EasyChair Preprint

№ 10561

A Deep Learning Model for Human Blood Cells Classification

M Pramodha, S Ansith, J V Bibal Benifa, Mohammed Al-Sarem, Hanumanthappa Jayappa Davanagere, A A Bini, Emmanuel Ndagijimana, Faisal Saeed, Md Belal Bin Heyat, Abdulrahman Alqaraf, Abdullah Y Muaad and Channabasava Chola

July 14, 2023

A Deep Learning model for Human Blood Cells Classification

Pramodha M[1, ☆ [0000-0002-4325-7135]] Ansith S [2, ☆ [0000-0002-8242-3239]] J V Bibal Benifa [3[0000-0002-8242-3239]] , Mohammed Al-Sarem [4,5[0000-0001-7172-8224]] , Hanumanthappa J [1[0000-0002-6031-6993]] , Bini A A [2[0000-0002-8242-3239]] , Emmanuel Ndagijimana[1[0000-0002ch-2822-1708]] , Faisal Saeed[5,[0000-0002-2822-1708]] , Md Belal Bin Heyat[7[0000-0001-5307-9582]], Abdulrahman Alqarafi [5[0000-0001-7172-8224]] , Abdullah Y Muaad [1,8,*[0000-0001-8304-9261]] and Channabasava Chola[1,3,*[0000-0002-7509-9354]]

[1] Department of Studies in Computer Science, University of Mysore, Manasagangothri, Mysore 570006, India
[2] Department of Electronics and Communication Engineering, Indian Institute of Information Technology Kottayam, Kerala, India
[3] Department of Computer Science and Engineering, Indian Institute of Information Technology Kottayam, Kerala, India
[4] Department of Computer Science, Saba'a Region University, Mareb, Yemen
[5] College of Computer Science and Engineering, Taibah University, Medina 42353, Saudi Arabia
[6]DAAI Research Group, Department of Computing and Data Science, School of Computing and Digital Technology, Birmingham City University, Birmingham B4 7XG, UK
[7] IoT Research Center, College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, Guangdong 518060, China
[8] IT department, Sana'a Community College, Sana'a 5695, Yemen
{abdullahmuaad9, channabasavac7}@gmail.com

Abstract.
Microscopic imaging is gaining focus in recent days, especially in the part of histopathological image analysis. Blood cells plays a critical role in assessment of health status of patients, especially given the rising frequency of infectious diseases. Automated blood analysis can aid in detecting early stages of diseases. In this work, we present study classification of Blood cell using different transfer learning approaches, MobileNetV2 based model designed for the accurate multi classification of blood cells. Deep Learning(DL) models require more time when training on big data sets, to overcome the computation complexity with light weight model MobileNetV2 is considered. In this paper we present the comparison among different transfer learning models such as VGG16, VGG19, Resnet50 with MobileNetV2. The performance evaluated with Accuracy, Precision, recall and F-Score. MobileNetV2 outperform all other model with accuracy of 97.89%. The proposed model has improved accuracy for classification blood cell for eight class.

Keywords: Blood Cell Classification, Microscopic Imaging, Convolutional Neural Network, Deep Learning, Transfer Learning.

1 Introduction

Medical imaging is the process of providing visual information of the human body to help the doctors and radiologists in an efficient diagnostic and treatment. There are certain modalities like CT, X-ray, MRI, microscopic blood smear images, PET, and ultrasound upon which experts and doctors can rely for diagnosing diseases and prescribing treatment [1]–[6]. Different computer-aided diagnosis (CAD) systems can automatically diagnose numerous hematic types, such as AIDS and blood cancer (Leukaemia), by analysing the microscopic images of blood cells with the help of DL and ML algorithms. WBCs (leukocytes), RBCs (erythrocytes), and platelets (thrombocytes) are major classifications of blood cell[7]–[9]. Among these, leukocytes are further divided into five subcategories: monocyte, lymphocyte, neutrophil, basophil, and eosinophil.

Classification and segmentation of WBCs can be done by conventional techniques using manual analysis of WBCs in blood smear images. This conventional method consumes more time and also a challenging task too. Sometimes errors can happen. Different ML and DL techniques have been deployed in the WBC classification over the past two decades[10]. In this proposed work we would like to use CNN architecture for the classification purpose. The use of DL models to classify WBC which can decrease the workload on haematologists and provide quick, efficient, and accurate results to assist medical experts in the diagnostic process at the same time DL can be used in different domino such as text, image and video[8], [11]–[13].

The main contributions of this work explore in the following:

- A deep learning model is proposed to classify blood cells to multiclass of blood cell.
- Compression of experiments for different model with latest work have been explored.

This paper is organized as we mention in this section. The introduction for related work is presented in section 2. In Section 3 the explanation of proposed models. The sections 4 and 5 discuss and explain the result and conclusion of this paper with future work are summarized in section 6.

2 Related work

There are many works has been done for blood classification. In this section we are going to explore some of them starting by authors in [14]. They proposed a system to descry WBCs. Their model detected WBCs from microscopic images using a simple relationship of R, G, B colours and some morphological operations. They applied SVM and CNN for the detection and they got good accuracy equal to 92.80%. In [15] the authors have used a convolutional neuronal network to detect eight groups of peripheral blood cells. They have designed CNN-based transfer learning model and they have got 96.20% in term of accuracy. The authors in [16] designed a new DL based on a two

stage which can detect and classify infected WBCs and RBCs. The result was 72% in term of accuracy. In [17] authors introduced a CNN method for classifying white blood cells into their corresponding five groups, which are monocyte, lymphocyte, neutrophil, basophil, and eosinophil. The validation set indicated an overall accuracy of 91%, while the training set recorded an accuracy of 99%. The proposed system's sensitivity and specificity were 91 and 97 percent, respectively. The authors in [18] proposed R-CNN technique to detect WBCs and RBCs. They got an average accuracy of 83.25% for RBCs while 99% for eosinophil and 66% for lymphocytes. The lowest testing accuracy was always above 66% for this method. In [19] The authors explored the identification of WBCs in three steps utilizing image processing and machine learning techniques (segmentation, feature extraction and classification). When the number of samples in a dataset is small, the accuracy obtained is rather low. As a result, a deep learning method is used instead of the traditional machine learning method. Capsule Networks (encoder-decoder) approaches obtained 96.86 % accuracy using LISC datasets and Deep Learning methods. When it comes to segmentation and classification of blood cells, deep learning methods outperform traditional machine learning methods. With small datasets, Caps Net has achieved a good accuracy rate. In [20], authors introduced the procedure of white blood cell identification, categorization, and quantitative analysis can be done both manually and automatically, according to the authors in [21]. The classification of WBCs manually takes a long time and is highly reliant on the hematologist's knowledge. CapsNet and other convolutional neural networks like AlexNet, VGG16, ResNet-18, and InceptionV3 are used in clinical practise to classify WBCs utilising an automated leukocyte analyzer. All IDB2, BCCD, and CellaVision from the Hospital Clinic of Barcelona are included in the samples, which are from a public dataset made available by Anna et al[22]. The accuracy rates are consistent with BloodCaps (99.3%), AlexNet (81.5%), VGG16 (97.8%), ResNet-18 (95.9%), and Incep-tionV3 (98.4%). Even with small datasets, BloodCaps, a proposed model capsule network for the categorization of human peripheral blood cells, gives precise classification. In [11] the authors investigated the automatic classification of white blood cells using rapid, accurate and cost-effective data enhancement approaches using DNNs as an alternative to the laboratory environment. Image processing methods and GAN image generation are the approaches used in data augmentation. Utilizing cutting-edge DNNs like VGG, ResNet, and DenseNet that have been trained on the CIFAR-100 datasets are used to examine the classification of leukocytes into their respective classes namely neutrophils, eosinophils, lymphocytes, monocytes, and basophils. On the CIFAR-100 dataset, they applied pre-trained weights. This approach uses the acquired images immediately, in contrast to other approaches that require complex image pre-processing and manual feature extraction prior to classification. It has been demonstrated through extensive testing that the suggested method can accurately classify white blood cells. With a validation accuracy of 98.8%, DenseNet-169, the best DNN model, is ranked first. The proposed approach performs better than existing approaches, which rely on manual feature engineering and sophisticated image processing. In [23] according to authors the accuracy of the measurement of White Blood Cells (WBC) in the blood is 95.89% for categorization and 97.40% for WBC detection. The ability to extract is a crucial sign of pathological conditions, and the suggested method also can detect WBC in raw microscopic

pictures. Computer vision-based methods for differential counting of WBC are becoming more prominent due to their benefits over conventional methods. However, many of these techniques aren't applicable to raw microscopic images with multiple WBC because they are made for single WBC images that have already been processed. In addition, the techniques fail to recognise WBC in images. This study provides a K-Means clustering-based image processing strategy for identifying and localising WBC in raw microscopic images, as well as a VGG-16 classifier to categorise those cells. The recommended method may remove dye particles, nucleated RBC, and overlapping cells from multiple-cell pictures in addition to extracting WBC from them. The recommended algorithm can be immediately merged with the initial microscopic images produced by an automated microscope platform to perform WBC differential counting. The suggested algorithms may also be easily calibrated for use with human samples, and they can take the place of expensive haematology analyzers and inefficient manual counting techniques. The authors in [24] presented a computer vision-based strategy for developing a malaria parasite detection algorithm. In order to classify blood cells that have been infected by parasites, the deep learning capsule network is updated, and a supervised pixel-based blood cell segmentation model is developed. They used ANN and CapsNet models to segment and classify blood cells using the Local (DB1) and CDC (DB2) datasets, respectively. Over ANN and CapsNet, the suggested architecture obtained an overall accuracy of 99.1% and 98.7%, respectively. In[25], the authors found that it is difficult to extract discriminate features from unprocessed microscopic images, therefore the commonly used SVM classifier simply requires minor adjustments. These techniques, however, are unable to handle fine-grained classification situations involving up to 40 different types of white blood cells. In addition to a leukocyte classifier, the author has proposed a model based on deep residual learning theory that employs RNN and CNN techniques (Call3Net). A standard dataset of about 100,000 leukocytes from 40 different groups was developed for training and testing. The accuracy gained is 76.84 %. Multiple training strategies are chosen and integrated to improve the classifier's generalization performance. This method can also easily handle fine-grained classification situations with up to 40 kinds of white blood cells. The authors in [26] introduced a new automated peripheral blood smear analysis method. It includes software for processing images and an automated microscope for collecting images of a blood sample at the microscopic level. The CNN(U-net) method utilising microscopic images as a dataset yielded a 99.5 percent sensitivity for WBC extraction. In the last applied segmentation for blood cell which will be future work[27].

3 Proposed Methodology

The proposed approach can be separated into three sections: data preprocessing, feature extraction and classification.

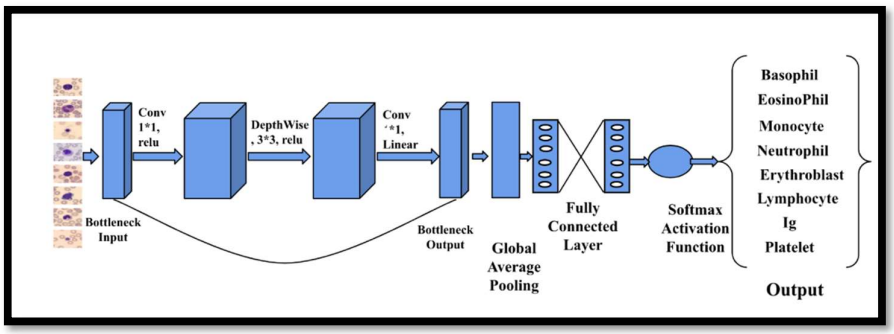

Figure 1: Architecture of blood cell classification

3.1 A Transfer Learning

Transfer Learning (TL) is defined as the transfer of knowledge, which was learned earlier in one area, to another for the purpose of classifying and extracting features. TL has been utilized in numerous deep learning applications because refining a pre-formed CNN model is usually a lot quicker and easier than forming a CNN model with randomly initialized weights from scratch. Here MobileNetV2 has been utilized for feature extraction which is one the deep CNN based model [15].

3.2 MobileNetV2

Here MobileNetV2[28], has been utilized for feature extraction which is one the deep CNN based model. The MobileNetV2 features 53 layers of depth and one Avg pool. It has two primary components, inverted residual block and bottleneck residual block. The traditional VGG architecture, which consists of building a network by stacking convolution layers to enhance accuracy, is referenced in the design of MobileNetV1. However, when too many convolutional layers are stacked, the issue of gradient vanishing arises. Residual block in ResNet facilitates information propagation, allowing for feature reuse in forward propagation and preventing gradient vanishing in back propagation. As a result, MobileNetV2 additionally makes use of ResNet's residual structure in addition to continuing to use MobileNetV1's depth separable convolution. The implementation of a linear bottleneck and an inverted residual block in the network are MobileNetV2's two primary enhancements over MobileNetV1 in comparison.

3.3 Data Preprocessing

The most crucial component of the model is the preprocessing method. In this work, we used raw data that has been resized to 224*224, the predefined size of our model. To reduce over fitting, we implemented the data augmentation and class-balancing technique, which sped up the training process. Later, we split data into 70% for training,

20% for testing, and 10% for validation. For every class, the dataset was chosen randomly.

3.4 Feature Extraction

The Deep Convolutional neural network demonstrated improved performance in all areas, particularly in medical imaging. The images of the blood cells are fed into the convolution layer to extract the features[29]. The output of the convolution layer will pass through the relu layer that converts the data into a non-linear form, then this output is inserted into the pooling layer which removes any redundant data which is acquired in the convolution process.

3.5 Classification

The extracted features are inserted into a fully connected layer to classify images into their respective classes. Dropout layer is utilized to decrease overfitting. These layers cut the part of the data flowing between the fully connected layers, making a model for accurate data classification. The output is computed using the SoftMax function.

3.6 Dataset

Table 1: Cell type and a number of images in each group

Class Name	Number of Images Per class
Basophil	1218
Eosinophil	3117
Erythroblast	1551
Ig	2895
Lymphocyte	1214
Monocyte	1420
Neutrophil	3329
Platelet	2348
Total	17,092

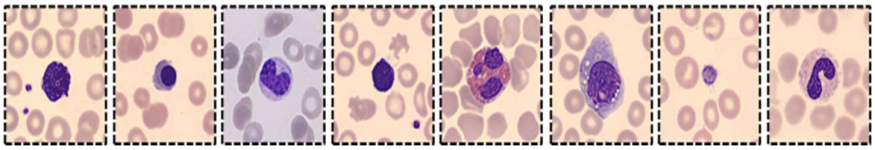

***Figure 1*:** Data Set Sample Example

The images utilized in this work are part of a public dataset that was made available by Anna et al. [15][22] from the hospital clinic in Barcelona. A total of 17,092 pictures of healthy people's human peripheral blood cells are included in the dataset. There are 8 other subtypes of these cells: basophil, eosinophil, erythroblast, Ig, lymphocyte, monocyte, neutrophil, and platelet. Table 1 presents the dataset's statistics in detail. The dataset contains 360 * 363 pixel JPG photos that are all in the RGB color space and format. Clinical pathologists from the Barcelona hospital clinic, annotated each image.

3.7 Evaluation metrics

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}},$$

$$\text{Precision} = \frac{\text{TN}}{\text{TN} + \text{FP}},$$

$$\text{F1} - \textit{score} = \frac{2 \cdot \text{TP}}{2 \cdot \text{TP} + \text{FP} + \text{FN}},$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}},$$

Where TP is true positive, TN is true negative, FP is false positive and FN is false negative for the overall test samples N.

4 Results and Discussion

We performed experiments on human peripheral blood cell dataset [22] with eight class multiclass scenario with state-of-the-art deep learning models namely VGG16, VGG19, ResNet-50 and MobileNetV2 to perform classification. In this section with 4 DL models were validated the eight-class classification.

As we can see in Table 2, the comparison of different model with four different metrics have done.

Table 2: Results on blood cell classification with state-of-the-art deep learning models.

Models	Accuracy	Precision	Recall	F-Measure
VGG16	90.53	92	91	91
VGG19	89.01	90	89	89
ResNet50	96.34	97	97	97
MobileNetV2	97.89	98	98	98

Table 2 present the result of different classification models to classify blood cell were evaluated on testing dataset. The authors in [22] collected data set. The training dataset, as described in section 2.6. The result of the proposed models is 97.89%, 98%, 98% and 98% terms of accuracy, F1 score, precision, and recall. From the table 1, we can see that the proposed model has got good result using MobileNetV2 model which was outperform all other models for all four metrics.

5 Conclusions

Classification of Blood Cells has emerged as one of the major research focusing on microscopic images which is subfield of the medical imaging. Blood cell classification is particularly plays and important role in CAD based diagnosis system with the rise in infectious diseases and due lack of facilities or knowledge led to difficulties in early detection. In this work, we suggested to investigate categorization of blood cells using various transfer learning methods, such as modified MobileNetV2. In parallel, we compare proposed model with different transfer learning models. The proposed MobileNetV2 model is more efficient and has less parameters, requires less computation burden compared to other transfer learning models. We achieved improved accuracy, which is equivalent to 97.89%, outperforming the compared models on the data. With eight classes, the categorization of blood cells using our suggested technique yields good performance.

References

[1] M. A. Al-masni, M. A. Al-antari, M. T. Choi, S. M. Han, and T. S. Kim, "Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks," *Comput. Methods Programs Biomed.*, vol. 162, pp. 221–231, 2018, doi: 10.1016/j.cmpb.2018.05.027.

[2] M. A. Al-antari, S. M. Han, and T. S. Kim, "Evaluation of deep learning detection and classification towards computer-aided diagnosis of breast lesions in digital X-ray mammograms," *Comput. Methods Programs Biomed.*, vol. 196, p. 105584, 2020, doi: 10.1016/j.cmpb.2020.105584.

[3] C. Chola and J. V. B. Benifa, “Detection and Classification of Sunspots via Deep Convolutional Neural Network,” *Glob. Transitions Proc.*, pp. 0–7, 2022, doi: 10.1016/j.gltp.2022.03.006.

[4] M. A. Al-antari, M. A. Al-masni, M. T. Choi, S. M. Han, and T. S. Kim, “A fully integrated computer-aided diagnosis system for digital X-ray mammograms via deep learning detection, segmentation, and classification,” *Int. J. Med. Inform.*, vol. 117, no. May, pp. 44–54, 2018, doi: 10.1016/j.ijmedinf.2018.06.003.

[5] M. A. Al-masni, M. A. Al-antari, H. Min, N. Hyeon, and T. Kim, “2nd IEEE Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability 2020, ECBIOS 2020,” *2nd IEEE Eurasia Conf. Biomed. Eng. Healthc. Sustain. 2020, ECBIOS 2020*, pp. 95–98, 2020.

[6] M. A. Al-masni, W. R. Kim, E. Y. Kim, Y. Noh, and D. H. Kim, “Automated detection of cerebral microbleeds in MR images: A two-stage deep learning approach,” *NeuroImage Clin.*, vol. 28, p. 102464, 2020, doi: 10.1016/j.nicl.2020.102464.

[7] X. Li, W. Li, X. Xu, and W. Hu, “Cell Classification Using Convolutional Neural Networks in Medical Hyperspectral Imagery,” pp. 501–504, 2017.

[8] C. Chola *et al.*, “Gender Identification and Classification of Drosophila melanogaster Flies Using Machine Learning Techniques,” vol. 2022, 2022.

[9] L. M. Mestetskiy, D. S. Guru, J. V. B. Benifa, H. S. Nagendraswamy, and C. Chola, “Gender identification of Drosophila melanogaster based on morphological analysis of microscopic images,” *Vis. Comput.*, 2022, doi: 10.1007/s00371-022-02447-9.

[10] Y. Y. Baydilli and Ü. Atila, “Classification of white blood cells using capsule networks,” *Comput. Med. Imaging Graph.*, vol. 80, p. 101699, Mar. 2020, doi: 10.1016/J.COMPMEDIMAG.2020.101699.

[11] A. Y. Muaad, J. Hanumanthappa, M. A. Al-antari, B. B. J. V, and C. Chola, “AI-based Misogyny Detection from Arabic Levantine Twitter Tweets,” *Proc. 1st Online Conf. Algorithms, 27 Sept. Oct. 2021, MDPI Basel, Switzerland, doi10.3390/IOCA2021-10880*, pp. 4–11, 2021.

[12] A. Y. Muaad, H. J. Davanagere, M. A. Al-antari, J. V. B. Benifa, and C. Chola, “AI-based Misogyny Detection from Arabic Levantine Twitter Tweets,” in *Computer Sciences \& Mathematics Forum*, 2021, vol. 2, no. 1, p. 15.

[13] A. Y. Muaad *et al.*, “An Effective Approach for Arabic Document Classification Using Machine Learning,” *Glob. Transitions Proc.*, pp. 0–5, 2022, doi: 10.1016/j.gltp.2022.03.003.

[14] J. Zhao, M. Zhang, Z. Zhou, J. Chu, and F. Cao, “Automatic detection and classification of leukocytes using convolutional neural networks,” *Med. Biol. Eng. Comput.*, vol. 55, no. 8, pp. 1287–1301, 2017, doi: 10.1007/s11517-016-1590-x.

[15] A. Acevedo, S. Alférez, A. Merino, L. Puigví, and J. Rodellar, “Recognition of peripheral blood cell images using convolutional neural networks,” *Comput. Methods Programs Biomed.*, vol. 180, p. 105020, 2019, doi: 10.1016/j.cmpb.2019.105020.

[16] J. Hung *et al.*, “Applying Faster R-CNN for Object Detection on Malaria Images,” pp. 1–7, 2018, [Online]. Available: http://arxiv.org/abs/1804.09548.

[17] D. Bani-Hani, N. Khan, F. Alsultan, S. Karanjkar, and N. Nagarur, “Classification of Leucocytes Using Convolutional Neural Network Optimized Through Genetic Algorithm,” no. November, pp. 1–7, 2018.

[18] R. R. Tobias *et al.*, "Faster R-CNN Model with Momentum Optimizer for RBC and WBC Variants Classification," *LifeTech 2020 - 2020 IEEE 2nd Glob. Conf. Life Sci. Technol.*, no. January 2021, pp. 235–239, 2020, doi: 10.1109/LifeTech48969.2020.1570619208.

[19] M. Sajjad *et al.*, "Leukocytes Classification and Segmentation in Microscopic Blood Smear: A Resource-Aware Healthcare Service in Smart Cities," *IEEE Access*, vol. 5, pp. 3475–3489, 2017, doi: 10.1109/ACCESS.2016.2636218.

[20] F. Long, J. Peng, W. Song, X. Xia, and J. Sang, "Computer Methods and Programs in Biomedicine BloodCaps : A capsule network based model for the multiclassification of human peripheral blood cells," vol. 202, 2021, doi: 10.1016/j.cmpb.2021.105972.

[21] X. Zheng, Y. Wang, G. Wang, and J. Liu, "Fast and robust segmentation of white blood cell images by self-supervised learning," *Micron*, vol. 107, no. January, pp. 55–71, 2018, doi: 10.1016/j.micron.2018.01.010.

[22] A. Acevedo, A. Merino, S. Alférez, Á. Molina, L. Boldú, and J. Rodellar, "A dataset of microscopic peripheral blood cell images for development of automatic recognition systems," *Data Br.*, vol. 30, p. 105474, 2020, doi: 10.1016/j.dib.2020.105474.

[23] C. B. Wijesinghe, D. N. Wickramarachchi, I. N. Kalupahana, L. R. De Seram, I. D. Silva, and N. D. Nanayakkara, "Fully Automated Detection and Classification of White Blood Cells," *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. EMBS*, vol. 2020-July, pp. 1816–1819, 2020, doi: 10.1109/EMBC44109.2020.9175961.

[24] M. Maity, A. Jaiswal, K. Gantait, J. Chatterjee, and A. Mukherjee, "Quantification of malaria parasitaemia using trainable semantic segmentation and capsnet," *Pattern Recognit. Lett.*, vol. 138, pp. 88–94, 2020, doi: 10.1016/j.patrec.2020.07.002.

[25] F. Qin, N. Gao, Y. Peng, Z. Wu, S. Shen, and A. Grudtsin, "Fine-grained leukocyte classification with deep residual learning for microscopic images," *Comput. Methods Programs Biomed.*, vol. 162, pp. 243–252, 2018, doi: 10.1016/j.cmpb.2018.05.024.

[26] D. Mundhra, B. Cheluvaraju, J. Rampure, and T. Rai Dastidar, "Analyzing microscopic images of peripheral blood smear using deep learning," *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10553 LNCS, no. December, pp. 178–185, 2017, doi: 10.1007/978-3-319-67558-9_21.

[27] E. Tavakoli, A. Ghaffari, Z. M. Kouzehkanan, and R. Hosseini, "New Segmentation and Feature Extraction Algorithm for Classification of White Blood Cells in Peripheral Smear Images," 2021.

[28] Q. Xiang, G. Zhang, X. Wang, J. Lai, R. Li, and Q. Hu, "Fruit image classification based on Mobilenetv2 with transfer learning technique," *ACM Int. Conf. Proceeding Ser.*, 2019, doi: 10.1145/3331453.3361658.

[29] M. Pramodha, A. Y. Muaad, B. B. J. V, J. Hanumanthappa, C. Chola, and A. Mugahed, "A Hybrid Deep Learning Approach for COVID-19 Diagnosis via CT and X - R ay Medical Images," pp. 1–10, 2021.