2023 4th International Conference for Emerging Technology (INCET)
Belgaum, India. May 26-28, 2023

A Deep Learning-based Convolutional Neural Networks Model for White Blood Cell Classification

Archana Saini[1]
[1]*Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura 140401,* Punjab, India
saini.archana@chitkara.edu.in

Kalpna Guleria[1,*]
[1]*Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura 140401,* Punjab, India
guleria.kalpna@gmail.com

Shagun Sharma[1]
[1]*Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura 140401,* Punjab, India
shagunsharma7098@gmail.com

2023 4th International Conference for Emerging Technology (INCET) | 979-8-3503-3575-0/23/$31.00 ©2023 IEEE | DOI: 10.1109/INCET57972.2023.10170666

***Abstract*— White blood cells, or leukocytes, are indispensable for the optimal functioning of the immune system. They play a critical role in protecting the body against infections, diseases, and other foreign invaders by identifying and fighting harmful bacteria and pathogens that can cause illness. Additionally, they contribute to the elimination of dead and damaged cells from the body and facilitate tissue healing and repair processes. The absence of white blood cells would render the body defenceless against infections and diseases, exposing it to a variety of harmful pathogens. This could result in significant health issues and potentially even lead to death in severe instances. White blood cell classification is an important task in medical diagnosis and treatment because healthcare professionals diagnose and treat a variety of immune system-related diseases and conditions, including autoimmune disorders, infections, and cancers by identifying the structure, characteristics and functions of white blood cells. In this work, a convolutional neural network (CNN) model has been trained to classify white blood cells. The proposed model has achieved an accuracy of 88.78%, which has been identified as the highest among all the models implemented by various authors in the literature review. This implies that the proposed model has correctly classified white blood cells in almost 9 out of 10 cases. Moreover, the error rate of the model is only 0.108967 which indicates that the model is very reliable and consistent in its predictions. Additionally, this work shows the promising result for white blood cell classification using deep learning techniques. Furthermore, with improvements and refinements in the future, it can be possible to achieve higher levels of accuracy and precision, which could have a significant impact on medical diagnosis and treatment.**

Keywords*— deep learning, white blood cell, leukocytes, convolutional neural networks, healthcare

I. INTRODUCTION

Platelets, white blood cells, and red blood cells are the three main subtypes of human blood cells. Red blood cells are crucial for carrying carbon dioxide away from tissues and going to deliver oxygen from the heart. White blood cells (WBCs), which serve as the body's main line of defence against infections and diseases, have significant roles in the immune system as well. Our body is protected from millions of disease-causing bacteria, parasites, and viruses by the immune system, which is a complex network of cells, tissues, and organs [1]. White blood cells, or leukocytes, are the most important part of our immune system and are divided into five main subtypes: neutrophils (50–70%), lymphocytes (25–30%), monocytes (3–9%), eosinophils (0–5%), and basophils (0–1%). The presence of an infection with bacteria may be indicated by an elevated monocyte and eosinophil count [2]. A rise in lymphocyte counts may be an early sign of AIDS (Acquired Immune Deficiency Syndrome). While an elevated neutrophil count may be a sign of cancer [3].

The bone marrow and lymphoid tissues produce white blood cells, which are a component of the immune system. They safeguard our bodies from bacterial, viral, and fungal infections [4]. A white blood cell surplus or deficiency can result in a number of diseases. Blood tests are used to diagnose these diseases. Leukopenia is a condition in which the number of WBC is lower than the reference value. In cases of hormonal causes, metabolic disorders, hemolysis, and bleeding, there is an increase in neutrophils in the blood [5]. Hematology professionals have generally categorized and counted WBCs manually under a magnifying glass. But as the process is so complicated, it might take a while and be prone to mistakes [6]. WBCs develop following body requirements, but in the case of leukaemia, they are produced abnormally and degenerate [7]. The analysis and further processing become extremely difficult due to the variability in shape and texture, even though they are frequently identified by their dark purple-like appearance. Cells with a wide range of differences fall under the category of leukocytes [8]. Even though their size and shape can distinguish them, the fact that WBCs are surrounded by other blood cells like red blood cells and platelets makes it difficult to identify them. The automatization's ability to identify and categorize these images is subject to biases or misjudgments [9]. These will have an immediate effect on the diagnosis, drive up the cost of care, and have a harmful effect on patient recovery and survival [10][11]. As opposed to conventional machine learning algorithms, the convolution neural network algorithm we used for this study is based on deep learning and is capable of doing aside from the difficult process of segmentation and extraction characteristics.

II. RELATED WORK

Authors in [12] proposed a technique that automatically categorizes WBCs using CNN. The network architectures ResNet50, Inception V3, VGG 16, VGG 19, and Xception were used by the researchers. The accuracy achieved by the suggested method was 88.5%.The use of SVM and Deep learning to classify abnormalities based on images of deformed RBCs was proposed in [13]. To forecast which will perform the best on RBCs to achieve the highest level of classification accuracy. This study made the claim that SVM classifiers outperformed deep

learning classifiers because SVM can classify the cells in all conditions regardless of whether the dataset is small or large, whereas deep learning only excels on large datasets.In [14] author proposes a machine-learning technique for automatically classifying and counting three different blood cell types. The presented algorithm counts platelets, red blood cells, and white blood cells precisely.

III. Material & Methods

This section explains the dataset and offers a technique for categorizing white blood cells automatically.

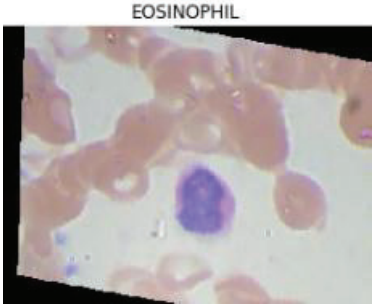
A. Dataset

The knowledge required to make predictions about the white blood cells was made available by the open Kaggle repository. There are several attributes in the dataset. There are 4 categories: Lymphocyte, Monocyte, Eosinophil, and Neutrophil [15].

TABLE I. Different categories of White Blood Cells

Type	% In blood	Nucleus	Cytoplasm
Eosinophils	3%	It has 2 lobes that each stain purple, and is difficult to be seen	It's pale pink-tan but contains large purple/blue-black granules which obscure the cell nucleus
Lymphocytes	30%	It's large, round, or oval, and is dark staining	It is not present or very small, is pale blue in color, and occasionally has purple-reddish granules
Monocytes	6%	It's singular and is kidney-shaped (convoluted shape), bean-shaped or horseshoe-shaped with a deep indentation	It stains a blue-gray color and is "ground glass" with tiny granules, Vacuoles are sometimes present in it
Neutrophils	60%	It's divided into 2 to 5 segments and stains dark purple (multi-lobed)	It's pale pink to tan with pink-purple granules

Table. I illustrates the different categories of white blood cells along with their percentage in blood and other attributes showing the nucleus and cytoplasm.



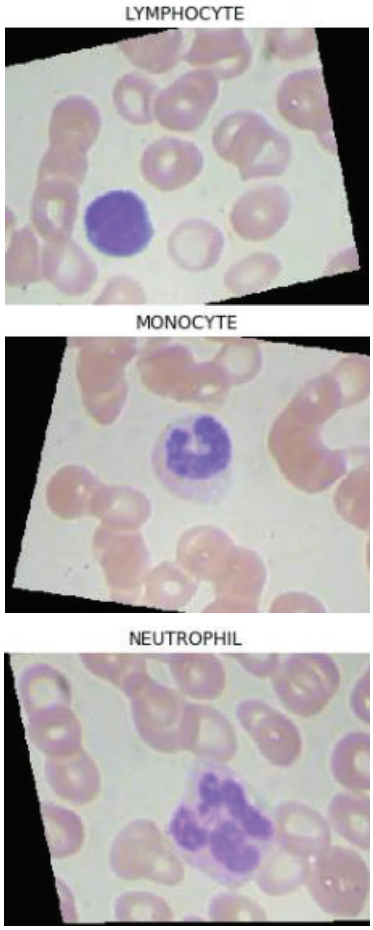


Fig. 1. White blood cell classification

B. Methodology

CNN model with each layer. The basic CNN architecture with some typical layers. Input layer takes in the input image and passes it on to the next layer [16]. The input layer dimensions depend on the size of the image being processed. The convolutional layer is where the actual computation happens. It applies a set of filters to the input image to detect features such as edges, corners, and other patterns. Each filter produces a feature map, which is a transformed version of the input image. The number of filters in this layer depends on the complexity of the problem being solved [17]. ReLU layer applies a non-linear activation function to the feature maps generated by the convolutional layer. The ReLU function zeros out all negative values and leaves positive values unchanged. This introduces non-linearity into the model and helps in better feature extraction [18]. The pooling layer is used to downsample the feature maps generated by the convolutional layer. It reduces the spatial dimensions of the feature maps and makes the model more robust to variations in the input image.The dropout layer is used to prevent overfitting in the model. It randomly drops out some of the neurons in the network during training, which forces the model to learn more robust features. The flatten layer is used to convert the output of the convolutional and pooling layers into a one-dimensional vector. This vector can then be passed on to the

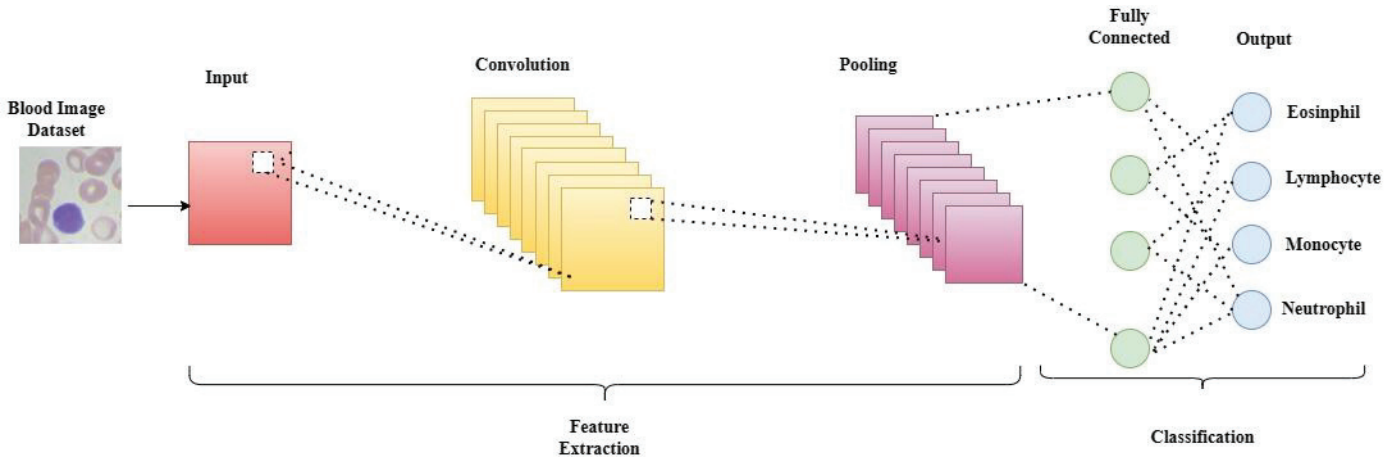


Fig. 2. Methodology of White blood cell classification

fully connected layers. The fully connected layer is a traditional neural network layer that takes in the flattened vector from the previous layer and applies weights to it to produce an output. This layer is used to classify the input image into one of the target classes.

Output Layer: This layer produces the final output of the model. In the case of classification, it produces a probability distribution over the target classes.These are the main layers in a typical CNN model. Of course, there can be many variations and modifications to this architecture depending on the specific problem being solved.

IV. Results & Discussion

This section discusses the performance outcomes for the suggested work.

Two parameters are used to show the performance of the model, one is accuracy and the other is error. When we are making changes in the learning rate the accuracy is increasing and the error is decreasing.

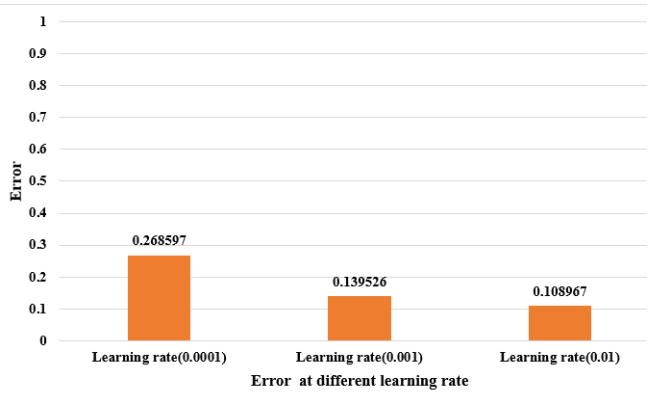


Fig. 3. Comparative error analysis with different learning rates

Figure 3 illustrates error analysis at various rates of learning. The highest error in Figure is 0.268597 at a learning rate of 0.0001, the second-highest error is 0.139526 at a learning rate of 0.001, and the lowest error is 0.108967 at a learning rate of 0.01.

TABLE II. Comparative Error Analysis with Different Learning Rates

Sr.No	No. of Epochs	Learning Rate	Error Rate
1	4	0.0001	0.268597
2	4	0.001	0.139526
3	4	0.01	0.108967

Table. II demonstrates the error analysis at different learning rates. In the figure, at a learning rate of 0.0001, the highest error occurred 0.268597, at a learning rate of 0.001 the error is 0.139526 and the lowest error of 0.108967 occurred at a learning rate of 0.01.

Fig. 4. Illustrates the confusion matrix at a learning rate of 0.0001. From this matrix, we get the results that eosinophil correctly classified accuracy is 63.72%, lymphocyte accuracy is 63.72%, monocyte accuracy is 88.22%, and neutrophil accuracy is 77.72%. The overall accuracy is 73.14%.

Fig. 5. Illustrates the confusion matrix at a learning rate of 0.001. From this matrix, we get the results that eosinophil correctly classified accuracy is 78.49%, lymphocyte accuracy is 98.54%, monocyte accuracy is 75.16%, and neutrophil accuracy is 91.98%. The overall accuracy is 86.04%.

Fig. 6. Illustrates the confusion matrix at a learning rate of 0.01. From this matrix, we get the results that eosinophil correctly classified accuracy is 85.71%, lymphocyte accuracy is 100%, monocyte accuracy is 75.16%, and neutrophil accuracy is 94.23%. The overall accuracy is 88.78%.

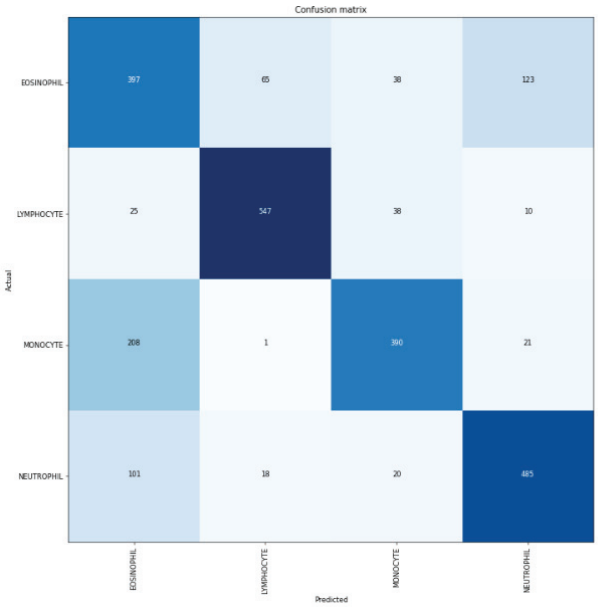


Fig. 4. Confusion matrix at a learning rate (0.0001)

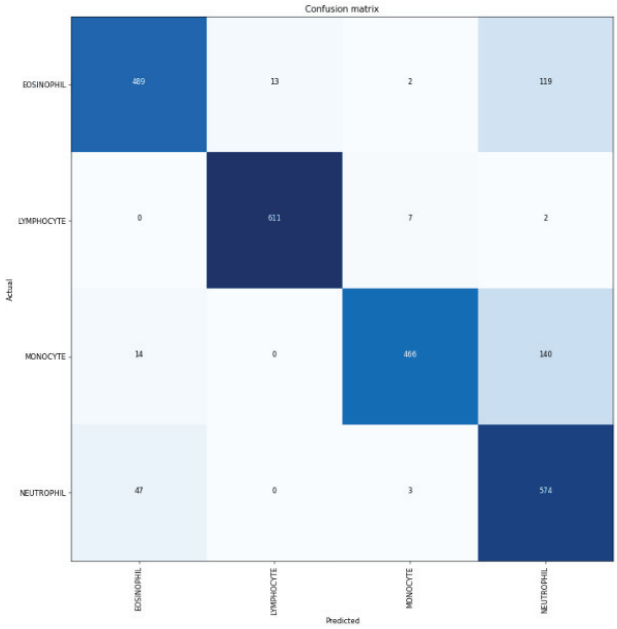


Fig. 5. Confusion matrix at learning rate 0.001

TABLE III. CATEGORY-WISE ACCURACY ANALYSIS WITH DIFFERENT LEARNING RATES

Category/Learning rate	0.0001	0.001	0.01
Eosinophil	63.72	78.49	85.71
Lymphocyte	88.22	98.54	100
Monocyte	62.90	75.16	75.16
Neutrophil	77.72	91.98	94.23

Table. III demonstrates the comparative analysis on the basis of accuracy at different learning rates. In the table, there are 3 learning rates (0.0001,0.001,0.01) that show the accuracy comparisons of different types of white blood cells. The results of each class namely, Lymphocyte, Monocyte, Eosinophil, and neutrophil have been shown at different learning rates. The results show that for the Lymphocyte class, the learning rate of 0.01 shows the highest accuracy of 100% whereas for Monocyte, Eosinophil, and neutrophil the best accuracy has been achieved as (75.16% at 0.01 and 0.001), (85.71% at 0.01), and (94.23% at 0.01), respectively.

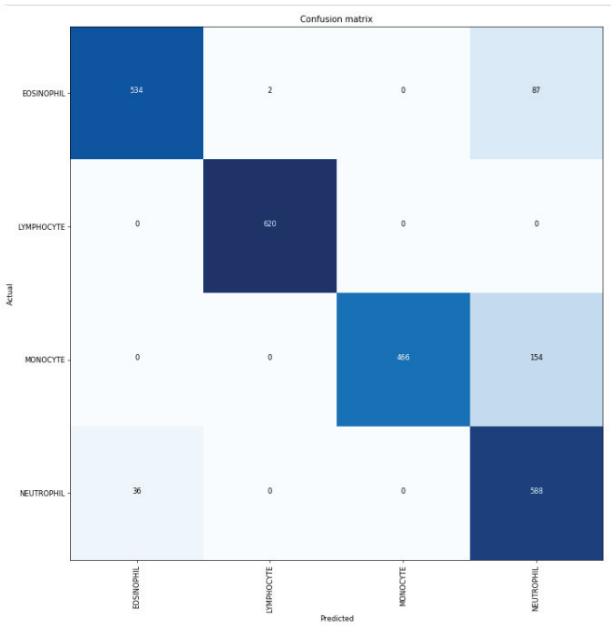


Fig. 6. Confusion matrix at learning rate 0.01

The comparative analysis based on accuracy at various learning rates is shown in Fig.7. The figure contains three learning rates (0.0001, 0.001, and 0.01) that compare the accuracy of various types of white blood cells.

Table. IV illustrates the overall accuracy of the model at different learning rates. The accuracy is 73.14%, 86.04%, and 88.78% at learning rates of 0.0001, 0.001, and 0.01 respectively. These results show that when the learning rate has been kept at 0.01, the model shows the best accuracy of 88.78%, however, with the decrease in learning rate the accuray also starts to decrease.

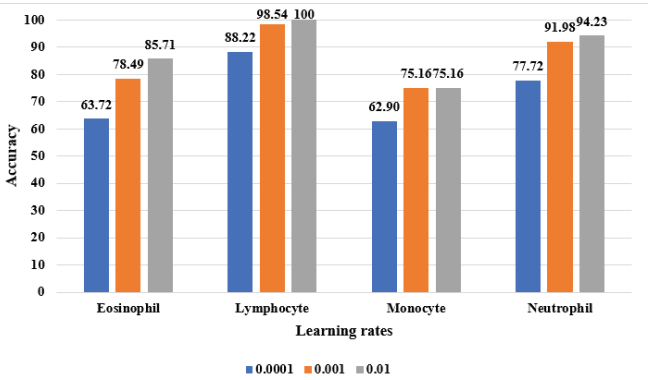


Fig. 7. Category-wise accuracy analysis with different learning rates

TABLE IV. ACCURACY ANALYSIS WITH DIFFERENT LEARNING RATES

Sr. No.	Learning Rate	Accuracy
1	0.0001	73.14%
2	0.001	86.04%
3	0.01	88.78%

Fig. 8. Illustrates the overall accuracy comparison of the model on the basis of different learning rates. With a 0.0001 learning rate, the accuracy is 73.14%, with 0.001 the accuracy is 86.04% and with 0.01 the learning rate is 88.78%.

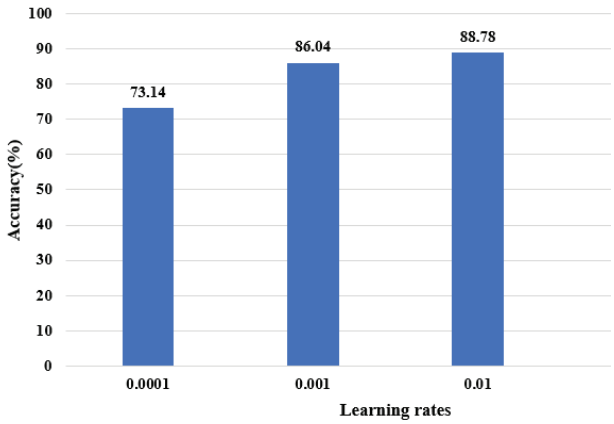


Fig. 8. Accuracy analysis with different learning rates

Additionally, it has been also observed that for using the model in practice, better results (accuracy) are required. The dataset can be increased and the model can be reimplemented for better performance to classify the white blood cell accurately.

V. CONCLUSION & FUTURE SCOPE

White blood cell play a vital role in defending the body against pathogens and also contributes in regulating the immune responses and the control of inflammation. They also participate in the processes of tissue repair and wound healing. Medical diagnosis and treatment heavily rely on accurate white blood cell classification. Recent advancements in deep learning techniques have enabled the automation of this process, resulting in high accuracy. In the proposed work, a convolutional neural network (CNN) model has been developed and trained to classify white blood cells into four categories: eosinophil, lymphocyte, monocyte, and neutrophil. To optimize the accuracy and minimize the error rate, the model's learning rate was adjusted using different values of learning rates as 0.0001, 0.001, and 0.01. Furthermore, by extensively training of the CNN model, it has been identified that the proposed model shows an accuracy of 88.78% and the lowest error rate of 0.108967 at a learning rate (0.01). Additionally, the model can be trained further with the increased dataset to achieve the highest accuracy rates.

REFERENCES

[1] M. Sharma, A. Bhave, and R. R. Janghel, "White blood cell classification using convolutional neural network," in Advances in Intelligent Systems and Computing, Singapore: Springer Singapore, 2019, pp. 135–143.

[2] M. J. Macawile, V. V. Quinones, A. Ballado, J. D. Cruz, and M. V. Caya, "White blood cell classification and counting using convolutional neural network," in 2018 3rd International Conference on Control and Robotics Engineering (ICCRE), 2018.

[3] H. Kutlu, E. Avci, and F. Özyurt, "White blood cells detection and classification based on regional convolutional neural networks," Med. Hypotheses, vol. 135, no. 109472, p. 109472, 2020.

[4] M. Jiang, L. Cheng, F. Qin, L. Du, and M. Zhang, "White blood cells classification with deep convolutional neural networks," Intern. J. Pattern Recognit. Artif. Intell., vol. 32, no. 09, p. 1857006, 2018.

[5] A. T. Sahlol, P. Kollmannsberger, and A. A. Ewees, "Efficient classification of white Blood Cell Leukemia with improved swarm optimization of deep features," Sci. Rep., vol. 10, no. 1, p. 2536, 2020.

[6] C. Cheuque, M. Querales, R. León, R. Salas, and R. Torres, "An efficient multi-level convolutional neural network approach for white blood cells classification," Diagnostics (Basel), vol. 12, no. 2, p. 248, 2022.

[7] A. M. Patil, M. D. Patil, and G. K. Birajdar, "White blood cells image classification using deep learning with canonical correlation analysis," IRBM, vol. 42, no. 5, pp. 378–389, 2021.

[8] S. Shamas, S. N. Panda, and I. Sharma, "Review on lung nodule segmentation-based lung cancer classification using machine learning approaches," in Artificial Intelligence on Medical Data, Singapore: Springer Nature Singapore, 2023, pp. 277–286.

[9] S. Sharma and K. Guleria, "Pneumonia detection from chest X-ray images using transfer learning," in 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), 2022.

[10] S. Sharma and K. Guleria, "A Deep Learning based model for the Detection of Pneumonia from Chest X-Ray Images using VGG-16 and Neural Networks," Procedia Comput. Sci., vol. 218, pp. 357–366, 2023.

[11] S. Gulati, K. Guleria, and N. Goyal, "Classification of Migraine Disease using Supervised Machine Learning," in 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Oct. 2022, pp. 1–7. doi: 10.1109/ICRITO56286.2022.9964524.

[12] W. Yu et al., "Automatic classification of leukocytes using deep neural network," in 2017 IEEE 12th International Conference on ASIC (ASICON), 2017.

[13] H. A. Aliyu, R. Sudirman, M. A. Abdul Razak, and M. A. Abd Wahab, "Red blood cell classification: Deep learning architecture versus support vector machine," in 2018 2nd International Conference on BioSignal Analysis, Processing and Systems (ICBAPS), 2018.

[14] M. M. Alam and M. T. Islam, "Machine learning approach of automatic identification and counting of blood cells," Healthc. Technol. Lett., vol. 6, no. 4, pp. 103–108, 2019.

[15] M. Nickparvar, "White blood cells dataset." Mar. 26, 2022. Accessed: Apr. 25, 2023. [Online]. Available: https://www.kaggle.com/datasets/masoudnickparvar/white-blood-cells-dataset.

[16] S. Sharma, K. Guleria, S. Kumar, and S. Tiwari, "Benign and Malignant Skin Lesion Detection from Melanoma Skin Cancer Images," in 2023 International Conference for Advancement in Technology (ICONAT), Jan. 2023, pp. 1–6. doi: 10.1109/ICONAT57137.2023.10080355.

[17] A. Kaur, K. Guleria and N. Kumar Trivedi, "A Deep Learning based Model for Rice Leaf Disease Detection," 2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), Noida, India, 2022, pp. 1-5, doi: 10.1109/ICRITO56286.2022.9964487

[18] R. Sharma, V. Kukreja and V. Kadyan, "Hispa Rice Disease Classification using Convolutional Neural Network," 2021 3rd International Conference on Signal Processing and Communication (ICPSC), Coimbatore, India, 2021, pp. 377-381, doi: 10.1109/ICSPC51351.2021.9451800.