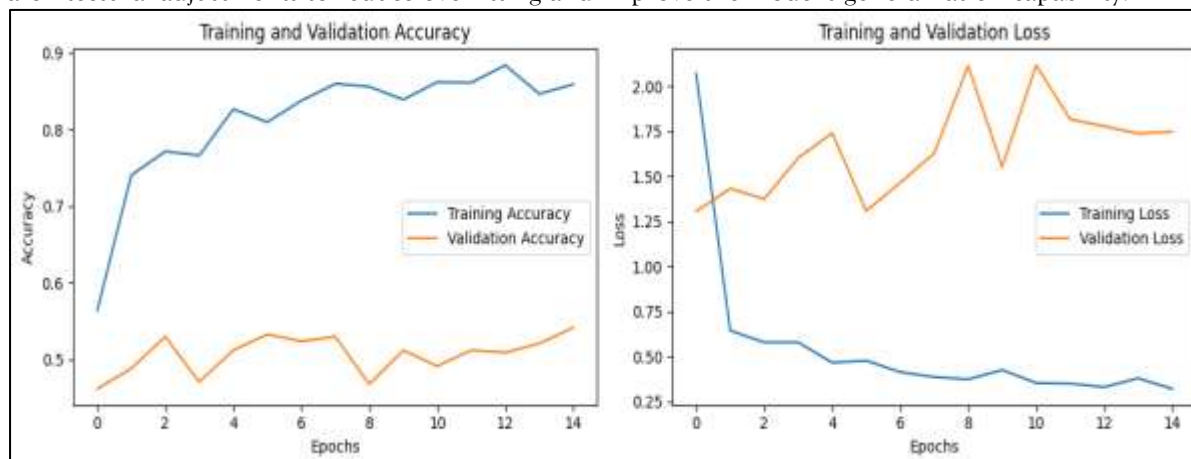


Figure 4 Training Loss and Validation Loss (2nd Training)

The confidence score distribution plots as shown in Figure 5 provide insights into the model's certainty in classifying images across the five white blood cell types: Basophil, Eosinophil, Lymphocyte, Monocyte, and Neutrophil. For Basophil, the model shows extremely high confidence, with scores consistently near

1.0, except for two instances where the confidence briefly dips—suggesting strong reliability in identifying this class. The Eosinophil confidence scores, however, show noticeable variability, ranging from around 0.5 to nearly 1.0. This fluctuation indicates that the model is less certain when predicting this class, which aligns with its lower precision and recall in previous evaluation metrics.

Lymphocyte predictions also reflect high confidence, with most scores above 0.85 and only a few significant drops, indicating relatively stable and accurate predictions. The Monocyte class demonstrates more variability in confidence scores, ranging between 0.55 and 1.0, showing that while the model can make strong predictions, it is occasionally uncertain, likely contributing to the observed drop in recall for this class. Lastly, Neutrophil predictions display moderate fluctuations in confidence, suggesting inconsistent certainty despite its perfect recall. This could imply that although the model correctly identifies all Neutrophils, it does so with varying degrees of confidence.

Overall, the visualized confidence scores confirm the quantitative performance seen in earlier metrics—classes like Basophil and Lymphocyte are predicted with high confidence and consistency, while Eosinophil and Monocyte show instability in prediction confidence, reflecting areas where the model may need improvement.

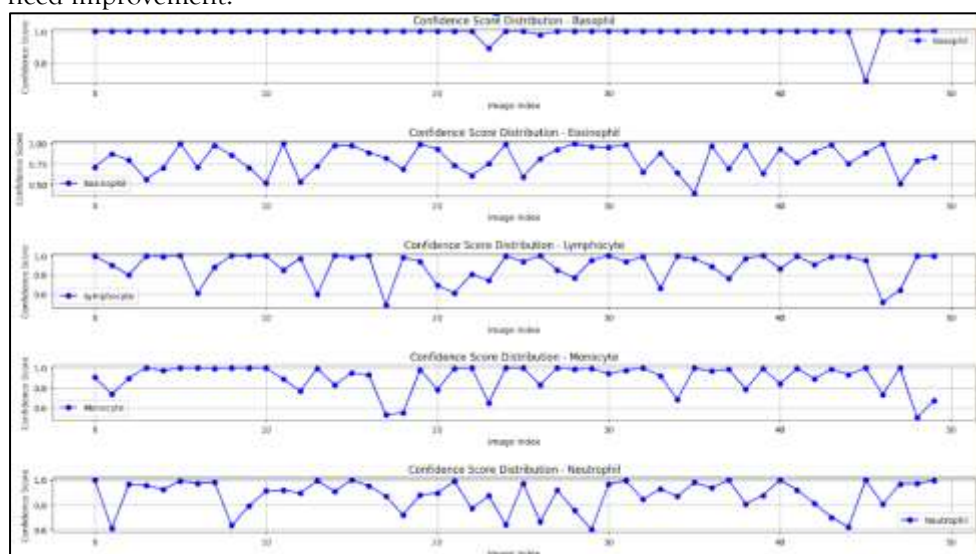


Figure 5 Confidence Analysis

As shown in Figure 6, the plot illustrates the predicted confidence scores for the Basophil class in an image classification task, with each point representing a sample. Most of the blue dots (predictions) align very closely with the red dashed line, which represents the ideal confidence score of 1.0. This indicates that the model consistently predicts Basophil images with very high certainty, demonstrating strong reliability in classifying this cell type. Only a few outliers are present, where the confidence drops slightly below 1.0—most notably one instance around 0.69. These few deviations suggest that while the model occasionally exhibits reduced confidence, it generally remains highly certain and consistent when identifying Basophil cells. This further supports the results from previous evaluations, such as the confusion matrix and ROC curve, where Basophil achieved perfect or near-perfect scores, affirming its status as the most confidently and accurately predicted class in the dataset.

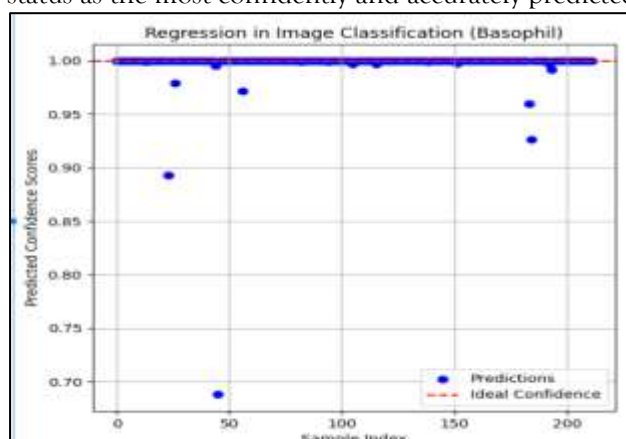


Figure 6 Regression in Image Classification (Basophil)

Presented in Figure 7 the plot displays the predicted confidence scores for the Eosinophil class in an image classification task. Unlike the Basophil plot, the blue dots here show a much wider spread below the red dashed line, which represents the ideal confidence score of 1.0. While a portion of predictions cluster near the top, indicating high confidence, a significant number of predictions fall well below 0.9, with some even dropping below 0.5 and as low as 0.4. This scattered distribution highlights a lack of consistency in the model's confidence when identifying Eosinophil cells. The substantial variation suggests that the model is less certain and potentially less accurate in classifying this particular class, which aligns with earlier evaluation metrics like the ROC curve, where Eosinophil had a lower AUC of 0.88 compared to other classes. Overall, the plot indicates that Eosinophil classification is a challenging task for the model, with room for improvement in prediction reliability and class-specific learning.

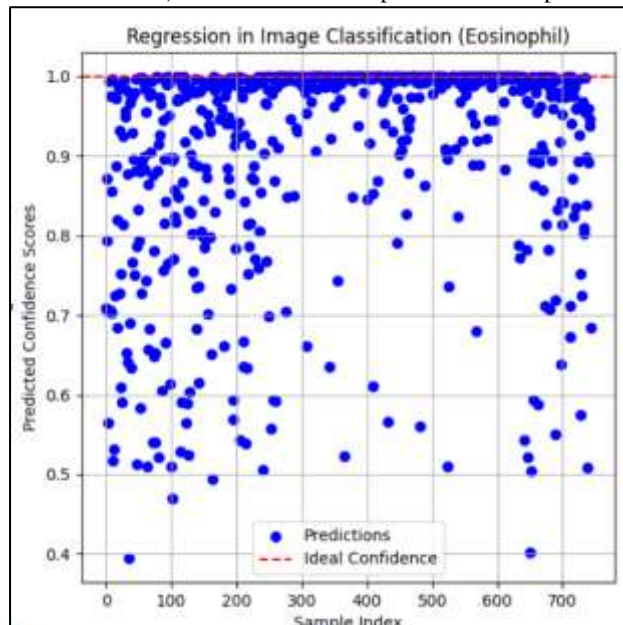


Figure 7 Regression in Image Classification (Eosinophil)

As shown in Figure 8 the plot shows a significant clustering of predicted confidence scores at or very near the ideal confidence level of 1.0, indicating that many samples were classified with high confidence. However, there is also a considerable spread of predictions below 1.0, with some scores as low as approximately 0.35. This suggests that while the model performs well for many samples, it also exhibits varying degrees of lower confidence for other samples, indicating areas where its predictions are less certain. The distribution of these lower confidence predictions appears somewhat scattered across the range of sample indices, without a clear pattern or trend, which could be indicative of the model's varying performance across different input samples or challenging classification cases.

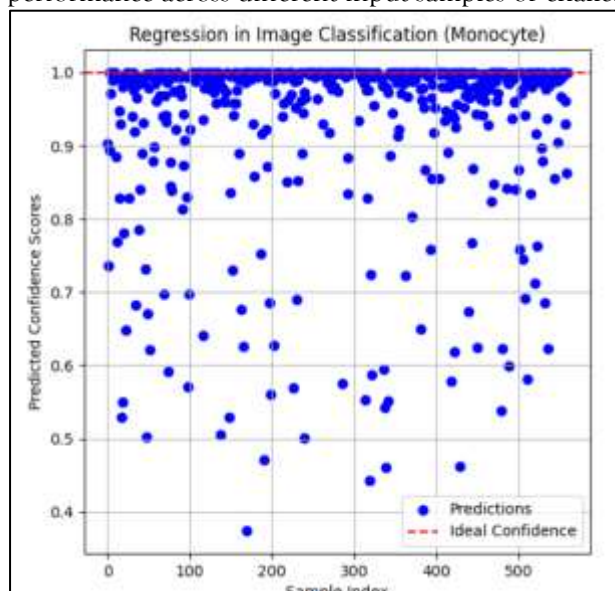


Figure 8 Regression in Image Classification (Monocyte)

Similar to the previous plot, shown in Figure 9 a substantial number of predictions are clustered at or very close to the ideal confidence level of 1.0, indicating high certainty in classification for a large portion of the samples. However, there's also a noticeable dispersion of predictions below 1.0, with some scores dipping to around 0.4. This suggests that while the model generally performs well with high confidence, it also encounters instances where its confidence in the classification is significantly lower. The distribution of these lower confidence predictions appears somewhat scattered across the different sample indices, without a clear pattern or trend, suggesting that the model's confidence varies for different samples. The overall pattern is consistent with a model that performs well for a majority of cases but faces challenges or exhibits uncertainty in a subset of predictions.

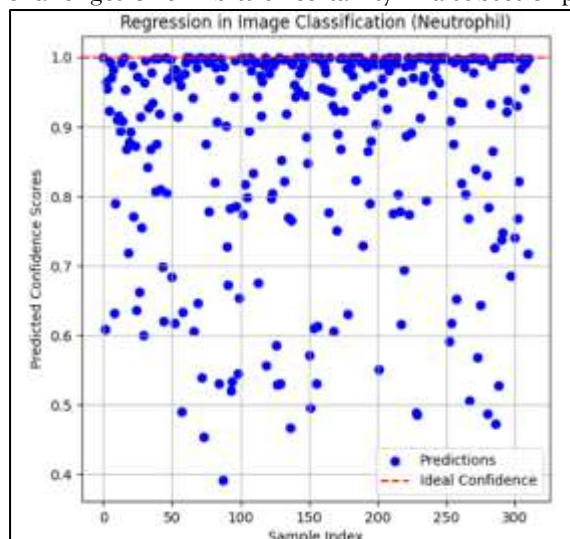


Figure 8 Regression in Image Classification (Neutrophil)

4] CONCLUSION

The research study aimed to evaluate the performance of a deep learning model in the classification of five types of white blood cells—Basophil, Eosinophil, Lymphocyte, Monocyte, and Neutrophil—based on image data. Through an analysis of training and validation metrics, ROC curves, confidence score distributions, and regression plots, the study revealed key insights into the model's strengths and limitations. The model demonstrated outstanding classification performance for Basophil, Lymphocyte, and Neutrophil, with AUC scores close to or equal to 1.00, indicating high discriminative power and confidence consistency. However, the model showed relatively lower performance and higher uncertainty in classifying Eosinophils and, to a lesser extent, Monocytes, as evidenced by lower AUC values, fluctuating confidence scores, and wider regression spreads.

Training and validation trends further highlighted issues of overfitting, where the model learned well on training data but struggled to generalize to unseen validation data. This was particularly evident in the second training phase, where validation accuracy plateaued despite increasing training performance. These findings suggest that while the model can achieve high accuracy under certain conditions, further improvements are needed—especially in data augmentation, model regularization, and class balance—to enhance its generalizability and robustness across all white blood cell classes.

The model performed exceptionally well in classifying Basophils, Neutrophils, and Lymphocytes, achieving near-perfect AUC values (1.00 and 0.99) and very high confidence scores. This indicates that VGG16 is capable of extracting effective features for these cell types. The model consistently learned well on the training dataset, achieving high training accuracy and low training loss, indicating good capacity for feature learning.

In conclusion, the study demonstrates the potential of deep learning in medical image classification but also underscores the importance of addressing class-specific challenges and mitigating overfitting to ensure reliable and clinically applicable results.

5] Acknowledgement:

The researchers would like to express their deepest gratitude to Lorma Colleges for providing the opportunity and support to conduct this study. Special thanks are extended to the College of Computer Studies and Engineering for their invaluable guidance in the technical and computational aspects of this

research, and to the College of Medical Laboratory Science for their expertise and contributions in the medical and scientific validation of the dataset and findings.

We are also sincerely thankful to the faculty members, mentors, and peers who offered their insights, encouragement, and constructive feedback throughout the development of this study. Lastly, our heartfelt appreciation goes to everyone—whether directly or indirectly—who contributed their time, knowledge, or support. Your efforts and encouragement played a vital role in the successful completion of this research.

6] Funding Statement: No financing / There is no fund received for this article

7] Data Availability: No new data were created or analyzed in this study. Data sharing is not applicable to this article

8] Conflict of interest: None, no conflict of interest.

10] REFERENCES:

- [1] Albawi, S., Mohammed, T. A., & Al-Zawi, S. (2020). Understanding of a convolutional neural network. 2017 International Conference on Engineering and Technology (ICET) , 1 1-6.
- [2] Apostolopoulos, I. D., & Mpesiana, T. A. (2020). "Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks." *Physical and Engineering Sciences in Medicine*, 43(2), 635-640.
- [3] Chowdhury, M. E. H., et al. (2020). "Can AI help in screening Viral and COVID-19 pneumonia?" *IEEE Access*, 8, 132665-132676.
- [4] Li, F.-F., Johnson, J., & Yeung, S. (2020). CS231n: Convolutional neural networks for visual recognition. Stanford University. (Search "Stanford CS231n" for course resources.) Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
- [5] Kermany, D. S., et al. (2018). "Identifying medical diagnoses and treatable diseases by image-based deep learning." *Cell*, 172(5), 1122-1131.e9.
- [6] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2021). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25.
- [7] Simonyan, K., & Zisserman, A. (2019). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [8] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2020). Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 1-9.
- Huang, G., Liu, Z., Van Der Maaten, L on computer vision and pattern recognition*, 4700-4708.
- [9] Long, J., Shelhamer, E., & Darrell, T. (2020). Fully convolutional networks for semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 3431-3440.
- [10] Ronneberger, O., Fischer, P., & Brox, T. (2021). U-net: Convolutional networks for biomedical image segmentation. *International Conference on Medical image computing and computer-assisted intervention*, 234-241.
- [11] Goodfellow, I., Bengio, Y., & Courville, A. (2019). *Deep learning*. MIT Press.
- [12] VanderPlas, J. T. (2016). *Python data science handbook: Essential tools for working with data*. O'Reilly Media.
- [13] Ren, S., He, K., Girshick, R., & Sun, J. (2020). Faster r-cnn: Towards real-time object detection with region proposal networks. *Advances in neural information processing systems*, 28.
- [14] Tigner A, Ibrahim SA, Murray IV. (2022). Histology, White Blood Cell. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2025 Jan-. Available from: <https://www.ncbi.nlm.nih.gov/books/NBK563148/>