

Article

White Blood Cell Classification Using Multi-Attention Data Augmentation and Regularization

Nasrin Bayat ¹, Diane D. Davey ², Melanie Coathup ² and Joon-Hyuk Park ^{1,*}¹ Department of Electrical and Computer Engineering, University of Central Florida, Orlando, FL 32816, USA² College of Medicine, University of Central Florida, 6850 Lake Nona Blvd, Orlando, FL 32827, USA

* Correspondence: joonpark@ucf.edu

Abstract: Accurate and robust human immune system assessment through white blood cell evaluation require computer-aided tools with pathologist-level accuracy. This work presents a multi-attention leukocytes subtype classification method by leveraging fine-grained and spatial locality attributes of white blood cell. The proposed framework comprises three main components: texture-aware/attention map generation blocks, attention regularization, and attention-based data augmentation. The developed framework is applicable to general CNN-based architectures and enhances decision making by paying specific attention to the discriminative regions of a white blood cell. The performance of the proposed method/model was evaluated through an extensive set of experiments and validation. The obtained results demonstrate the superior performance of the model achieving 99.69 % accuracy compared to other state-of-the-art approaches. The proposed model is a good alternative and complementary to existing computer diagnosis tools to assist pathologists in evaluating white blood cells from blood smear images.

Keywords: attention mechanism; medical image analysis; deep learning; blood cell detection; convolutional neural networks



Citation: Bayat, N.; Davey, D.D.; Coathup, M.; Park, J.-H. White Blood Cell Classification Using Multi-Attention Data Augmentation and Regularization. *Big Data Cogn. Comput.* **2022**, *6*, 122. <https://doi.org/10.3390/bdcc6040122>

Academic Editors: Nadav Rappoport, Yuval Shahar and Hyojung Paik

Received: 18 September 2022

Accepted: 19 October 2022

Published: 21 October 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The general health condition of a patient can be learned through a quantitative and qualitative examination of blood components, such as cell counts. Blood cells are primarily classified into two categories: leukocytes or White Blood Cells (WBCs) and erythrocytes or Red Blood Cells (RBCs). WBCs are further divided into four nucleated subtypes, namely eosinophils, lymphocytes, monocytes, and neutrophils, as shown in Figure 1, [1]. WBC counts and their subtype proportions contain critical information about the status of infectious diseases and chronic processes, e.g., inflammatory, leukemia, malnutrition, and white cell proliferative conditions [2].

The traditional WBC analysis includes differentiation of subtypes through microscopic observation of the blood smear and assessment of the morphological characteristics of the cell nucleus and cytosol. Such techniques is highly dependent on the experience level of the analyst and, at the same time, it could be labor intensive and time consuming [3]. Additionally, a completely automatic blood cell analyzer has been used to perform WBC analysis. However, they frequently have high requirements for test samples and are expensive, which prevents them from being widely used at point-of-care settings or in township hospitals [4].

Therefore, researchers in the community have devised automatic yet faster approaches for analysis of leukocytes leveraging computer vision techniques [5–9]. Given the recent advancement of machine learning and computer vision, several approaches have been proposed for leukocyte classification and segmentation, ranging from more conventional machine learning models such as support vector machine [10] and Naïve Bayesian [11] to more advanced deep learning methods [12,13]. Within deep learning methods, Convolutional Neural Networks (CNNs) have shown exemplary performance in medical image

processing [14,15], while computer-aided approaches allow a faster, economic and reproducible means for WBC classification, automating the computational process to reach the clinical level of accuracy and reliability in WBC classification is still in development.

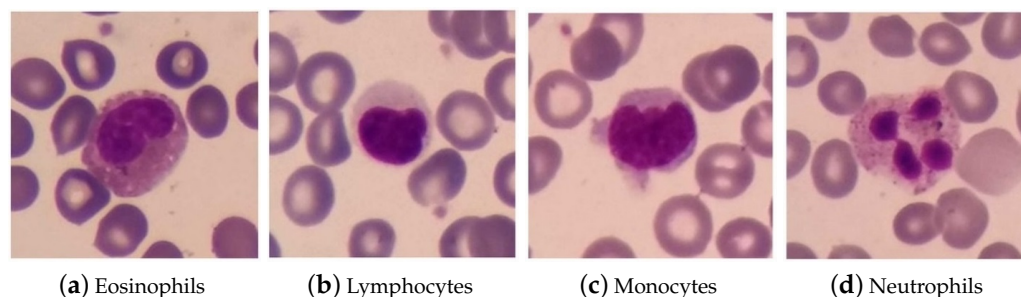


Figure 1. Example of different white blood cell types.

In this study, we demonstrate an advanced white blood cell classification by approaching it as a fine-grained visual classification problem, where the main goal was to identify the subordinate-level categories of WBC by tackling few challenges as the following. First, there is a substantial variance in the characteristics associated with cell morphology, i.e., size, shape, texture, nucleus, etc., [5] of each cell subtype. Second, there is a small variance between images of different cell types, making it a challenging classification task. Such subtle differences between different cell types hinder accurate leukocytes classification. Therefore, it is desirable to capture more discriminative regions of the cell to access more enriched feature space which, in turn, can improve the classification accuracy. By imposing extra supervision on instance interpretation during the learning process using an attention-based data augmentation method, the model is compelled to pay more attention to the regions of interest in order to accomplish this goal [16,17].

This work presents a data augmentation and regularization framework based on multi-attention mechanism to force CNN-based models to extract more discriminative features to enhance leukocyte subtype recognition. The presented framework is specifically designed to produce an enriched feature space by extracting texture-related information and deep features. Specifically, the proposed model employs attention-based augmentation and regularization to focus on various regions within the WBC image to learn more discriminative features. The presented framework is applicable to other CNN-based backbone architectures to achieve better performance. The effectiveness of the proposed method is assessed through a large number of WBC microscopic image samples, and the classification performance was compared with other state-of-the-art methodologies.

The proposed model is a good alternative and complementary to existing computer-aided diagnosis tools to assist pathologists in evaluating white blood cells from blood smear images. The primary contributions of this work are summarized as follows:

- The WBC classification task is considered as a fine-grained visual classification problem for which a multi-attention framework for efficient WBC classification has been developed. The presented method captures texture-aware information from shallow layers and deep features from deep layers to ensure that the model learns only discriminative features through attention-based augmentation and regularization mechanisms.
- The presented attention-based mechanism is composed of three main components: texture-aware/attention map generation blocks, attention regularization and attention-based data augmentation. The presented multi-attention framework is applicable to all other existing CNN-based models for WBC classification.
- An extensive set of experiments are conducted to assess the performance of the model from different perspectives. The obtained results demonstrated the surpassing performance of the model, achieving 99.69% classification accuracy, compared to existing state-of-the-art approaches.

The rest of the paper is organized as follows. Recent related studies on white blood cell classification are discussed in Section 2. Section 3 presents the outline of the proposed attention-based WBC classification approach. Model evaluation settings, including implementation specifics, evaluation metrics, and the employed WBC dataset are described in Section 4. The obtained WBC subtype detection results are presented and discussed in Section 5, with their implications in comparison with existing methods and results from other studies. Finally, concluding remarks are drawn in Section 6.

2. Related Work

Various deep learning models have been developed and used to perform WBC classification of automatic detection of leukocytes [18,19]. For example, Togacar et al., presented a WBC subclass separation framework based on the AlexNet model [20]. Wang et al., proposed to learn spectral and spatial features from microscopy hyperspectral images using deep convolution networks [21]. A CNN model with loss enhancement with regularization was presented that reduced the processing time [22]. Further, Jiang et al., employed residual convolution structure with batch normalization to improve activation function for enhancing feature extraction in the WBC classification [23]. Furthermore, Yao et al., introduced weighted optimized deformable CNN for WBC classification [6] while Khan et al., proposed multi-layer convolutional features with an extreme-learning machine for a similar WBC identification task [24].

In addition, using hybrid approaches such as an ensemble of several models have been studied. For example, Çınar and Tuncer [7] employed two feature extraction models, namely AlexNet and GoogleNet, for white blood cell feature extraction and classification using support vector machine model. Özyurt [25] used several well-known pre-trained models as a feature extractor and used Extreme Learning Machines (ELM) classifiers to classify the fused features. Patil et al., [26] proposed the extraction of overlapping and multiple nuclei patches using a combination of CNN and recurrent neural networks. Baghel et al., [27] presented a two-stage classification approach to perform mononuclear and polymorphonuclears identification and associated subtypes based on a CNN model.

Table 1 summarizes the literature in chronological order to provide a better understanding of the current status of the WBC classification methods along with the model architectures employed. As can be seen from the table, most previous methods highly relied on CNN-based architectures, such as AlexNet, MobileNet, etc., due to their efficiency in analyzing images, while these approaches have shown good performance in the WBC classification [8,24,28], extracting the features associated with distinct regions of the cell is still difficult to achieve. There exist subtle discrepancies among different cell types, which tend to be retained in textural information of shallow features. On the other hand, different regions of WBC images have different textural patterns, which should be maintained as important discriminative information throughout the pooling operation. Hence, identification and intensification of such a small difference between cell types and the associated features are critically important to achieving more accurate and reliable classification with greater efficiency (shorter processing time). This requires the model to focus more on the distinctive regions within the cell. To address this limitation, we proposed an attention-based data augmentation and regularization approach which was implemented and validated for WBC classification. In addition, recent studies [29] show that deep layers of network capture high-level semantic information but messy details, while it is the opposite for shallow layers. In our experiments, we noticed that incorporating texture features besides the deep features improves the overall model performance.

Table 1. Summary of WBC classification methods in chronological order.

Year	Authors	Model Description
2017	Razzak [30]	CNN combined with ELM
2017	Yu et al. [31]	Ensemble of CNN's
2018	Jiang et al. [23]	Residual convolution architecture
2018	Liang et al. [32]	Combination of Xception-LSTM
2019	Hegde et al. [33]	AlexNet and CNN model
2019	Huang et al. [34]	MFCNN CNN with hyperspectral imaging
2019	Togacar et al. [20]	AlexNet with QDA
2020	Abou et al. [35]	CNN model
2020	Banik et al. [36]	CNN with feature fusion
2020	Basnet et al. [22]	DCNN model with modified loss
2020	Baydilli et al. [37]	capsule networks
2020	Kutlu et al. [28]	Regional CNN with a Resnet50
2020	Özyurt [25]	Ensemble of CNN models with ELM classifier.
2021	Baghel et al. [27]	CNN model
2021	Çinar et al. [7]	Ensemble of CNN models and SVM
2021	Khan et al. [24]	AlexNet model and ELM
2021	Yao et al. [6]	Deformable convolutional neural networks.
2022	Cheuque et al. [8]	Faster R-CNN with MobileNet model
2022	Girdhar et al. [9]	CNN model

3. Methodology

This section provides a detailed description of the above-mentioned attention-based white blood cell classification framework, while attention-based approaches can improve the performance of the backbone models in various vision tasks, a dual-attention mechanism was employed to enhance the accuracy and efficiency of WBC classification. The motivation behind using the attention mechanism for WBC classification is that all parts of the WBC image may not carry distinguishing information, rather they are mutual across different cell types. Therefore, it is important to mimic cognitive attention and utilize the most relevant parts of the input WBC image. The attention mechanism enables the traditional deep learning networks to have the flexibility to utilize different regions of the input image in the run-time using a weighted combination of all the encoded input images. The most relevant regions scored the highest weights. The presented framework is applicable to CNN-based backbone models and is composed of three main components: an attention generation module, an attention regulation module, and an attention-based data augmentation module. The general pipeline of the presented attention-based white blood cell detection approach is illustrated in Figure 2. While attention-based data augmentation methods can improve the performance of the model by enhancing discriminative feature space, it could also lead to performance degradation if multiple attention maps focus on a single region and ignore other discriminative regions. Therefore, each attention map was made sure to be non-overlapping and cover only a specific region from all input blood smear images. The generalizability of the proposed approach and its impact on improving the classification accuracy and efficiency (computational time) were demonstrated, which supports its validity and applicability for use in the WBC classification.

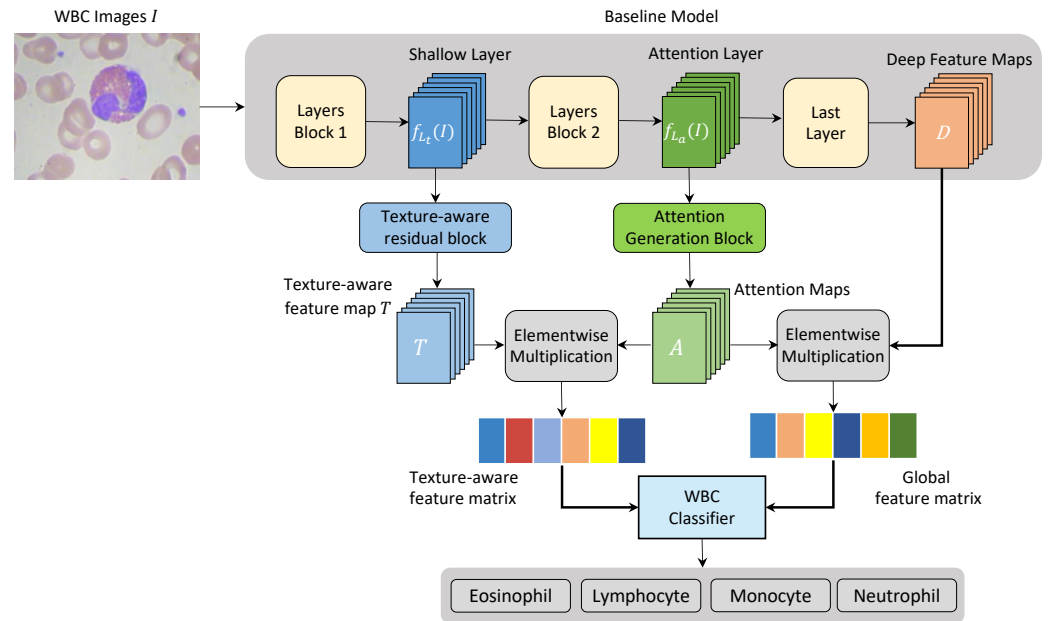


Figure 2. Overall framework of the proposed attention-based white blood cell classification approach. It is composed of three main components, including texture-aware residual block, attention generation, and attention-based data augmentation through element-wise multiplication and normalized average pooling. The presented framework is generalizable to different backbone models. The attention-based data augmentation mechanism helps the model not only focus on more robust features but also forces the model to pay attention to different parts of the input image to obtain more discriminative features from texture-aware shallow features.

3.1. Attention Generation

For every given input WBC image I , the feature map from the n^{th} layer of the backbone model $f^b(\cdot)$ can be represented as $F = f_n^b(I) \in R^{C_n \times H_n \times W_n}$, where the number of channels, height, and width of the feature map are represented by C_n , H_n , and W_n , respectively. Then, the extracted feature maps from particular layers are used to generate attention maps (A) from mutually exclusive regions of the input image using attention generator block $f^g(\cdot)$ as described in Equations (1) and (2).

$$A = f^g(F) = \bigcup_{k=1}^M A_k, \quad F = f_n^b(I) \quad (1)$$

$$f^g(\cdot) = \text{Linear}(\text{Norm}(\text{Conv}_{1D}(\cdot))) \quad (2)$$

where $A_k \in R^{H_n \times W_n}$ represents one attention map corresponding to k^{th} discriminative region of the input image from a predefined attention layer L_a of the model, that is selected for attention map generation. As aforementioned, it is important to preserve textural information of shallow features to capture subtle discrepancies among different cell types. To maintain and intensify those subtle differences, a feature-level residual block along with densely connected convolution layers are utilized to obtain feature maps as depicted in Figure 3. Shallow layer $n = L_t$ is specifically selected to extract feature maps that represent textural information of different cell types. The obtained texture-aware feature map contains critical discriminative information about subtle differences in cell-types that could boost the performance of the backbone model.

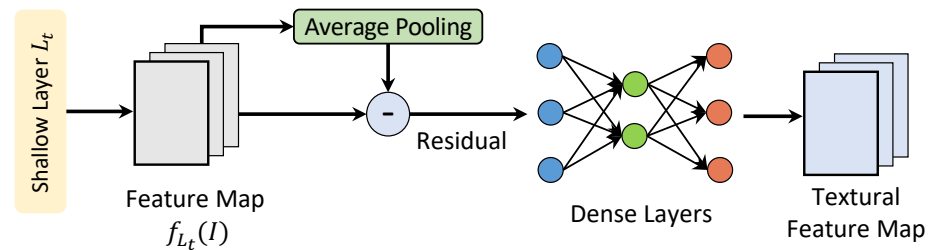


Figure 3. Texture-aware residual block helps preserve and enhance the texture information of shallow feature maps at layer L_t through average pooling, feature-level residuals, and densely connected convolution layers.

Having generated attention maps from attention layer $f_{L_a}(I)$ and texture-aware feature maps from shallow layer $f_{L_t}(I)$, two sets of attention-based representative feature could be obtained, i.e., texture-aware feature matrix T and global feature matrix G . Texture-aware feature matrix and global feature matrix could be calculated through element-wise multiplication of attention maps with texture-aware feature maps from the shallow layer and network's last layer feature map, respectively. The process of element-wise multiplication of texture-aware feature maps from shallow layer $f_{L_t}(I)$ with specific attention map and normalized average pooling $g(\cdot)$ is shown in Figure 4. The obtained discriminative features are concatenated and fed into the classifier.

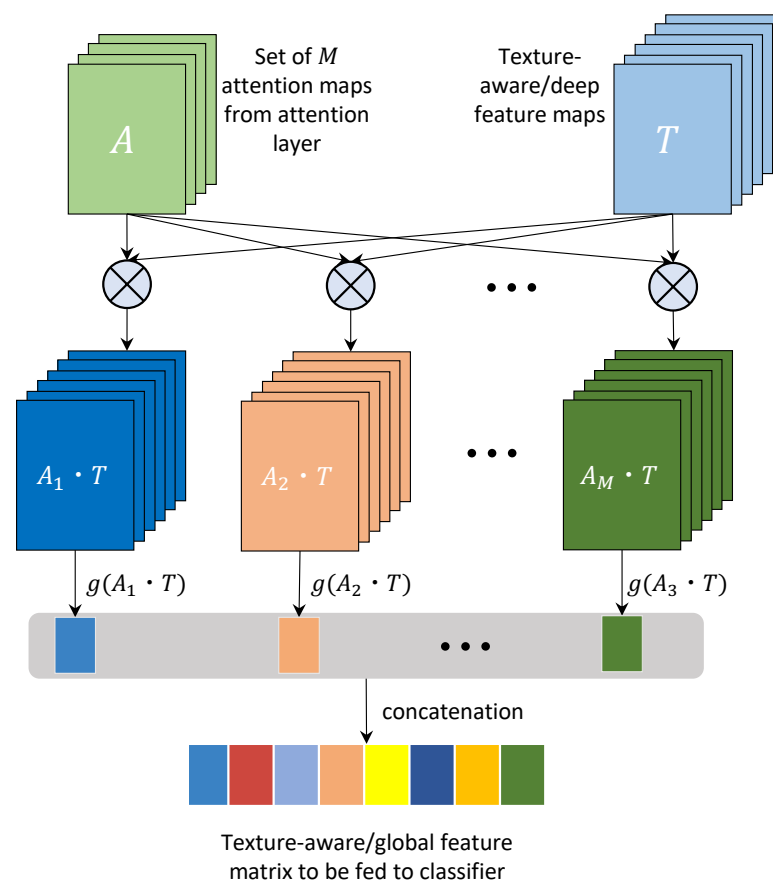


Figure 4. Texture-aware discriminative feature extraction through attention analysis and normalized average pooling. Discriminative features are pooled using localized feature maps, which are the product of element-wise multiplication of texture-aware feature maps with unique attention maps.

3.2. Attention Regularization

In the attention-based data augmentation process, if all attention maps focus on the same regions and ignore exploring different regions of the image, the network may fail to capture the necessary information. Furthermore, it is expected that each attention map always refers to the same semantic region, rather than random parts of the input image. Inspired by [38] and to keep attention maps non-overlapping and forcing them to focus on specific regions of the input image, an attention-based loss function $\mathcal{L}_{AL}$ is utilized, as shown in Equation (3).

$$\mathcal{L}_{AL} = \sum_{i=1}^B \sum_{j=1}^M \max\left(\|V_j^i - c_j^t\|_2^2 - m_{in}(y_i), 0\right) + \sum_{i,j \in (M,M), i \neq j} \max\left(m_{out} - \|c_i^t - c_j^t\|_2^2, 0\right) \quad (3)$$

where $V \in R^{M \times N}$ is a semantic feature vector obtained through element-wise multiplication of pooled feature map, y_i indicates class label, M denotes the number of attentions, m_{in} indicates feature and feature center's margin, m_{out} is the margin between feature centers, and c is the feature center. Feature centers are updated in each iteration using Equation (4).

$$c^t = c^{t-1} - \alpha \left(c^{t-1} - \frac{1}{B} \sum_{i=1}^B V^i \right) \quad (4)$$

where α denotes the feature center update rate at each iteration and B represents the batch size. The first component of Equation (3), i.e., $\sum_{i=1}^B \sum_{j=1}^M \max\left(\|V_j^i - c_j^t\|_2^2 - m_{in}(y_i), 0\right)$ is responsible for reducing intra-class loss through pulling V closer to feature center c , whereas the inter-class loss i.e., $\sum_{i,j \in (M,M), i \neq j} \max\left(m_{out} - \|c_i^t - c_j^t\|_2^2, 0\right)$, is responsible for increasing the distance between feature centers. Ultimately, the final loss function is a combination of attention-based loss function $\mathcal{L}_{AL}$ and the traditional cross-entropy loss $\mathcal{L}_{CE}$ as written in Equation (5).

$$\mathcal{L} = \mathcal{L}_{CE} + \mathcal{L}_{AL} \quad (5)$$

3.3. Attention-Based Data Augmentation

While random data augmentation techniques generate high background noise, the obtained attention maps from different layers of the model can be helpful for better data augmentation. The attention-based data augmentation mechanism makes sure that the model gets exposed to additional variations of the original input within the training process. This helps the model to not only learn the original representation of a given input but also learn additional variations of the input through the augmentation process [39,40]. For each sample from the training WBC image set, a unique attention map A_k is randomly selected and normalized as k^{th} augmentation map, A_k^* , as shown in (6).

$$A_k^* = \frac{A_k - \min(A_k)}{\max(A_k) - \min(A_k)} \quad (6)$$

The augmentation map is utilized as a regulation weight between the degraded image I_d , which is generated through Gaussian blur, and the original image as $I' = I_d \times A_k^* + (1 - A_k^*) \times I$. The augmentation map can be employed from two different perspectives to help train the model. First, it can pay more attention to regions with high attention scores through input image cropping, which forces the model to learn more robust features from the most discriminative parts of the image. Second, it can be utilized to allow the model to produce different attention maps focusing on different regions by discarding regions with higher attention scores. Figure 5 shows some examples

of attention-based cropping and dropping methods for a sample input image from different white blood cell classes.

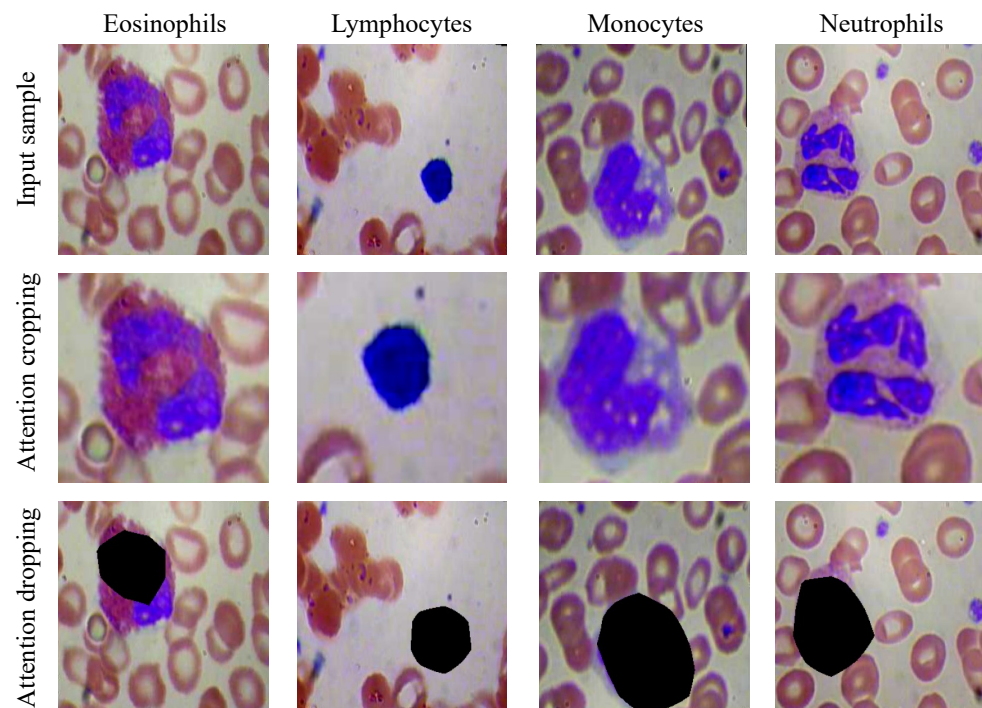


Figure 5. The obtained attention maps could be utilized to force the model to focus on different regions of the input image for more discriminative feature extraction. First, it is forced to pay more attention to regions with high attention scores through input image cropping. Second, the model is encouraged to explore different regions of the image by dropping regions with high attention scores.

4. Evaluation Settings

In this section, general evaluation settings, e.g., white blood cell datasets, preprocessing steps, implementation specifics, and evaluation metrics are described in detail.

4.1. Dataset

This study uses a publicly available dataset consisting of four different cell categories, i.e., Lymphocytes, Monocytes, Eosinophil, and Neutrophils [41]. The dataset contains 12,444 images of white blood cells with approximately equal distribution across each class Table 2. Different experiments are carried out with different number of blood smear images in train and test sets. This experiment will demonstrate how well the model performs even through training on smaller training sets. Train and test sets are randomly selected from each cell type separately to ensure the data distribution is intact.

Table 2. Statistical specifics of WBC dataset utilized in this study. Three different experiments with different train/test split ratios are designed to evaluate the generalizability of the proposed method.

Cell Type	Distribution (%)	Exp. 1 (60/40)		Exp. 2 (70/30)		Exp. 3 (80/20)	
		Train	Test	Train	Test	Train	Test
Eosinophil	25.10	1872	1248	2184	936	2496	624
Lymphocytes	24.93	1862	1240	2174	930	2482	620
Monocytes	24.84	1855	1236	2164	927	2473	618
Neutrophils	25.10	1874	1249	2187	936	2499	624
Total	100	7463	4973	8707	3729	9950	2486

4.2. Baseline Architectures

The presented attention-based white blood cell identification approach is applicable to different baseline models. In the following, three state-of-the-art deep learning networks used in this study are explained, and refer interested readers to the original references. In this study, these three models are utilized as baseline models.

ResNet Structure. A type of deep convolutional neural network called Residual Networks (ResNets) [42] that skip convolutional layer blocks while utilizing shortcut connections. The downsampling procedure in this architecture occurs at the convolutional layers with a stride of 2, followed by batch normalization and a ReLU activation function. The architecture consists of 101 layers in total, including a fully connected layer with softmax activation at the end of the network [42].

Xception Structure. Xception is a convolutional neural network with residual connections based on separable convolutions. This model has 71 deep layers. The feature extraction base of the network in the Xception architecture is composed of 36 convolutional layers. With the exception of the first and last modules, the 36 convolutional layers are structured into 14 modules which contain linear residual connections around them [43].

EfficientNet Structure. EfficientNet is a convolutional neural network design and scaling technique that uses a compound coefficient to consistently scale all depth, width, and resolution dimensions. The goal, which may be expressed as an optimization problem, is to maximize the model accuracy for any given resource constraints. Model scaling attempts to increase the network length (L_i), width (C_i), and/or resolution (H_i, W_i) without altering the baseline network's predefined F_i . This is in contrast to standard ConvNet designs, which primarily focus on identifying the ideal layer architecture F_i [44]. The EfficientNets family of models are created using neural architecture search [45] to develop a new baseline network, and scaling it up. The 8 models in the EfficientNet model range from B0 to B7, with each model number denoting a version with additional parameters and greater accuracy. Transfer learning is a technique used by the EfficientNet design to speed up the process. As a result, it offers higher accuracy than other competitor models. This is a result of the ingenious depth, width, and resolution scaling used [46].

4.3. Implementation Specifics

All baseline models along with associated attention-analysis are implemented using the *PyTorch* machine learning library and trained using Stochastic Gradient Descent *SGD* optimizer [47] with a learning rate of 5×10^{-4} , a momentum value of 0.9, and 10^{-4} weight decay. The model training is performed for 15 epochs using a mini-batch size of 64 to minimize the predefined loss function. A Lambda Quad deep learning workstation is used to implement, train and test the models. The machine is equipped with Ubuntu 20.04.3 LTS operating system, Intel Core™ i7-6850K CPU, 64 GB DDR4 RAM, and 4 NVIDIA GeForce GTX 1080 Ti Graphics Processing Units (GPUs).

4.4. Evaluation Metrics

The confusion matrix and associated evaluation metrics were computed to evaluate the performance of the proposed approach. A confusion matrix is composed of True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP) values. Performance of the model is evaluated against different evaluation metrics, including accuracy rate, recall, and F1-score.

5. Results & Discussion

The performance of the proposed attention-based white blood cell classification approach is investigated through an extensive set of experiments. The obtained results are presented and discussed as follows. The presented attention-based method for WBC classification is implemented on three different well-established CNN models. These models were then trained and tested using three different train/test split set sizes. The obtained results from these analyses are shown in Figure 6 which indicate a satisfactory WBC

classification accuracy above 99% even with the smallest training set (60/40 ratio) across all backbone models. For example, the detection rate has dropped only less than 1% when the training set is cut down from 80/20 to 60/40 in the Xception backbone model. and the classification performance of the proposed method using the aforementioned backbone architectures for three different train/test ratios at each epoch is illustrated in Figure 7. As can be observed all three backbone architectures achieve a high classification accuracy after only 15 epochs. For example, a configuration of the model with EfficientNet architecture offers state-of-the-art classification performance, i.e., 99.69%, only after 15 epochs in Exp. 3. To provide additional insight into the class-specific performance of the proposed approach, confusion matrix of different configurations of the presented WBC detection model are illustrated in Figure 8. Each confusion matrix demonstrates the classification performance of the model on the test set. It can be seen that while Lymphocytes and Monocytes have been classified more accurately, most of the mislabeled samples belong to Eosinophils and Neutrophils.

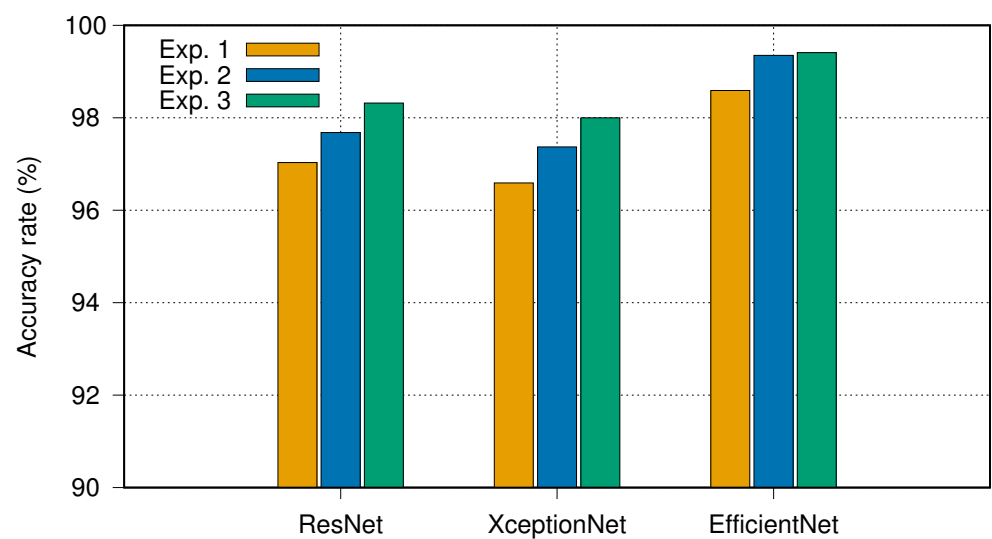


Figure 6. Performance comparison between different architectures used in the presented attention-based white blood cell detection, with varying train/test split sizes. Here, Exp. 1, Exp. 2, and Exp. 3 represent 60/40, 70/30, and 80/20 split sizes for train/test sets, respectively.

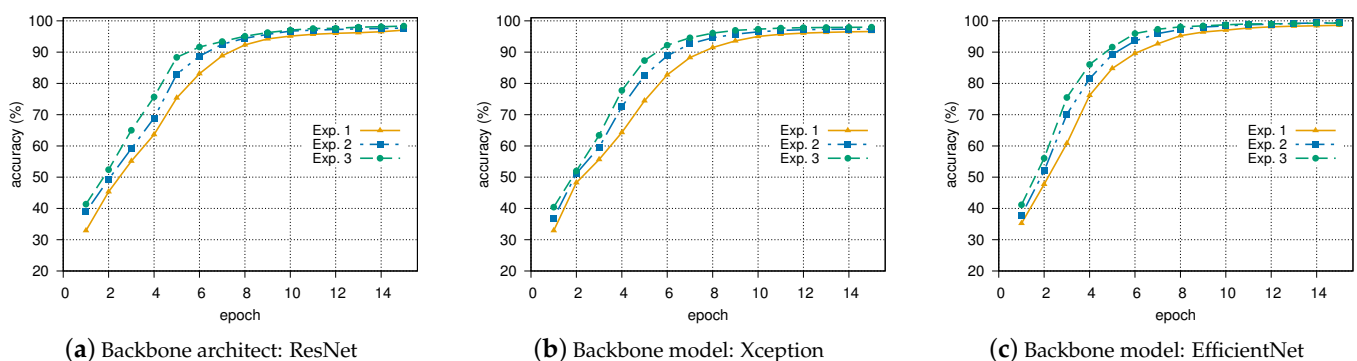


Figure 7. Performance of the proposed attention-based WBC detection approach while using aforementioned backbone architectures and three different ratios of train/test splits for the test set.

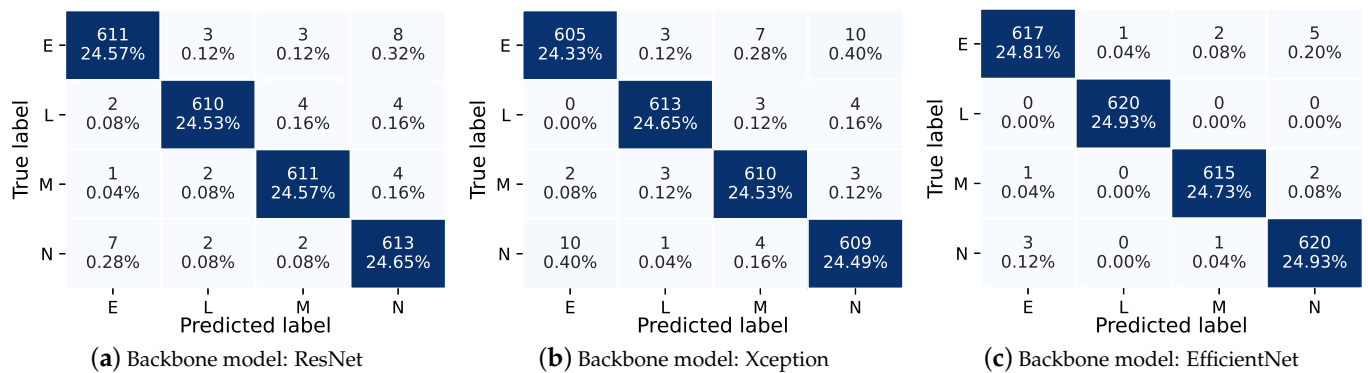


Figure 8. Confusion matrix of the presented WBC classification model using different backbone configurations. Note that E: Eosinophils, L: Lymphocytes, M: Monocytes, and N: Neutrophils.

5.1. Attention-Based Data Augmentation

To investigate the impact of the proposed attention-based data augmentation framework on the overall performance of the backbone models are compared with and without attention-based data augmentation Figure 9. To be in line with the literature and for comparability purposes, the rest of the experiments are conducted with a train/test set of 80/20 split ratios. It was seen that the presented attention-based framework evidently improve the performance of the WBC classification. For instance, the WBC classification model using EfficientNet architecture is able to achieve a classification accuracy of 99.69% using the proposed attention-based data augmentation mechanism. It should be noted that integration of the presented attention-based data augmentation approach with each of the backbone models results in the improvement of their performance, showing its generalizability to potentially enhance the classification performance in other applications Table 3.

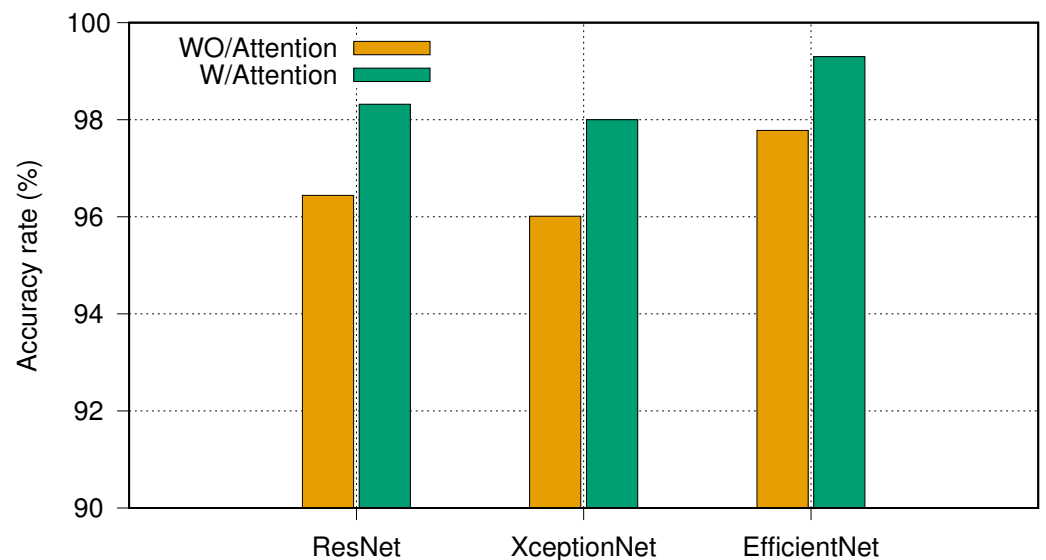


Figure 9. Performance of the presented attention-based white blood cell detection method compared with not using attention.

Table 3. Comparison of classification performance from three CNN backbones. The best performance was achieved using EfficientNet as the backbone with 99.69 % accuracy.

Backbone	Metrics	Class Specific Performance (%)				Ave.
		Eosinophils	Lymphocytes	Monocytes	Neutrophils	
Xception	ACC	98.71	99.43	99.11	98.71	98.99
	Recall	96.80	98.87	98.70	97.59	97.99
	F1 score	97.42	98.87	98.22	97.44	97.99
ResNet	ACC	99.03	99.30	99.35	98.91	99.15
	Recall	97.76	98.38	98.86	98.23	98.31
	F1 score	98.07	98.62	98.70	97.84	98.31
EfficientNet	ACC	99.51	99.95	99.75	99.55	99.69
	Recall	98.72	100.00	99.51	99.35	99.40
	F1 score	99.03	99.91	99.51	99.12	99.39

5.2. Comparison with Other SOTA Approaches

The performance of the proposed WBC classification method was compared with existing SOTA approaches. Table 4 summarizes the comparison of the obtained results in this work with that of other studies. It can be concluded that all configurations of the presented attention-based WBC detection approach presented in this study outperform other previous SOTA approaches used for WBC classification. In particular, the presented method was able to achieve superior detection rates even with a smaller number of training samples and fewer training epochs compared to other studies in the literature [9,26,48]. For example, a configuration of the presented approach using EfficientNet backbone architecture could achieve 98.59% and 99.69% accuracy rates after only 15 epochs of training with 60% and 80% of the samples, respectively. These results demonstrate that the proposed method offers not only better accuracy but also time and computational efficiency compared to other SOTAs considered in WBC classification.

Table 4. A quantitative comparison of the performance of the presented WBC classification approach with that of existing SOTA methods. NI: Not Indicated.

Authors	Accuracy (%)	Recall (%)	F1 Score (%)
Abou et al. [35]	96.8	NI	NI
Baghel et al. [27]	98.9	97.7	97.6
Baydilli et al. [37]	96.9	92.5	92.3
Banik et al. [36]	97.9	98.6	97.0
Basnet et al. [22]	98.9	97.8	97.7
Çinar et al. [7]	99.7	99	99.0
Hegde et al. [33]	98.7	99	99
Huang et al. [34]	97.7	NI	NI
Jiang et al. [23]	83.0	NI	NI
Khan et al. [24]	99.1	99.0	99
Kutlu et al. [28]	97	99.0	98
Liang et al. [32]	95.4	96.9	94
Özyurt [25]	96.03	NI	NI
Patil et al. [26]	95.9	95.8	95.8
Razzak [30]	98.8	95.9	96.4
Togacar et al. [20]	97.8	95.7	95.6
Wang et al. [21]	97.7	NI	NI
Yao et al. [6]	95.7	95.7	95.7
Yu et al. [31]	90.5	92.4	86.6
Cheque et al. [8]	98.4	98.4	98.4
Authors	Accuracy (%)	Recall (%)	F1 Score (%)
Xception (Ours)	98.99	97.99	97.99
ResNet (Ours)	99.15	98.31	98.31
EfficientNet (Ours)	99.69	99.40	99.39

5.3. Limitation and Future Work

Over recent years, the use of deep learning has increasingly shown significant potential to improve healthcare. We are now able to perform many tasks that were once the sole

domain of humans. Theoretical advantages to this include accurate and early detection of anomalies, increased diagnostic and therapeutic efficacy, and a reduction in medical error while also decreasing administrative workload and costs. This study focused on the differential count of WBCs as it is one of the most common laboratory tests used. Future work will enhance the framework to include other cells found within the peripheral bloodstream, such as progenitor cells, immature/neoplastic/dysplastic cells; key cells that also act as important indicators of many pathological conditions. The presented work has further implications for other areas of cell and molecular biology where the detection and classification of different types and conditions are needed through microscopy. The presented framework has demonstrated a surpassing classification accuracy rate after only 15 training epochs, even with a relatively small number of training samples, its performance and transferability to other datasets need further exploration. In future work, the authors would like to train the model on a WBC dataset and test its transferability on other datasets with different distributions. In addition, the presented framework in this study is evaluated against CNN-based backbone architectures. The extension of the proposed framework to other deep learning architectures needs to be investigated in future work.

6. Conclusions

This work investigates the white blood cell type classification task and provides an attention-based approach to improve the classification rate and efficiency of the classifier. More specifically, the proposed approach is composed of Attention regularization, texture-aware/attention map generating blocks, and attention-based data augmentation. The proposed approach helps the model to explore various regions of a given WBC image to discover more distinguishing visual representations. Through this process the model learns even tiny differences across different WBC types, leading to higher accuracy rate. The generalizability of the presented method to other CNN-based architectures have been demonstrated through three well-established networks. An extensive set of experiments are carried out to evaluate the performance of the model. The obtained results demonstrate that it could achieve state-of-the-art classification performance 99.69% after only 15 epochs, surpassing its existing counterparts. The transferability of the proposed method to other WBC datasets will be investigated in the future study.

Author Contributions: N.B. came up with the idea, ran the experiments, and wrote the manuscript. M.C. and D.D.D. provided technical feedback. J.-H.P. provided technical feedback and revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: Article processing charges were provided in part by the UCF College of Graduate Studies Open Access Publishing Fund.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: (<https://www.kaggle.com/datasets/paultimothymooney/blood-cells>, accessed on 1 May 2022).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Adewoyin, A. Peripheral blood film—a review. *Ann. Ib. Postgrad. Med.* **2014**, *12*, 71–79. [PubMed]
2. Bonilla, M.A.; Menell, J.S. Disorders of white blood cells. In *Lanzkowsky's Manual of Pediatric Hematology and Oncology*; Elsevier: Amsterdam, The Netherlands, 2016; pp. 209–238.
3. Gurcan, M.N.; Boucheron, L.E.; Can, A.; Madabhushi, A.; Rajpoot, N.M.; Yener, B. Histopathological image analysis: A review. *IEEE Rev. Biomed. Eng.* **2009**, *2*, 147–171. [CrossRef] [PubMed]
4. Dong, N.; Zhai, M.D.; Chang, J.F.; Wu, C.H. A self-adaptive approach for white blood cell classification towards point-of-care testing. *Appl. Soft Comput.* **2021**, *111*, 107709. [CrossRef]
5. Xing, F.; Yang, L. Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: A comprehensive review. *IEEE Rev. Biomed. Eng.* **2016**, *9*, 234–263. [CrossRef] [PubMed]
6. Yao, X.; Sun, K.; Bu, X.; Zhao, C.; Jin, Y. Classification of white blood cells using weighted optimized deformable convolutional neural networks. *Artif. Cells Nanomed. Biotechnol.* **2021**, *49*, 147–155. [CrossRef] [PubMed]

7. Çınar, A.; Tuncer, S.A. Classification of lymphocytes, monocytes, eosinophils, and neutrophils on white blood cells using hybrid Alexnet-GoogleNet-SVM. *SN Appl. Sci.* **2021**, *3*, 503. [\[CrossRef\]](#)
8. Cheuque, C.; Querales, M.; León, R.; Salas, R.; Torres, R. An Efficient Multi-Level Convolutional Neural Network Approach for White Blood Cells Classification. *Diagnostics* **2022**, *12*, 248. [\[CrossRef\]](#)
9. Girdhar, A.; Kapur, H.; Kumar, V. Classification of White blood cell using Convolution Neural Network. *Biomed. Signal Process. Control.* **2022**, *71*, 103156. [\[CrossRef\]](#)
10. Hegde, R.B.; Prasad, K.; Hebbar, H.; Singh, B.M.K.; Sandhya, I. Automated decision support system for detection of leukemia from peripheral blood smear images. *J. Digit. Imaging* **2020**, *33*, 361–374. [\[CrossRef\]](#)
11. Gautam, A.; Singh, P.; Raman, B.; Bhadauria, H. Automatic classification of leukocytes using morphological features and naïve Bayes classifier. In Proceedings of the 2016 IEEE Region 10 Conference (TENCON), Singapore, 22–25 November 2016; pp. 1023–1027.
12. Acevedo, A.; Alférez, S.; Merino, A.; Puigví, L.; Rodellar, J. Recognition of peripheral blood cell images using convolutional neural networks. *Comput. Methods Programs Biomed.* **2019**, *180*, 105020.
13. Hegde, R.B.; Prasad, K.; Hebbar, H.; Singh, B.M.K. Feature extraction using traditional image processing and convolutional neural network methods to classify white blood cells: A study. *Australas. Phys. Eng. Sci. Med.* **2019**, *42*, 627–638. [\[CrossRef\]](#) [\[PubMed\]](#)
14. Ullah, A.; Muhammad, K.; Hussain, T.; Baik, S.W. Conflux LSTMs network: A novel approach for multi-view action recognition. *Neurocomputing* **2021**, *435*, 321–329. [\[CrossRef\]](#)
15. Mellado, D.; Saavedra, C.; Chabert, S.; Torres, R.; Salas, R. Self-improving generative artificial neural network for pseudorehearsal incremental class learning. *Algorithms* **2019**, *12*, 206. [\[CrossRef\]](#)
16. Li, J.; Jin, K.; Zhou, D.; Kubota, N.; Ju, Z. Attention mechanism-based CNN for facial expression recognition. *Neurocomputing* **2020**, *411*, 340–350. [\[CrossRef\]](#)
17. Niu, Z.; Zhong, G.; Yu, H. A review on the attention mechanism of deep learning. *Neurocomputing* **2021**, *452*, 48–62. [\[CrossRef\]](#)
18. Khan, S.; Sajjad, M.; Hussain, T.; Ullah, A.; Imran, A.S. A Review on Traditional Machine Learning and Deep Learning Models for WBCs Classification in Blood Smear Images. *IEEE Access* **2020**, *9*, 10657–10673. [\[CrossRef\]](#)
19. Deshpande, N.M.; Gite, S.; Aluvalu, R. A review of microscopic analysis of blood cells for disease detection with AI perspective. *PeerJ Comput. Sci.* **2021**, *7*, e460. [\[CrossRef\]](#)
20. Togacar, M.; Ergen, B.; Sertkaya, M.E. Subclass separation of white blood cell images using convolutional neural network models. *Elektron. Elektrotehnika* **2019**, *25*, 63–68. [\[CrossRef\]](#)
21. Wang, Q.; Wang, J.; Zhou, M.; Li, Q.; Wen, Y.; Chu, J. A 3D attention networks for classification of white blood cells from microscopy hyperspectral images. *Opt. Laser Technol.* **2021**, *139*, 106931. [\[CrossRef\]](#)
22. Basnet, J.; Alsadoon, A.; Prasad, P.; Aloussi, S.A.; Alsadoon, O.H. A novel solution of using deep learning for white blood cells classification: Enhanced loss function with regularization and weighted loss (ELFRWL). *Neural Process. Lett.* **2020**, *52*, 1517–1553. [\[CrossRef\]](#)
23. Jiang, M.; Cheng, L.; Qin, F.; Du, L.; Zhang, M. White blood cells classification with deep convolutional neural networks. *Int. J. Pattern Recognit. Artif. Intell.* **2018**, *32*, 1857006. [\[CrossRef\]](#)
24. Khan, A.; Eker, A.; Chefranov, A.; Demirel, H. White blood cell type identification using multi-layer convolutional features with an extreme-learning machine. *Biomed. Signal Process. Control.* **2021**, *69*, 102932. [\[CrossRef\]](#)
25. Özyurt, F. A fused CNN model for WBC detection with MRMR feature selection and extreme learning machine. *Soft Comput.* **2020**, *24*, 8163–8172. [\[CrossRef\]](#)
26. Patil, A.; Patil, M.; Birajdar, G. White blood cells image classification using deep learning with canonical correlation analysis. *IRBM* **2021**, *42*, 378–389. [\[CrossRef\]](#)
27. Baghel, N.; Verma, U.; Nagwanshi, K.K. WBCs-Net: Type identification of white blood cells using convolutional neural network. *Multimed. Tools Appl.* **2021**, *4*, 1–17. [\[CrossRef\]](#)
28. Kutlu, H.; Avci, E.; Özyurt, F. White blood cells detection and classification based on regional convolutional neural networks. *Med. Hypotheses* **2020**, *135*, 109472. [\[CrossRef\]](#)
29. Chen, S.; Tan, X.; Wang, B.; Lu, H.; Hu, X.; Fu, Y. Reverse attention-based residual network for salient object detection. *IEEE Trans. Image Process.* **2020**, *29*, 3763–3776. [\[CrossRef\]](#)
30. Imran Razzak, M.; Naz, S. Microscopic blood smear segmentation and classification using deep contour aware CNN and extreme machine learning. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Honolulu, HI, USA, 21–26 July 2017; pp. 49–55.
31. Yu, W.; Chang, J.; Yang, C.; Zhang, L.; Shen, H.; Xia, Y.; Sha, J. Automatic classification of leukocytes using deep neural network. In Proceedings of the 2017 IEEE 12th International Conference on ASIC (ASICON), Guiyang, China, 25–28 October 2017; pp. 1041–1044.
32. Liang, G.; Hong, H.; Xie, W.; Zheng, L. Combining convolutional neural network with recursive neural network for blood cell image classification. *IEEE Access* **2018**, *6*, 36188–36197. [\[CrossRef\]](#)
33. Hegde, R.B.; Prasad, K.; Hebbar, H.; Singh, B.M.K. Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images. *Biocybern. Biomed. Eng.* **2019**, *39*, 382–392. [\[CrossRef\]](#)
34. Huang, Q.; Li, W.; Zhang, B.; Li, Q.; Tao, R.; Lovell, N.H. Blood cell classification based on hyperspectral imaging with modulated Gabor and CNN. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 160–170. [\[CrossRef\]](#)

35. Abou El-Seoud, S.; Siala, M.; McKee, G. Detection and Classification of White Blood Cells Through Deep Learning Techniques. *LearnTechLib* **2020**, 94–105. [[CrossRef](#)]
36. Banik, P.P.; Saha, R.; Kim, K.D. An automatic nucleus segmentation and CNN model based classification method of white blood cell. *Expert Syst. Appl.* **2020**, 149, 113211. [[CrossRef](#)]
37. Baydilli, Y.Y.; Atila, Ü. Classification of white blood cells using capsule networks. *Comput. Med. Imaging Graph.* **2020**, 80, 101699. [[CrossRef](#)] [[PubMed](#)]
38. Hanselmann, H.; Yan, S.; Ney, H. Deep Fisher Faces. BMVC. 2017. Available online: <https://d-nb.info/1194238424/34> (accessed on 17 September 2022).
39. Behera, A.; Wharton, Z.; Hewage, P.R.; Bera, A. Context-aware attentional pooling (cap) for fine-grained visual classification. In Proceedings of the AAAI Conference on Artificial Intelligence, Virtual, 2–9 February 2021; Volume 35, pp. 929–937.
40. Guo, M.H.; Xu, T.X.; Liu, J.J.; Liu, Z.N.; Jiang, P.T.; Mu, T.J.; Zhang, S.H.; Martin, R.R.; Cheng, M.M.; Hu, S.M. Attention mechanisms in computer vision: A survey. *Comput. Vis. Media* **2022**, 8, 331–368. [[CrossRef](#)]
41. Mooney, P. Blood Cell Image. Available online: <https://www.kaggle.com/datasets/paultimothymooney/blood-cells> (accessed on 1 May 2022).
42. Zagoruyko, S.; Komodakis, N. Wide residual networks. *arXiv* **2016**, arXiv:1605.07146.
43. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.
44. Tan, M.; Le, Q. Efficientnet: Rethinking model scaling for convolutional neural networks. In Proceedings of the International Conference on Machine Learning, Nanchang China, 21–23 June 2019; pp. 6105–6114.
45. Zoph, B.; Le, Q.V. Neural architecture search with reinforcement learning. *arXiv* **2016**, arXiv:1611.01578.
46. Marques, G.; Agarwal, D.; de la Torre Díez, I. Automated medical diagnosis of COVID-19 through EfficientNet convolutional neural network. *Appl. Soft Comput.* **2020**, 96, 106691. [[CrossRef](#)]
47. Sutskever, I.; Martens, J.; Dahl, G.; Hinton, G. On the importance of initialization and momentum in deep learning. In Proceedings of the International Conference on Machine Learning, Atlanta, GA, USA, 16–21 June 2013; pp. 1139–1147.
48. Şengür, A.; Akbulut, Y.; Budak, Ü.; Cömert, Z. White blood cell classification based on shape and deep features. In Proceedings of the 2019 International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, 21–22 September 2019; pp. 1–4.