JCBD
JOURNAL OF COMPUTERS AND DIGITAL BUSINESS
Volume 4, Number 2, May 2025, pp. 64-75
Homepage: https://jurnal.delitekno.co.id/index.php/jcbd

Machine Learning (ML) Algorithms for Diagnosing Blood Cancer in Blood Smear Images

Amira Hassan Abed

Business Information Systems Department, Faculty Of Business Administration, Al Ryada University for Science And Technology, El-Sadat City, Egypt

amira.Abed@rst.edu.eg

https://doi.org/10.56427/jcbd.v4i2.756

ARTICLE INFO

Article History
Received: February 21, 2025
Revised: April 25, 2025
Accepted: May 14, 2025

Keywords
Blood Cancer
Deep Learning
Medical Imaging
White Blood Cells
Classification

ABSTRACT

Artificial intelligence (AI), particularly deep learning (DL), has significantly advanced medical image analysis, including the detection and classification of blood cancer through blood smear images. This review explores the state-of-the-art data mining (DM) and DL techniques applied in the identification and classification of white blood cells (WBCs), with a focus on leukemia diagnosis. By systematically analyzing relevant literature from 2014 to 2024, the study highlights key AI algorithms, including traditional machine learning models such as SVM, KNN, and ANN, as well as modern DL architectures like CNN, RCNN, ResNet, and hybrid models. The review evaluates their performance, clinical applicability, and implementation challenges. Particular attention is given to the strengths of DL in feature extraction and classification accuracy, which often surpass traditional DM approaches. Despite these advances, issues such as data scarcity, computational cost, and the need for medical expertise remain major challenges. The study also outlines future directions involving lightweight DL models, transfer learning, and open-access datasets to enhance clinical deployment. Ultimately, this work provides a comprehensive foundation for researchers and developers aiming to improve blood cancer diagnosis through automated medical imaging systems powered by AI.

1. Introduction

Artificial intelligence (AI) approaches have become commonplace for many purposes, including the interpretation of images from hospitals [1]. Medical imaging evaluation has become a crucial component of contemporary healthcare systems, providing intelligent support to medical specialists. Image analysis in medicine processes many image modalities, including CT scans, MRIs, blood smear pictures, and ultrasounds, and is essential in the diagnosis of numerous disorders [2]. Regarding diagnosis and academic objectives, imaging paradigms are essential in medical imaging evaluation because they help identify and categorize both soft and hard tissue of various human organs [3]. Professionals in computer vision have a lot to contribute to the analysis of medical images. AI is important for leukocytes identification of cancer, healthcare data commentary, and retrieving images into computer-assisted diagnostics. Since the efficacy of AI algorithms immediately impacts medical diagnosis and therapy processes, the use of computers in detection depends on them [4], [5]. It facilitates how doctors use conventional working methods in addition helps them even more with diagnosis and therapy. Computer-aided design is greatly impacted by contemporary technological advancements in device fashion, capacity for storage, as well as high rates of processing power. Leukocyte characterization using MRI scans were previously important applications for the Computer-aided design system using AI.

This gives medical professionals crucial knowledge that aids in the diagnosis of many hematic issues, including blood-related cancers. Images processing's primary goal is to efficiently support clinicians during the diagnosis procedure. Within the medical community, leukocytes or WBCs are seen to be a good indicator of most human disorders. Diverse individual illnesses can be identified by variations within the dimensions, form, and hue of leukocytes as well as by changes in their geometrical structure, as seen in smear imaging. Different blood cell types exist, including leukocytes, which are further classified into five subtypes. Different AI methods have been developed for recognizing classifying white blood cells in images of micro blood smears. Traditional methods rely on the laborious, difficult, and error-prone manual examination of WBCs in images of smear blood [6]–[9]. Diagnostic procedures along with suitable therapy are greatly aided by automatic systems [10]. As a result, automated Leukocyte detection throughout microscopically scanned blood sample data has become increasingly prevalent since it may minimize physicians' load and offer precise findings that help professionals during the procedure of diagnosis [11]. Computerized Leukocyte classification in blood sampling imaging could be accomplished primarily through two methods: data mining (DM) and deep learning (DL) approaches. Both offer the possibility of helping produce computerized systems that can improve the efficiency of clinical hematologic [11]–[13]. Figure 1 displays an extensive outline of the DM versus DL Approaches for the classification of white blood cells. By examining leucocytes, several computer-aided diagnosis techniques may automatically identify a variety of hematic conditions, including cancer of blood [12]. DM involves a series of interrelated stages, like ROI segmentation, the extraction of features, and appropriate classifications. For determining the ROI from an image, a range of methods for data mining can be employed [13]. Using the DM techniques, extraction of features is an additional process.

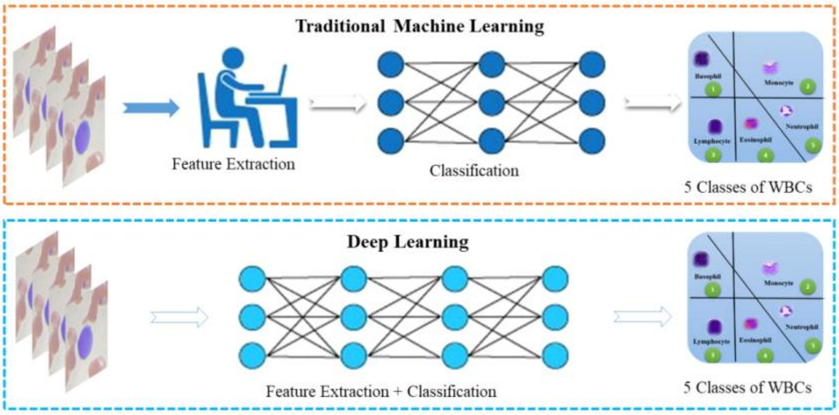

Figure 1. Summary of traditional ML and DL models for the classification of leukocytes.

Nonetheless, since feature dynamics are different, choosing the best feature extractor might be difficult. Exceptional image processing approaches are now available because to the widespread adoption of DL methods, particularly in clinical hematology when blood pictures are used [14]–[24]. This work aims to classify WBCs in blood pictures and offers an extensive overview of the various deep learning algorithms that are currently offered and their medical implications. Microscopic blood smears remain an issue that is rarely addressed, despite the fact that there have been numerous studies on image processing employing methods such as DL along with future trends aimed at CT, MRI, or a chest X [25], [14]. Thus, by examining cutting-edge AI techniques—specifically, leucocytes categorization methods—our study aims to close this gap. The main goal of the researched topic has to present an extensive analysis of the application of DL.

The most popular DM and DL techniques are identified in this suggested study using unconventional classifying methods, which is then evaluated in different groups based on the approach and research topic. This work also contributes to the direction of future study by using DL methods to the classification of leucocytes in images from blood smears. Among the important contributions of the suggested review work are the following ones:

a. This paper explores the numerous possibilities and uses employing AI algorithms in the blood cancer classifications.
b. It also identifies the DL methods that currently exist for leukocytes identification and evaluates their accuracy.
c. Discussing the main obstacles and needs for AI algorithms, then we discuss their future prospects and possibilities for computational imaging research.

The rest of the article is organized in the following manner: Part two goes into great length about the work examination process and reviewing technique. A short overview to the processing of medical images is provided within Part three. The comprehensive overview along with uses regarding the deep learning in leucocyte identification based blood images are presented throughout the fourth part. The prerequisites and present concerns are covered by part five. The sixth part outlines the planned review study's prospects for the future. The seventh part concludes with a discussion of the latest advances in models of DL and the recommendations for further research.

2. Research Methodology

A thorough explanation of the digital libraries employed in the planned review study's formal research process is given in this section. To identify the literature that meets the search requirements and make deliberate use of the accessible digital resources, a planned search process is necessary [15]. Considering the Recommended Items for Reporting over Systematic Reviewing, we used both human and automated searches within the suggested study to find the most pertinent research papers [16]. We used both automatic and manual search functions to retrieve the most pertinent content. Our search approach starts with a machine-generated search across digital records to find pertinent information. Then, image processing and leucocytes classification specialists verify the results. We look seeking papers published between 2014 and 2024 for the suggested survey. Given the selected keywords, both a computerized and customized search is conducted across all the available sources. The authors choose "Medical Imaging processing," "Classification of Leucocytes," OR "WBC's Identification in Bloody Smears Images" as some of these keywords. These steps are used for establishing search keywords:

a. Research questions serve as the basis for the main terms.
b. For the main terms, alternate equivalents or spelling are found.
c. Publications and academic papers that are pertinent are used to find keywords.
d. To switch up equivalents, use the Boolean operator's OR.
e. The AND operator links the important phrases, and the query is created when the key phrases are analyzed to obtain pertinent data within the sources.

The aforementioned terms are examined in every record and recourse, and their pattern is adjusted in accordance with the retrieving of pertinent results. Several terms related to the main topic of the study are derived from the four questions being investigated, which are formulated with consideration for the Patients, Interventions, and Result frameworks [17].

RQ1: Which AI methods are available for classifying leukocytes?
RQ2: How are DL methodologies used in medical research, specifically in the classification of leukocytes?
RQ3: In medical imaging processing area, specifically for leukocyte identification, how are DL approaches applied?
RQ4: What kind of DL is useful and effective for identifying white blood cells in images of blood smears?

3. Results and Discussion

3.1 Evaluation Standards for Extracted Studies

The PRISMA rules are followed to apply inclusion/exclusion criteria to the papers that were initially obtained. The requirements for being included and excluded for weeding out items that are irrelevant are shown in table 1 There are three phases involved in the selecting of articles for research. [18] First, redundant documents are eliminated. Subsequently, the relevance of the article's heading, abstractions, as well search terms is examined, and following a comprehensive examination, the additional studies were considered. Bias and conflict analysis are eliminated by the application of criteria for inclusion and excluded. Throughout the publication's process of selection, an overall of 1893 studies are gathered for assessing the existing research based on the study emphasis. 1319 publications are retrieved after the first process of selection, which includes personal sorting. The works have been reduced employing the pertinent title. After reviewing the headline and summary, these 1319 publications have been reduced down to 894 studies. Following the next step, these studies are selected based on approach and findings, yielding 633 publications. After reviewing the entire contents, the articles are filtered, leaving 286 articles. We assessed the technique, findings, impact factor of the journal, along with citations in order to determine the worth of the remaining papers. Eighty papers are selected for the planned study after each of these parameters has been verified. The quality assessment is carried out following the conclusion of the article's scrutiny procedure (paper inclusion and exclusion). Every research article is evaluated in light of the scrutinization standards. Every research paper is examined, and its quality in relation to every research issue is evaluated. The authors carefully examine and assess each of the chosen papers.

Table 1. Qualifications for papers inclusion and exclusion

Qualifications for inclusion	**Qualifications for exclusion**
The academic publications authored in the English language.	The academic publications were not in English language.

The papers adequately addressed the central idea of the subject under investigation.	Papers that don't adequately clarify the suggested study's primary subject are disqualified.
Just appropriate academic publications with definite conclusions in addition evidence.	The articles that lacked sufficient comparison or unambiguous outcomes.
Current works published between 2016 and 2024.	The academic works that are duplicates.

3.2 Medical Image Processing

Medical imaging is the procedure that can give visual data about the body of a person that helps doctors and physicians diagnose and treat patients effectively [19]. Medical professionals and specialists use a variety of imaging techniques to diagnose conditions and provide treatments. These types of imaging comprise ultrasound, PET, MRI, CT, X-ray, and pictures from samples of blood [25], [19], [20]. These techniques for imaging are crucial to medical image processing because they enable medical professionals automatically identify and diagnose a variety of chronic illnesses through image analysis. In order to do study, they are able to observe various parts of the body [21]. Over the past twenty years, there has been a significant increase in the use of medical imaging in computer-aided diagnosis. Examples of these applications include classification of leukocytes, the classification of tumors, cancers of the breast identification [22], [23], images guiding treatment [24], as well imaging annotations [25], [26]. As a result, it is now an essential component of contemporary healthcare systems [27].

3.3 Blood Cancer Classification Based on Data Mining (DM)

There are plenty newly released scholarly reviews on AI methods for medical image processing in the literature. The following sections cover the most current and pertinent research on DM and DL techniques in medical imaging, namely for identifying the types of leucocytes in blood smear pictures [28]. The most recent and pertinent research is found in the proposed study by filtering the recent publications using the keywords "leucocytes detection" or "leucocytes classification." During our search, we discovered that the use of DM and DL techniques for the study of leukocytes in blood sample images is expanding exponentially. Images initial processing, segmentation; extracting features, selecting features, and classification are all interrelated processes in DM methods. [29] Image enhancing techniques including contrast modification, noise reduction, and sharpening are part of the pre-processing procedure. [30] Prior to image segmentation, all of these procedures have been applied to an input image [31]. There are several different pre-processing methods, including Gabor, low pass, high pass, and median filters. [32] Prior to picture segmentation, these are typically employed for image sharpening, contrast modification, and noise reduction. [33] Several studies have focused on DM in order to detect and categorize leucocytes.

Nevertheless, these methods have made it difficult and time-consuming to recognize nuclei accurately, segment ROI, separate borders to recover overlapping cells, extract robust features, and choose the best features [34]. Using that method, following separating ROI, the following process is feature extracting. Classification in classic supervised learning approaches relies on selecting the optimal features selection algorithm and a robust features descriptor [34], which are the most important elements in ensuring the technique's accuracy and efficiency. Fig. 2 provides a general overview of TML. Leucocytes in microscopic blood sample images have been classified using a variety of traditional supervised learning techniques, including SVM [35], NB [36], KNN [37], and ANN [31].

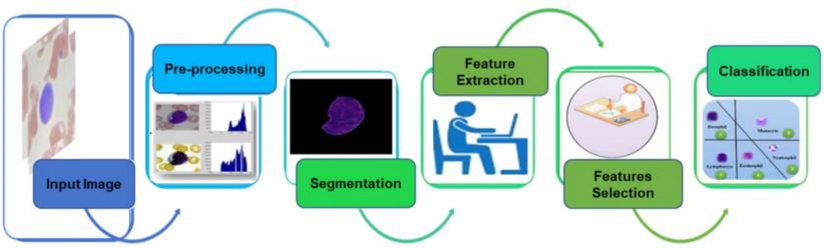

Figure 2. TML based leucocytes segmentation and classification.

Leucocyte classification can be handled by a variety of supervised learning methods, including SVM, ANN, and Decision Trees. Hegde et al. [38] presented a novel method wherein the authors classified WBC cells as either normal or leukemic cells by first segmenting the WBCs and then using SVM. A unique method for segmenting and classifying leukocytes in smear of blood pictures was proposed by Zhao et al. [39]. In order to categorize WBCs under its five subclasses, color correlation and morphological-based segmentation were used, followed by the extraction of texture data and classification utilizing SVM [40]. ANNs and SVMs are important components of medical image processing, but hybrids, Bayesian, Ensembles, KNN, and Tree models are also being utilized to solve problems in various sub-domains of medical imaging, including brain tumor detection, lung tumor identification, leukocyte classification, etc. [41], [42] Abdulkadir et al. [43] suggested a combination of techniques for WBC classification in blood smear images, and Sajjad et al. [12] provided a smartphone-based superior system for healthcare regarding smart cities, in which WBCs are classified using an ensemble multi-class SVM.

Tantikitti et al. [44] suggested an automated technique for diagnosing dengue fever sickness. In blood smear images, leukocytes are segmented using a multi-level threshold approach. This study uses two models of decision trees for classification. The initial model was used to determine if white blood cells either lymphocytes or phagocytes. The second approach classifies the infection with dengue virus as either negative or positive. Work in [45] proposed a novel technique for segmenting WBC nuclei and cytoplasm using simple thresholding. Upon segmentation, some morphological procedures are carried out with elliptical curve fitting, subsequent to feature extraction. The sequentially forward selection strategy is utilized to pick features, and the NB classifier is then used to categorize WBCs. Vogado et al. [46] employed a hybrid technique to classify and segment leukocytes. In their suggested method, CNN features are employed given data for training the SVM classification.

3.4 Blood Cancer Classification Based on Deep Learning

With the use of DL, we may create a framework where a broadly applicable learning technique is used to extract features from data rather than having to be created by individuals [47]. Deep learning produced good results around the domain of image processing for medical imaging, and building a complete network with CNN was very simple [48]. To accomplish various categorization tasks, DM algorithms are trained by using physically extracted characteristics or by using other basic machine learning approaches to acquire features. [49] As a result, researchers are becoming more interested in investigating the advantages of DL approaches for WBC categorization. These days, deep learning is a potent research instrument in the fields of AI, voice analysis, the processing of human language, and imaging in medicine [50]. In the realm of medical image processing, the application of DL as a pattern detection tool is also becoming increasingly important [51]. A recent review on deep learning (DL) oriented image processing [52] claims that numerous individuals now choose to examine healthcare data using DL methods, especially convolution networks. These techniques work especially well in fields where analyzing vast amounts of data requires intelligence comparable to that of a person. Furthermore, for obtaining valuable characteristics given a large number of initially collected data, solid understanding is required [53]. Nevertheless, when a large amount of data needs to be managed effectively, this work becomes difficult and taking time. As illustrated in **Fig. 3**, DL offers complete learning and removes all unnecessary overheads associated with choice of features as well as features descriptor selection.

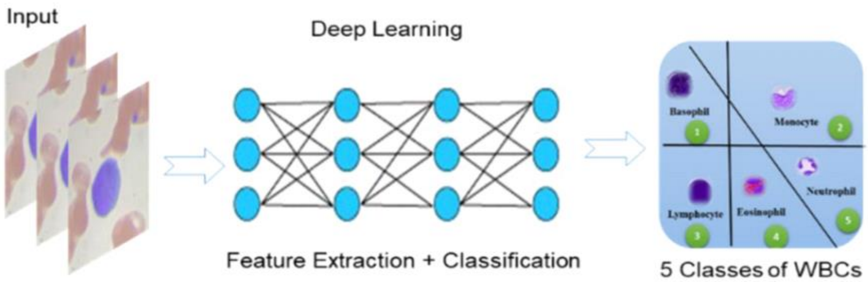

Figure 3. deep learning for Blood Cancer classification

The key benefit of deep learning techniques is their ability to automatically identify and extract linguistically valuable characteristics from unprocessed input. That's the primary distinction among the DL versus DM approaches. [54] A sizable field of study has been drawn to employ DL-based techniques in analyzing medical images due to its unparalleled advantages. DL models fall into various categories, including CNN, DBN, LSTM, as well RNN. CNN is a popular tool for imaging in medicine. Numerous convolutional, collecting, completely linked, and activation function-equipped layers make up CNN. It is trained similarly like every other regular ANN by employing the methods of gradient descent and back-propagation (as seen in **Figure 4**) [55]. A completely interconnected tier comes after the layers of pooling and convolution of an ordinary CNN. The results from the networks classify WBCs into their five separate groups using the Softmax method. Using the BCCD database, Banik al. [56] developed a unique CNN model for classifying WBCs by combining the properties of the initial and final convolutional layers. The CNN structure comprising eight distinct levels was proposed by Choi et al. [57] for the classification of WBCs. An LSM-TIDC technique was presented by Karthikeyan et al. [58] to categorize white blood cells using blood photos. Images are processed beforehand initially, followed by a multidirectional approach is used to extract characteristics of geometry and texture. Ultimately, DCN receives the extracted features in order to effectively and early detect White blood cells. The authors of [11] suggested a methodology that classified White blood cells in blood photos employing CNN. **Table 2** provides a summary of a few recently published studies that used DL in blood cancer identification.

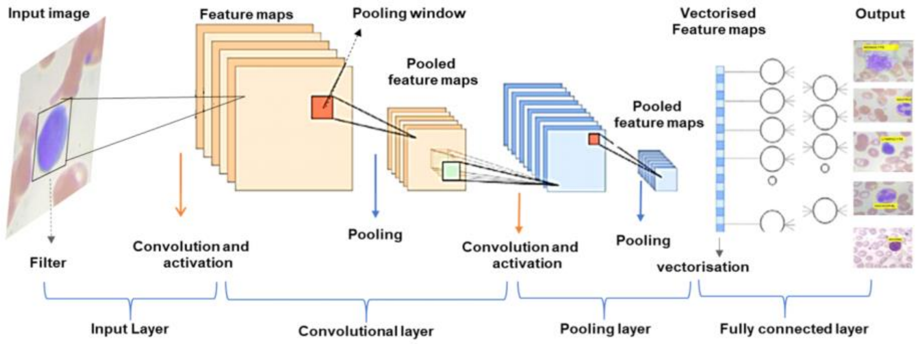

Figure 4. Building blocks of general CNN architecture for leukocytes classification.

Table 2. the major contributions to the Blood Cancer classification using DL models

Ref.	Technique/s	key contribution
[11]	RCNN	A DL-based computer-aided method has been proposed that can quickly identify and categorize the different types of WBCs in blood sample data.
[28]	CNN& RCNN	They used CNN-based intensive machine learning for classification of WBCs, with a precision of 94.71% for RBC & its anomalies detection and 98.68% for WBC. They also used an effective contouring aware CNN as well as a quicker R-CNN centred segmented strategy for dividing RBC and WBCs.
[39]	CNN & RF	With an accuracy of 92.8, an automatic technique utilizing CNN features and a random forest classifier is suggested to separate and categorize WBCs of central blood scans into the appropriate subclasses.
[44]	CNN & SVM	After incorporating a transfer learning strategy into CNN architecture to extract features, an SVM algorithm is utilized to categorize blood cells as normal and leukemic categories achieving overall accuracy of 99.20%.
[59]	CNN	SSD and a YOLOV3 Detector Incremental Improvement Version are used for CNN-based leukocyte recognition. The mean accuracy and mean precision of the procedure were 90.09% and 93.10%, respectively.
[60]	ResNet50	Prior to network training, leukocyte-mask is first applied at the pixel level. Next, the ROI is segmented, and lastly, the trained model is applied to classify leukocytes in blood smear images with 0.98 of accuracy and 0.99 precision on datasets 1 and 2.
[61]	AlexNet	White blood cell categorization achieved a standard sensitivities and accuracy of 99% through a comparative analysis of CNN-based features extraction and classical learning.
[62]	CNN	With an average accuracy in classification of 96.2%, the primary contribution of this study is an approach to classification built around a trained CNN that can identify eight distinct categories of cells in the blood floating in microscopic blood samples images.
[63]	DRNN	Using a fine-grained method based on the deep residual learning theory, leukocytes are classified from microscopic blood test images. The typical accuracy during testing was 76.84%, the top-1 accuracy was 77.80%, and the top-5 accuracy was 98.75%.
[64]	fine-tuned CNN	A refined CNN method which is capable of classifying blood smear images into their subgroups of cells in blood with a mean accuracy of 0.97 was proposed by the authors of this paper.
[65]	RCNN & AlexNet	RBCs and other cells, including infected and WBCs, can be automatically detected and classified using a two-stage DL-based approach, which has a 72% overall accuracy rate for non-difficult infected cells.
[66]	ResNet & Inception	Using a sequentially proposed DL framework, WBCs were classified into four main types. ResNet 101 and ResNet V1 152 achieved an accuracy rate of 99.46% and 99.84%, respectively, after fine-tuning all layers over 3000 epochs.
[67]	CNN	A region of interest identification technique based on CNN and super pixels is suggested for cervical cancer cell segmentation. The experimental results showed a 94.50% accuracy rate.

[68]	CNN	CNN models that have been pre-trained are utilized as feature extractors. The extracted features from completely connected layers were then combined, and MRMR features selection and an extraordinary leaning machine are added to classify WBCs with reported accuracy rates of 99% and 97.37%.
[69]	DCNN	A brand-new method for classifying ALL in colored bone marrow pictures is put forth. After segmenting the lymphocytes using a straightforward threshold method, the segment lymphocytes are subjected to CNN-based DL algorithms, which successfully categorize ALL into L1, L2, L3, and normal lymphocytes with an accuracy of 97.78%.
[70]	DCNN	In order to categorize WBCs into their five distinct groups, one study suggested using CNN. The validation set has an overall accuracy of 91%, whereas the training set achieves an accuracy of 99%. The reported sensitivity and specificity are 91% and 97%, respectively.
[71]	AlexNet	After pre-processing the blood smear image, characteristics are extracted using a CNN model called AlexNet. Once AlexNet has been fine-tuned, it are capable of accurately identifying blood cells as normal and pathological cells.
[72]	DCGAN & ResNet	A comprehensive detection framework built on DC-GAN and ResNet was proposed by the authors for blood cell images, and its accuracy was 91.7%. Additionally, transferable learning is used to increase the accuracy of the model.
[73]	capsule network	In order to classify WBCs, a novel deep learning technique is proposed. First, data is augmented using various methods of augmentation to increase the amount of data. Next, a DL-based capsule network serves for classifying WBCs into its five subclasses, reaching 96.86% accuracy.
[74]	R-CNN	In order to help medical professionals diagnose various hematic diseases automatically, a deep learning-based method called Faster R-CNN is proposed to detect WBCs and RBCs. The accuracy of the test is 83.25% for RBCs and 99% for Eosinophils, with the lowest testing accuracy not falling below 66%.
[75]	FPM & SOYOLO DL	This work built a hybrid method for WBC identification by combining the SOYOLO DL framework with FPM. The high-resolution WBC blood smear images are obtained using the FPM approach, and SOYOLO detects WBCs in these images with a 100% recall and 100% precision rate.
[76]	DCNN	Lightweight DL models which effectively categorize RBCs into three groups—sickle cells, normal, and other blood content—were provided in this study.
[77]	hybrid DL	A hybrid DL-based approach showing 95.17% sensitiveness, 98.58% specificity, and 96.17% overall accuracy is suggested for the classification of adolescent leukemic blast and regular cells.
[78]	AlexNet	Blob detection and DL approach are used to identify WBCs in tiny blood smear images. The accuracy of this new technique is 100% for SMC-IDB and IUMS-IDB datasets, 99.7% for ALL-IDB in white cell detection, and 94.1% for leukemidical image analyzing identification.

3.5 Current Challenges and Requirements

We discovered the primary demands and research problems, a number of important characteristics, uses, and benefits of DL approaches for image processing, especially for blood cancer detection, thorough analysis of the existing literature. Some common and potent machine learning algorithms have been established in the past couple of decades for healthcare image processing, including the identification of brain tumors from MRIs, the detection of blood cancer using smear visuals, or the cancer of lungs detection using CT scans [79]. However, there are still a few major obstacles that the scientific community must either accept or work towards overcoming. Specialized medical expertise, efficient DL approaches, and superior data sets that are publicly available are some of these issues.

A number of the difficulties stem from the conceptual and mathematical foundations of numerous deep learning approaches. Unsupervised and supervised approaches have to be developed for addressing these obstacles [80]. To prevent these problems, the effectiveness of unsupervised or supervised image processing techniques can be undermined. Transitioning from supervised to unsupervised learning methodologies has difficulties as well, as it impacts system performance and accuracy. Applications and platforms for image processing that use deep learning techniques remain not exactly ideal, with plenty of opportunity for development. The lack of publicly available data represents the main issue facing the medical imaging industry. In order to solve this problem, scholars must persuade health institutions to provide healthcare data; high-quality data that is available to the general public to feed research purposes can be fascinating. Efforts that promote open datasets from various health organizations across the globe are additionally welcomed; however, certain operations (such as clinical data and limited availability of data) are needed as well. In each of these scenarios, incentive systems may be linked to monetary gains, amusement, or services rendered to these organizations in exchange for high quality data. Whenever large volumes of data are available, the problem gets more intriguing for research, similar to other domains (such as summarization of videos, Internet of Things, controlling energy). For certain image processing applications, gathering large, accurate data sets having truth labels is essential. Such data sets can also be utilized for evaluating and organizing other tournaments. [81]

The training of predictive algorithms is the main problem in image processing and blood cancer identification and classification. To solve this issue, an optimal learning method with a more balanced capacity for generality and a computationally effective heuristics approach need to be developed. For the development of an algorithm with impressive generality skills, an approach to learning which employs either real or arbitrary labels, offers helpful instruments to deal with accessible datasets, along with effective algorithms for training is required. In the past few years, training utilizing DNNs achieved tremendous experimental success in a range of tasks associated with image processing. Even though this is a difficult, non-convex optimization issue, straightforward techniques like random gradient descent may uncover workable solutions that reduce the error generated by training. Regardless of the total quantity of parameters is far greater than the quantity of training data [49], the networks learnt in this manner show remarkable capacity to generalize. Merely reducing the training errors has become insufficient when learning a machine learning model. Selecting the incorrect global minima may also cause the estimator to behave poorly when it comes to generalization. Under such circumstances, the strategy employed to reduce the error in training essentially determines the behavior of generalization. Various global minimums with varying capacities for generality will result of numerous optimization algorithmic decisions, including beginning, updating guidelines, training rate, and terminating condition. Without medical competence, the current DM and DL approaches are not reliable enough to be used in real-world health diagnosing platforms [32]. To develop a model for learning to assist with medical imaging processing and classifying blood cancer, it needs both technical and expert knowledge. We must investigate highly accurate and reliable techniques that can be applied in practical healthcare situations without the requirement for medical specialists.

3.6 Future Research Directions

A significant amount of work is required from the healthcare business and research society to contribute to MAI, and specifically leukocytes that processing in smear of blood pictures, given the considerable obstacles facing the image processing field as described in section V. We have concentrated on the most commonly brought up issue in this study, which is the lack of processed medical images and leucocyte classification datasets. To enhance image processing and WBC identification and categorization in blood smear images, a comprehensive data enhancement method and transfer learning models are advised. To enhance the available data, a variety of data augmentation techniques are employed, such as generative adversarial networks, style transfer, and traditional picture alterations. The latest developments in deep learning models will be helpful for computer-aided diagnostic applications in the future.

These models are currently accessible on several open-source software platforms. However, due to a lack of clinical and medical competence, choosing and training a suitable machine-learning algorithm for a particular MAI problem can be difficult. Recent advancements in DL have led to improved performance in tasks involving the detection of brain tumors, the classification of leukocytes, the detection of breast cancer, and other MIDICAL IMAGE PROCESSING tasks. These include GANs, the R-CNN quicker R-CNN, as well as the fusion of TML and DL models. High memory requirements and computational costs are, however, their main worries. Therefore, highly computational AI algorithms have to be discovered for leukocytes identification in blood test data. Moreover, these lightweight models are easily implementable on devices with limited resources. DNNs Models can take the place of conventional methods. Due to the fact that they can classify leucocytes into five groups and their ability to simplify the model-building process, full models are also becoming more and more popular with the recent progress of CNNs [47]. These models compete with complex DNN-based models and depend on data-driven learning techniques. Various end-to-end architectures, including CNN-based models and attention-based methods [30], are also widely used for leucocyte detection in blood images. The research community mostly uses subjective evaluation methodologies in the field of image processing. But this work is difficult, time-consuming, and error-prone at times. Therefore, more investigation is needed to examine general assessment methods that can automatically gauge how well ML models for processing images perform from various angles.

3.7 Research Contribution

The current research included a thorough analysis of artificial intelligence methods for classifying leukocytes in blood smear images. We looked at various DM and DL methods for identifying WBCs in blood sample pictures. The information was gathered from primary research projects that were released between 2014 and 2024. 87 primary research (articles published in publications, texts, seminars, and online sources) outlining artificial intelligence approaches for leucocyte categorization in blood smear photos and their applications in medical diagnostics are found in the literature reviewed for this study. Both DM and DL (as AI techniques) techniques have done similarly regarding overall contributions in the processing of medical images, according to our evaluation of the studies. The collection of all of the data collected during the present investigation will aid the scientific community by determining where to begin with subsequent studies using DM and DL algorithms for the processing of medical images. In the future, these techniques will make significant contributions to the advancement in clinical imaging, natural language processing, and speech analysis. Aside from WBCs, DM and DL approaches are utilized to identify and classify several medical image processing domains, such as MRI, CT, X-ray, and ultrasound image analysis. In this study, we studied various DM and DL algorithms used to analyze blood smear images, including SVM ANNs, Ensembles, Bayesians, neuro-fuzzy, hybrids, DL, and CNNs [12], [42].

4. Conclusion

This work aims to identify various uses of AI in the processing of medical scans and leucocyte classifications in blood sample scans. The goal of this work is to obtain insight into complicated aspects of AI by collecting and analyzing publications with the objective to help further studies in the discipline of the processing of medical images. This study demonstrates that much more research is needed to explore the use of AI approaches for meaningful clinical scans processing as well as leucocyte categorization in blood smear images. This study also sought to explore uses for advanced DL models outside blood cancer classification. However, it has been discovered that practically all other diagnostic applications are directly or indirectly tied to AI. In the field of medical imaging, smear of blood images are an emerging topic that has received significant interest from the research community over the last three decades. This article presents the typical contributions and potential uses of deep learning and machine learning in medical image processing. Furthermore, we identified existing obstacles, future goals, and solutions for the advancement of DL models in medical image processing, including WBC categorization in blood smear pictures. In the future, we plan to broaden our survey to include diverse processed medical image areas that include MRI, CT, Ultrasound, and X-ray pictures, using the potential of DM and DL approaches.

References

[1] P. Pandey, S. Pallavi, and S. C. Pandey, ''Pragmatic medical image analysis and deep learning: An emerging trend,'' in Advancement of Machine Intelligence in Interactive Medical Image Analysis. Singapore: Springer, Jan. 2020, pp. 1–18.

[2] S. Asgari Taghanaki, K. Abhishek, J. P. Cohen, J. Cohen-Adad, and G. Hamarneh, ''Deep semantic segmentation of natural and medical images: A review,'' Artif. Intell. Rev., vol. 4, pp. 1–42, Jun. 2020.

[3] A. Elshewey et al., ''Enhancing heart disease classification based on greylag goose optimization algorithm and long short-term memory,'' Scientific Reports, vol. 15, p. 1277, 2025.

[4] H. P. Chan, L. M. Hadjiiski, and R. K. Samala, ''Computer-aided diagnosis in the era of deep learning,'' Med. Phys., vol. 47, no. 5, pp. e218–e227, 2020.

[5] H. Amira and N. Mona, ''Diabetes Disease Detection through Data Mining Techniques'', International Journal of Advanced Networking and Applications (IJANA), vol. 11, no. 1, pp. 4142-4149, 2019.

[6] M. Sajjad et al., ''Computer aided system for leukocytes classification and segmentation in blood smear images,'' in Proc. Int. Conf. Frontiers Inf. Technol. (FIT), Dec. 2016, pp. 99–104.

[7] N. S. M. Raja et al., ''Contrast enhanced medical MRI evaluation using tsallis entropy and region growing segmentation,'' J. Ambient Intell. Humanized Comput., vol. 2018, pp. 1–12, May 2018.

[8] Y. Duan et al., ''Leukocyte classification based on spatial and spectral features of microscopic hyperspectral images,'' Opt. Laser Technol., vol. 112, pp. 530–538, Apr. 2019.

[9] H. Cao, H. Liu, and E. Song, ''A novel algorithm for segmentation of leukocytes in peripheral blood,'' Biomed. Signal Process. Control, vol. 45, pp. 10–21, Aug. 2018.

[10] E. Abdulhay et al., ''Computer aided solution for automatic segmenting and measurements of blood leucocytes using static microscope images,'' J. Med. Syst., vol. 42, no. 4, p. 58, Apr. 2018.

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

[12] 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.

[13] F. Xing and L. Yang, ''Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: A comprehensive review,'' IEEE Rev. Biomed. Eng., vol. 9, pp. 234–263, Jan. 2016.

[14] T. Zhou, S. Ruan, and S. Canu, ''A review: Deep learning for medical image segmentation using multi-modality fusion,'' Array, vol. 3–4, Sep. 2019, Art. no. 100004.
[15] T. Hussain, K. Muhammad, W. Ding, J. Lloret, and S. W. Baik, ''A comprehensive survey of multi-view video summarization,'' Pattern Recognit., vol. 109, Jan. 2020, Art. no. 107567.
[16] I. Balki et al., ''Sample-size determination methodologies for machine learning in medical imaging research: A systematic review,'' Can. Assoc. Radiol. J., vol. 70, no. 4, pp. 344–353, Nov. 2019.
[17] A. Littlewood and D. Kloukos, ''Searching the literature for studies for a systematic review. Part 1: Identifying search concepts in a question,'' Amer. J. Orthodontics Dentofacial Orthopedics, vol. 155, no. 2, pp. 299–301, Feb. 2019.
[18] S. Marwa, A. H. Abed, and M. A. Mahmoud, ''A systematic review for the determination and classification of the CRM critical success factors supporting with their metrics'', Future Computing and Informatics Journal, vol. 3, pp. 398–416, 2018.
[19] S. M. Anwar et al., ''Medical image analysis using convolutional neural networks: A review,'' J. Med. Syst., vol. 42, no. 11, p. 226, Nov. 2018.
[20] P. Lakhani et al., ''Hello world deep learning in medical imaging,'' J. Digit. Imag., vol. 31, no. 3, pp. 283–289, Jun. 2018.
[21] H. A. Amira and M. Essam, ''Modeling Deep Neural Networks for Breast Cancer Thermography Classification: A Review Study'', International Journal of Advanced Networking and Applications (IJANA), vol. 13, no. 2, pp. 4939–4946, 2021.
[22] H. Amira, ''Deep Learning Techniques for Improving Breast Cancer Detection and Diagnosis'', International Journal of Advanced Networking and Applications, vol. 13, no. 06, pp. 5197–5214, 2022.
[23] H. Amira, M. Essam, O. P. Jena, and A. Elngar, ''A Comprehensive Survey on Breast Cancer Thermography Classification Using Deep Neural Network'', in Machine Learning and Deep Learning in Medical Data Analytics and Healthcare Applications. Routledge, CRC Press, Taylor and Francis Group, 2022, pp. 250–265.
[24] H. A. Amira, ''The Applications of Deep Learning Algorithms for Enhancing Big Data Processing Accuracy'', International Journal of Advanced Networking and Applications, vol. 16, no. 02, pp. 6332–6341, 2024.
[25] H. A. Amira, N. Mona, L. A. Laila, and E. Laila, ''Applications of IoT in Smart Grids Using Demand Respond for Minimizing On-peak Load'', International Journal of Computer Science and Information Security (IJCSIS), vol. 19, no. 8, 2021.
[26] H. Amira and F. Hany, ''The Evaluation of Electronic Human Resources (eHR) Management based Internet of Things using Machine Learning Techniques'', International Journal of Advanced Networking and Applications, vol. 16, no. 03, pp. 6437–6452, 2024.
[27] A. Abed and N. Mona, ''Business Intelligence (BI) Significant Role in Electronic Health Records - Cancer Surgeries Prediction: Case Study'', International Journal of Advanced Networking and Applications, vol. 13, no. 06, pp. 5220–5228, 2022.
[28] M. I. Razzak and S. Naz, ''Microscopic blood smear segmentation and classification using deep contour aware CNN and extreme machine learning,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops (CVPRW), Jul. 2017, pp. 49–55.
[29] S. Marwa, M. A. Mahmoud, and A. H. Abed, ''The Success Implementation CRM Model for Examining the Critical Success Factors Using Statistical Data Mining Techniques'', International Journal of Computer Science and Information Security (IJCSIS), vol. 15, no. 1, pp. 455–475, 2017.
[30] M. Ahmed et al., ''Optimized Deep Learning for Potato Blight Detection Using the Waterwheel Plant Algorithm and Sine Cosine Algorithm'', Potato Res., 2024. doi: 10.1007/s11540-024-09735-y.
[31] S. Safuan, M. Md Tomari, and W. Wan Zakaria, ''White blood cell (WBC) counting analysis in blood smear images using various color segmentation methods,'' Measurement, vol. 116, pp. 543–555, Feb. 2018.
[32] A. H. Amira, ''Enhancing Big Data Processing Performance using Cutting-Edge Deep Learning Algorithms'', Al-Ryada Journal for Computational Intelligence and Technology (ARJCIT), vol. 1, no. 1, pp. 39–53, 2024.
[33] A. Mohamed and A. H. Amira, ''A comprehensive investigation for Quantifying and Assessing the Advantages of Blockchain Adoption in banking industry'', in 2024 6th International Conference on Computing and Informatics (ICCI), IEEE, pp. 322–333.
[34] S. Kar, K. Das Sharma, and M. Maitra, ''Optimised feature selection for identification of carcinogenic leukocytes using weighted aggregation based transposition PSO,'' IETE J. Res., vol. 4, pp. 1–14, Nov. 2019.

[35] J. Li et al., ''Identification of leukemia stem cell expression signatures through Monte Carlo feature selection strategy and support vector machine,'' Cancer Gene Therapy, vol. 27, no. 1–2, pp. 56–69, Feb. 2020.
[36] H. Gao, X. Zeng, and C. Yao, ''Application of improved distributed naïve Bayesian algorithms in text classification,'' J. Supercomput., vol. 75, no. 9, pp. 5831–5847, Sep. 2019.
[37] E. Purwanti and E. Calista, ''Detection of acute lymphocyte leukemia using k-nearest neighbor algorithm based on shape and histogram features,'' J. Phys., Conf. Ser., vol. 853, May 2017, Art. no. 012011.
[38] R. B. Hegde et al., ''Automated decision support system for detection of leukemia from peripheral blood smear images,'' J. Digit. Imag., vol. 7, pp. 1–14, Nov. 2019.
[39] J. Zhao et al., ''Automatic detection and classification of leukocytes using convolutional neural networks,'' Med. Biol. Eng. Comput., vol. 55, no. 8, pp. 1287–1301, Aug. 2017.
[40] M.-C. Su, C.-Y. Cheng, and P.-C. Wang, ''A neural-network-based approach to white blood cell classification,'' Sci. World J., vol. 2014, Jan. 2014, Art. no. 796371.
[41] A. H. Abed, ''Recovery and Concurrency Challenging in Big Data and NoSQL Database Systems'', International Journal of Advanced Networking and Applications, vol. 11, no. 04, pp. 4321–4329, 2020.
[42] S. Naglaa and H. Amira, ''Big Data with Column Oriented NOSQL Database to Overcome the Drawbacks of Relational Databases'', International Journal of Advanced Networking and Applications (IJANA), vol. 11, no. 5, pp. 4423–4428, 2020.
[43] A. Sengur, Y. Akbulut, U. Budak, and Z. Comert, ''White blood cell classification based on shape and deep features,'' in Proc. Int. Artif. Intell. Data Process. Symp. (IDAP), Sep. 2019, pp. 1–4.
[44] S. Tantikitti and W. Premchaiswadi, ''Image processing for detection of dengue virus based on WBC classification and decision tree,'' in Proc. 13th Int. Conf. ICT Knowl. Eng., Nov. 2015, pp. 84–89.
[45] J. Prinyakupt and C. Pluempitiwiriyawej, ''Segmentation of white blood cells and comparison of cell morphology by linear and naïve Bayes classifiers,'' Biomed. Eng. OnLine, vol. 14, no. 1, p. 63, Dec. 2015.
[46] L. Vogado et al., ''Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification,'' Eng. Appl. Artif. Intell., vol. 72, pp. 415–422, Jun. 2018.
[47] A. Amira, N. Mona, and L. A. Laila, ''A conceptual Framework for Minimizing Peak Load Electricity using Internet of Things'', Int. J. of Computer Science and Mobile Computing, vol. 10, no. 8, pp. 60–71, 2021.
[48] A. H. Amira, N. Mona, and W. S. Walaa, ''The Future of Internet of Things for Anomalies Detection using Thermography'', International Journal of Advanced Networking and Applications (IJANA), vol. 11, no. 03, pp. 4294–4300, 2019.
[49] A. Amira, ''Internet of Things (IoT) Technologies for Empowering E-Education in Digital campuses of Smart Cities'', International Journal of Advanced Networking and Applications, vol. 13, no. 2, pp. 4925–4930, 2021.
[50] A. Amira and M. Bahloul, ''Authenticated Diagnosing of COVID-19 using Deep Learning-based CT Image Encryption Approach'', Future Computing and Informatics Journal, vol. 8, no. 2, Art. 4, 2023.
[51] H. Greenspan, B. van Ginneken, and R. M. Summers, ''Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique,'' IEEE Trans. Med. Imag., vol. 35, no. 5, pp. 1153–1159, May 2016.
[52] G. Wang, ''A perspective on deep imaging,'' IEEE Access, vol. 4, pp. 8914–8924, 2016.
[53] H. Amira, H. Ahmed, and M. Mohamed, ''Authentication in Cloud Computing Environments'', Multicriteria Algorithms with Applications, vol. 5, pp. 59–67, 2024.
[54] H. Amira et al., ''The Applications of Digital Transformation towards Achieving Sustainable Development Goals: Practical Case Studies in Different Countries of the World'', Journal of Artificial Intelligence & Metaheuristics, vol. 07, no. 01, pp. 53–66, 2024.
[55] A. Ullah, K. Muhammad, T. Hussain, and S. W. Baik, ''Conflux LSTMs network: A novel approach for multi-view action recognition,'' Neurocomputing, Dec. 2020.
[56] P. P. Banik, R. Saha, and K.-D. Kim, ''An automatic nucleus segmentation and CNN model based classification method of white blood cell,'' Expert Syst. Appl., vol. 149, Jul. 2020, Art. no. 113211.
[57] J. W. Choi et al., ''White blood cell differential count of maturation stages in bone marrow smear using dual-stage convolutional neural networks,'' PLoS ONE, vol. 12, no. 12, Dec. 2017, Art. no. e0189259.
[58] M. Karthikeyan and R. Venkatesan, ''Interpolative leishman-stained transformation invariant deep pattern classification for white blood cells,'' Comput., vol. 5, pp. 1–11, Jan. 2020.

[59] J. Liu et al., ''NTU RGB+D 120: A large-scale benchmark for 3D human activity understanding,'' IEEE Trans. Pattern Anal. Mach. Intell., vol. 42, no. 10, pp. 2684–2701, Oct. 2019.

[60] H. Fan et al., ''LeukocyteMask: An automated localization and segmentation method for leukocyte in blood smear images using deep neural networks,'' J. Biophotonics, vol. 12, no. 7, Jul. 2019, Art. no. e201.

[61] R. B. Hegde et al., ''Feature extraction using traditional image processing and convolutional neural network methods to classify white blood cells: A study,'' Australas. Phys. Eng. Sci. Med., vol. 42, no. 2, pp. 627–638, Jun. 2019.

[62] A. Acevedo et al., ''Recognition of peripheral blood cell images using convolutional neural networks,'' Comput. Methods Programs Biomed., vol. 180, Oct. 2019, Art. no. 105020.

[63] F. Qin, N. Gao, and A. Grudtsin, ''Fine-grained leukocyte classification with deep residual learning for microscopic images,'' Comput. Methods Programs Biomed., vol. 162, pp. 243–252, Aug. 2018.

[64] P. Tiwari et al., ''Detection of subtype blood cells using deep learning,'' Cognit. Syst. Res., vol. 52, pp. 1036–1044, Dec. 2018.

[65] J. Hung and A. Carpenter, ''Applying faster R-CNN for object detection on malaria images,'' in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, vol. 2017, pp. 56–61.

[66] M. Habibzadeh and M. Totonchi, ''Automatic white blood cell classification using pre-trained deep learning models: Resnet and inception,'' in Proc. 10th Int. Conf. Mach. Vis. (ICMV), 2018, vol. 10696, Art. no. 106960Z.

[67] Y. Song et al., ''A deep learning based framework for accurate segmentation of cervical cytoplasm and nuclei,'' in Proc. 36th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc., Aug. 2014, pp. 2903–2906.

[68] F. Özyurt, ''A fused CNN model for WBC detection with MRMR feature selection and extreme learning machine,'' Soft Comput., vol. 24, pp. 8163–8172, Jun. 2019.

[69] A. Rehman et al., ''Classification of acute lymphoblastic leukemia using deep learning,'' Microsc. Res. Techn., vol. 81, no. 11, pp. 1310–1317, Nov. 2018.

[70] D. Bani-Hani et al., ''Classification of leucocytes using convolutional neural network optimized through genetic algorithm,'' in Proc. 7th Annu. World Conf. Soc. Ind. Syst. Eng., Oct. 2018, pp. 1–6.

[71] M. Loey, M. Naman, and H. Zayed, ''Deep transfer learning in diagnosing leukemia in blood cells,'' Computers, vol. 9, no. 2, p. 29, Apr. 2020.

[72] L. Ma et al., ''Combining DC-GAN with ResNet for blood cell image classification,'' Med. Biol. Eng. Comput., vol. 4, pp. 1–14, Mar. 2020.

[73] Y. Y. Baydilli and Ü. Atila, ''Classification of white blood cells using capsule networks,'' Comput. Med. Imag. Graph., vol. 80, Mar. 2020, Art. no. 101699.

[74] R. R. Tobias et al., ''Faster R-CNN model with momentum optimizer for RBC and WBC variants classification,'' in Proc. IEEE 2nd Global Conf. Life Sci. Technol. (LifeTech), Mar. 2020, pp. 235–239.

[75] X. Wang et al., ''SO-YOLO based WBC detection with Fourier ptychographic microscopy,'' IEEE Access, vol. 6, pp. 51566–51576, 2018.

[76] L. Alzubaidi et al., ''Deep learning models for classification of red blood cells in microscopy images to aid in sickle cell anemia diagnosis,'' Electronics, vol. 9, no. 3, p. 427, 2020.

[77] S. Hosseinzadeh Kassani et al., ''A hybrid deep learning architecture for leukemic B-lymphoblast classification,'' 2019, arXiv:1909.11866.

[78] C. Di Ruberto, A. Loddo, and G. Puglisi, ''Blob detection and deep learning for leukemic blood image analysis,'' Appl. Sci., vol. 10, no. 3, p. 1176, Feb. 2020.

[79] M. I. Razzak, S. Naz, and A. Zaib, ''Deep learning for medical image processing: Overview, challenges and the future,'' in Classification BioApps. Ithaca, NY, USA: Cornell Univ., 2018, pp. 323–350.

[80] H. Abed, Amira, ''The Techniques of authentication in the Context of Cloud Computing'', Int. J. Advanced Networking and Applications, vol. 16, no. 04, pp. 6515–6522, 2025.

[81] A. H. Amira, N. Mona, and B. S. Basant, ''The Principle Internet of Things (IoT) Security Techniques Framework Based on Seven Levels IoT's Reference Model'', in Proceedings of Internet of Things—Applications and Future ITAF 2019. Springer, 2019, vol. 114, pp. 131–145.