Mustansiriyah Journal of Pure and Applied Sciences, Print ISSN: 2957-9910. Online ISSN: 2957-9929, Vol. 2, No. 3, 2024, Pages 33-49
DOI:

MJPAS MUSTANSIRIYAH JOURNAL OF PURE AND APPLIED SCIENCES

Journal homepage: https://mjpas.uomustansiriyah.edu.iq/index.php/mjpas

RESEARCH ARTICLE - Computer

Leukemia Detection Using Machine Learning Algorithms: Current Trends and Future Directions (Literature Survey)

Amal Abdulbaqi Maryoosh [1,2,*], Saeid Pashazadeh [2]

[1] Department of Computer Science, Faculty of Education, Mustansiriyah University, Iraq.

[2] Department of Computer Engineering, Faculty of Electrical and Computer Engineering, University of Tabriz, Iran.

[*] Corresponding author E-mail: amalmaryoosh@uomustansiriyah.edu.iq

Article Info.

Article history:

Received
10 December 2023

Accepted
07 February 2023

Publishing
30 June 2024

Abstract

The process of medical diagnosis of leukemia cases is a complex process and requires diligent efforts. It is done by examining samples under a microscope and distinguishing the number, shape, size, and morphological features of white blood cells. This process takes a long time to predict leukemia. The professional skills and experience of the pathologist may also influence this procedure as a large number of overlapping structures and conditions, distractions, fatigue, and limitations in the human visual system can lead to an inappropriate diagnosis. Machine and deep learning approaches have been found to be effective strategies for enhancing the precision and efficiency of diagnosis and classification of medical images, including those of microscopic blood cells. In this paper, we will review the recent studies of leukemia detection and/or classification in the period of (2015 - 2023) in machine and deep learning and the combining of them. The review of these schemes is carried out in a systematic manner. In order to achieve the desired objective, segmentation schemes can be classified into four main categories: supervised machine learning techniques, unsupervised machine learning techniques, deep learning approaches, and traditional image processing techniques. Classification approaches can be categorized into three main groups: classical machine learning, deep learning, and a combination of both. CNN-based categorization systems can be further classified into three categories: conventional CNN, transfer learning, and additional developments in CNN. This paper also provides a concise analysis of these strategies and their significance in the categorization of leukemia. Ultimately, a comprehensive examination is conducted to elucidate the current state of research in this particular domain.

The official journal published by the College of Education at Mustansiriya University

Keywords: Leukemia, Deep learning, Machine learning, Classification, Detection.

1. Introduction

Leukemia is a particular kind of cancer that impacts the hematopoietic system, specifically the blood and bone marrow. The bone marrow of individuals diagnosed with leukemia demonstrates a notable and uncontrolled increase in the replication of atypical cells. In contrast to alternative medical conditions, leukemia does not typically manifest as an X-ray-detectable mass (tumor). It occurs when the bone marrow produces new blood cells. Hematopoietic stem cells (HSCs) are a type of precursor cell responsible for generating all blood cell types. These cells progress through multiple stages of development before reaching their mature form. Hematopoietic cells undergo proliferation and division within the bone marrow, resulting in the production of red blood cells (RBCs), white blood cells (WBCs), and platelets (PLTs).

However, in the case of an individual afflicted with Leukemia, there is an abnormal and uncontrolled proliferation of one of these specific types of blood cells. The cells referred to as Leukemia cells exhibit an aberrant behavior by occupying the intracellular space of the Bone Marrow [1]. In the domains of pathology and hematology, microscopic analysis of white blood cells, or leukocytes, in human blood smears is a crucial issue. This study or research is especially crucial for diagnosing immune-related diseases such as autoimmune anemia, allergies, and polycythemia, as well as blood disorders like leukemia, anemia, and polycythemia. As part of the diagnosis process, a differential count is usually done, which compares the amounts of the five different types of white blood cells: basophils, lymphocytes, monocytes, eosinophils, and neutrophils (see Figure 1 and Table 1) [2, 3].

Table1. Characteristics of white blood cells [2, 4]

WBCs	% In blood	Description	Size (µm)
Basophil	1%	A bilobed nucleus that each stains purple, contains a deep purple basophilic granule that obscures the cell nucleus.	14-16
Lymphocyte	30%	It is abundant, round, or oval, and is dark staining. The cytoplasm is pale blue in color, scarce, basophilic, and sometimes has purple-reddish granules.	8-15
Monocyte	6%	A kidney bean-shaped (convoluted shape) or horseshoe-shaped nucleus. Abundant cytoplasm stains a blue-gray color and presents some fine pink/purple granules, vacuoles are sometimes present in it.	14-20
Neutrophil	60%	U-shaped nucleus divided into 2 to 5 lobed and stains dark purple. Chromatin condensed, coarse, and clumped. The cytoplasm is moderate and pale pink to tan with pink-purple granules.	12-16
Eosinophil	3%	Blue bilobed nucleus. Chromatin condensed and clumped. Cytoplasm pink tan stained with large orange and red granules.	14-16

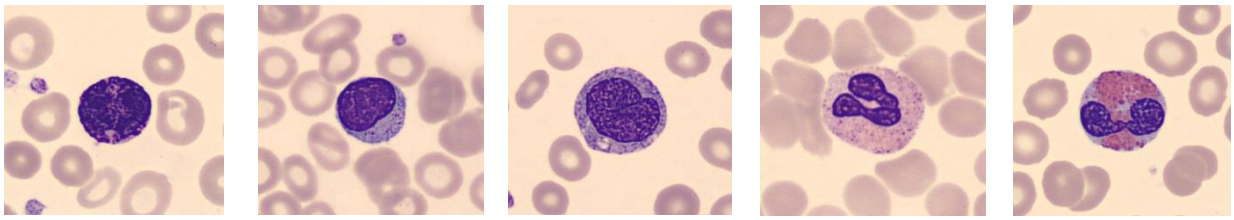

Fig. 1. The five types of white blood cells [4].

Leukemia is primarily categorized into two distinct categories based on the specifics impacted and the rate of disease progression: acute and chronic. Acute leukemia manifests itself abruptly and advances quickly over days or weeks, reaching its most severe phase in a relatively shorter period than chronic leukemia. Untreated acute leukemia has the potential to result in fatal outcomes within a short span of several months. Chronic leukemia exhibits a gradual growth pattern, characterized by a protracted duration that may span several months or even years. The categorization of acute leukemia, as per the French-American-British (FAB) classification scheme, involves two distinct subtypes: acute lymphoblastic leukemia (ALL) and acute myeloid leukemia (AML). In a similar vein, chronic leukemia can be classified into two distinct subtypes, namely chronic lymphocytic leukemia (CLL) and chronic myeloid leukemia (CML). Therefore, leukemia can be classified into four distinct forms, namely acute lymphoblastic

leukemia (ALL), acute myeloid leukemia (AML), chronic lymphocytic leukemia (CLL), and chronic myeloid leukemia (CML) [5, 6]. The most common types of leukemia can be identified on stained slides; as shown in figure 2. There are describe four main types as follows:

- Acute Lymphoblastic Leukemia (ALL) exerts a significant impact on B-lymphoblasts present in the bone marrow, blood, and Extra-medullary locations. The main symptom of this condition is an elevated count of B-lymphocytes relative to T-lymphocytes. B lymphocytes prevent germ infection, whereas T cells kill the infected cells. The rapid growth of atypical lymphocytes is referred to as lymphoblastic. When the process of full maturation of WBCs occurs, this particular form of leukemia is observed to be more prevalent among pediatric populations. Patients with this disease who don't get medication make a lot of lymphoblasts, which can end their lives. The frequency of ALL is higher in B cells compared to T-cells. The French-American-British (FAB) group conducted the classification of acute lymphoblastic leukemia (ALL) using morphological criteria. This classification system categorizes ALL into three distinct groupings. The variables L1, L2, and L3 [5, 6].
- Acute Myeloid Leukemia (AML) is a type of leukemia that occurs in the marrow and is characterized by the proliferation of myeloid blasts, red blood cells, or abnormal platelets and grows rapidly and usually occurs in both adults and children. Eight subtypes of AML are identified, namely M0, M1, M2, M3, M4, M5, M6, and M7 [7, 8].
- Chronic lymphoblastic leukemia (CLL) is the most common form of leukemia, with higher rates in adults while being relatively infrequent in children. This type affects B lymphocytes or B cells. B cells help the human body fight against infection; however, the cancerous B cells cannot fight infections. CLL develops when the immune system is weakened and normal blood cells are pushed out by an excess of aberrant lymphocytes. The symptoms of CLL include weight loss, fever, sleep sweats, and frequent infection [5, 7, 9].
- Chronic myeloid leukemia (CML) is a kind of cancer that mostly affects white blood cells. Increases in bone marrow myeloid cell proliferation, concentration, and blood density cause this condition. CML is mostly affecting adults. The rate of illness progression is characterized by a gradual and manageable course. Patients diagnosed with CML typically experience a normal life and frequently reveal no symptoms[5, 7, 9].

Early diagnosis of leukemia symptoms in individuals can significantly enhance their chances of survival. It is currently advised to observe blood cells using cytogenetic and immunophenotype diagnostic procedures due to their excellent accuracy. One challenge associated with these methods is their inherent slowness and lack of standardization since they rely on the operator's individual talents and fatigue levels. The utilization of microscopic imaging techniques for the detection of leukemia in human blood samples is particularly suitable for cost-effective and geographically distant diagnostic systems. This is where new approaches are implemented. Researchers have successfully employed deep learning, machine learning, and neural networks to devise systems capable of effectively detecting and categorizing many forms of blood cancer with remarkable precision [1].

Machine learning (ML) currently stands as one of the most significant fields within the realm of artificial intelligence (AI). The aforementioned approach is widely employed in the field of image categorization. The field of image categorization has consistently remained a prominent subject of research. The availability of the output data label is a prerequisite in the context of supervised machine learning, necessitating the utilization of a supervised machine learning algorithm. The images that are included within the WBC databases are appropriately labeled. Hence, the classification of acute leukemia and white

blood cells (WBCs) falls under the category of monitored problems. In the domain of supervised learning, it is customary to utilize both standard and deep learning algorithms for classification in a classified dataset [10]. This study aims to evaluate leukemia's automatic diagnostic and classification literature to provide a comprehensive overview.

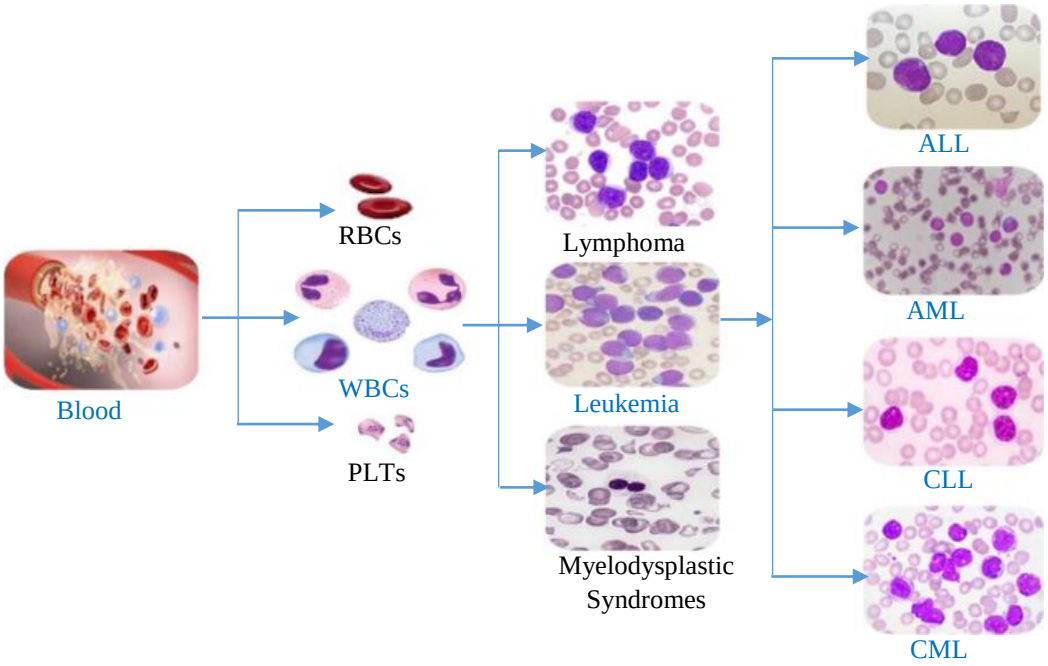

Fig. 2. The four main types of Leukemia [10].

The subsequent section describes the structural elements present in this work. This paper provides a comprehensive overview of preprocessing techniques, segmentation methods, feature extraction and selection approaches, and classification techniques described in Section 2. A brief analysis of the datasets was presented in Section 3. In Section 4, a comparative analysis of ML algorithms was discussed. Finally, Section Five has ended with the conclusions of the current study.

2. Methodology for leukemia detection and classification

A generalized methodology for microscopic blood cell analysis for people who are suspected of having leukemia is shown in Figure 3. The procedure encompasses various stages, including image acquisition, image preprocessing, image segmentation, feature extraction and selection, and diagnosing a disease. A pathologist collects a blood sample from the patient. The slide is then processed to generate a blood smear. The slide is seen through a high-quality microscope, which produces a clear and detailed image. Either a camera or an adaptor connected to a microscope directly captures the image. The present image is being considered for subsequent analysis. The obtained images may exhibit undesired regions and overlapping of distinct blood components. The visual quality of this image is improved through the application of an appropriate image enhancement technology. As a result, a high-quality image is now available for study. Following the pre-processing stage, the procedure involves separating distinct blood constituents, including the isolation of red blood cells (RBCs), white blood cells (WBCs), plasma, and platelets. Segmentation is carried out by taking into account the generalized properties of blood components. The region of interest will be isolated during this phase for further classification. Additional components, including RBC and WBC, are subdivided into subclasses. This approach aids in the identification of a specific subclass image that can be utilized to extract features for the purpose of

analyzing blood cells. Subsequently, disease detection is performed based on the outcomes of this study. Following the segmentation process, various features are extracted by taking into account distinct components of blood. Size, shape, color, and count of various blood components, such as WBC and RBC counts, are among the features. By analyzing these characteristics, the disease can be detected or the number of cells can be determined. The decision regarding the disease might be made based on the several retrieved features. In order to facilitate decision-making processes, many classifiers might be developed.

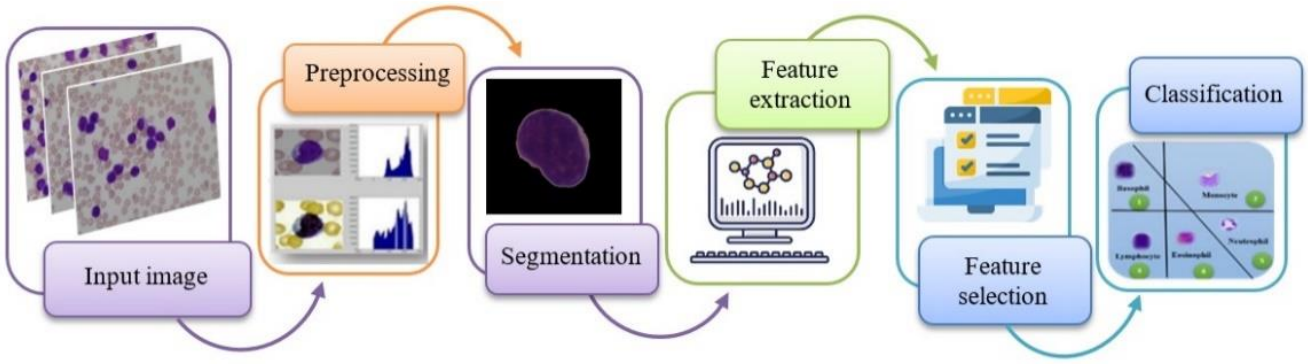

Fig. 3. General Structure of Leukemia Classification

2.1. Preprocessing

The type of microscope, camera, camera angle, light source, variation in illumination, and noise are all have an influence on how well microscopic images turn out. Preprocessing is a set of operations used to prepare an image for segmentation, such as image enhancement and de-nosing [6, 11]. Histogram equalization is a straightforward solution to this problem. Chand and Vishwakarma [12] for smear image preprocessing firstly convert the image to gray scale then perform contrast stretching on gray image then implement histogram equalization to enhance it. Rawat et al. [13] used histogram equalization in the preprocessing step to improve the quality of the image and minimize noise by employing a 2-D order statistical filter (replace each pixel of the input image with a non-zero pixel from a sorted set of neighbors). In [14] Kumar et al. utilized a Wiener filter which adequately reduced the blurriness without reducing the image sharpness. Further histogram equalization technique is issue to enhance the contrast in the image. The primary objective of preprocessing is to enhance the quality of images. This technique is utilized for the purposes of noise reduction, image deblurring, and edge improvement.

Stain intensity generally has the potential to impact both segmentation and clustering. In order to address this issue, it is recommended to use an image re-coloring technique during the pre-processing phase. This approach serves to normalize both the provided picture dataset and the subsequent analysis of the images. In practical applications, there is significant variation in hue and intensity between blood cells and the background image. Numerous factors, including different lighting conditions, age stains, and camera settings, could contribute to this. The RGB smear picture is converted to the CIELAB or YCbCr color space using an adaptive technique by [15-18] in order to make the cell segmentation robust with regard to these fluctuations.

Normalization is an important preprocessing procedure that adjusts all stain slides to various capture situations, especially those with different lighting conditions. It reduces the lighting and color differences caused by diverse capture conditions of microscopic images obtained from various facilities, and it avoids the problem of vanishing by stabilizing the gradient descent step, allowing us to use larger learning rates or help models converge faster for a given learning rate, thus improving the classification [6, 11]. Many research employed nomalization as a preprocessing such as [19-21].

Some of studies used median and wiener filters with image resizing like [21-24]. Furthermore, data augmentation is used as a first stage in machine and deep learning algorithms to mitigate the problem of overfitting. Data augmentation is often used during the training phase to increase the amount of the training data by making modest modifications to the current data. These modifications may include re-scaling, padding, vertical or horizontal flipping, random rotation, translation, cropping, and zooming. Properly training a system leads to enhanced segmentation or classification. In [9, 25-28] the authors used image augmentation for preprocessing step. Also, many studies used more than one of the method that is mentioned above for smear image preprocessing like [8, 20, 21, 29].

2.2. Image Segmentation

Segmentation plays a crucial role in the overall performance of the diagnosis of diseases. The desired WBCs are extracted by eliminating platelets and RBCs and separating the overlapped cells. Image segmentation includes many techniques that some of these based on traditional image processing techniques, some are based on machine learning techniques, and others are based on deep learning techniques. It is worth noting that segmentation methods based on deep learning and machine learning give better results than traditional processing methods. The segmentation stage is one of the most critical and difficult processes in detecting and categorizing leukemia since the effectiveness of the next steps is entirely dependent on it.

Many WBC segmentation approaches based on image processing will be discussed in this section such as morphological operations, thresholding-based techniques, watershed-based techniques. The most basic segmentation approach is threshold-based segmentation. It is appropriate when the cells do not touch or overlap and there is a distinct depth difference between the items and the background [13]. The Ostu-based threshold approach has been used by many studies to segment WBC. The key problem is determining an adequate threshold value. Damayanti et al. [8] used thresholding with morphological opening for WBC extracted from background. In [26] Sampathila et al. employed HIS (Hue, Intensity, Suturation) color space with thresholding to segment WBC regions. Dasariraju et al. [30] segmented the nucleus and cytoplasm using multi-Otsu thresholding and morphological procedures. In [13] Rawat et al. implement Otsu thresholding with morphological opening on smear images to detect precise edge of nucleus in leukocyte cell. Iswarya et al. [22] used zack algorithm for WBC detection and lymphocytes are discuss together another researchers have employed the watershed algorithm to segment overlapped and touched cells like Al Mamun et al. [15], they used the watershed algorithm with morphological operation to segment the particular lymphocytes cell. However, owing to the complex structure of cells, overlapping cells, and intensity inhomogeneities, most image processing-based approaches are unable to give more accurate segmentation. Furthermore, more precise segmentation may be obtained by combining more than one approach to reach the best results.

Segmentation methods based on machine learning are widely used and classified into two types: supervised and unsupervised. In supervised machine learning algorithms trains the machine by labeled data. ANN (Artificial Neural Network), SVM (Support Vector Machine), Decision trees, and other types of supervised machine used by researchers to detect ROI area. Al-jaboriy et al. [31] employed ANN on WBC image to extract the ROI. In [32] Abdulhay et al. used SVM for leucocyte segmentation. A decision tree (DT) machine learning algorithm used by Sharma et al. [33] to segment WBC efficiently.

In unsupervised learning, machine learning algorithms are used to train unlabeled or unclassified data. The two most common unsupervised machine learning-based approaches are K-means clustering and Fuzzy C-means clustering (FCM). The goal of these clustering algorithms is to find comparable areas (clusters) within an image. K-means clustering to segment nuclei images [18, 34]. Some of studies used K-means

with another techniques such as, Kumar et al. [17] employed K-means clustering with color based segmentation and Watershed filtering. Kumar et al. [14] used k-means with texture based segmentation and color based segmentation. One study [35] used a special mix of k-means clustering, thresholding, and modified watershed algorithms to separate overlapping cells and nuclei, extract nuclei from images of cells, and separate WBCs. Patel and Mishra [36] used K-means for WBC detection and applied histogram equalization with zack algorithm for grouping WBC. Dese et al. [21] applied K-means for separate WBC from RBC, and used marker-controlled watershed segmentation algorithms, with erosion and dilation features to separate the overlapped nucleated WBC. Gayathri and Jyothi [37] used Adaptive k-means for image segmentation with image cleaning operations (morphological opining and dilation). Despite being slower than k-means, the k-medoid approach delivers superior segmentation performance and is more resilient than k-means. A color-based image segmentation using K-medoids algorithm used in [16]. When using the k-means algorithm for the segmentation of nuclei yields an empty cluster at times. To avoid this, the fuzzy c-means clustering (FCM) technique was used in [38]. In [39] Zheng et al. used k-means clustering to extracts the overall foreground region from the cell image then uses the SVM classifier to improve segmentation result.

Due to the accuracy advantage of deep learning and a recent advancement, it becomes a referred approach for image segmentation. Alagu et al. [40] employed U-Net convolutional neural network to extract the nucleus from healthy and blast cell. In [41] Balasubramanian et al. developed a modified U-Net to segment WBC. For automatic nuclei segmentation, Shiv et al. [42] suggested an Encoder-Decoder based Convolutional Neural Network with Nested-Feature Concatenation (EDNFC-Net). To yield more accurate WBC segmentation, Reena and Ameer [43] employed a transfer learning-based semantic segmentation approach. They used a DeepLabv3+ algorithm.

2.3. Feature Extraction and selection

The feature extraction stage is used to extract and identify characteristics obtained from either the segmented region or the entire image. In this section, the researchers discuss several ways of extracting texture or shape properties during the segmentation process. Additionally, the researchers aim to decrease the image size by eliminating unnecessary information from the original image. Consequently, it enhances the processing velocity and reduces the execution duration at this stage. The items in the picture may yield many characteristics, including form features, texture features, statistical features, geometrical features, color features, etc… . Deep learning networks automatically do feature extraction [5, 31].

Many studies extracted geometrical, color, and statistical texture features to get efficient Leukemia classification [13-15, 17, 21, 23, 24, 29, 36, 44] . In [45] a cytoplasm and nucleus shape features were extracted like area, Euler number, parameter, solidity, diameter, minor axis, major axis, eccentricity, convex area, and orientation. The textural properties obtained from the Gray Level Co-occurrence Matrix (GLCM), including correlation, contrast, homogeneity, energy, and entropy statistics, are extracted from the GLCM matrix of the nucleus and cytoplasm. Region and border shape characteristics are retrieved. For the detection of the overlapping nuclei, Ghane et al. [35] extracted the area and solidity features of the nuclei. Gayathri and Jyothi [37] used shape and area features for Leukemina detection and classification. Mirmohammadi et al. [38] have extracted some of the geometric and statistical features from the nuclei to classify the cells as ALL and determine their subtype, and the best features selected using principal component analysis (PCA). Windows feature extraction method was used by Abdulhay et al. [32], they measured the texture of the obtained microscope smear images using a local binary pattern (LBP). Zheng et al. [39] introduced a new type of feature for each pixel based on a weak edge enhancement operator (WEEO), which can enhance the weak boundaries between the CROI (cell region of interest) and the non-CROI and support better segmentation. Three types of features were extracted by Damayanti et al. [8]

these features WBC diameter, nucleus ratio, and nucleus roundness. Three types of feature selection was used information gain, the gain ratio, and ReliefF to select the most influence features. Acharya and Kumar [16] extracted a shape features, visual features, and texture features for diagnosis acute lymphoblastic leukemia, and for feature selection, they used InfoGainAttributeEval and the Ranker Search method. In [30] Dasariraju et al. extracted 16 cytomorphological features from each image, these features were: nucleus size, shape, elliptical features, and color features. Also, they proposed two new color features: the average and standard deviation of the nucleus in the B channel of LAB color space, and for important feature selection the used Gini importance. Two cytoplasm color features conceived by Ghane et al. [46] were also used for the diagnosis of AML leukemia. After using PCA, Analysis of Variance (ANOVA), and boxplot, it turns out that these features is not very appropriate and they proposed three types of features: Number of nucleus lobes (Num), Area of Nucleus (AoN), and Average Color of Cytoplasm (ACoC) .

Recently, most researchers have turned towards using Deep Learning for feature extraction, because they found promising results for improving classification. Shahin et al. [47] used either (AlexNet and VGGNet) or OverfeatNet for extracting features. For feature selection, they used either chi-squared or PCA technique. Shafique and Tehsin [28] employed Deep convolutional neural networks (DCNNs) based on pre-trained AlexNet for the feature extraction, detection, and classification of ALL and its subtypes. For minimize the features vector, they added another fully connected layer with 1024 neurons followed by ReLU layer. In [25] Habibzadeh et al. used pre-trained Inception and ResNet architectures with all their versions for feature extraction and leukocyte classification. Shaheen et al. [7] used AlexNet for feature extraction and AML detection. Alagu et al. [40] fused al the features together from AlexNet, GoogleNet, and SqueezeNet. They selected the distinct features using mutual information (MI), minimum recursive maximal relevance (mRmR), and recursive feature elimination (RFE) based methods. Furthermore, they used a heatmap to obtain the prominent deep features. Then, they employed ANOVA to do statistical analysis utilizing reliable and consistent feature sets. Vieira and Valle [48] proposed eight hypercomplex-valued convolutional neural networks (HvCNNs) for feature extraction and ALL detection. Sampathila et al. [26] proposed a custom ALLNET for feature extraction and ALL detection. Hagar et al. [9] employed VGG16 architecture and DenseNet-121 for leukemia classification. Sulaiman et al. [19] proposed approach uses seven deep-learning models: VGG16, ResNet152, DenseNet121, InceptionV3, MobileNetV2, EfficientNetB0 and ResNet50 for deep feature extraction from blood smear images, and for the best feature selection, they used Random Forest, principal component analysis (PCA), and analysis of variance (ANOVA).

2.4. Classification

Numerous techniques are used for the classification phase. Supervised and unsupervised algorithms are employed for classification purposes during this stage. It was discovered that supervised algorithms were utilized in the vast majority of work during the classification phase to diagnose acute leukemia and WBCs. Many researchers used SVM for Leukemia classification and detection [13, 15, 17-19, 21, 34, 36, 38, 39, 45, 47], random forest algorithm was used in [8, 30], and AlexNet was used in [7, 28, 43].

Other researchers used SVM with other types of machine learning like: Daqqa et al. [49] used three types of machine learning algorithms, k-nearest Neighbor (kNN), DT, and SVM for blood cancer classification. In the experimental results, it was noticed that the DL algorithm had the highest percentage of 77.30% compared with the other two techniques. Gayathri and Jyothi [37] employed three types of machine learning algorithms ANN, SVM, and CNN (Convolutional Neural Network). Then they made a comparison among the results of these algorithms. A recognition efficiency of 89.47% with SVM and 92.10% with ANN and the efficiency of CNN is 93%. Chand and Vishwakarma [12] used SVM Classifier

and Extreme Learning Machine (ELM) for Leukemia diagnosis. The results showed that ELM with an accuracy of 92.24% outperforms SVM with an accuracy of 86.36%. Hegde et al. [24] used SVM classifier for white blood cell classification into normal and abnormal, and also for detection of leukemic WBCs from the abnormal class. Also, they used the neural network (NN) classifier for the normal white blood cell classification into five sub-types. An overall classification accuracy of 98.8% was obtained using the combination of SVM and NN. Balasubramanian et al. [41] used a modified U-Net and SVM for WBC classification. the experiments indicated that these mixture algorithms U-Net-SVM can recognize WBCs in three datasets with an accuracy of 99.45%, 98.62%, and 98.81%, respectively. Further investigation concerning on the leukemia dataset, ALL-IDB2, revealed an accuracy of 99.42%.

Kumar et al. [14] proposed an automated system that was tested with kNN and Naive Bayes Classifier for ALL diagnoses, the accuracy achieved is 92.8%. The different supervised classifiers were compared by Acharya and Kumar [16] for ALL detections, they used random forest, kNN, naive Bayes, and DT. The proposed algorithm achieved an overall accuracy of 98.6%. The experimental results have shown that the accuracy of the random forest was 98.16%, kNN 96.8%, and naive Bayes 96%. The promising results showed that it can be used as a diagnostic tool by the pathologists. Hossain et al. [50] used seven different machine learning models based on rule generation and association analysis by the Apriori algorithm: DT classifier, Random Forest, kNN, Adaboost, Logistic Regression, Naive Bayesian, and ANN. DT algorithm accuracy value was 97.74%, Random Forest accuracy was 97.74%, kNN accuracy 87.07%, Adaboost accuracy value was 92.22%, Logistic Regression was 89.37%, and Naive Bayes model gave fewer accuracy values than other models 85.69%. Ghane et al. [46] utilized a decision tree classifier for CML cell classification into eight groups. The accuracy value was 99.0 %, specificity 99.4 %, and sensitivity 98.3 %.

DenseNet-121 and ResNet-34 were used in [27] and [40] for Leukemia classification. In [25] Habibzadeh et al. used pre-trained Inception in four versions and ResNet in three versions for leukocyte recognition. The results showed that the best accuracy gained when using ResNet V1 50 was 100% with 3000 epochs and fine-tuning all layers. Vieira and Valle [48] used eight hypercomplex-valued convolutional neural networks (HvCNNs) in conjunction with real-valued convolutional networks for ALL classifications. Their findings demonstrated that HvCNNs outperformed the real-valued model, exhibiting superior accuracy using fewer parameters. Additionally, it was discovered that HvCNNs, which used Clifford algebras to analyze HSV-encoded pictures, had the highest recorded levels of accuracy. Specifically, their Hierarchical Convolutional Neural Network (HvCNN) achieved an average accuracy rate of 96.6%. A lightweight CNN proposed by Genovese [51] called ALLNet for ALL detections, the proposed architecture was based on fixed binary kernels that replicate the Local Binary Patterns and that use only nearly 1.6% of the learnable parameters of a traditional CNN. The accuracy results of this proposal based on ResNet-34 was 98.46%. In [33] Sharma et al. employed the DenseNet121 model to classify the different types of WBC. This model yielded an accuracy of 98.84%, a precision of 99.33%, a sensitivity of 98.85%, and a specificity of 99.61%. Pre-trained ResNet-50 used by Shalini and Viji [20] for Leukemia detection and classification. The accuracy gained from implementing this model was 99.61%. Hagar et al. [9] utilized two models for Leukemia classification. The first model used the VGG16 architecture, while the second model used DenseNet121. The results indicated that DenseNet-121 achieved a lower accuracy compared to VGG16 with an accuracy of 98.2%. For Leukemia detection in children, Talaat and Gamel [29] proposed A2M-LEUK algorithm based on CNN. The performance metrics gained from implementing this algorithm were: precision of 99.97%, recall of 100%, F1-score of 99.98%, and Accuracy of 99.98%.

3. Datasets

Analyzing microscopic medical blood images, which starts with the analysis of peripheral blood images, helps find leukemia disease. Researchers may use several datasets to test and evaluate models in image-processing tasks, which can be very beneficial for their work. Table 2 represents various publicly available standard datasets that were used in the different studies within the domain of our study.

Table 2. Leukemia datasets.

Datasets	Description
ALL-IDB [52]	It has two versions the first one is the ALL-IDB1 dataset. It contains 108 leucocyte images. The second one is the ALL-IDB2 dataset. It contains 260 images and 50% of these represent lymphoblast.
ASH [53]	It consists of a total of 240 images, with 100 images depicting healthy people, 60 images representing patients with ALL, and 80 images depicting patients with AML. A set of 420 sub-images is obtained by extracting 140 sub-images from each class, comprising 100 images of healthy patients, 60 images of patients with ALL, and 80 images of patients with AML.
JMC [21]	It comprises a collection of 250 images sourced from the Department of Hematology at Jimma University Medical Center (JMC). A comprehensive collection of 520 images was obtained from the provided slides. Among these, 460 images were of leukemic cells, with each leukemic kind comprising 115 images, while the remaining 60 images depicted normal cells.
CNMC2019 [54]	It consists of a total of 10,661 images, with 7,272 images depicting individuals diagnosed with leukemia and 3,389 images representing patients without leukemia.
The Munich AML Morphology Dataset [55]	The dataset comprises 18,365 single-cell images that have been expertly labeled. These images were captured from peripheral blood smears of 100 patients who were diagnosed with AML at Munich University Hospital during the period from 2014 to 2017. Additionally, the dataset includes images from 100 patients who did not exhibit any indications of hematological malignancy.
LISC [56]	It collected 400 samples from 100 microscope slides from 8 normal patients' peripheral blood. The images have a combined resolution of 720 × 576 pixels. All color images were collected from the Imam Khomeini Hospital's Hematology-Oncology and BMT Research Center in Tehran, Iran.
BCCD [57]	The dataset comprises a collection of 12,500 augmented images of blood cells in JPEG format, accompanied by corresponding cell type annotations in CSV format.

4. Comparative Analysis of ML Algorithms

This section presents a comparative performance study of several machine and deep learning techniques that have been utilized for the detection and classification of Leukemia. In this study, a comparative analysis is conducted by utilizing the findings as reported in the individual research publications. Table 2

presents a comparative study of the latest developments in machine learning and deep learning-based methods for the detection and classification of Leukemia.

Table 3. A comparison among the latest developments studies of machine learning and deep learning-based methods for Leukemia detection and classification.

Ref.	Year	Segmentation	Feature		Classifier	Dataset	Accuracy
			Extraction	Selection			
[45]	2015	—	Shape, texture, and GLCM	—	SVM	ALL-IDB (ALL-IDB1, ALL-IDB2)	89.8 %
[18]	2016	k-means	color, texture, shape, and Hausdorff dimension	—	SVM	ALL-IDB1	> 94%
[14]	2016	k-means with texture and color-based segmentation	color, textural, geometrical, and statistical	—	KNN and Naive Bayes	Private dataset	92.8%
[47]	2017	—	AlexNet and VGGNet or OverfeatNet	chi-squared or PCA	WBCsNet	Private dataset	96.1%
[13]	2017	Otsu thresholding with morphological	shape, statistical, Laws' mask, FPS and 2-D Gabor	Genetic algorithm	SVM with Multilayer perceptron (MLP) kernel	ASH	97.1%
[49]	2017	—	—	—	kNN DT SVM	Private dataset	60% 67% 70%
[28]	2018	—	DCNNs based on pre-trained AlexNet	—	DCNNs based on pre-trained AlexNet	ALL-IDB (ALL-IDB1, ALL-IDB2)	96.06%
[37]	2018	k-means	area, perimeter, major axis, minor axis and number of nucleuses	—	ANN SVM CNN	ALL-IDB1	92.10% 89.47% 93%
[8]	2019	thresholding with morphological opening	WBC diameter, nucleus ratio, and nucleus roundness	information gain, the gain ratio, and ReliefF	Random Forest	Private dataset	89.6%
[12]	2019	histogram processing	Geometrical and shape	—	SVM ELM	ALL-IDB1	86.36% 92.24%

[16]	2019	K-medoids	shape, visual, and texture	InfoGainAt-tributeEval and Ranker Search method	random forest, kNN, naive Bayes, and DT	ALL-IDB, Atlas, and from internet	98.6%
[46]	2019	CMYK, Otsu thresholding, morphological operations	Num, AoN, and ACoC	PCA, ANOVA, and boxplot	DT	Private dataset	99%
[24]	2020	Thresholding, morphological operations, and area filter	color, texture, and shape	—	SVM with NN	Private dataset	98.8%
[30]	2020	multi-Otsu thresholding, and morphological operations	size, shape, elliptical, color, and average and standard deviation of the nucleus in the B channel of LAB color space	Gini importance	Random Forest	Private dataset	92.99% for detection 93.45% for classificati on
[43]	2020	DeepLabv3+	Transfer learning using AlexNet	—	Transfer learning using AlexNet	LISC	98.87%
[27]	2020	—	ResNet-34 DenseNet-121	—	ResNet-34 DenseNet-121	ALL-IDB, ASH	99.56% 99.91%
[21]	2021	k-means	shape, color, texture and statistical Features	—	SVM	JMC	97.69%
[7]	2021	—	AlexNet	—	AlexNet	Private dataset	96.25%
[15]	2022	watershed algorithm with morphological	statistical features with YCbCr color space	—	SVM	ALL-IDB	99.21%
[48]	2022	—	HvCNNs	—	HvCNNs	ALL-IDB2	96.6%
[51]	2022	—	LBP	—	ALLNet	ALL-IDB2	98.46%
[33]	2022	decision tree	DenseNet121	—	DenseNet121	BCCD	98.84%
[20]	2022	—	Pre-trained ResNet-50	—	Pre-trained ResNet-50	ALL-IDB	99.61%

[19]	2023	—	ResNet152, VGG16, DenseNet121, MobileNetV2, InceptionV3, EfficientNetB0 and ResNet50	ANOVA, PCA, and Random Forest	SVM	CNMC2019	90%
[9]	2023	—	VGG16 DenseNet121	—	VGG16 DenseNet121	Collected from three dataset	98.2% 98.1%
[29]	2023	—	edge detection, texture analysis, and shape analysis	—	A2M-LEUK	CNMC2019	99.98%

This section presents an analysis of the previous approach employed for the detection and classification of leukemia. The methodology involves a combination of conventional techniques and artificial neural network (ANN) and/or convolutional neural network (CNN) methods for the purpose of classification. The current funding indicate that mixture classification techniques have received less attention by researchers. The present study provides a comprehensive review of the existing literature pertaining to the combined methodologies of leukemia and is described as follows:

This section presents an analysis of the performance of leukemia classification using three different types of classifiers: conventional, CNN, and mixed classifiers. The SVM classifier is commonly employed in traditional network approaches for classification tasks due to its satisfactory accuracy. In numerous instances, CNNs have shown higher levels of accuracy in comparison with traditional networks. In the context of the experiment, it was observed that both SVM and CNN demonstrated favorable outcomes. Furthermore, it is worth noting that in recent years, the majority of endeavors pertaining to the identification and categorization of leukemia have relied on CNNs. These approaches have garnered significant interest and have demonstrated exemplary outcomes.

5. Conclusions

This study provides a concise analysis of current progress in the field of deep and machine learning-based techniques for the identification and classification of leukemia and white blood cells (WBC). The researchers have conducted an analysis of several current methods for segmentation, feature extraction and selection, and classification that are utilized in the efficient detection of leukemia. Based on previous review, It is discovered that unsupervised schemes are preferable for segmentation tasks, but supervised schemes are preferred for classification tasks. Nonetheless, the favored method for automatic and more reliable detection and classification is deep learning, especially transfer learning. It has also been noticed that using a mix of deep learning and machine learning techniques in hybrid schemes makes it very accurate at classifying leukemia, especially when CNN algorithms and SVM are used together. The highest achieved performance in traditional machine learning was observed with the Support Vector Machine (SVM) algorithm, achieving an accuracy of 99.21% when trained on the ALL-IDB dataset. In the field of deep learning, the DenseNet-121 model achieved the best performance, attaining an accuracy of 99.91% when trained on both the ALL-IDB and ASH datasets. Additionally, the pre-trained ResNet-50 model achieved an accuracy of 99.61% when trained solely on the ALL-IDB dataset. Finally, in hybrid

schemes, the combination of SVM with MLP yielded an accuracy of 97.1% when trained on the ASH dataset. In addition to encouraging researchers to conduct more studies, this publication will assist researchers in analyzing recent developments concerning leukemia detection and classification.

References

[1] Naina Sharma, Ankit Mukopadhyay, Aman Shrivastava, and A. Garg, "A Brief Survey on Leukemia Detection Systems," *International Research Journal of Engineering and Technology*, vol. 8, no. 12, pp. 411-413, 2021.

[2] Z. M. Kouzehkanan *et al.*, "A large dataset of white blood cells containing cell locations and types, along with segmented nuclei and cytoplasm," *Sci Rep*, vol. 12, no. 1, p. 1123, 2022.

[3] S. Tsutsui, Z. Su, and B. Wen, "Benchmarking White Blood Cell Classification under Domain Shift," in *ICASSP 2023 - 2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2023, pp. 1-5.

[4] A. Acevedo, S. Alferez, A. Merino, L. Puigvi, and J. Rodellar, "Recognition of peripheral blood cell images using convolutional neural networks," *Comput Methods Programs Biomed*, vol. 180, p. 105020, 2019.

[5] M. Zolfaghari and H. Sajedi, "A survey on automated detection and classification of acute leukemia and WBCs in microscopic blood cells," *Multimedia Tools and Applications*, vol. 81, no. 5, pp. 6723-6753, 2022.

[6] P. K. Das, D. V. A, S. Meher, R. Panda, and A. Abraham, "A Systematic Review on Recent Advancements in Deep and Machine Learning Based Detection and Classification of Acute Lymphoblastic Leukemia," *IEEE Access*, vol. 10, pp. 81741-81763, 2022.

[7] M. Shaheen *et al.*, "Acute Myeloid Leukemia (AML) Detection Using AlexNet Model," *Complexity*, vol. 2021, pp. 1-8, 2021.

[8] F. Damayanti, S. Herawati, I. Imamah, F. A. M, and A. Rachmad, "Classification of blast cell type on acute myeloid leukemia (AML) based on image morphology of white blood cells," *TELKOMNIKA (Telecommunication Computing Electronics and Control)*, vol. 17, no. 2, 2019.

[9] M. Hagar, F. K. Elsheref, and S. R. Kamal, "A New Model for Blood Cancer Classification Based on Deep Learning Techniques," *International Journal of Advanced Computer Science and Applications*, vol. 14, no. 6, 2023.

[10] F. A. J. Al-sudani and H. A. F. Al-Afare, "Age estimation from face image using Hybrid Representation for Deep Learning," *Mustansiriyah journal of pure and Applied Sciences(MJPAS)*, vol. 23, no. 2, pp. 35-48, 2022.

[11] A. Anghel *et al.*, "A High-Performance System for Robust Stain Normalization of Whole-Slide Images in Histopathology," *Front Med (Lausanne)*, vol. 6, p. 193, 2019.

[12] S. Chand and V. P. Vishwakarma, "Leukemia Diagnosis using Computational Intelligence," in *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*, 2019, vol. 1, pp. 1-7.

[13] J. Rawat, A. Singh, B. Hs, J. Virmani, and J. S. Devgun, "Computer assisted classification framework for prediction of acute lymphoblastic and acute myeloblastic leukemia," *Biocybernetics and Biomedical Engineering*, vol. 37, no. 4, pp. 637-654, 2017.

[14] S. Kumar, S. Mishra, P. Asthana, and Pragya, "Automated Detection of Acute Leukemia using K-mean Clustering Algorithm," *ArXiv*, vol. abs/1803.08544, 2016.

[15] A. Al Mamun, M. J. Hossen, A. Tahabilder, A. Musha, R. Hasnat, and S. K. Saha, "Acute lymphoblastic leukemia detection approach from peripheral blood smear using color threshold and morphological techniques," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 12, no. 4, 2022.

[16] V. Acharya and P. Kumar, "Detection of acute lymphoblastic leukemia using image segmentation and data mining algorithms," *Med Biol Eng Comput*, vol. 57, no. 8, pp. 1783-1811, Aug 2019.

[17] P. Kumar and S. M. Udwadia, "Automatic detection of acute myeloid leukemia from microscopic blood smear image," in *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*, 2017, pp. 1803-1807: IEEE.

[18] S. Agaian, M. Madhukar, and A. T. Chronopoulos, "A new acute leukaemia-automated classification system," *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, vol. 6, no. 3, pp. 303-314, 2016.

[19] A. Sulaiman *et al.*, "ResRandSVM: Hybrid Approach for Acute Lymphocytic Leukemia Classification in Blood Smear Images," *Diagnostics (Basel)*, vol. 13, no. 12, Jun 20 2023.

[20] S. V and A. V. K S, "Integration of Convolutional Features and Residual Neural Network for the Detection and Classification of Leukemia from Blood Smear Images," *International Journal of Engineering Trends and Technology*, vol. 70, no. 9, pp. 176-184, 2022.

[21] K. Dese *et al.*, "Accurate Machine-Learning-Based classification of Leukemia from Blood Smear Images," *Clin Lymphoma Myeloma Leuk*, vol. 21, no. 11, pp. e903-e914, Nov 2021.

[22] M. Iswarya, S. A, A. M. R, and K. A, "Detection of Leukemia using Machine Learning," presented at the 2022 International Conference on Applied Artificial Intelligence and Computing (ICAAIC), 2022.

[23] P. P. S. Patil, U. Kunwar, M. Gandalwad, N. Gurme, and A. Shinge, "Leukemia Diagnosis Using Deep Learning," *International Journal for Research in Applied Science and Engineering Technology*, vol. 10, no. 11, pp. 1079-1082, 2022.

[24] R. B. Hegde, K. Prasad, H. Hebbar, B. M. K. Singh, and I. Sandhya, "Automated Decision Support System for Detection of Leukemia from Peripheral Blood Smear Images," *J Digit Imaging*, vol. 33, no. 2, pp. 361-374, Apr 2020.

[25] M. Habibzadeh, M. Jannesari, Z. Rezaei, H. Baharvand, and M. Totonchi, "Automatic white blood cell classification using pre-trained deep learning models: ResNet and Inception," presented at the Tenth International Conference on Machine Vision (ICMV 2017), 2018.

[26] N. Sampathila *et al.*, "Customized Deep Learning Classifier for Detection of Acute Lymphoblastic Leukemia Using Blood Smear Images," *Healthcare (Basel)*, vol. 10, no. 10, Sep 20 2022.

[27] N. Bibi, M. Sikandar, I. Ud Din, A. Almogren, and S. Ali, "IoMT-Based Automated Detection and Classification of Leukemia Using Deep Learning," *J Healthc Eng*, vol. 2020, p. 6648574, 2020.

[28] S. Shafique and S. Tehsin, "Acute Lymphoblastic Leukemia Detection and Classification of Its Subtypes Using Pretrained Deep Convolutional Neural Networks," *Technol Cancer Res Treat*, vol. 17, p. 1533033818802789, Jan 1 2018.

[29] F. M. Talaat and S. A. Gamel, "A2M-LEUK: attention-augmented algorithm for blood cancer detection in children," *Neural Computing and Applications*, vol. 35, no. 24, pp. 18059-18071, 2023.

[30] S. Dasariraju, M. Huo, and S. McCalla, "Detection and Classification of Immature Leukocytes for Diagnosis of Acute Myeloid Leukemia Using Random Forest Algorithm," *Bioengineering (Basel)*, vol. 7, no. 4, Oct 1 2020.

[31] S. S. Al-jaboriy, N. N. A. Sjarif, S. Chuprat, and W. M. Abduallah, "Acute lymphoblastic leukemia segmentation using local pixel information," *Pattern Recognition Letters*, vol. 125, pp. 85-90, 2019.

[32] E. Abdulhay, M. A. Mohammed, D. A. Ibrahim, N. Arunkumar, and V. Venkatraman, "Computer Aided Solution for Automatic Segmenting and Measurements of Blood Leucocytes Using Static Microscope Images," *J Med Syst*, vol. 42, no. 4, p. 58, Feb 17 2018.

[33] S. Sharma *et al.*, "Deep Learning Model for the Automatic Classification of White Blood Cells," *Comput Intell Neurosci*, vol. 2022, p. 7384131, 2022.

[34] D. Goutam and S. Sailaja, "Classification of acute myelogenous leukemia in blood microscopic images using supervised classifier," in *2015 IEEE International Conference on Engineering and Technology (ICETECH)*, 2015, pp. 1-5.

[35] N. Ghane, A. Vard, A. Talebi, and P. Nematollahy, "Segmentation of white blood cells from microscopic images using a novel combination of K-means clustering and modified watershed algorithm," *Journal of medical signals and sensors*, vol. 7, no. 2, p. 92, 2017.

[36] N. Patel and A. Mishra, "Automated leukaemia detection using microscopic images," *Procedia Computer Science*, vol. 58, pp. 635-642, 2015.

[37] S. Gayathri and R. Jyothi, "An automated leucocyte classification for leukemia detection," *International Research Journal of Engineering and Technology (IRJET)*, 2018.

[38] P. Mirmohammadi, A. Rasooli, M. Ashtiyani, M. M. Amin, and M. R. Deevband, "Automatic recognition of acute lymphoblastic leukemia using multi-SVM classifier," *Current Science*, vol. 115, no. 8, pp. 1512-1518, 2018.

[39] X. Zheng, Y. Wang, G. Wang, and J. Liu, "Fast and robust segmentation of white blood cell images by self-supervised learning," *Micron*, vol. 107, pp. 55-71, Apr 2018.

[40] S. Alagu, A. P. N, K. G, and B. B. K, "Automatic Detection of Acute Lymphoblastic Leukemia Using UNET Based Segmentation and Statistical Analysis of Fused Deep Features," *Applied Artificial Intelligence*, vol. 35, no. 15, pp. 1952-1969, 2021.

[41] K. Balasubramanian, K. Gayathri Devi, and K. Ramya, "Classification of white blood cells based on modified U-Net and SVM," *Concurrency and Computation: Practice and Experience*, vol. n/a, no. n/a, p. e7862, 2023.

[42] S. Gehlot, A. Gupta, and R. Gupta, "EDNFC-Net: Convolutional Neural Network with Nested Feature Concatenation for Nuclei-Instance Segmentation," in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1389-1393.

[43] M. R. Reena and P. M. Ameer, "Localization and recognition of leukocytes in peripheral blood: A deep learning approach," *Comput Biol Med*, vol. 126, p. 104034, Nov 2020.

[44] S. Tavakoli, A. Ghaffari, Z. M. Kouzehkanan, and R. Hosseini, "New segmentation and feature extraction algorithm for classification of white blood cells in peripheral smear images," *Sci Rep*, vol. 11, no. 1, p. 19428, Sep 30 2021.

[45] J. Rawat, A. Singh, H. S. Bhadauria, and J. Virmani, "Computer Aided Diagnostic System for Detection of Leukemia Using Microscopic Images," *Procedia Computer Science*, vol. 70, pp. 748-756, 2015.

[46] N. Ghane, A. Vard, A. Talebi, and P. Nematollahy, "Classification of chronic myeloid leukemia cell subtypes based on microscopic image analysis," (in eng), *Excli j*, vol. 18, pp. 382-404, 2019.

[47] A. I. Shahin, Y. Guo, K. M. Amin, and A. A. Sharawi, "White blood cells identification system based on convolutional deep neural learning networks," *Comput Methods Programs Biomed*, vol. 168, pp. 69-80, Jan 2017.

[48] G. Vieira and M. E. Valle, "Acute lymphoblastic leukemia detection using hypercomplex-valued convolutional neural networks," in *2022 International Joint Conference on Neural Networks (IJCNN)*, 2022, pp. 1-8: IEEE.

[49] K. A. A. Daqqa, A. Y. Maghari, and W. F. Al Sarraj, "Prediction and diagnosis of leukemia using classification algorithms," in *2017 8th international conference on information technology (ICIT)*, 2017, pp. 638-643: IEEE.

[50] M. A. Hossain, A. K. M. M. Islam, S. Islam, S. Shatabda, and A. Ahmed, "Symptom Based Explainable Artificial Intelligence Model for Leukemia Detection," *IEEE Access*, vol. 10, pp. 57283-57298, 2022.

[51] A. Genovese, "ALLNet: Acute Lymphoblastic Leukemia Detection Using Lightweight Convolutional Networks," in *2022 IEEE 9th International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, 2022, pp. 1-6: IEEE.

[52] R. D. Labati, V. Piuri, and F. Scotti, "All-IDB: The acute lymphoblastic leukemia image database for image processing," in *2011 18th IEEE international conference on image processing*, 2011, pp. 2045-2048: IEEE.

[53] A. S. o. Hematology. (2023/11/13). *ASH Image Bank*. Available: http://imagebank.hematology.org./

[54] S. Mourya, Kant, S., Kumar, P., Gupta, A., & Gupta, R., "ALL Challenge dataset of ISBI 2019 (C-NMC 2019) (Version 1) [dataset]," T. C. I. Archive, Ed., ed, 2019.

[55] C. Matek, Schwarz, S., Marr, C., & Spiekermann, K, "A Single-cell Morphological Dataset of Leukocytes from AML Patients and Non-malignant Controls [Data set]," T. C. I. Archive, Ed., ed, 2019.

[56] S. H. Rezatofighi, Soltanian-Zadeh, H., "Automatic recognition of five types of white blood cells in peripheral blood," C. M. I. a. Graphics, Ed., ed, 2011.

[57] P. Mooney, "Blood Cell Images," Kaggle, Ed., ed, 2018.