

# Automated Diagnosis of Leukemia: A Comprehensive Review

AFSHAN SHAH<sup>1,2</sup>, SYED SAUD NAQVI<sup>1</sup>, KHURAM NAVEED<sup>1</sup>, NEMA SALEM<sup>3</sup>,  
MOHAMMAD A. U. KHAN<sup>3,4</sup>, AND KHURRAM S. ALIMGEER<sup>1</sup>

<sup>1</sup>Department of Electrical and Computer Engineering, COMSATS University Islamabad (CUI), Islamabad 45550, Pakistan

<sup>2</sup>Department of Electrical Engineering, Sukkur IBA University (SIBAU), Sukkur, Sindh 65200, Pakistan

<sup>3</sup>Department of Electrical and Computer Engineering, Effat College of Engineering, Effat University, Jeddah 22332, Saudi Arabia

<sup>4</sup>Department of Electrical Engineering, Namal Institute Mianwali, Mianwali 42250, Pakistan

Corresponding author: Khuram Naveed (khurram.naveed@comsats.edu.pk)

This work was supported by the Effat University, Jeddah, Saudi Arabia.

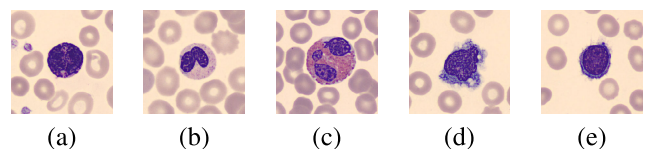
**ABSTRACT** Leukemia is the rapid production of abnormal white blood cells that consequently affects the blood and damages the bone marrow. The overproduction of abnormal and immature white blood cells leads to the damage of the immune system due to the reduced production of red blood cells and platelets by the bone marrow of the body. This hematological malignancy is generally diagnosed by manual methods such as complete blood count (CBC), bone marrow aspiration, or microscopic examination of the blood smear. Nevertheless, the manual methods of leukemia diagnosis are economical but are found to be less reliable, time-consuming, and hectic. Technological advancement in the medical field has effectively addressed these issues in the past. The problems in the manual diagnosis of leukemia detection have been overcome by the development of automated methods using the computer-aided diagnostic (CAD) systems for efficient and reliable leukemia diagnosis. Since the last decade, multiple approaches have been proposed for the CAD systems regarding pre-processing, segmentation, feature extraction, feature selection, and for the improvement of the classification accuracy of the CAD system for the leukemia detection. This paper presents a comprehensive review of the CAD systems for the detection of the various types of leukemia. The review presented here entails the details of various CAD systems for the automated diagnosis of various types of leukemia and analyses their methodologies in terms of their efficiency in pre-processing, segmentation, feature extraction and selection, and overall classification accuracy of the CAD system.

**INDEX TERMS** Acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML), chronic lymphocytic leukemia (CLL), chronic myeloid leukemia (CML), computer-aided diagnostic (CAD), leukemia.

## I. INTRODUCTION

The Leukocytes, commonly known as WBCs (white blood cells) are a prime component of our immune system that protects the human body from viruses, infections, bacteria and other foreign bodies. These nucleated cells are formed in the bone marrow and circulate through the bloodstream, so they can be found throughout the lymphatic system [1]. Leukocytes differ from the other blood cells such as red blood cells and platelets and therefore are majorly classified in two categories based on cell heredity as lymphoid and myeloid and cell structure as agranulocytes and granulocytes. These categories of leukocytes are further divided into five subcategories i.e. neutrophils, basophils, eosinophils,

The associate editor coordinating the review of this manuscript and approving it for publication was Ravibabu Mulaveesala<sup>1b</sup>.



**FIGURE 1.** Image sample of (a) Basophils; (b) Neutrophils; (c) Eosinophils; (d) Monocytes; and (e) Lymphocytes.

lymphocytes and monocytes as given in Figure 1. Lymphocytes are further divided into three cell types as thymus (T cells), bone marrow (B cells) and natural killer (NK or K) cells [2]. However, the cancerous production of abnormal and immature leukocytes affects the immune system leading to the reduced ability of bone marrow to produce RBCs and platelets [3]. The proliferation of leukocytes, i.e., blasts in the bone marrow and blood is specified by a disease

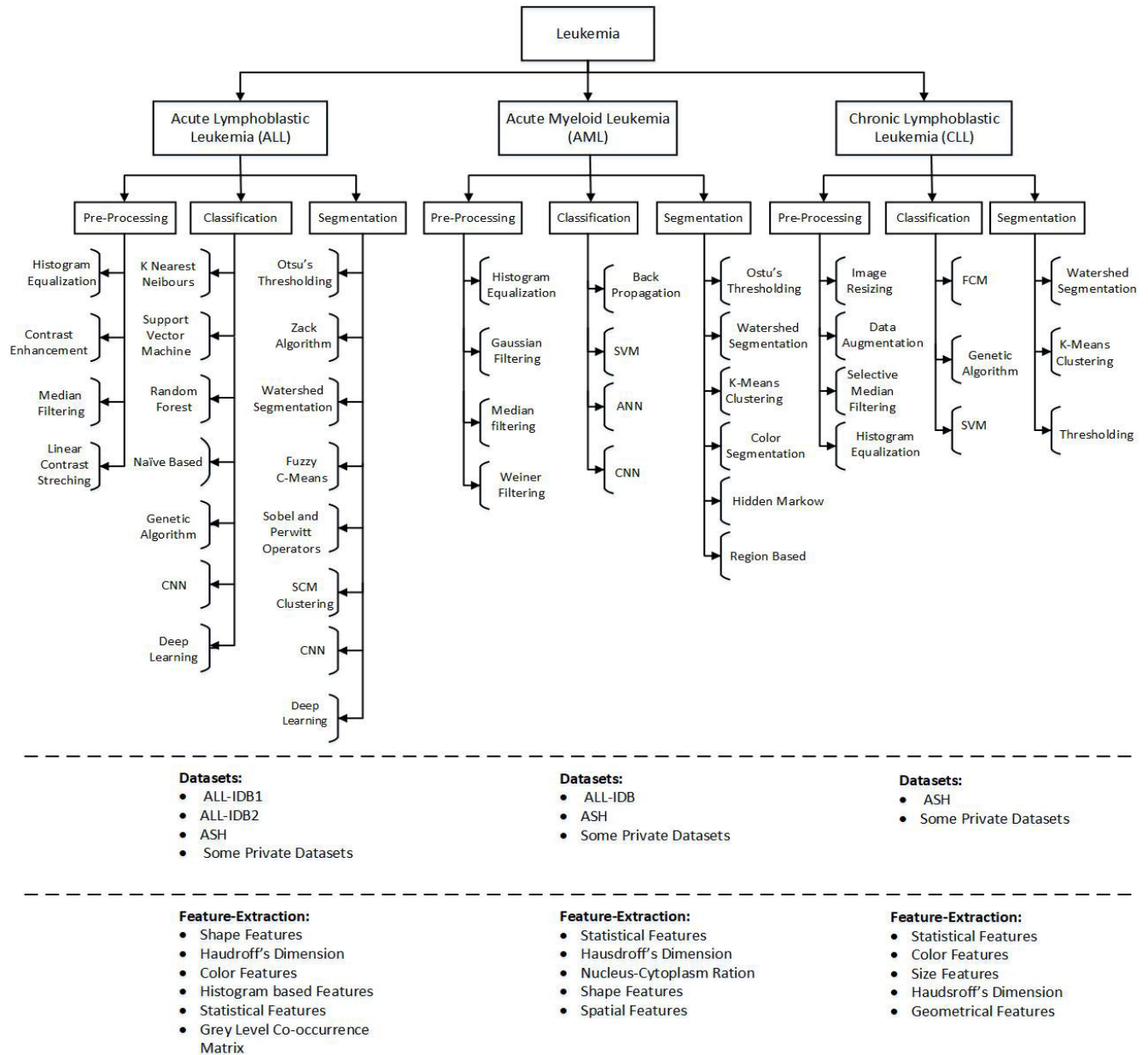


FIGURE 2. A review of computer aided diagnosis of Leukemia and its types.

commonly referred as leukemia. The abnormal production of malignant immature leukocytes termed as blast results/cells reduces the production of normal blood cells [4]. As it is crucial to maintain a required level of leukocytes in the blood, hence complete blood count i.e. CBC is often an indicator of disease. The causes of leukemia have not been known precisely perhaps due to the inherited and environmental factors. However, the diagnosis of leukemia at an earlier stage is mandatory as the abnormal leukocytes rapidly spreads through the blood stream and affects the other body organs [5]. The manual diagnoses of leukemia are less reliable, time consuming, tedious and inter-observer variation dependent. Therefore, automated computerized methods were introduced for Leukemia diagnosis in early 90s which

quickly became popular owing to their low cost, high accuracy and less or no manual involvement. Consequently, in the subsequent part of this section we shed some light on the origin and evolution of automated diagnosis of Leukemia. However, before that it is only natural to talk about the types or classes of Leukemia first.

### A. CLASSES OF LEUKEMIA

The types or classes of leukemia depends upon the type of affected leukocytes. If the immature leukocytes are lymphocytes then leukemia is classified as Lymphocytic Leukemia, and if the immature leukocytes are monocytes and granulocytes then it is considered as myelogenous Leukemia [6]. In general, leukemia can be categorised as acute (when

affected leukocytes are unable to perform as normal leukocytes) or chronic (when affected leukocytes can perform as normal cells), making chronic leukemia extreme [7]. An illustration of famous Leukemia types and their diagnoses is presented in Figure 2. Detailed description of each of these types is given below.

1) ACUTE LYMPHOBLASTIC LEUKEMIA (ALL)

Acute lymphoblastic or lymphocytic leukemia is the type of cancer occur in bone marrow and blood. Lymphoblast is the immature blast cell that has altered morphology and is activated by antigen which transforms it into a mature lymphocyte [8]. The French American British (FAB) classifies ALL into three classes i.e. L1, L2 and L3 according to their morphological differences. The L1 cells are homogeneous, vacuoles-less, round nucleus and contain cytoplasm. L2 cells are larger in size than L1 cells. They have irregular nucleus and varying cytoplasm. L3 cells are of homogeneous shape, with round nucleus and cytoplasm containing vacuoles [9]. Figure 3 depicts the image samples of healthy leukocytes and four leukemia types. Although numerous studies have been proposed for the automated diagnosis of acute lymphoblastic leukemia, only fewer studies have proposed the methods to diagnose the subtypes (L1, L2 & L3) of ALL [10].

2) ACUTE MYELOID LEUKEMIA (AML)

It is the most common type of acute leukemia that is particularly common in elderly people and is known as acute myelocytic leukemia, acute myelogenous leukemia, acute non-lymphocytic leukemia and acute granulocytic leukemia [12]. It is characterized by the formation of immature blast cells which are other than lymphocytes [13]. In AML, the bone marrow may also make abnormal red blood cells and platelets [14]. The symptoms of AML include anaemia, fever, ulcers of mucous membranes and granulocytic insufficiency [15], [16]. Because of the cells that form myeloid cells, AML differs from other types of leukaemia. FAB has classified the blast cells into two types as Type I and Type II. Both of the blast types have prominent central nucleoli which are usually three to five in number with well-defined and uncondensed chromatin. Type I has lesser number of cytoplasmic granules and bigger number of primary granules, compared to Type II blasts. The major discriminating factor of AML from other leukemia types is its eight sub types, according to FAB classification. These subtypes are shown in Table 1 along with their characteristics. For M1 to M6 AML diagnosis, typically >30% blasts proportion is needed in the bone marrow smear. M0-M5 AML subtype start in the primitive WBCs, whereas M6 and M7 start in primitive RBCs and platelets respectively [15].

3) CHRONIC LYMPHOCYTIC LEUKEMIA (CLL)

It is common in adults, particularly in people who are over the age of 60, and common in men than women. This type of leukemia is different from others in terms of symptoms: inflamed lymph nodes, anorexia, weight loss, fatigue, and

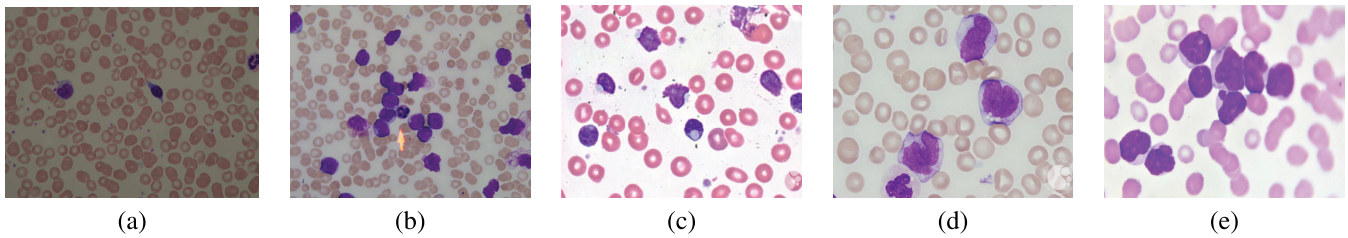
TABLE 1. Description of AML subtypes as per FAB classification [4].

AML Subtypes	Name	Characteristics
M0	Acute Minimally Differentiated Leukemia	Minimally differentiated leukemia with less than 50% erythroblast and 30% blasts cells in bone marrow smear
M1	Acute Myelogenous leukemia without maturation	Existence of Auer rods and lack of granulocytic maturation with myeloblasts dominate in the bone marrow
M2	Acute myeloblastic leukemia with maturation	Auer rods are found with the maturation of the neoplastic cells is beyond the myeloblastic and promyelocytic stages.
M3	Acute promyelocytic leukemia	It is known as the hypergranularpromyelocytic leukemia in which Auer rods are found with Promyelocytes being predominant in the bone marrow
M4	Acute myelomonocytic leukemia	Presence of myelocytic and myeloblastic components upto 20% of the total cellular components. It is the type of acute myelomonocytic leukemia (AMML) with folded, twisted on indented nuclei and lower ratio of cytoplasm in the nucleus. There is another type of M4 called acute myelomonocytic leukemia with eosinophilia (M4 eos).
M5	Acute monocytic leukemia	Presence of myelocytic and myeloblastic components upto 20% of the total cellular components and it is the case of pure monocytic leukemia with folded, twisted on indented nuclei and lower ratio of cytoplasm in the nucleus
M6	Acute erythroid leukemia	It is the known as erythroleukemia characterized by the abnormal proliferation of erythroid as well as granulocytic precursors.
M7	Acute megakaryoblastic leukemia (AMKL)	Found in the early platelets.

weakness. It is characterized by the accumulation of B-cell lymphocytes [12]. In CLL, almost 90% of leukocytes appear to be similar to normal cells. However, comparable to normal lymphocytes, the para chromatin is clearly separated and condensed from nuclear chromatin. Occasionally, it is less condensed relative to normal lymphocytes. That is why the diagnosis of CLL is a difficult task. In many lymphocytes, the nucleoli are clearly visible with moderate cytoplasm [16].

4) CHRONIC MYELOID LEUKEMIA (CML)

It is predominant in young and adults of the middle age. In CML, there is an excessive production of immature granulocytes that spill out of bone marrow and travel in the blood stream. Moreover, the percentage of myeloblasts is less than 10% with the excessive production of neutrophils, monocytes, myelocytes, and eosinophil myelocytes. In most of the cases, myelocytes and neutrophils exceed in quantity than the percentage of the normal leukocytes. The symptoms of CML include anaemia, weight loss, progressive enlargement of



**FIGURE 3.** Image sample of (a) Healthy leukocytes [11]; (b) acute lymphoblastic leukemia (ALL); (c) Chronic lymphocytic leukemia (CLL); (d) Chronic myeloid leukemia (CML); and (e) Acute myeloid leukemia (AML) [7].

spleen, fever and night sweats. Over time, a CLL patient gets a decline in the quantity of RBCs and platelets is associated with anaemia, bruising and blood loss [15].

This review paper entails a detailed review of CAD for the various types of leukemia. The research reviews different datasets and approaches used for various types of leukemia diagnosis. The review analyses and compares different approaches and presents a comprehensive review of the studies done in the recent past.

The organization of paper is as follows: The subsequent subsection presents authors take on the origin, progress, challenges and developments in automated diagnosis of Leukemia. Subsequently, a detailed discussion on materials and methods available in the literature for automated leukemia diagnostic systems is provided in Section II. That also presents detailed notes on image acquisition, publicly available datasets and stages involved in computer aided Leukemia detection, i.e., pre-processing, segmentation, feature extraction, and leukemia classification. Section III describes performance measures which have been used in the literature while Section IV presents a discussion on the state of the art of the automated diagnosis of Leukemia that not only describes the traditional methods but reports on modern machine learning and deep learning methods. Section V presents analysis and discusses on the reported work while Section VI concludes the research giving it a future dimension.

## B. BACKGROUND OF AUTOMATED DIAGNOSIS OF LEUKEMIA

This section gives a brief overview of the origin and evolution of the automated diagnosis of Leukemia. In this regard, major challenges faced by researchers and scientist along with their solutions, are also discussed separate subsections.

### 1) ORIGIN

One of the conventional methods of the diagnosis of leukemia is to analyse the medical history of a person and his family. However, that requires a series of appointments and blood test examinations. Cytogenetic analysis is another type of leukemia diagnosis method where abnormality in individual chromosomes is observed [17]. Complete blood count (CBC) is the conventional method of leukemia diagnosis which is performed by doctors as they check the complete count of WBCs, RBCs and platelets. The CBC identifies the leukemia

cells but is not reliable for the confirmation of leukemia in the patient. For this reason, bone marrow aspiration and microscopic examination of blood smear are the conventional manual methods used for leukemia diagnosis [18], [19].

Generally, Leukemia is diagnosed by the manual examination of microscopic blood smear images by physicians and experts. Immunohistochemistry is another method of analysis of the blood samples where the antigens of the cells in a tissue are analysed for the diagnosis of leukemia. Other methods for the leukemia diagnosis include interventional radiology techniques such as biopsy, catheter drainage, and percutaneous aspiration and other methods like Long Distance Inverse Polymerase Chain Reaction (LDI-PCR), Molecular Cytogenetics, and Array based Comparative Genomic Hybridization (aCGH). However, radiological methods suffer from resolution and sensitivity of the radio images and requires extensive labour work to diagnose various leukemia types [20].

The problem with these manual methods of Leukemia diagnosis is their labour intensive and time consuming nature that leads to lesser efficiency. Moreover, the accuracy of diagnosis is subjective as it requires highly trained medical professionals to perform sophisticated examination. These practical challenges paved the way for the advent of automated diagnostic systems for accurate and effective leukemia diagnosis. Apart from overcoming these problems of manual diagnosis, computerized diagnosis systems are found to be more reliable, efficient and accurate than manual detection techniques [10].

### 2) PROGRESS

Technological advancements have led to the progressive and improved methods of leukemia detection. Over the years, automated methods have been increasingly used in the medical for the diagnosis of cancer, malaria, skin diseases, and many others [21]. However, since last decade, numerous automated methods for various types of leukemia diagnostics have been proposed which are efficient and have improved accuracy of classification. Dozens of researches have been presented in this field related to WBC segmentation, blood counting, blood smear analysis and leukemia diagnosis using Image Processing Techniques [22]. In recent years, various studies have proposed computer vision, machine learning (ML) and deep learning (DL) frameworks that effectively classify the leukemia blood cells with better accuracy. One of the biggest advantages of computer vision-based leukemia detection methods is the ease in the diagnosis: instead of

analysis of blood and bone marrow smear, the image of these samples can be remotely transported [23]. These automated methods of leukemia detection using microscopic blood smear images have utilized variety of methods that are efficient and effective. This can help hematologists and pathologists to effectively and efficiently diagnose leukemia and provide the patient with the appropriate treatment in less time.

### 3) MAJOR CHALLENGES AND KEY DEVELOPMENTS

As automated diagnosis of Leukemia using microscopic images had increased since decades, but still there are multiple open problems for researchers to address or improve upon. Initially, noise, artifacts and poor quality of the acquired images were prominent challenges which reduced the efficiency of computer aided diagnosis (CAD) methods for Leukemia. These issues were progressively addressed by introducing various pre-processing techniques in the subsequent literature. To enhance the quality of the acquired image, various pre-processing methods were developed to enhance the quality of the medical image, e.g., Histogram Equalization, Contrast Stretching etc. However, the drawback of these methods was their inability to distinguish between noise and image pixels leading to poor image quality. Another technique used in literature is unsharp masking but it is vulnerable to noise and hence making image blur and difficult to diagnose Leukemia cells. That prompted researchers to employ variants of mean and median filters to filter out noise prior to image enhancement. However, these classical noise removal methods are known eat up important image details that adds to lack of image quality. Hence, development of reliable and efficient pre-processing techniques is still a major challenge in automated diagnosis of Leukemia.

The techniques used for detection of Leukemia cells from an image are generally based on Histogram based thresholding techniques for segmentation of region of interest. Despite their ease of implementation, these techniques require the selection of a suitable threshold that is a challenging task. In order to overcome this issue, researchers Watershed Segmentation is used despite its yielding over segmentation. A key issue with these classical techniques is that these methods may yield good results for one specific problem but may decrease accuracy for another.

Once the region of interest is segmented, a number of classifier are developed to classify the presence of absence of Leukemia cells. For this purpose, support vector machines (SVM) are most commonly used as a classifier. However, its efficacy is limited because it is a binary classifier hence it is not suitable for the classification of subtypes of Leukemia. It can only be used for classifying healthy cells and Leukemia Blast cells. Lately, Machine learning and Deep learning methods are being actively utilized for segmentation and classification of healthy and malignant cells owing to their increased accuracy. Nevertheless, this superior performance comes with challenges of its own which include computational cost and availability of large datasets to learn

properly and help in diagnosis efficiently. Key results in the automated diagnosis of Leukemia are presented pictorially in Figure. 2.

### 4) PRESENT DAY METHODS

Automated diagnosis of Leukemia has seen various innovative solutions which involve classical image processing methods and ML/ DL methods. These techniques follow a pipeline of specially designed stages generally termed as pre-processing, segmentation, feature extraction and classification of the disease. Here, success of every stage leads to success of previous stage. Hence, high accuracy of classification depends upon favourable outcome of all the stages, and is problem dependent. A key development in automated diagnosis of Leukemia is the use of innovative DL networks which have achieved higher amount of accuracy in segmentation and classification. The recent innovation, in this regard, is the concept of transfer learning that tunes a DL network with better accuracy for another problem to detect Leukemia. These networks includes AlexNet, ResNet etc. With the development of ML and DL techniques the accuracy of the Automated diagnosis of Leukemia increased but at the cost of vast resources spent on computing power and dataset collection.

## II. MATERIALS AND METHODS

An automated procedure for detecting lymphoblasts from colored microscopic images comprises of four successive steps which are segmentation, leucocytes identification, lymphocytes identification, and identification of potential lymphoblasts. The process of automated diagnosis of Leukemia start with acquisition of images from Leukemia patients which may be subject to noise, artifacts and poor contrast due to the imaging modality. Consequently, the acquired datasets are preprocessed to enhance their quality. Next, segmentation techniques are utilized to extract the foreground from the background pixels based on their distinguishable characteristics, e.g., color, texture and shape [24]–[26]. Lastly, WBCs are distinguished from RBCs through classifiers whereby the presence of nucleus is first determined from the color information of the cells. In the next phase, Lymphocytes are discriminated from the other WBCs by examining the nucleus structure and presence of cytoplasm. The final phase of classification examines the morphological deformation in the lymphocytes for the determination of potential lymphoblasts. This general methodology of automated CAD system using microscopic images and computer vision methods for leukemia detection and classification is depicted in Figure 4 and it consists of five distinct stages [27] that not only depict all the steps involved in automated diagnosis pipeline of Leukemia but also list key results in each stages of the pipeline. Here, we aim to present a graphical overview of the literature on leukemia diagnosis where divide listed techniques into their respective categories to impart better understanding to the reader.

In the following subsections, we present a thorough discussion on each stage which involves discussion and

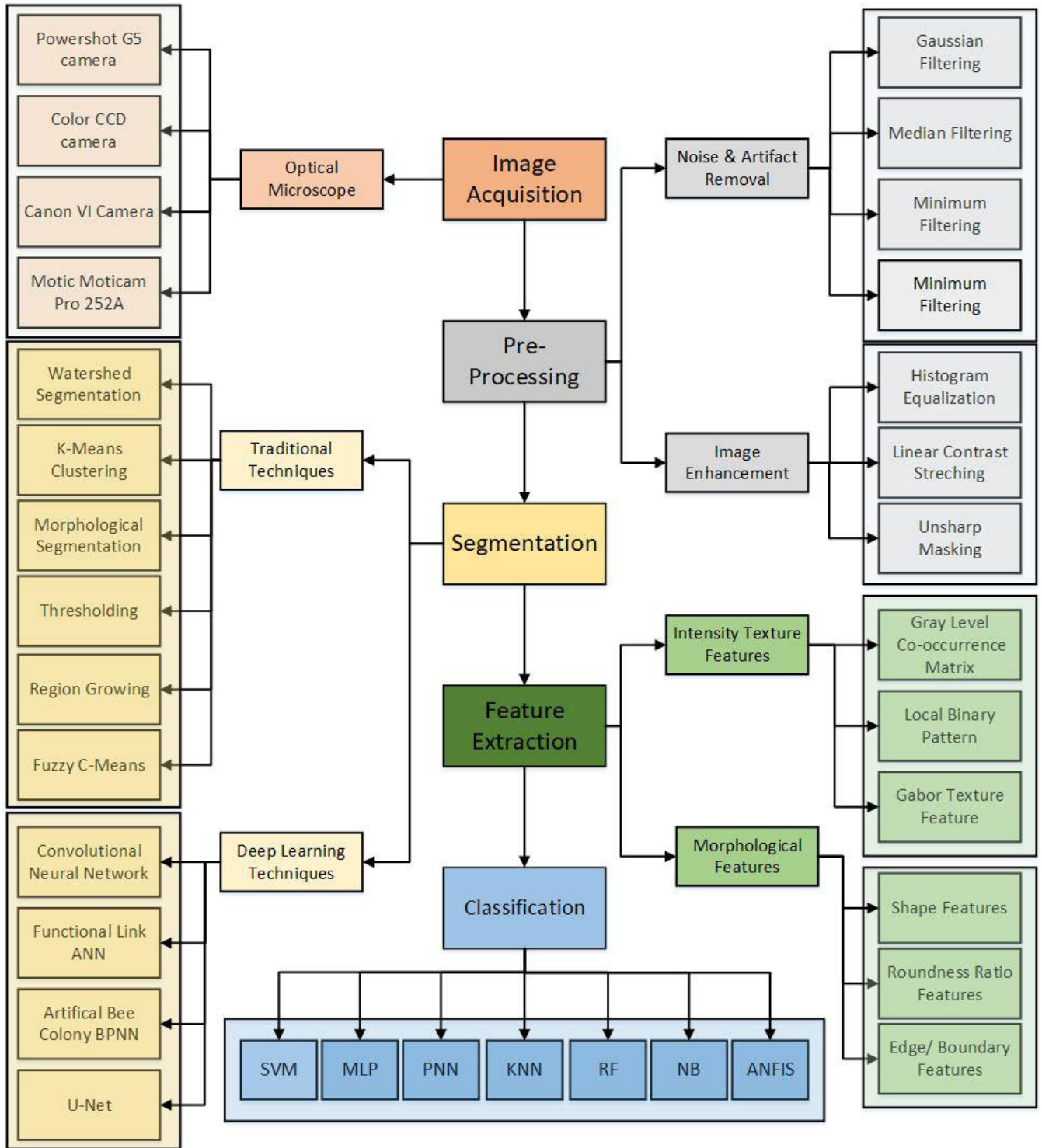


FIGURE 4. Illustration of the pipeline involved in automated diagnosis of Leukemia along with the list of key developments in each stage.

critical analysis of the key pre-processing, segmentation, feature extraction and classification algorithms.

**A. IMAGE ACQUISITION**

In recent years, many researches have proposed variety of techniques for leukemia diagnosis. For the purpose of

investigation, variety of public and private leukemia datasets have been formulated from the blood samples collected from healthy and leukemia patients. These blood sample images are acquired and digitized using different types of cameras and microscopes. In figure 4 we depicted types of Cameras and microscopes used for Image acquisition and we also give

brief overview of the Datasets that are Publicly available and can be used for further research, this section is normally ignored in literature. The datasets which are publicly available are used by many researchers for investigating the integrity of their proposed framework as given in figure 2. However, many researches have collected their own datasets from local hospitals and laboratories which are kept private and can be requested on demand. Furthermore, since DL based approaches require a large data repository for the training purpose, various data augmentation methods have been presented in the literature that makes the data large enough to be suitable for training the classification models. Various data augmentation techniques include image transformation methods such as image rotation, translation, shearing, blurring, mirroring, Gray scale image transformation and histogram equalization [28]–[30] have been used.

### 1) PUBLICLY AVAILABLE DATASETS

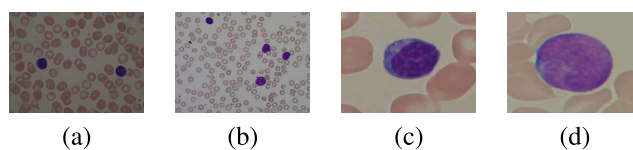
After reviewing literature, it has been inferred that 17% of the studies that have performed automated analysis of blood smear images have used public databases. Two of the most common used databases include ALL-IDB and ASH [14]. Other datasets are also available that are Bloodseg, Cellavision, JTSC, SMC-IDB and HUMS-IDB and discussed below.

**ALL-IDB** only contains annotated (done by oncologists) blood cell images of ALL subtype of leukemia and other images for the subtypes such as AML, CLL, CML does not exist. All of the images in the datasets were taken with a PowerShot G5 camera and kept in JPG (Joint Photographic Experts Group) format with a colour depth of 24 bits and native resolution of 2592 x 1944 pixels. The photos correspond to various microscope magnifications (from 300 to 500). This database is available in two forms (ALL-IDB1 and ALL-IDB2), both of which are based on segmentation and categorization and is available for free [31].

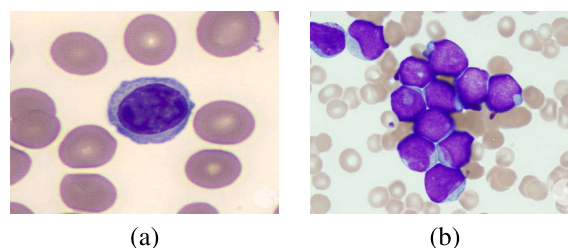
**ALL-IDB1:** It comprises approximately 39000 blood elements, with lymphocytes identified by oncologists, and available 510 lymphoblasts. Only the lymphoblasts in the image that are fully described have been considered and listed. The image files in this dataset are termed as “ImABC\_X.jpg,” for which ABC is a three-digit counter and X is a Boolean digit value of 0 when there is no any blast cells in the image and 1 if there is atleast one blast cell. A classification text file “ImABC\_X.xyc” is provided for each image.

**ALL-IDB2:** This image collection was created for evaluating the precision of classification structures. The ALL-IDB2 contains cropped areas of interest of blast and normal cells included in the ALL-IDB1 dataset. There are 260 in total images, with lymphoblasts accounting for half of them. The dataset is open to the public and free to use. If the significant cell is a possible blast, the images in the ALL-IDB2 dataset are called “ImABC\_1.jpg,” and “ImABC\_0.jpg” in the other instances. Figure 5 shows the Images of ALL-IDB1 and ALL-IDB2 dataset.

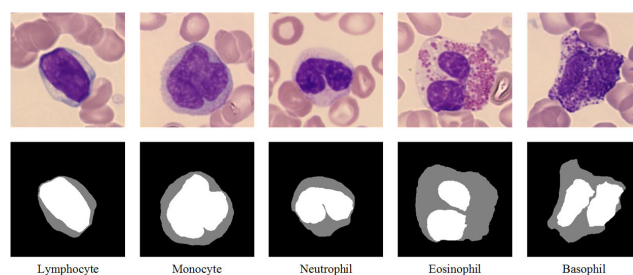
**ASH (The American Society of Hematology)** offers a large collection of highest-quality haematological images



**FIGURE 5.** ALL-IDB1 Images of: (a) lymphocytes (b) lymphoblast, ALL-IDB2 Images of: (c) Healthy cell from non-ALL patient (d) Lymphoblast from all patient.



**FIGURE 6.** ASH dataset Images of: (a) Healthy cell (b) Leukemia [33].



**FIGURE 7.** Cellavision dataset images of: colored images (First row), ground truth images (second row) [35].

taken with various microscopes at various resolutions in a web-based image library. The ASH images bank is a benchmark dataset that includes leukemic and normal (AML and ALL) blood photographs that have been labelled by specialist pathologists. The database consists of 240 images, including 100 images of healthy patients, 80 images of AML and 60 images of ALL. Around 420 of the sub-images were extracted from 100 healthy patients’ images, 80 AML patients’ images, and 60 ALL patients’ images [32]. The majority of the data was obtained between 2003 and 2015 from a variety of patients ranging in age from 4 to 73 years. Figure 6 shows the image samples obtained from ASH dataset.

**BloodSeg** image dataset contains 367 blood smear images that have ground truth segmentation done by an expert for the nucleus of the cell and captured at  $640 \times 480$  pixels resolution. All images are acquired with an optical microscope coupled to a color CCD camera where the magnification of the microscope is 100 times the objective of the lens.

**CellaVision** dataset is made up of 100  $300 \times 300$  colour images that were obtained from [34]. The images of the cells are usually purple color with possibility finding red blood cells surrounding the white blood cells as given in Figure 7.

**JTSC** dataset was obtained from Jiangxi Tecom Science Corporation in China [36]. It comprises 300 WBC images with a colour depth of 24 bits and a resolution of  $120 \times 120$  pixels. The blood smears in the dataset were treated with

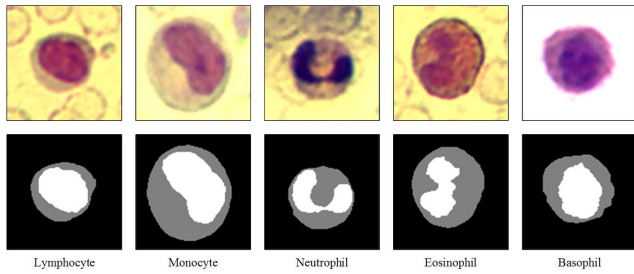


FIGURE 8. JTSC dataset images of colored images (first row), ground truth images (second row) [35].

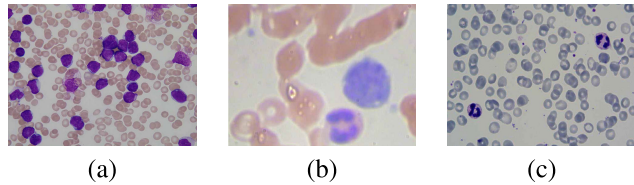


FIGURE 9. Image examples of (a) ALL-IDB1 (b) SMC-IDB (c) IUMS-IDB [33].

a new developed haematology substance for quick staining of WBCs, and the images taken were by a Motic Moticam Pro 252A optical microscope camera using an auto-focus microscope N800-D motorised. The majority of the images have a yellow background as shown in Figure 8.

**SMC-IDB**, [37] proposed this dataset. This image dataset of peripheral blood smears is obtained from the Giemsa-stained slides. The images in this dataset are distinct as they were captured by different cameras and microscopes. The dataset contains 367, 640 x 480 pixel representations of peripheral blood and is accessible at [38].

**IUMS-IDB** The Iran University of Medical Science [39] dataset, available at [40] contains 195 microscopic photographs taken from the peripheral blood of around 10 patients. WBC images are taken by Canon V1 camera that is attached on Canon optical microscope. The images are stored in jpg format with 300 dpi and 3872 x 2592 pixels resolution. The pictures are in RGB color format. Visually they are different from ALL-IDB and SMC-IDB photos. Figure 9 shows the difference of image samples from ALL-IDB1, SMC-IDB and IUMS-IDB. Despite the availability of the public datasets, some studies have also used blood smear images from Google or internet [28], [41]–[43]. Table 2 summarizes the publicly available leukemia datasets along with their specifications which are used in the automated detection of leukemia.

2) PRIVATE DATASETS

Various studies have used ASH and ALL-IDB public leukemia dataset. However, ALL-IDB dataset does not contain blood cell images of other leukemia subtypes. Furthermore, for ML and DL based frameworks considerable amount of data is needed to train the model [30]. For this reason, various researchers have collected their own dataset of blood smear images from pathologists and local leukemia

TABLE 2. Publicly available datasets for automated detection of leukemia.

Datasets	Details
Acute Lymphoblastic Leukemia Image Database for Image Processing (ALL-IDB) [11]	<p><b>Versions:</b> ALL-IDB1 and ALL-IDB2 fixated on segmentation and classification.</p> <p><b>ALL-IDB 1:</b> 109 images, 2592 x 1944 resolution, 39000 blood elements, 510 candidate lymphoblasts</p> <p><b>ALL-IDB 2:</b> 260 images, 257 x 257 resolution, 260 blood elements, 130 candidate lymphoblasts</p> <p><b>Color Depth:</b> 24 bits</p> <p><b>Camera Details:</b> All images are captured with a Canon PowerShot G5 camera</p> <p><b>Format:</b> JPG</p> <p><b>Magnification:</b> 300 -500</p> <p><b>Availability:</b> Can be downloaded from [31]</p>
American Society of Hematology (ASH) Image Bank	<p>ASH image bank contains variety of blood smear images including leukemia subtypes</p> <p><b>Availability:</b> Can be downloaded from [44]</p>
BloodSeg [45]	<p><b>Images:</b> 367 blood smear images that have ground truth segmentation done by an expert for the of nucleus of the cell</p> <p><b>Resolution:</b> 640 x 480 pixels</p> <p><b>Camera Details:</b> All images are acquired with an optical microscope coupled to a color CCD camera.</p> <p><b>Magnification:</b> 100</p>
CellaVision	<p><b>Images:</b> 100 color images that have ground truth segmentation done by a hematologist for the of nucleus and cytoplasm of the cell</p> <p><b>Resolution:</b> 300 x 300 pixels</p> <p><b>Availability:</b> Can be downloaded from [35]</p>
JTSC	<p><b>Images:</b> 300 WBC images that have ground truth segmentation done by a hematologist</p> <p><b>Color Depth:</b> 24 bits</p> <p><b>Camera Details:</b> Motic Moticam Pro 252A optical microscope camera with a N800-D motorized auto-focus microscope</p> <p><b>Resolution:</b> 120 x 120 pixels</p> <p><b>Availability:</b> Can be downloaded from [35]</p>
SMC-IDB [37]	<p><b>Images:</b> 367 peripheral blood images</p> <p><b>Resolution:</b> 640 x 480 pixel</p> <p><b>Availability:</b> can be downloaded from [38]</p>
IUMS-IDB [39]	<p><b>Images:</b> 195 microscopic photographs taken from the peripheral blood of 10 patients.</p> <p><b>Camera details:</b> Canon V1 camera attached on Canon optical microscope.</p> <p><b>Format:</b> JPG</p> <p><b>Resolution:</b> 300 dpi and 3872 x 2592 pixels</p> <p><b>color space:</b> RGB.</p> <p><b>Availability:</b> can be downloaded from [40]</p>

hospitals [27], [46]–[48]. The details of the datasets used by the past studies are given in the next sections.

B. COLOR/STAIN NORMALIZATION

Numerous cancer types diagnosis can be performed by analysing histological image samples [49]. The histopathological image sample diagnosis based absolutely on the visually analysing tissue images under a microscope. Though, the appearance of histological tissue may have several color intensities depends upon the ability of operator, scanner specifications and staining process. Due to this variability in stain the diagnosis is affected, and accuracies of CAD systems also decreases. For this reason, the color/stain normalization method has been introduced. This technique proved to be an effective tool, letting to normalize of the color/ stain appearance of a original image with respect to a reference



image. Many of the researchers in literature have proposed techniques for Staining/ Color normalization. In [50] a novel ACD (adaptive color deconvolution) algorithm is proposed for separating stain and normalizing color of hematoxylin-eosin-stained WSIs (whole slide images). Four datasets of WSI were used to evaluate this technique comprising lung, cervix, and breast cancers and comparison between 6 state-of-the-art techniques was performed. The technique proposed attained the most reliable performance in color normalization as per the quantitative metrics. In [51] an automated novel stain normalization and separation technique for H&E stained slides of histological images is proposed. The suggested techniques Stain Color Adaptive Normalization (SCAN) can enhance contrast between background and histological cells/tissue while preserving local formation. Discussed methods were validated both quantitatively and qualitatively on multiple datasets, providing extremely satisfactory outcomes with lowest possible computational costs. In [52] a proposed technique takes H&E (Hematoxylin and Eosin) stained color digital images as an input to recognize lymph cell. The method includes cells segmentation from extracellular matrix, feature extraction, categorization and overlap resolution. We can observe that since 2008 there is increasing number of techniques proposed for normalization of H&E-stained histological images. Yet, drawbacks/weaknesses of these proposed methods and future possibilities in research encourage the analysis of new methods that may provide pertinency in medical applications and promises to provide greater performance of Computer Aided Diagnosis systems.

### C. PRE-PROCESSING

Images are often distorted by variety of factors that distort the image quality [53]–[57]. This includes the light source, poor contrast, background clutter, blurriness, or any other type of noise [58]–[61]. Pre-processing is the step intended to remove such kind of unwanted entities from the images [62]–[65]. It is one of the crucial steps in a variety of domains particularly medical [66]–[69]. Thus, it is one of the important steps to pre-process the medical images using variety of simple and complex image processing and computer vision techniques so as to improve the accuracy of the diagnostic systems [70]–[72]. A variety of image pre-processing methodologies have been presented in literature that pre-processes the blood images to enhance the region of interest (ROI) to be utilized further for diagnosis of leukemia [73]. These techniques are of two main categories, one is noise and artifact removal, another is image enhancement. For each category we also enlisted the techniques commonly used in literature for Noise and Artifact removal as well as for Image enhancement as depicted in figure 4. These two main categories are further divided into different types and most common of them are the Histogram Equalization, Linear Contrast Stretching, Unsharp Masking, and application of filters like Gaussian, Median, Minimum, and Weiner Filters. Occasionally, the results of filter application or contrast adjustment on the blood images are not satisfactory and it is better to convert the image to

another domain as from Red, Green, Blue (RGB) domain to Hue, Saturation, and Value (HSV) or Hue, Saturation, and Lightness (HSL) domains for better identification and efficient detection of region of interest [74]–[76]. An overview of some of the excessively used pre-processing techniques is as follows.

#### 1) HISTOGRAM EQUALIZATION

It is among the simplest image processing techniques that enhances the contrast of the image by normalizing its histogram [61], [77]. It is useful for images which have either darker background or brighter foreground and vice versa. Thus, it is suitable to improve the background and the blood image contrast [78]. However, it is not a robust technique since it is unable to distinguish between noisy and true image and thus reduce the image quality.

#### 2) LINEAR CONTRAST STRETCHING

It enhances the contrast of blood image by spreading its intensity values to a dynamic range. Thus, it normalizes the intensity values present in an image. It is a simple and useful technique suitable for low contrast images [79]. However, it is highly vulnerable to noise and its effectiveness is reduced if the image contains outliers.

#### 3) UNSHARP MASKING

It is a simple method to produce sharp images and remove blurriness from the image. It can be applied to a whole image or onto selective parts depending upon the need and the required adjustment in the image. Yet, this technique is not robust as it is highly sensitive to noise.

#### 4) GAUSSIAN FILTERING

It is a simple filter for removing blurriness and noise from images [80]. However, it is not beneficial in the cases when it is used alone for removing noise as it can often blur the fine details such as edges and may affect the contrast of the image.

#### 5) MEDIAN FILTERING

It is one of the famous nonlinear filters for removing salt and pepper noise. Unlike Gaussian filter it preserves the fine details and sharp edges in the image. However, its efficacy is limited to the images having low density of salt and pepper noise and does not perform well for the images which have a percentage of salt and pepper noise.

#### 6) MINIMUM FILTERING

It is useful for removing salt or the positive outlier noise. It highlights the lighter objects to be easily recognized after the process of segmentation [81], [82]. However, it is not robust enough for highly noisy blood images as it can remove details present in an image.

#### 7) WEINER FILTERING

It produces the linear estimation of the original image by minimising the mean square error of the ideal and original

image. It removes the blurriness and additive noise from the image making it approximately closer to the original image.

Despite the benefits of all of the aforementioned techniques, all of them lack in one aspect or the other and in some cases lead to the loss of the image details. Furthermore, these are feasible for some low contrast blood images that do not have high density of noise. Thus, researchers have emphasized the need of efficient and reliable pre-processing techniques to enhance the image quality for the better diagnosis of leukemia [83].

#### D. SEGMENTATION

In automated diagnosis system of disease, Image segmentation is a main part that analyse bone marrow and blood smear images [84]–[86]. In this step an image is subdivided into its constituent objects or regions [87], [88]. The process of leukocytes segmentation separates the cell from its background. This is frequently done by the distinction of nucleus from cytoplasm of the cell [89]. This can be conveniently done through variety of image processing techniques and also available in various medical softwares. A variety of techniques for image processing have been proposed in literature which produces a binary image of leukocytes that is the mask of the original color image [42]. Reference [17] proposed a framework for blood image segmentation that segments nucleus of the leukocyte cell and produces good results. For Segmentation techniques, again there are two main categories, one is Traditional methods that were initially used for Automated diagnosis of Leukemia along with latest trends as Deep Learning Methods utilized nowadays in order to provide better accuracy with less computational time as shown in figure 4. These categories further divided into famous techniques as cauterization, morphological filtering, contrast stretching, thresholding, color space conversion and watershed segmentation and various Deep learning networks [90], [91]. Reference [14] has reviewed various methods of segmentation that have been used for segmenting bone marrow and peripheral blood smear images for classifying various types of leukemia. It is observed that segmentation of leucocytes is performed using clustering particularly K-means clustering and neural networks during the years 2016-2020. Contrast Stretching is found to be the best method for enhancing the nuclei of the WBC followed by the application of morphological filter that averages the diameters of the WBC. Various studies have looked for the morphological alterations in the grayscale microscopic blood smear image of the WBC [92]. Some studies have performed segmentation of the peripheral blood smear and bone marrow images based on Neural Networks along with the classification of leukemia [7], [36], [93]–[97]. The common methods of segmentation of peripheral blood smear and bone marrow images that are used in literature for automated detection of leukemia are discussed as follows.

##### 1) THRESHOLDING

It is simple yet productive method for image segmentation used in variety of medical studies and particularly in leukemia

detection. It converts the gray scale image into binary image by segmenting foreground image from background regions. Depending upon the threshold value, if the intensity value of the image region is less than the predefined threshold value, then the resultant image will have zero value (black) or if the intensity value of the image region is greater than the predefined threshold value, then the resultant image will have one value (white color). Thus, a resultant image produced is a binary image made up of black and white pixels. Foreground and background are not always fixed and depends upon the region that is desired to be extracted. It is an effective technique for high contrast images and gives good results of segmentation [98]. Global thresholding or Otsu's thresholding is an extensively used methods in leukemia studies but it blurs the edges of the blood cells so after performing Otsu thresholding edge preserving filters are applied [99].

Lymphocytes detection is extensively been done by Zack's algorithm. This algorithm segments the blood smear image using a threshold value that forms a line among the lowest to the highest histogram value [32]. Thus, this method is called as triangle oriented thresholding method [100].

##### 2) REGION GROWING

Type of region-based segmentation that effectively selects the region of interest from the predefined criteria. Such criteria can be fine details such as edges or pixels as seed points [101]. This is actually a pixel-based segmentation where initially a single pixel is selected and as per criteria the neighbouring pixels are selected and connected in the region of interest of the initial pixel. It is a widely used method in cancerous detection [102].

##### 3) WATERSHED SEGMENTATION

It is a form of region growing segmentation technique that starts from marker (initial pixel) and based on the predefined criteria it includes the neighbouring pixels [103]. It is extensively used segmentation technique in the medical images because of the efficacy of production of its results as it extracts the closed and connected regions relative to other segmentation techniques that distorts the boundaries and disconnected regions are produced for which morphological operations need to be performed. Since local minima can produce over segmentation of the blood smear images, thus variants of watershed segmentation are used. One such variant is the marker controlled watershed segmentation that uses external and internal markers for segmentation of the region of interest [104].

##### 4) MORPHOLOGICAL SEGMENTATION

This extracts different components from the blood smear image based on the set theory in mathematics. It has the basic idea of moving a predefined structuring element over the image like a sliding window and specific region will be extracted from the image depending on the morphological operation performed. The common morphological operations that are used in leukemia detection are erosion, dilation,

opening, closing, and hole filling [105]. This type of segmentation is effective for regions and shape-based feature descriptors and for the same reason it is used in the segmentation of microscopic blood images [106], [107].

#### 5) K-MEANS CLUSTERING

A method that segments the image regions into clusters of similar objects. It is a semi supervised learning technique for the unlabelled data. It divides the images into K clusters using the total number of blood smear images so that images having the similarity can be grouped into one cluster. It is an extensively used method for blood image segmentation that extracts the WBCs and blasts from the image [108].

#### 6) FUZZY C-MEANS

It is general form of K-means clustering that allows each data point to have same cluster. This technique can effectively handle outliers and works well for the noisy medical images unlike the segmentation done by K-means clustering. It gives better results for the images that have high illumination or percentage of noise but it is a time consuming method because of the long calculations [109].

Since last decade, many studies have performed artificial intelligence (AI) based segmentation of the WBCs. These studies utilize some sort of neural network for the segmentation purpose, especially for the medical images. Some of the extensively used are Convolutional Neural Network (CNN), Support Vector Machine (SVM), Functional Link Artificial Neural Network (FLANN), and Artificial Bee Colony Back-Propagation Neural Network (ABCPNN). These methods segment the microscopic peripheral blood and bone marrow images and are also used for classification of leukemia cells.

### E. FEATURE EXTRACTION

Feature extraction is a phase of prime importance in image processing, pattern recognition and machine learning. It is the phase intended for dimensionality reduction; where large datasets are represented by distinct set of values that are non-redundant and significant enough to be used for recognition and classification purpose. In the context of leukemia detection, the various type of features that are used in my studies include shape, color, texture, fractal dimension, statistical, and many others [110], [111]. For Feature Extraction, we also enlisted the Key features that are commonly extracted for Automated Leukemia Diagnosis. These Features mainly falls into two categories, one is Intensity and Texture Features another is Morphological Features. Also the main techniques used to extract both the Feature categories are given in figure 4. The features extracted at this phase are used in the next phase for leukemia detection and classification. The accuracy of the detection of leukemia highly depends upon the features that are used for classification purpose. Thus, selecting the best and non-redundant features is of utter importance for the accurate diagnosis of leukemia. The detection of acute

lymphoblastic leukemia is highly dependent on the feature extraction phase as the characteristics of the nucleus and cytoplasm of the blast cells [110]. The brief overview of the features that have been used in the literature for leukemia detection is presented as follows.

#### 1) INTENSITY AND TEXTURE FEATURES

Texture plays an important role in the medical images because of the ease of selection of region of interest (ROI). Additionally, color and intensity also plays the significant role in the ROI selection. In particular, for the identification and classification of blast cells, texture and intensity of the blood smear features are important and significant for leukemia detection. Gray Level Co-occurrence Matrix (GLCM) is a statistical method for feature extraction that examines the texture by analysing the spatial association between the pixels. In the context of leukemia detection, it is a useful method to extract features from the blood smear image based on its texture and intensity and thus classify the blast cells [110]. Gabor texture feature is another important method to extract the texture features [99]. Local Binary Pattern (LBP) is a fast and efficient texture feature that is widely used for leukemia detection particularly for the detection of blasts through the variance in the illumination [112], [113]. Reference [114] proposed a variant of LBP that is promising for the detection of leukemia cells. Entropy is the texture feature that can be used for measuring the randomness of the nucleus from the blood smear image, leading to a reliable feature for detection of the acute leukemia [115]. Fractal dimension is an extensively used feature to measure quantitative information from the blood and bone marrow smear images. The fractal geometry of the nucleus of the WBC is measured to analyse the roughness of the cell and to identify it as either blast or normal [116]. Hausdorff dimension is used with the fractal dimension as a feature to measure the roughness of nucleus and is widely used for the analysis of microscopic blood smear images [117]. First order statistical features are important features that are based on the original pixels of the blood smear images. These features are based on the histogram and include mean, standard deviation, energy, entropy, skewness and kurtosis [118]. Since blood cells are darker than the background so color features are extensively used for the classification of blasts. Usually mean values of different color models are used as features to the classifiers [119]. Usually color features are used by extracting features in the HSV domain.

#### 2) MORPHOLOGICAL FEATURES

Morphology is the important factor to be considered for medical image processing. In leukemia detection and classification, morphology of the blast cells is an important feature since they have area, perimeter, and circularity that is different from the normal cells, thus providing ease in the classification of blasts [120].

Shape based features such as circulatory, solidity, eccentricity, area, perimeter, and many others can help in the

classification of the blast cells. These features have significantly been used for the acute leukemia detection [121]. Bending energy is another shape-based feature that is used for acute leukemia detection particularly acute lymphoblastic leukemia through the boundary of the cell and its curvature [122]. Roundness Ratio is also an important feature that is extensively used for leukemia detection and WBCs counting. It is an efficient feature for the improved classification of leukemia cell and its subtypes because of the variance in the circular shape of the blasts that makes it a distinctive among other features [123]. Chain code features are widely used in acute leukemia detection. These features trace out the nucleus and cytoplasm boundary thus separating the nucleus and cytoplasm of blasts [96]. Some studies have used morphological and texture features in conjunction while some have utilized NN based features [105]. Reference [27] used a variety of features for lymphoblasts classification such as Hausdorff dimension, signature contour, shape features that include eccentricity, perimeter, compactness, elongation, form factor, area, and solidity, color feature as homogeneity, energy, correlation, and entropy. The combination of features produces a better accuracy of lymphoblasts classification. Reference [73] utilized the basophilia intensity texture features to characterize the cytoplasmic profile of the leukemia cell. This type of texture feature is obtained by thresholding segmentation method to the green component of the RGB blood smear image followed by counting the pixels of this region. Reference [74] obtained features through histogram and used local direction operator (LDP) to extract features from the nucleus of the leukemia cell. Reference [124] proposed a method for efficient leukemia detection. Their method utilized Fisher's Discrimination Ratio (FDR) followed by Exhaustive Search to get three distinctive features which are diameter, cluster prominence of the nucleus, and the minor axis of the ellipse bounding.

Since a lot of information is present in the blast cells, so a larger feature set for classification might not produce good results. Thus, feature selection should be performed to reduce the dimensionality of the feature set and for the selection of non-redundant features prior to the classification phase. The common algorithms used for feature reduction are Local Direction Pattern (LDA) and Principal Component Analysis (PCA). Some other feature selection methods that have been used in the literature include Fisher's Discrimination Ratio followed by application of Exhaustive Search [4]. Reference [125] used PCA to effectively reduce the redundancy in the feature set for leukemia classification [110]. Probabilistic Principal Component Analysis (PPCA) and Genetic Algorithm (GA) are also efficient methods for feature reduction [110], [126].

## F. FEATURE CLASSIFICATION

Classification of leukemia is normally through supervised learning where model is trained on the labelled data and tested on the new data that is different from the training data [89]. Classification model for leukemia can be partially

or fully automated. For Classification we also enlisted the key Classifier types Which have been used for Classification purpose in literature in figure 4. Machine learning and deep learning models are also used for the classification of leukemia cells. The common classifiers used in the literature are SVM (Support Vector Machine), Multilayer Perceptron (MLP), Random Forest (RF), K-Nearest Neighbor (KNN), Probabilistic Neural Network (PNN), Naive Bayes (NB), and Adaptive Neurofuzzy Inference System (ANFIS). Brief overview of the different classifiers used for the leukemia detection is given as follows.

### 1) SUPPORT VECTOR MACHINE (SVM)

It is one of the extensively used algorithms for blood cell analysis, leukemia detection and classification. It is extensively used for binary classification as normal vs blast or normal vs acute leukemia classification. However, for subtype classification of any type of leukemia a hybrid approach is used [127], [128].

### 2) MULTILAYER PERCEPTRON (MLP)

It is an artificial neural network model that is simple and has powerful computational capability. Its design is layered and each layer is connected to the subsequent layer through a network of nodes. which are neurons that map inputs to the outputs through the activation function [129].

### 3) PROBABILISTIC NEURAL NETWORK (PNN)

It is an extensively used classifier for leukemia detection and classification [130]. It approximates the probability distribution function (PDF) of each class of the leukemia cell by application of Bayes theorem.

### 4) K-NEAREST NEIGHBOR (KNN)

A simplest and efficient method that is extensively used for classification and regression scenarios. It uses the lazy learning and nonparametric method to classify the acute leukemia [131].

### 5) RANDOM FOREST (RF)

It is a classification model that uses ensemble learning to classify an object from the input feature vector. It is an efficient classifier that is not prone to overfitting and contains diverse groupings of trees that perform voting for the selection of the output class and the class with maximum votes is selected. In the context of leukemia detection and classification, RF efficiently classifies between the normal and blast cells [132], [133].

### 6) NAIVE BAYES (NB)

It is a probabilistic classifier that is efficient and simple. It is based on the Bayes' theorem that assumes independence in the feature values and it is extensively used for general leukemia classification and for acute lymphoblastic leukemia detection from the microscopic blood images [134], [135].

7) ADAPTIVE NEUROFUZZY INFERENCE SYSTEM (ANFIS)

It is a hybrid method that is a combination of fuzzy logic and ANN. It is a robust classifier that is extensively used in the medical images and performs better than the RF algorithm even if the feature set is not strong. Thus, it is extensively used for the classification of blast and normal cells from the microscopic blood smear images [136], [137].

III. PERFORMANCE MEASURES

This part explains the metrics that number of researchers consider in order to evaluate the efficiency of the suggested approaches. A system that can detect the existence of lymphoblast cells in given images could operate with a variety of module structures and the classification is based on the correctly classified blast cell or no blast. As a result, ( $T_P$ ) (True Positives) can be calculated as the number of cells properly categorized by the test as positive; ( $T_N$ ) (True Negatives) can be calculated as the quantity of cells classified correctly by the test as negative; False Positive ( $F_P$ ) shows to the quantity of cells that the test classifies as positive but aren't; False Negative ( $F_N$ ) indicates to the number of elements that the test classifies as negative but aren't.as. These measures can be calculated as shown in Eq.1, 2, 3 and 4, respectively. Whereas  $Seg_{img}$  is the Image segmented, and  $GT_{img}$  is the Image Ground truth. These expressions are used to describe how well system performs.

$$T_P = \sum_{i=1}^N [(Seg_{img}(i) = 1) \& (GT_{img}(i) = 1)] \quad (1)$$

$$T_N = \sum_{i=1}^N [(Seg_{img}(i) = 0) \& (GT_{img}(i) = 0)] \quad (2)$$

$$F_P = \sum_{i=1}^N [(Seg_{img}(i) = 1) \& (GT_{img}(i) = 0)] \quad (3)$$

$$F_N = \sum_{i=1}^N [(Seg_{img}(i) = 0) \& (GT_{img}(i) = 1)] \quad (4)$$

True positive rate ( $TPR$ ) is the correctly classified elements' probability with leukemia and can be calculated as shown in Eq.5. False positive rate ( $FPR$ ) is the correctly classified elements' probability without Leukemia and can be determined as shown in Eq. 6.

$$TPR = \frac{T_P}{T_P + F_N} \quad (5)$$

$$FPR = \frac{F_P}{F_P + T_N} \quad (6)$$

Sensitivity ( $Sn$ ), called as a recall, is the fraction of the true positive results to the total results that are supposed to be as shown in Eq.7. Specificity ( $Sp$ ) represents the ability to identify background pixels as shown in Eq.8. Accuracy reflects the model's ability to achieve correct predictions and can be calculated using Eq. 9. Precision also termed a  $PPV$  (positive predicted value) is fraction of true positive results to

the total positive results as given in Eq.10.

$$Sensitivity/Recall = \frac{T_P}{T_P + F_N} \quad (7)$$

$$Specificity = \frac{T_N}{T_N + F_P} \quad (8)$$

$$Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \quad (9)$$

$$Precision/PPV = \frac{T_P}{T_P + F_P} \quad (10)$$

F1-score identifies the not correctly classified cases by using precision and recall as given in Eq.11.  $AUC$  (The area under the curve) used to estimate the performance of binary classification model as used in [138]. It is the area underneath the receiver operating characteristic curve that plots two parameters; sensitivity over 1-specificity as given in Eq. 12. Classification Error( $CE$ ) as the total error in an analyser calculated as given in Eq. 13

$$F_1Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

$$AUC = \frac{1}{2} \left( \frac{T_P}{T_P + F_N} + \frac{T_N}{T_N + F_P} \right) \quad (12)$$

$$CE = \frac{T_P + T_N}{F_N + F_P} \quad (13)$$

The jaccard similarity is the fraction of the intersection and the union of ground truth and segmented image presented in Eq. 14 and the overlapping error can be calculated using Eq. 15

$$Jaccard = \left\{ \frac{S_{img}(Area) \cap G_{img}(Area)}{S_{img}(Area) \cup G_{img}(Area)} \right\} \quad (14)$$

$$Error = 1 - \left\{ \frac{S_{img}(Area) \cup G_{img}(Area)}{S_{img}(Area) \cap G_{img}(Area)} \right\} \quad (15)$$

where  $S_{img}$  and  $G_{img}$  represent the segmented region and ground truth respectively. A receiver operating characteristic ( $ROC$ ) curve illustrates the clinical specificity and sensitivity relationship for every single probable cut-off. It is a graph with x-axis displaying  $FPR$  and y-axis displaying  $TPR$ . The Matthews Correlation Coefficient ( $MCC$ ) is used to calculate of the quality of binary classifiers was presented by Brian W. Matthews biochemist [139] and can be calculated as depicted in Eq. 16.

$$MCC = \frac{(T_P \times T_N) - (F_P \times F_N)}{\sqrt{(T_P + F_P)(T_P + F_N)(T_N + F_P)(T_N + F_N)}} \quad (16)$$

Pearson Correlation Coefficient ( $PCC$ ) is image quality index and is used to measure linear correlation between source image and the normalized image, and its value is between 0-1, where result greater than 0 shows the correlation of two said images and value 0 indicates that two images are unsimilar.  $PCC$  can be calculated using Eq.17.where  $\chi_i$  and  $y_i$  indicates source image and normalized image and  $\mu_\chi$  &  $\mu_y$  are the source image and normalized image means.

$$PCC = \frac{\sum_i(\chi_i - \mu_\chi)(y_i - \mu_y)}{\sqrt{\sum_i(\chi_i - \mu_\chi)^2} \sqrt{\sum_i(y_i - \mu_y)^2}} \quad (17)$$

Quaternion Structure Similarity Index Metric (*QSSIM*) is another image quality metric and its range lies between 0-1 where value 0 depicts poor color normalization and value tends towards 1 indicates that the method used for color normalization is better. (*QSSIM*) can be calculated by given below Eq. 18. Where  $\mu_{qref}$  &  $\mu_{qnorm}$  indicate sample image and normalized images; and the standard deviations are denoted by  $\sigma_{qref}$  &  $\sigma_{qnorm}$ .

$$QSSIM = \left| \left( \frac{2\mu_{qref}\mu_{qnorm}}{\mu_{qref}^2 + \mu_{qnorm}^2} \right) \left( \frac{2\sigma_{qref}\sigma_{qnorm}}{\sigma_{qref}^2 + \sigma_{qnorm}^2} \right) \right| \quad (18)$$

Structural Similarity Index Metric (*SSIM*) is a perceptual measure for calculating degradation in image quality, and its value lies between 0-1 where near to zero indicates poor normalization method and near 1 indicated that the method for color normalization is better. Three factors are involved in calculating *SSIM*, i.e. Luminance, contrast and Structure. and mathematical formula to calculate *SSIM* is given below in Eq.19

$$SSIM = \left\{ \left( \frac{2\mu_x\mu_y + c_1}{\mu_x^2 + \mu_y^2 + c_1} \right) \left( \frac{2\sigma_{xy} + c_2}{\sigma_x^2 + \sigma_y^2 + c_2} \right) \right\} \quad (19)$$

where  $\mu_x$  and  $\mu_y$  are the original and normalized images and  $\sigma_x$  &  $\sigma_y$  are their standard deviations and the constants are  $c_1$  and  $c_2$ . Reference [140] used *PCC*, *QSSIM* and *SSIM* image quality metrics to calculate the performance of its purposed stain normalization method. Absolute error is an error quantity between true and normalized image. This error can be evaluated by using formula given in Eq.20

$$(\Delta) = |x_i - x| \quad (20)$$

#### IV. REVIEW OF THE LEUKEMIA DETECTION METHODS FROM PERIPHERAL BONE MARROW AND BLOOD SMEAR IMAGES

This section discusses the research performed for blood diseases, especially Leukemia identification and Classification. In this regard, we first discuss traditional image processing methods for automated diagnosis of Leukemia which are followed by a detailed discussion on recent state of the art ML and DL-based frameworks.

##### A. TRADITIONAL METHODS

These methods employ conventional image processing algorithms for segmentation, feature extraction and classification of Leukemia. A class of methods for ALL and AML leukemia diagnosis employ classical thresholding techniques for segmentation and extraction of morphological and texture features which are subsequently fed to support vector machines (SVM) for classification [32], [141]–[143]. SVM classifies between the normal and blast cells while further classifying them into its sub-types: L1, L2 and L3. A framework for automated diagnosis of ALL employs geometry and KNN for acute lymphocyte classification in peripheral

blood microscopic images [144]. Similarly, in [145] geometric, color, statistical and textural features were used to classify between malignant and benign cell using Naive Bayes and K-NN. Here, 60 image samples of blood were used for evaluation that provided accuracy of 92.8% for classification. A system for classification of two acute leukemia types, i.e., AML and ALL, was presented in [146] which employed twelve manually obtained features from images. Here, again classification was done using K-NN whereby experiments were run on 1500 image dataset, providing an accuracy of 86%.

A mobile-cloud-assisted segmentation and classification framework for leukocytes into sub-classes was proposed in [147]. This method used k-means clustering algorithm for removing irrelevant components via morphological operations on a number of features, i.e., statistical, texture and geometric. Subsequently, an ensemble multi-class SVM was used for classification at the end. Here, the dataset used for evaluation consist of 1030 images of WBCs and 98.6% accuracy was obtained.

A system for classification of AML and its subtypes M4, M5 and M7 was introduced by Setiawan *et al.* [148]. Initially, the cell segmentation was performed using a color k-means algorithm. Classification was done on six statistical features using multi-class SVM. That led to 87% accuracy for segmentation and 92.9% accuracy for classification. A three-layered feature extraction framework, coding and classification [149] uses blood smear images to identify leukemia and its type. For this purpose, feature extraction was done using DSI feature transform and classification was performed using multi-class SVM. Experiment evaluation was done on 400 image samples providing an accuracy of 79.38% for classification. An interesting approach that used CD marker, texture, shape and color as features, K-means clustering for segmentation and SVM, Deep Learning for feature extraction and classification achieved an amazing accuracy of more than 99% [142].

##### B. DEEP LEARNING BASED METHODS

Lots of studies used deep-learning techniques for WBCs and Leukemia Segmentation and classification. As [95] suggested a pipeline which employs deep learning for WBCs classification. Number of pretrained networks like Overfeat-Net, AlexNet, VGG were checked for feature extraction, additionally, a novel architecture from scratch was trained and tested. The maximum accuracy obtained was 96.1% employing the new architecture. The dataset used, contains 2551 images. Reference [150] suggested CAD-system. The given CAD system comprises of two major steps i.e. recognition & classification. For recognition SSD [151] is applied, whereas for classification multiple networks pretrained were analysed. The network architectures applied were VGG, AlexNet [152], ResNet and GoogleNet. Maximum average accuracy attained was 97% for AlexNet when checked on dataset of 7500 image. A DCLNN architecture in [153] is recommended. The given architecture is quite simple in

comparison to pretrained architectures available and is also trained from scratch. This methodology attained an average accuracy of 88% for classification, when tested on image dataset of 13,000 images. In recent study [154], CAD-system is recommended for WBCs classification by applying capsule networks. These networks architecture majorly contains encoder and decoder. The encoder encodes image to a vector with 16-dimensions containing information required to depict the image, and decoder learns decoding the instantiation parameters provided in the image of the item it identifies. Euclidean distance is used in the decoder for measuring the loss function for calculating resemblance between the features reconstructed and the trained feature. The key advantage while applying the capsule networks is getting images reconstructed which can be inspected visually. By applying this network architecture, average obtained accuracy was 92.5% for 263 images.

Reference [155] proposed a deep learning network architecture for WBCs and RBCs segmentation. The recommended architecture is SegNet architecture [156]. The proposed architecture employs encoder and decoder pair for creating feature maps for segmentation. By using architecture, segmentation of 42 images was performed, and mean intersection over union achieved was 79%. Yu *et al.* [157] offered WBCs classification method by employing multiple deep learning networks. The proposed system was assessed on 2000 microscopic images dataset of 7 WBCs types and compared to various conventional techniques. An average accuracy obtained was 88.5%, which confirms the dominance of employing a CNN. Shafique and Tehsin [30] proposed a CNN based framework for diagnosis of ALL leukemia subtypes and achieved 99% accuracy for binary classification (healthy vs. ALL) cells and 96% accuracy for diagnosis of ALL subtypes. Another CNN based framework for ALL subtype classification was proposed by Thanh *et al.* [158] that achieved 96.6% accuracy. However, Nizar *et al.*, have proposed a CNN based framework that is capable to diagnose all subtypes of leukemia and perform better than other ML based binary leukemia classification approaches [7].

References [159] and [160] have proposed a DL based framework for ALL classification and its subtypes using the bone marrow imagery. Their proposed framework achieves a reliable accuracy of 97.78% and can be a viable technique for pathologists. An efficient NN based method has been proposed by Kumar *et al.*, for the identification of ALL and its subtypes [22]. The diagnosis of Leukemia using AB Colony (ABC) for the training of Back Propagation Neural Network (BPNN), initially, Principal Component Analysis (PCA) used for the dimensionality for the leukemia dataset. Then the optimum feature set is obtained by ABC algorithm followed by the classification phase. Their results suggest that the ABC-BPNN system based on PCA is more accurate than the genetic algorithm based BPNN (GA-BPNN) with achieved accuracy of 98.72%. However, their proposed framework of ALL and subtype classification is tested on a small dataset.

Vogado *et al.* [161] presented a diagnosis system for leukemia used on 377 images. The proposed system uses CNN network and transfer learning for extracting discriminant features. Later, feature selection employed information gain. Lastly, Classification was performed by SVM classifier. Classification accuracy of approximately 99% was obtained with three heterogeneous datasets. Zhao *et al.* [162] offered a detection and classification system for WBCs. Later, for extracting features, CNN was used. Lastly, RF and SVM classifier were used together for classifying WBCs. The system suggested was assessed on datasets. The average accuracy for classification obtained was 92.8%. Habibzadeh *et al.* [163] proposed a system for WBCs classification based on deep learning and transfer learning both. The method anticipated starts from pre-processing step to transfer learning for extracting features. Lastly, Inception and ResNet were used for classification. 1244 images of WBCs were used to evaluate the performance of proposed method. Accuracy of 99.84% was achieved which is the top most case. Lin *et al.* [164] leukocyte proposed a system for classification of leukocytes. Initially, advanced technique was used to extract complex leukocytes by k-means algorithm. A CNN model performed classification. The recommended method was assessed on 368 images dataset. The achieved accuracy results for classification was 98.96%.

Rehman *et al.* [105] suggested a system for Classification for ALL and its types. Initially, the ROI of lymphoblast from bone marrow aspirants was segmented with thresholding method. For classification CNN i.e., AlexNet was used. 330 Images dataset was used for assessment. The accuracy achieved for classification was 97.78%. Wang *et al.* in [165] offered a technique for detection and classification of WBCs using ensemble classifier in which outputs are fused for a multiple CNNs. An average accuracy for classification produced was 99.37%, using a 3000 images dataset for each category.

The advancement in deep neural networks impacted in evolution of another type of automated leukemia diagnosis methods labelled as end-to-end learning-based methods. The main concept behind this type is to construct and develop a DNN which gets input microscopic blood smear image and returns the output category of the image instead of undergoing the multiple challenging steps followed normally in traditional methods like ROI (region of interest) for segmentation and feature extraction.

Table 3 gives an overview of the methods that have performed WBC segmentation. It summarizes the major work in leukocytes segregation, counting, blood smear analysis and highlights the studies that have performed automated detection of leukemia and have utilized some kind of image processing, computer vision, machine learning or deep learning framework. Table 4 provides survey of the studies that are carried out on ALL and its sub-types classification using image processing and Table 5 depicts a survey about work carried out for WBCs and ALL diagnosis using Deep Learning methods. Table 6 summarizes the studies that have proposed

**TABLE 3. Overview of the studies performed for white blood cells segmentation and classification.**

Author	Year	Area of research	Dataset	Segmentation	Feature Extraction	Classification	Accuracy (%)
Hazwani et al., [181]	2011	Blast cells counting	From HUSM	Thresholding	Morphological, Color(HSV)	N/A	97.8%
Benazzouz et al., [97]	2013	Blood smear	From local hospital	Pixel classification	Color	SVM	95.02%
Biswas et al., [182]	2016	Blood Cell detection	30 blood images	Watershed	contour	N/A	93%
Byoung et al., [183]	2011	WBC segmentation	CellaVision	Active Contour	N/A	N/A	88.4%
F. Ismail Sholeh, [184]	2013	WBC segmentation	From Universitas Gadjah Mada, Indonesia	Otsu's thresholding, region growing	Morphological, Color	N/A	89-92%
Li et al., [185]	2013	WBC segmentation	N/A	Active Contour, (ISAM) [96]	N/A	N/A	91.06%
Ravikumar et al., [186]	2014	WBC segmentation	From Indian pathologist	Modified fuzzy possibilistic C-means	N/A	N/A	94.56%
Jati et al., [187]	2014	WBC segmentation	From MM College, India	Intuitionistic fuzzy divergence-based thresholding	Morphological	N/A	98.52%
Liu et al., [188]	2016	WBC segmentation	CellaVision & Jiashan	Adaptive location Iteration	N/A	N/A	98.18%
Ananthi et al., [189]	2016	WBC segmentation	Cellavision & Sathiya Sai Diagnostic center	Fuzzy thresholding	Morphological	N/A	98.72%
Cao et al., [190]	2018	WBC segmentation	From Zhongnan Hospital	Step wise averaging method, Intervals valued fuzzy sets	Morphological (Shape contour)	N/A	93.75%
Abdulhay et al., [94]	2018	WBC segmentation	From Iranian pathologist	SVM	Texture features	SVM	95.3%
Xin Zheng et al., [36]	2018	WBC segmentation	JTSC, CellaVision	Self-supervised learning	Topological, color	SVM	N/A
Benomer et al., [191]	2019	WBC segmentation	From hospital in Algeria	Watershed	Color, texture, morphological	RF	95.8%
Wang et al., [174]	2019	WBC segmentation	From People's Hospital, Sichuan	Thresholding	Color	N/A	97.4%
Nazlibilek et al., [192]	2014	WBC classification	ALL-IDB	Otsu's thresholding	Morphological (size and shape)	MLP	95%
Rawat et al. [41]	2018	WBC classification	Online images	Region based	Morphological, texture	ANFIS	98.7%
Hegde et al., [193]	2019	WBC classification	Private dataset	Active Contour segmentation, thresholding	Morphological, color, texture	CNN	99%
Zang et al., [180]	2020	WBC classification	From a local hospital	Adversarial residual network	CNN-HOG	SVM	94.57%
Pramit Ghosh et al., [194]	2016	WBC counting	From Kolkata Municipal Corporation, India	Region based segmentation	Morphological, texture	Fuzzy classifier	95.8%
Nabilah et al., [195]	2018	WBC counting	ALL-IDB	Otsu's thresholding	Morphological (Color)	N/A	96.92%
Shahin et al., [95]	2019	WBC identification	ALL-IDB	Deep Learning	Morphological, texture	SVM	96.1%



**TABLE 4. Overview of the studies performed on ALL and its subtypes diagnosis using image processing, ML and DL networks.**

Author	Year	Dataset	Pre-processing	Segmentation	Feature Extraction	Classification	Accuracy (%)
Chatap et al., [128]	2014	ALL-IDB1 and ALL-IDB 2	Histogram equalization, contrast stretching	Otsu’s thresholding	Shape, area, circularity, perimeter	KNN	93%
Mohapatra et al., [47]	2014	From IG Hospital, India	Contrast enhancement, selective median filtering	SCM clustering	Fractal features (Hausdorff dimension)	NB, KNN, MLP, RBF NN, SVM	94.73%
Putzua et al., [74]	2014	ALL-IDB1 and ALL-IDB 2	Histogram equalization, contrast stretching	Zack algorithm	Shape, color, texture	SVM	92%
Moradi Amin et al., [48]	2015	From Isfahan Al-Zahra and Omid Hospital	Histogram equalization	Fuzzy C-means, watershed segmentation	Statistical, area, solidity, perimeter	SVM	96.33%
Amin et al., [196]	2015	From Isfahan Al-Zahra and Omid Hospital	Histogram equalization, Linear Contrast Stretching	K-means clustering	Histogram-based features, shape-based features	SVM	95.6% (ALL subtype), 97% (blast vs normal cell)
Shankar et al., [197]	2016	ALL-IDB	Contrast stretching, histogram equalization	Zack algorithm, watershed segmentation	Shape	SVM	90%
Li et al., [118]	2016	ALL-IDB	Contrast stretching, histogram equalization	Zack algorithm, watershed segmentation	Shape	SVM	97.85%
Amin et al., [198]	2016	From Isfahan Al-Zahra and Omid Hospital	Histogram equalization	Fuzzy C-means, watershed segmentation	Geometric or shape based, 1st and 2nd-order statistical features	SVM	97.52%
Mishra et al., [110]	2017	ALL-IDB 1	Histogram equalization, Weiner filtering	Marker based watershed segmentation, Sobel and Perwitt operators	GLCM, PPCA	RF	96.29%, 99.004%
Karthikeyan et al., [28]	2017	From Google	Histogram equalization, adaptive median filter	K-means clustering, Fuzzy c-means clustering	Color features	SVM	90%
Rawat et al., [111]	2017	ALL-IDB2	Histogram equalization, order statistic filter	Morphological opening, global thresholding	chromatic, statistical texture, geometrical features	KNN, PNN, SVM, SSVM, and ANFIS	89.8%
Rawat et al., [32]	2017	ASH	N/A	Otsu’s thresholding	Color, texture, shape	GA & SVM	>97%
Khosrosereshki et al., [199]	2017	N/A	N/A	Fuzzy based	Morphological	Fuzzy based	93.75%
Srisukham, et al., [200]	2017	ALL-IDB2, Sonar (UCI ML Repository)	N/A	Marker-controlled watershed segmentation	Shape, GLCM texture, color	SVM	94.94% (BBPSO-1), 96.25% (BBPSO-2)
Rehman et al., [105]	2018	From Amreek Clinical Laboratory KP Pakistan	N/A	Thresholding	Color	KNN, naïve Bayesin, and SVM	97.78%
Nizar et al., [7]	2020	ALL-IDB & ASH	Data augmentation	N/A	N/A	CNN	88.25% (leukemia) 81.74% (subtypes)

**TABLE 5. Studies performed on WBCs and leukemia diagnosis using deep learning frameworks.**

Author	Year	Dataset	Pre-processing	Segmentation	Feature Extraction	Classification	Accuracy (%)
Escalante <i>et al.</i> , [201]	2012	Collected from Mexican Institute	N/A	Markov random fields	Geometry, statistical, texture	EPSMS	97.68%
Mohapatra <i>et al.</i> , [93]	2012	Ispat General Hospital, India	N/A	FLANN	N/A	FLANN	N/A
Mishra <i>et al.</i> , [177]	2015	ALL-IDB	Median and Wiener Filtering	K-means clustering, Zack algorithm	Statistical, geometric, texture, color	SVM, Deep CNN	93.57% (SVM), 97.78% (Deep CNN)
Neoh <i>et al.</i> , [202]	2015	ALL-IDB	N/A	Watershed, stimulating discriminant measures	Texture, color, shape	SVM, Ensemble, MLP	96.72%
Khalilabad <i>et al.</i> , [203]	2017	Private dataset	N/A	(TV + L1)	Morphological	J48 decision Tree	95.45%
Savita Dumyan, [204]	2017	36 blood images	Histogram equalization	Binarization, canny edge detection	Moment invariants, statistical features	ANN	97.8%
Moshavash <i>et al.</i> , [205]	2018	ALL-IDB	N/A	Decision Support System	Color, texture, shape	Ensemble	98.1%
Krishnakumar <i>et al.</i> , [206]	2019	ALL-IDB2	N/A	MI based hybrid model	Statistical, LDP	Deep Learning	98.7%
Kutluet <i>al.</i> [150]	2019	BCCD & LISC	N/A	CNN	CNN layers	multiple CNN	97%
Shahinet <i>al.</i> , [95]	2019	3 different datasets	Resize	N/A	Shape, texture, complex feature	CNN	96.1%
Tatdow Pansombut <i>et al.</i> , [207]	2019	ASH	N/A	Manual segmentation	Color, geometry, texture	CNN	80%
Kumar <i>et al.</i> , [43]	2019	ALL-IDB, Atlas	N/A	K-medoid	Shape, texture, visual	ANN	98.6%
Kumar <i>et al.</i> , [22]	2019	From mldata repository	N/A	N/A	optimal feature set	PCA based ABCBPNN	98.7%
Zhang <i>et al.</i> , [180]	2020	From Shandong Hospital and BCCD	N/A	Res-net, discriminator network	HOG features	SVM	95.9%
Nizar <i>et al.</i> , [7]	2020	ALL-IDB & ASH	Data augmentation	N/A	N/A	CNN	88.25% (leukemia) 81.74% (subtypes)

**TABLE 6. Studies performed for acute leukemia types (ALL & AML) classification from peripheral blood smear and bone marrow images.**

Author	Year	Dataset	Segmentation	Feature Extraction	Classification	Accuracy (%)
Poon et al., [141]	1992	From Children’s Hospital, Columbia	Thresholding	Morphological	N/A	80.6%-93.2%
Jyoti Rawat et al., [32]	2017	ASH	Otsu’s thresholding	Morphological i.e. Color, shape, and texture	Genetic algorithm, SVM	>97%
Laosai et al. [142]	2018	From Hospitals in Thailand	K-means clustering	CD marker, texture, shape, color	SVM, K-means clustering, Deep Learning	99.67%
Negm et al. [143]	2018	ALL-IDB	K-means	Color, texture, shape	Decision tree, Artificial Neural Network	99.52%

**TABLE 7. Studies performed for AML diagnosis from peripheral blood smear and bone marrow images.**

Author	Year	Dataset	Segmentation	Features	Classifier	Accuracy (%)
Nee et al., [211]	2012	U S Hospital Malaysia	Watershed	N/A	N/A	94.5%
Skrobanski et al., [212]	2012	Hospital U. S. Malaysia	Otsu’s thresholding	Statistical	ANN	88.69%
Agaian et al., [166]	2014	ASH	K-means clustering	Hausdorff dimension, color, texture, shape	SVM	98%
Kazemi et al., [213]	2016	Microscopic images of blood smears of patients with AML	Color segmentation	nucleus-cytoplasm ratio, irregularity, shape, color, Hausdorff dimension, texture	SVM	96%
Suryani et al., [214]	2017	Dr. Moewardi Hospital	Watershed distance transform	Statistical, shape	Back propagation	94.29%
Kumar et al., [175]	2017	ASH	K-means clustering	Shape & texture	SVM	95%
Su et al., [79]	2017	H. Lee Moffit Cancer center	K-means clustering, Hidden Markov random	Color, spatial	N/A	96%-98%
Matek et al., [215]	2019	Munich University Hospital	Region based	N/A	CNN	90%

**TABLE 8. Studies performed for CLL detection and classification from peripheral blood smear and bone marrow images.**

Author	Year	Dataset	Segmentation	Feature extraction	Classification	Accuracy (%)
Merino et al., [176]	2014	Collected from Hospital Clinic in Barcelona	Watershed	Geometry, granulometric, texture	Fuzzy C-means (FCM)	N/A
Rejintal et al., [170]	2016	Blood smear images collected from 374 patients in Spain	Watershed segmentation	Color, geometric, texture	SVM	Not mentioned
Cabrera et al., [172]	2017	N/A	Watershed segmentation	N/A	Genetic algorithm	90.33%
Preeti et al., [173]	2017	Collected from Goa medical college, India	K-means clustering	Size & statistical	SVM	N/A
Fakhouri et al., [42]	2018	Used online image bank from various resources	Thresholding	Shape	SVM	N/A
Anilkumar et al., [167]	2018	ASH	K-means	Hausdorff dimension (HD)	SVM	96%
Rodellar et al., [168]	2018	N/A	K-means	Texture	SVM	N/A

methods for ALL and AML classification from bone marrow and peripheral blood smear images. Table 7 summarizes the studies that have progressively performed AML diagnosis

and classification using computer vision, image processing, ML and DL frameworks. It is clearly observed that not many studies have performed myeloblast detection and

classification. References [166] and [79] have the highest accuracy of myeloid classification using Hausdorff dimension, color, texture, shape and spatial features, K-means clustering, and Hidden Markov random field based segmentation. It is also observed that either SVM, or NN are used for classification of the myeloid leukemia.

Table 8 summarizes the studies that have performed leukemia detection in general or any other type of leukemia using the image processing, computer vision, machine learning or deep learning approach. Reference [167] proposed framework for diagnosing types of leukemia whereas [42], [46], [168]–[173], did general leukemia classification.

Table 9 compare and analysis the accuracies of different methods for Segmentation and classification of WBCs and types of Leukemia (ALL, AML and CLL) using image processing, ML, DL approaches along with their Computational time. For WBC segmentation [174] provides 97.4% accuracy with 2.4 to 2.6 msec computational time per leukocyte, which is good. For ALL segmentation and Classification [110] technique provides better accuracy of almost 99% with computational time of just 0.346 sec (per Lymphocyte) but [118] gives less accuracy i.e. 97% with almost same computational time. For AML computational time with full feature set is 104 sec and with reduced feature it is reduced to 98 sec and providing 95% accuracy [175]. As compared to acute Leukemias little work has been done in past for Chronic Leukemia.

Although various studies have performed WBC analysis, segmentation, counting, blood smear analysis and acute leukemia identification and classification only handful studies have addressed the chronic leukemia diagnosis. In the literature, Alferez *et al.*, [176] has classified the CLL & Hairy Cell leukemia (HCL) using texture, geometry and granulometric features and classified through Fuzzy C-means (FCM).

## V. DISCUSSION

Automated methods for leukemia diagnosing using microscopic blood smear images increased since last decade and the increase in their number can be clearly observed. Many studies have proposed automated detection of acute type of leukemia (ALL and AML) but just a few research studies have proposed framework for the diagnosis of chronic type of leukemia (CML and CLL) using image processing methods [14]. Although numerous studies have been presented for the automatic diagnosis of acute lymphoblastic leukemia, only fewer studies have proposed the methods to diagnose the subtypes of ALL. In the case of AML classification, little work has been performed [10] while no study is found for CML type of leukemia detection. The diagnosis of leukemia into subtypes is a tough task and many studies have only performed binary classification such as any subtype of leukemia versus the healthy leukocytes [21], [28], [29], [158], [177]–[179]. There are just a handful studies which have proposed a framework that diagnosis all types of leukemia with optimal and acceptable accuracy. After reviewing literature, it has been found that 79% of the studies have utilized

private data that is obtained from either nearby hospitals or pathologists from the local laboratories. 17% of the studies have utilized publicly available datasets such as ASH and ALL-IDB in their methodologies. 14% studies did not mention the source of the dataset and 32% studies have used SVM for classification of leukemia cells [14]. Since automated frameworks require large repository of data to train the model, various data augmentation methods have been suggested in the literature that makes the data large enough to be suitable for training purpose. Most of them include image rotation, translation, shearing, blurring, mirroring, gray scale image transformation and histogram equalization [30].

Patel and Mishra [177] proposed an automated approach for the diagnosis of acute leukemia from the microscopic blood smear images and suggested some methods such as K-state grouping, Zack algorithm, and histogram alignment for filtering the blood smear images. Their proposed method achieves with an accuracy of 93.57% using SVM and 97.78% using Deep convolutional neural network. Rehman *et al.* [105] performed classification of ALL into its types from the peripheral bone marrow smear images. Their proposed framework that converts the smear image to HSV domain and applies operation on the saturation component to get lymphoblasts followed by thresholding and hole filling. They have used convolutional neural network to extract features and classification of ALL into subtypes. Their proposed framework works better than ML based approaches and achieves accuracy of 97.78%. An efficient neural network based method has been presented by Kumar *et al.*, for the identification of ALL into its subtypes [22]. The detection of Leukemia using AB Colony for the training of BPNN. Initially, Principal Component Analysis (PCA) used for the dimensionality for the leukemia dataset. Then the optimum feature set is obtained by ABC algorithm followed by the classification phase. Their results suggest that the AB Colony based Back-Propagation Neural Network (ABC-BPNN) system based on PCA is more accurate than the genetic algorithm based BPNN (GA-BPNN) with achieved accuracy of 98.72%. However, their proposed framework of ALL and subtype classification is tested on a small dataset. Zhang *et al.* [180] suggested a NN based framework for the classification of leukocyte that achieves accuracy of 94.57% using the CNN and HOG features and tested on a larger dataset i.e. 5000 images obtained from a local hospital. They also tested their results using CNN features and classified using SVM classifier and achieved accuracy of 94.23% that is not different from 94.57%. Thus, their results suggest that hybrid neural network framework achieves greater accuracy of leukocyte classification instead of utilizing only HOG features achieved only 85% accuracy. However, their outcomes were not compared with other state of the art techniques. Many studies have been proposed that can ideally identify and segment blast cells.

Out of the various segmentation methods, marker-based watershed technique is found to be the best that can segment the overlapping blasts cell with an accuracy of 96.29% [110].

**TABLE 9. Comparison of different CAD systems with respect to accuracy and known computational time.**

Author	Year	Segmentation method	Classifier	Accuracy (%)	Computational time
Cao et al., [190]	2018	Step wise averaging method, Interval-s valued fuzzy sets	N/A	93.75%	0.23 sec (Per image)
Nazlibilek et al., [192]	2014	Otsu’s thresholding	MLP	95%	556 sec (Classifier A), 11.21 sec (Classifier B)
Wang et al., [174]	2019	Thresholding	N/A	97.4%	(0.24-0.26)msec per Leukocyte
Ravikumar et al., [186]	2014	Modified fuzzy possibilistic C-means	N/A	94.56%	1.3974 sec (Per leukocyte)
Abdulhay et al., [94]	2018	SVM	SVM	95.3%	18 sec (per Image)
Byoung et al., [183]	2011	Active Contour	N/A	88.4%	2.6 msec (Per Leukocyte)
Liu et al., [188]	2016	Adaptive location Iteration	N/A	98.18%	< 10 sec (per Image)
Benazzouz et al., [97]	2013	Pixel classification using SVM	SVM	95.02%	2130 sec (segmentation time)
Zang et al., [180]	2020	Adversarial residual network	SVM	94.57%	3615 min (Training time)
Mohapatra et al., [47]	2014	SCM clustering	Ensemble method (NB, KNN, MLP, radial basis functional NN, SVM)	94.73%	0.1 sec (NB), 0.13 sec (KNN), 3.8 sec (MLP), 13.3 sec (RBFN), 0.05 sec (SVM), 3.92 sec (EOC)
Li et al., [118]	2016	Zack algorithm, watershed segmentation	SVM	97.85%	<0.5 sec (Per image)
Mishra et al., [110]	2017	Marker based watershed, Sobel and Perwitt operators	RF	96.29%, 99.004%	0.346 sec (per Lymphocyte)
Negm et al. [143]	2018	K-means	Decision tree, ANN	99.52%	0.3 sec per cell (using Decision tree), 2.47 sec (using NN)
Skrobanski et al., [212]	2012	Otsu’s thresholding	ANN	88.69%	Total Processing time 56709 sec (CPU 100), 4364 sec (GPU 100)
Suryani et al., [214]	2017	Watershed distance transform	Back propagation	94.29%	36855 sec (training time)
Kumar et al., [175]	2017	K-means clustering	SVM	95%	104 sec (with full feature set), 98 sec (with reduced features set)

Some studies have utilized the Neural Network (NN) for the segmentation of peripheral blood smear and bone marrow

images along with the classification of leukemia [7], [36], [93]–[97]. Additionally, many other segmentation techniques

such as HSV color based, thresholding, watershed, and clustering based methods have also been used in the past, but these methods lack in efficiency due to variation in contrast, light and noise. Thus, it is needed to develop more accurate methods to segment blast cells from peripheral blood smears images [83]. Rawat *et al.* [41] have improved the accuracy of blast cell classification to 99.517% using Histogram green color of RGB component for pre-processing of blood smear images and using K-means clustering technique for segmentation of blast cells followed by the extraction of geometry, statistical and textures features and classification by ANN. In the classification phase, different techniques were carried out by researchers to obtain the accurate results, where the ANN and SVM techniques were the most accurate. Moreover, the detection phase was dependent on the previous stages (segmentation and feature extraction). The studies that used Support Vector Machine (SVM) have presented higher accuracy compared to other classifiers. Moreover, other studies also found the Fuzzy logic-based classifier was greater than SVM in terms of accuracy. Therefore, it is indispensable to develop an effective classifier to detect blast cells from peripheral blood smears images. Makem and Tiedeu [208] proposed algorithm based on adaptive fusion to fuse color components obtained from CMYK and HSV color spaces to extract nuclei of the WBC without the applying morphological operators. Their method addressed the problem of contrast, color or brightness variation and does not requires the application of morphological operations. Srisukkhram *et al.* [200] proposed the framework for ALL classification with achieved accuracy of 94.94%. Their proposed method subdivides the blood smear images and pre-processes it followed by the application of marker-controlled watershed segmentation that segments the blast cells. Their propose framework extracts features from Particle Swarm Optimization (PSO), Cuckoo Search (CS) and Dragonfly Algorithm (DA) and classification was performed using SVM for detection and classification of blasted cells.

By reviewing literature, it has been observed that in studies related to blood smear analysis such as anaemia, or other, thresholding is found to be the common method of segmentation. In number of research studies that have performed leukemia diagnosis and classification, clustering-based segmentation i.e. K-means and Fuzzy C-means clustering technique is the extensively used method for the segmentation of blast cells with achieving accuracy that is more than 95% [4]. Watershed segmentation and thresholding are two other widely used methods in the literature. Thresholding, watershed segmentation, clustering are the methods that are extensively used in the studies related to WBC analysis i.e. identification, segmentation, and counting. References [32], [74], [105], [141], [144], [177], [197], [209], and [210] have detected acute leukemia using thresholding-based segmentation methods. Otsu's thresholding is found to be a better algorithm for segmentation than Zack's algorithm in case of acute leukemia classification [92].

For leukemia detection, morphological features have been extensively utilized that takes into account the shape of the blast cell. Other extensively features used are fractal dimension, Hausdorff dimension, color and texture based features. Since a lot of information is present in the blast cells, so using a larger feature set for classification might not produce the good results. Thus, feature selection should be performed for reduction of the feature set dimensions and for the selection of non-redundant features prior to the classification phase. The common algorithms used for feature reduction are Local Direction Pattern (LDA) and Principal Component Analysis (PCA). Some other feature selection methods that have been used in the literature include Fisher's Discrimination Ratio followed by application of Exhaustive Search to select the best feature set [4]. For the classification of leukemia cells, various ML and DL based frameworks have been proposed but amongst all ANN and SVM are found to be most accurate. One of the extensively used classifiers for the blast cell classification is the SVM followed by MLP and ensemble methods. However, the identification and classification phase is highly dependent on the segmentation outcomes and the features extraction from the peripheral blood smear images [89]. The studies which have performed analysis on multiple classifiers have found SVM to be highly accurate for ALL classification whereas other studies have found the accuracy obtained from Fuzzy logic-based classifier is greater than SVM. However, in few cases, MLP has performed better than the other algorithms for leukemia classification [4]. A study performed by Jakkrich *et al.* has achieved an accuracy of 99% using SVM and deep learning algorithms [142]. Moreover, neural network based frameworks have also performed better for ALL classification and have achieved an accuracy of 97.78% that is more than the ML algorithms [105]. Therefore, there is still a need to develop an effective classifier that is not prone to overfitting or underfitting, can detect blast cells from peripheral blood smears and addresses the challenges and complexities of blasts classification.

In future work, a combination of two or more segmentation techniques can be made to segment acute leukemia cells. The features selection and dimensional reduction techniques should be applied to achieve greater efficiency in the leukemia classification [94]. The available literature studies that performed microscopic blood smear image analysis using various image processing and computer vision techniques have performed their analysis using Image Processing Toolbox in MATLAB which is not a familiar software for a lot of medical professionals and pathologists. Thus, there is a need to develop specialized and user friendly medical software that can be readily used by medical professionals.

## VI. CONCLUSION AND FUTURE WORK

Leukemia is a fatal disease that weakens the immune system by affecting the WBCs and bone marrow. This review paper has critically analysed the dozens of studies that have conducted in the past. It has presented a survey of techniques used

in automatic diagnosis of leukemia. A survey of image processing, computer vision, ML and DL based framework has been presented that analyses and compares these methods in terms of datasets, pre-processing, segmentation, features and classification algorithms and classification accuracy. CBC, blood smear examination under microscope and bone marrow aspiration have been the conventional manual methods for the detection of leukemia. However, these methods have shortcomings and because of the ease provided by the technology, dozens of automated methods for CAD systems have been recommended for the Leukemia diagnosis. These automated image processing and computer vision-based methods have the ability to diagnose the leukemia readily with ease and accuracy. Thus, the patient can be given a proper treatment in due time and his life can be saved. Some of the studies presented in the past have a very high accuracy of leukemia classification. Thus, CAD systems have the potential to replace the manual diagnosis of leukemia. However, by reviewing literature it has been inferred that many of studies have proposed frameworks for detecting acute leukemia and its subtypes but just few studies have proposed methodologies for detecting chronic leukemia or all leukemia subtypes. Hence, more studies should be conducted to develop propose methodologies that can detect all subtypes of leukemia, thus helping the pathologists to diagnose leukemia efficiently and accurately. Furthermore, pre-processing and segmentation are important stages in leukemia detection and categorization, and accuracies of classifiers highly depends upon it. Thus, researches should propose more efficient pre-processing and segmentation algorithms that works better for high density noisy images and segments the blast cells more efficiently thus helps in achieving a higher accuracy of leukemia classification. Deep learning frameworks have been used by researchers for automated diagnosis of Leukemia, but a common challenge with these frameworks is to obtain big datasets which is difficult to get. So, big datasets must be created to make optimal usage ML & DL frameworks. If the data available is limited, then some methods must be proposed to give promising results with less available data, and avoiding the overfitting issue. Lastly, to implement meaningful ML & DL algorithms, there is need for the ML & DL tools to be available, also perform well technically, transparent and explainable to both providers and patients.

Although the review covered here is reliable and comprehensive and has attempted to cover all aspects concerned to automated detection of leukemia. However, the timeliness of this study should be regarded as one of limitations as the research in this domain is increasing rapidly increasing and variety of techniques have been proposed in the multifarious aspects.

## REFERENCES

- [1] A. Maton, *Human Biology and Health*. London, U.K.: Pearson, 1993.
- [2] M. LaFleur-Brooks, *Exploring Medical Language: A Student-Directed Approach*, 10th ed., D. L. Brooks, Ed. St. Louis, MO, USA: Elsevier, 2018, pp. 757–758.
- [3] *The American Society of Hematology*. Accessed: Oct. 2020. [Online]. Available: <https://www.hematology.org/>
- [4] R. G. Bagasjvara, I. Candradewi, S. Hartati, and A. Harjoko, "Automated detection and classification techniques of acute leukemia using image processing: A review," in *Proc. 2nd Int. Conf. Sci. Technol.-Comput. (ICST)*, Oct. 2016, pp. 35–43.
- [5] Health Publishing. *What is Leukemia?* Accessed: Oct. 4, 2020. [Online]. Available: <https://www.health.harvard.edu/cancer/leukemia>
- [6] L. Butcher, "Profiles in oncology social media: Dana-Farber cancer institute DanaFarber," *Oncol. Times*, vol. 37, pp. 69–70, May 2015.
- [7] N. Ahmed, A. Yigit, Z. Isik, and A. Alpkocak, "Identification of leukemia subtypes from microscopic images using convolutional neural network," *Diagnostics*, vol. 9, no. 3, p. 104, Aug. 2019.
- [8] *What is Lymphoblast?* Accessed: Oct. 17, 2020. [Online]. Available: <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/lymphoblast>
- [9] K. Seiter. *What is the French American British (FAB) Classification of Acute Lymphoblastic Leukemia (ALL)*. Accessed: Oct. 18, 2020. [Online]. Available: <https://www.medscape.com/answers/207631-105131/what-is-the-french-american-british-fab-classification-of-acute-lymphoblastic-leukemia-all>
- [10] S. Shafique and S. Tehsin, "Computer-aided diagnosis of acute lymphoblastic leukaemia," *Comput. Math. Methods Med.*, vol. 2018, pp. 1–13, Feb. 2018.
- [11] R. D. Labati, V. Piuri, and F. Scotti, "All-IDB: The acute lymphoblastic leukemia image database for image processing," in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 2045–2048.
- [12] J. Todd, A. Sanford, I. Davidsohn, and J. Henry, *Todd-Sanford-Davidsohn Clinical Diagnosis and Management by Laboratory Methods*. Philadelphia, PA, USA: Saunders, 1979. [Online]. Available: <https://books.google.com.pk/books?id=YMYhAQAAMAAJ>
- [13] K. Raymakers. *Blast Cells and Myeloblasts Overview*. Accessed: Oct. 2020. [Online]. Available: <https://www.verywellhealth.com/overview-of-blast-cells>
- [14] K. K. Anilkumar, V. J. Manoj, and T. M. Sagi, "A survey on image segmentation of blood and bone marrow smear images with emphasis to automated detection of Leukemia," *Biocybernetics Biomed. Eng.*, vol. 40, pp. 1406–1420, Sep. 2020.
- [15] L.L.Society. *Acute Promyelocytic Leukemia Facts*. Accessed: Oct. 18, 2020. [Online]. Available: [https://www.lls.org/sites/default/files/National/USA/Pdf/Publications/APL\\_FactSheet\\_10\\_15FINAL.pdf](https://www.lls.org/sites/default/files/National/USA/Pdf/Publications/APL_FactSheet_10_15FINAL.pdf)
- [16] C. R. UK. *About Various Cancer Types*. Accessed: Oct. 17, 2020. [Online]. Available: <https://www.cancerresearchuk.org/about-cancer/type>
- [17] C. Raje and J. Rangole, "Detection of leukemia in microscopic images using image processing," in *Proc. Int. Conf. Commun. Signal Process.*, Apr. 2014, pp. 255–259.
- [18] Leukemia for Childrens Cancer. *Leukemia in Children*. Accessed: Oct. 4, 2020. [Online]. Available: <https://curesearch.org/Leukemia-in-Children>
- [19] A. C. Society. *Leukemia*. Accessed: Oct. 4, 2020. [Online]. Available: <https://www.cancer.org/cancer/leukemia.html>
- [20] K. M. Garrett, F. A. Hoffer, F. G. Behm, K. W. Gow, M. M. Hudson, and J. T. Sandlund, "Interventional radiology techniques for the diagnosis of lymphoma or leukemia," *Pediatric Radiol.*, vol. 32, no. 9, pp. 653–662, Sep. 2002.
- [21] L. Boldú, A. Merino, S. Alférez, A. Molina, A. Acevedo, and J. Rodellar, "Automatic recognition of different types of acute leukaemia in peripheral blood by image analysis," *J. Clin. Pathol.*, vol. 72, no. 11, pp. 755–761, Nov. 2019.
- [22] R. Sharma and R. Kumar, "A novel approach for the classification of leukemia using artificial bee colony optimization technique and back-propagation neural networks: ICCCN 2018," in *Proc. 2nd Int. Conf. Commun., Comput. Netw.* Chandigarh, India: NITTTR, Jan. 2019, pp. 685–694, doi: 10.1007/978-981-13-1217-5\_68.
- [23] S. H. Shirazi, A. I. Umar, N. U. Haq, S. Naz, and M. I. Razzak, "Accurate microscopic red blood cell image enhancement and segmentation," in *Proc. Int. Conf. Bioinf. Biomed. Eng.*, 2015, pp. 183–192.
- [24] T. M. Khan, A. Robles-Kelly, and S. S. Naqvi, "A semantically flexible feature fusion network for retinal vessel segmentation," in *Neural Information Processing*, H. Yang, K. Pasupa, A. C.-S. Leung, J. T. Kwok, J. H. Chan, and I. King, Eds. Cham, Switzerland: Springer, 2020, pp. 159–167.

- [25] T. M. Khan, M. Alhussein, K. Aurangzeb, M. Arsalan, S. S. Naqvi, and S. J. Nawaz, "Residual connection-based encoder decoder network (RCED-Net) for retinal vessel segmentation," *IEEE Access*, vol. 8, pp. 131257–131272, 2020.
- [26] M. Tabassum, T. M. Khan, M. Arsalan, S. S. Naqvi, M. Ahmed, H. A. Madni, and J. Mirza, "CDED-Net: Joint segmentation of optic disc and optic cup for glaucoma screening," *IEEE Access*, vol. 8, pp. 102733–102747, 2020.
- [27] S. Mohapatra, D. Patra, and S. Satpathi, "Image analysis of blood microscopic images for acute leukemia detection," in *Proc. Int. Conf. Ind. Electron., Control Robot.*, 2010, pp. 215–219.
- [28] D. T. Karthikeyan and N. Poornima, "Microscopic image segmentation using fuzzy C means for leukemia diagnosis," *Leukemia*, vol. 4, no. 1, pp. 3142–3316, 2017.
- [29] S. Paswan and Y. Rathore, "Detection and classification of blood cancer from microscopic cell images using SVM KNN and NN classifier," *Int. J. Advance Res., Ideas Innov. Technol.*, vol. 3, no. 6, pp. 315–324, 2017.
- [30] 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, Sep. 2018, Art. no. 1533033818802789.
- [31] R. D. Labati, V. Piuri, and F. Scotti. *ALL-IDB (Acute Lymphoblastic Leukemia Image Database for Image Processing)*. Accessed: Oct. 2020. [Online]. Available: <https://homes.di.unimi.it/scotti/all/download>
- [32] 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 Biomed. Eng.*, vol. 37, no. 4, pp. 637–654, 2017.
- [33] 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.
- [34] *Cellavision*. Accessed: Oct. 18, 2020. [Online]. Available: <http://blog.cellavision.com/>
- [35] GitHub. *WBC Image Dataset*. Accessed: Oct. 18, 2020. [Online]. Available: [https://github.com/zxaoyou/segmentation\\_WBC](https://github.com/zxaoyou/segmentation_WBC)
- [36] 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.
- [37] M. Mohamed, B. Far, and A. Gually, "An efficient technique for white blood cells nuclei automatic segmentation," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, Oct. 2021, pp. 220–225.
- [38] *SMC-IDB*. Accessed: Oct. 2020. [Online]. Available: <https://www.mathworks.com/matlabcentral/fileexchange/36634-an-efficient-technique-for-white-blood-cells-nuclei-automatic-segmentation>
- [39] O. Sarrafzadeh, H. Rabbani, A. Talebi, and H. Banaem, "Selection of the best features for leukocytes classification in blood smear microscopic images," *Med. Imag., Digit. Pathol.*, vol. 9041, Feb. 2014, Art. no. 90410P.
- [40] *IUMS-IDB*. Accessed: Oct. 2020. [Online]. Available: <https://misp.mui.ac.ir/fa/download>
- [41] J. Rawat, A. Singh, H. S. Bhadauria, J. Virmani, and J. S. Devgun, "Leukocyte classification using adaptive neuro-fuzzy inference system in microscopic blood images," *Arabian J. Sci. Eng.*, vol. 43, no. 12, pp. 7041–7058, Dec. 2018.
- [42] H. N. Fakhouri and S. H. Al-Sharaeh, "A hybrid methodology for automation the diagnosis of leukemia based on quantitative and morphological feature analysis," *Modern Appl. Sci.*, vol. 12, no. 3, p. 56, Feb. 2018.
- [43] 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.
- [44] American Society Of Haematology. *ASH-DATASET*. Accessed: Oct. 2020. [Online]. Available: <http://imagebank.hematology.org/>
- [45] M. M. A. Mohamed and B. Far, "A fast technique for white blood cells nuclei automatic segmentation based on gram-Schmidt orthogonalization," in *Proc. IEEE 24th Int. Conf. Tools Artif. Intell. (ICTAI)*, vol. 1, Nov. 2012, pp. 947–952.
- [46] S. Arslan, E. Ozyurek, and C. Gunduz-Demir, "A color and shape based algorithm for segmentation of white blood cells in peripheral blood and bone marrow images," *Cytometry A*, vol. 85, no. 6, pp. 480–490, 2014.
- [47] S. Mohapatra, D. Patra, and S. Satpathy, "An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images," *Neural Comput. Appl.*, vol. 24, nos. 7–8, pp. 1887–1904, 2014.
- [48] M. MoradiAmin, N. Samadzadehaghdam, S. Kermani, and A. Talebi, "Enhanced recognition of acute lymphoblastic leukemia cells in microscopic images based on feature reduction using principle component analysis," *Frontiers Biomed. Technol.*, vol. 2, no. 3, pp. 128–136, Sep. 2015. [Online]. Available: <https://fbt.tums.ac.ir/index.php/fbt/article/view/56>
- [49] T. A. A. Tosta, P. R. de Faria, L. A. Neves, and M. Z. D. Nascimento, "Computational normalization of H&E-stained histological images: Progress, challenges and future potential," *Artif. Intell. Med.*, vol. 95, pp. 118–132, Apr. 2019.
- [50] Y. Zheng, Z. Jiang, H. Zhang, F. Xie, J. Shi, and C. Xue, "Adaptive color deconvolution for histological WSI normalization," *Comput. Methods Programs Biomed.*, vol. 170, pp. 107–120, Mar. 2019.
- [51] M. Salvi, N. Michielli, and F. Molinari, "Stain color adaptive normalization (SCAN) algorithm: Separation and standardization of histological stains in digital pathology," *Comput. Methods Programs Biomed.*, vol. 193, Sep. 2020, Art. no. 105506.
- [52] M. Kuse, T. Sharma, and S. Gupta, "A classification scheme for lymphocyte segmentation in H&E stained histology images," in *Proc. Int. Conf. Pattern Recognit.* Berlin, Germany: Springer, 2010, pp. 235–243.
- [53] M. T. Ibrahim, T. M. Khan, S. A. Khan, M. A. Khan, and L. Guan, "Iris localization using local histogram and other image statistics," *Opt. Lasers Eng.*, vol. 50, no. 5, pp. 645–654, May 2012.
- [54] T. M. Khan, M. A. Khan, S. A. Malik, S. A. Khan, T. Bashir, and A. H. Dar, "Automatic localization of pupil using eccentricity and iris using gradient based method," *Opt. Lasers Eng.*, vol. 49, no. 2, pp. 177–187, 2011.
- [55] T. M. Khan, M. A. U. Khan, and Y. Kong, "Fingerprint image enhancement using multi-scale DDFB based diffusion filters and modified Hong filters," *Optik*, vol. 125, no. 16, pp. 4206–4214, Aug. 2014.
- [56] T. M. Khan, M. A. U. Khan, Y. Kong, and O. Kittaneh, "Stopping criterion for linear anisotropic image diffusion: A fingerprint image enhancement case," *EURASIP J. Image Video Process.*, vol. 2016, no. 1, pp. 1–20, Dec. 2016.
- [57] I. U. Afridi, T. Bashir, H. A. Khattak, T. M. Khan, and M. Imran, "Degraded image enhancement by image dehazing and directional filter banks using depth image based rendering for future free-view 3D-TV," *PLoS ONE*, vol. 14, no. 5, May 2019, Art. no. e0217246.
- [58] M. A. Khan, A. Khan, T. Mahmood, M. Abbas, and N. Muhammad, "Fingerprint image enhancement using principal component analysis (PCA) filters," in *Proc. Int. Conf. Inf. Emerg. Technol.*, Jun. 2010, pp. 1–6.
- [59] T. M. Khan, D. G. Bailey, M. A. Khan, and Y. Kong, "Efficient hardware implementation strategy for local normalization of fingerprint images," *J. Real-Time Image Process.*, vol. 16, no. 4, pp. 1–13, 2016.
- [60] T. A. Soomro, T. M. Khan, M. A. U. Khan, J. Gao, M. Paul, and L. Zheng, "Impact of ICA-based image enhancement technique on retinal blood vessels segmentation," *IEEE Access*, vol. 6, pp. 3524–3538, 2018.
- [61] M. A. U. Khan, T. M. Khan, D. G. Bailey, and O. Kittaneh, "Deriving scale normalisation factors for a GLoG detector," *IET Image Process.*, vol. 12, no. 9, pp. 1673–1682, 2018.
- [62] M. A. U. Khan and T. M. Khan, "Fingerprint image enhancement using data driven directional filter bank," *Optik*, vol. 124, no. 23, pp. 6063–6068, Dec. 2013.
- [63] M. A. U. Khan, T. M. Khan, D. G. Bailey, and Y. Kong, "A spatial domain scar removal strategy for fingerprint image enhancement," *Pattern Recognit.*, vol. 60, pp. 258–274, Dec. 2016.
- [64] T. M. Khan, D. G. Bailey, M. A. U. Khan, and Y. Kong, "Efficient hardware implementation for fingerprint image enhancement using anisotropic Gaussian filter," *IEEE Trans. Image Process.*, vol. 26, no. 5, pp. 2116–2126, May 2017.
- [65] A. Khawaja, T. M. Khan, M. A. U. Khan, and S. J. Nawaz, "A multi-scale directional line detector for retinal vessel segmentation," *Sensors*, vol. 19, no. 22, p. 4949, Nov. 2019.
- [66] T. A. Soomro, M. A. U. Khan, J. Gao, T. M. Khan, and M. Paul, "Contrast normalization steps for increased sensitivity of a retinal image segmentation method," *J. Signal Image Video Process.*, vol. 11, no. 8, pp. 1509–1517, 2017.
- [67] T. A. Soomro, J. Gao, M. A. U. Khan, T. M. Khan, and M. Paul, "Role of image contrast enhancement technique for ophthalmologist as diagnostic tool for diabetic retinopathy," in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Gold Coast, QLD, Australia, Nov. 2016, pp. 1–8.
- [68] Z. U. Rehman, S. S. Naqvi, T. M. Khan, M. Arsalan, M. A. Khan, and M. Khalil, "Multi-parametric optic disc segmentation using superpixel based feature classification," *Expert Syst. Appl.*, vol. 120, pp. 461–473, Apr. 2019.



- [69] A. Khawaja, T. M. Khan, K. Naveed, S. S. Naqvi, N. U. Rehman, and S. J. Nawaz, "An improved retinal vessel segmentation framework using frangi filter coupled with the probabilistic patch based denoiser," *IEEE Access*, vol. 7, pp. 164344–164361, 2019.
- [70] E. U. Francis, M. Y. Mashor, R. Hassan, and A. A. Abdullah, "Screening of bone marrow slide images for leukemia using multilayer perceptron (MLP)," in *Proc. IEEE Symp. Ind. Electron. Appl.*, Sep. 2011, pp. 643–648.
- [71] M. A. U. Khan, T. A. Soomro, T. M. Khan, D. G. Bailey, J. Gao, and N. Mir, "Automatic retinal vessel extraction algorithm based on contrast-sensitive schemes," in *Proc. Int. Conf. Image Vis. Comput. New Zealand (IVCNZ)*, Nov. 2016, pp. 1–5.
- [72] S. S. Naqvi, N. Fatima, T. M. Khan, Z. U. Rehman, and M. A. Khan, "Automatic optic disk detection and segmentation by variational active contour estimation in retinal fundus images," *Signal, Image Video Process.*, vol. 13, no. 6, pp. 1191–1198, Sep. 2019.
- [73] E. A. Baquero, A. Merino, L. Delgado, M. R. Ordóñez, L. Bigorra, and J. R. Benedé, "Digital blood image processing and fuzzy clustering for detection and classification of atypical lymphoid B cells," *Jornades de recerca EUETIB*, Barcelona, Spain, Tech. Rep., 2013, pp. 1–12.
- [74] L. Putzu, G. Caocci, and C. D. Ruberto, "Leucocyte classification for leukaemia detection using image processing techniques," *Artif. Intell. Med.*, vol. 62, no. 3, pp. 179–191, Nov. 2014.
- [75] M. T. Ibrahim, T. M. Khan, M. A. Khan, and L. Guan, "Automatic segmentation of pupil using local histogram and standard deviation," in *Visual Communications and Image Processing 2010*, vol. 7744, P. Frossard, H. Li, F. Wu, B. Girod, S. Li, and G. Wei, Eds. Bellingham, WA, USA: SPIE, 2010, pp. 879–886, doi: [10.1117/12.863264](https://doi.org/10.1117/12.863264).
- [76] M. A. U. Khan, T. M. Khan, and S. A. Khan, "Coherence enhancement diffusion using multi-scale DFB," in *Proc. 7th Int. Conf. Emerg. Technol.*, Sep. 2011, pp. 1–6.
- [77] T. M. Khan, D. G. Bailey, M. A. U. Khan, and Y. Kong, "Real-time edge detection and range finding using FPGAs," *Optik*, vol. 126, no. 17, pp. 1545–1550, Sep. 2015.
- [78] Biji and D. Hariharan, "White blood cell segmentation techniques in microscopic images for leukemia detection," *IONS J. Dent. Med. Sci.*, vol. 15, no. 2, pp. 45–51, Feb. 2016.
- [79] J. Su, S. Liu, and J. Song, "A segmentation method based on HMRP for the aided diagnosis of acute myeloid leukemia," *Comput. Methods Programs Biomed.*, vol. 152, pp. 115–123, Dec. 2017.
- [80] T. A. Soomro, M. A. U. Khan, J. Gao, T. M. Khan, M. Paul, and N. Mir, "Automatic retinal vessel extraction algorithm," in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov. 2016, pp. 1–8.
- [81] R. Yogamangalam and B. Karthikeyan, "Segmentation techniques comparison in image processing," *Int. J. Eng. Technol.*, vol. 5, no. 1, pp. 307–313, 2013.
- [82] M. A. Mohammed, M. K. A. Ghani, N. Arunkumar, S. A. Mostafa, M. K. Abdullah, and M. A. Burhanuddin, "Trainable model for segmenting and identifying nasopharyngeal carcinoma," *Comput. Electr. Eng.*, vol. 71, pp. 372–387, Oct. 2018.
- [83] S. Chuprat, "Segmentation and detection of acute leukemia using image processing and machine learning techniques: A," Tech. Rep. REVISTA AUS 26-2 511–531, doi: [10.4206/aus.2019.n26.2.60](https://doi.org/10.4206/aus.2019.n26.2.60).
- [84] K. Naveed, F. Abdullah, H. A. Madni, M. A. U. Khan, T. M. Khan, and S. S. Naqvi, "Towards automated eye diagnosis: An improved retinal vessel segmentation framework using ensemble block matching 3D filter," *Diagnostics*, vol. 11, no. 1, p. 114, Jan. 2021.
- [85] M. Mehmood, T. M. Khan, M. A. Khan, S. S. Naqvi, and W. Alhalabi, "Vessel intensity profile uniformity improvement for retinal vessel segmentation," *Procedia Comput. Sci.*, vol. 163, pp. 370–380, Jan. 2019.
- [86] R. M. Abbas, T. M. Khan, M. A. U. Khan, and H. A. Khan, "Fast and accurate iris segmentation methods using LoG filters," in *Proc. Int. Conf. Comput., Electron. Commun. Eng. (iCCECE)*, Aug. 2019, pp. 124–129.
- [87] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Upper Saddle River, NJ, USA: Prentice-Hall, Mar. 2017.
- [88] A. K. Jain, *Fundamentals of Digital Image Processing*. London, U.K.: Pearson, 1989.
- [89] J. Rawat, H. S. Bhadauria, A. Singh, and J. Virmani, "Review of leukocyte classification techniques for microscopic blood images," in *Proc. 2nd Int. Conf. Comput. Sustain. Global Develop. (INDIACom)*, 2015, pp. 1948–1954.
- [90] H. Ramoser, V. Laurain, H. Bischof, and R. Ecker, "Leukocyte segmentation and classification in blood-smear images," in *Proc. IEEE Eng. Med. Biol. 27th Annu. Conf.*, Jan. 2005, pp. 3371–3374.
- [91] 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, pp. 1–9, Jan. 2014.
- [92] L. Putzu and C. D. Ruberto, "White blood cells identification and counting from microscopic blood image," in *Proc. World Acad. Sci., Eng. Technol.*, no. 73, 2013, p. 363.
- [93] S. Mohapatra, D. Patra, S. Kumar, and S. Satpathy, "Lymphocyte image segmentation using functional link neural architecture for acute leukemia detection," *Biomed. Eng. Lett.*, vol. 2, no. 2, pp. 100–110, Jun. 2012.
- [94] 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, pp. 1–12, Apr. 2018.
- [95] 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. 2019.
- [96] D. M. U. Sabino, L. da Fontoura Costa, E. G. Rizzatti, and M. A. Zago, "A texture approach to leukocyte recognition," *Real-Time Imag.*, vol. 10, no. 4, pp. 205–216, Aug. 2004.
- [97] M. Benazzouz, I. Baghli, and M. A. Chikh, "Microscopic image segmentation based on pixel classification and dimensionality reduction," *Int. J. Imag. Syst. Technol.*, vol. 23, no. 1, pp. 22–28, Mar. 2013.
- [98] A. S. Abdul Nasir, N. Mustafa, and N. M. Nasir, "Application of thresholding technique in determining ratio of blood cells for leukemia detection," in *Proc. Int. Conf. Man-Mach. Syst. (ICoMMS)*, Penang, Malaysia, Oct. 2009, pp. 1–6.
- [99] P. Sarkar, M. Ghosh, and C. Chakraborty, "Texture-based leukocyte image retrieval using color normalization and quaternion Fourier transform based segmentation," *J. Netw. Innov. Comput.*, vol. 1, no. 2013, pp. 126–138, 2013.
- [100] L. Putzu and C. D. Ruberto, "White blood cells identification and classification from leukemic blood image," in *Proc. Int. Work-Confer. Bioinf. Biomed. Eng.*, 2013, pp. 99–106.
- [101] A. N. A. Salihah, M. Y. Mashor, N. H. Harun, A. A. Abdullah, and H. Rosline, "Improving colour image segmentation on acute myelogenous leukaemia images using contrast enhancement techniques," in *Proc. IEEE EMBS Conf. Biomed. Eng. Sci. (IECBES)*, Nov. 2010, pp. 246–251.
- [102] A. Mehnert and P. Jackway, "An improved seeded region growing algorithm," *Pattern Recognit. Lett.*, vol. 18, no. 10, pp. 1065–1071, 1997.
- [103] R. Bhattacharjee and L. M. Saini, "Detection of acute lymphoblastic leukemia using watershed transformation technique," in *Proc. Int. Conf. Signal Process., Comput. Control (ISPCC)*, Sep. 2015, pp. 383–386.
- [104] Amanpreet and A. Verma, "The marker-based watershed segmentation—A review," *Int. J. Eng. Innov. Technol.*, vol. 3, no. 3, pp. 171–174, 2013.
- [105] A. Rehman, N. Abbas, T. Saba, S. I. U. Rahman, Z. Mehmood, and H. Kolivand, "Classification of acute lymphoblastic leukemia using deep learning," *Microsc. Res. Technique*, vol. 81, no. 11, pp. 1310–1317, Nov. 2018.
- [106] H. N. Lim, M. Y. Mashor, N. Z. Supardi, and R. Hassan, "Color and morphological based techniques on white blood cells segmentation," in *Proc. 2nd Int. Conf. Biomed. Eng. (ICoBE)*, Mar. 2015, pp. 1–5.
- [107] A. Ghosh, S. Singh, and D. Sheet, "Simultaneous localization and classification of acute lymphoblastic leukemic cells in peripheral blood smears using a deep convolutional network with average pooling layer," in *Proc. IEEE Int. Conf. Ind. Inf. Syst. (ICIIS)*, Dec. 2017, pp. 1–6.
- [108] A. Gautam and H. S. Bhadauria, "White blood nucleus extraction using K-mean clustering and mathematical morphing," in *Proc. 5th Int. Conf. Confluence Next Gener. Inf. Technol. Summit (Confluence)*, Sep. 2014, pp. 549–554.
- [109] S. Ghosh and S. Kumar, "Comparative analysis of K-means and fuzzy C-means algorithms," *Int. J. Adv. Comput. Sci. Appl.*, vol. 4, no. 4, pp. 35–39, 2013.
- [110] S. Mishra, B. Majhi, P. K. Sa, and L. Sharma, "Gray level co-occurrence matrix and random forest based acute lymphoblastic leukemia detection," *Biomed. Signal Process. Control*, vol. 33, pp. 272–280, Mar. 2017.
- [111] J. Rawat, A. Singh, H. S. Bhadauria, J. Virmani, and J. S. Devgun, "Classification of acute lymphoblastic leukaemia using hybrid hierarchical classifiers," *Multimedia Tools Appl.*, vol. 76, no. 18, pp. 19057–19085, Sep. 2017.
- [112] V. Singhal and P. Singh, "Local binary pattern for automatic detection of acute lymphoblastic leukemia," in *Proc. 20th Nat. Conf. Commun. (NCC)*, Feb. 2014, pp. 1–5.
- [113] A. Bharathivanan and T. N. R. Shankar, "Local binary pattern texture based method for segmentation of leukemia in blood microscopic images," *Int. J. Appl. Eng. Res.*, vol. 10, no. 20, pp. 16291–16296, 2015.

- [114] N. TP and T. Divya, "Discriminative robust LBP for leukemia detection," *Int. J. Innov. Res. Sci. Technol.*, vol. 3, no. 4, pp. 192–196, Sep. 2016.
- [115] G. M. PatilTejashri and V. B. Raskar, "A robust segmentation method for acute lymphoblastic leukemia detection," *Int. J. Adv. Res. Comput. Sci. Softw. Eng.*, vol. 5, no. 1, 2015. [Online]. Available: <http://www.ijarcsse.com>
- [116] A. Jacob and F. A. Mundackal, "Automated screening system for acute leukemia detection and type classification," *Int. J. Adv. Res. Elect., Electron. Instrum. Eng.*, vol. 5, pp. 2426–2432, Apr. 2016.
- [117] A. A. Regina, "Detection of leukemia with blood microscopic images," *J. Innov. Res. Comput. Commun. Eng.*, vol. 3, no. 3, pp. 27–30, 2015.
- [118] Y. Li, R. Zhu, L. Mi, Y. Cao, and D. Yao, "Segmentation of white blood cell from acute lymphoblastic leukemia images using dual-threshold method," *Comput. Math. Methods Med.*, vol. 2016, pp. 1–12, Apr. 2016.
- [119] A. A. Nasir, M. Mashor, and H. Rosline, "Detection of acute leukaemia cells using variety of features and neural networks," in *Proc. 5th Kuala Lumpur Int. Conf. Biomed. Eng.* Berlin, Germany: Springer, 2011, pp. 40–46.
- [120] D. S. A. Trupti Kulkarni, "A robust morphological analysis of normal and abnormal leukemic cells populations in acute lymphoblastic leukemia," *Int. J. Electr., Electron. Data Commun.*, vol. 2, no. 8, pp. 62–66, Aug. 2014.
- [121] V. R. Tejashri and G. Patil, "Automated leukemia detection using shape feature extraction," *Int. J. Comput. Eng. Appl.*, vol. 9, no. 2, pp. 45–51, Feb. 2015.
- [122] M. Subhan and M. Kaur, "Significant analysis of leukemic cells extraction and detection using KNN and Hough transform algorithm," *Int. J. Comput. Sci. Trends Technol.*, vol. 3, no. 1, pp. 27–33, 2015.
- [123] T. Hazra, M. Kumar, and D. S. S. Tripathy, "Automatic leukemia detection using image processing technique," *Int. J. Latest Technol. Eng., Manage., Appl. Sci.*, vol. 6, no. 4, pp. 42–45, 2017.
- [124] H. T. Madhloom, S. A. Kareem, and H. Ariffin, "A robust feature extraction and selection method for the recognition of lymphocytes versus acute lymphoblastic leukemia," in *Proc. Int. Conf. Adv. Comput. Sci. Appl. Technol. (ACSAT)*, Nov. 2012, pp. 330–335.
- [125] J. Li and G. R. Weng, "PCA for leukemia classification," *Appl. Mech. Mater.*, vol. 43, pp. 744–747, Dec. 2010.
- [126] D.-C. Huang, K.-D. Hung, and Y.-K. Chan, "An adaptive leukocyte nucleus segmentation using genetic algorithm," in *Proc. Int. Symp. Intell. Signal Process. Commun. Syst.*, Nov. 2012, pp. 559–563.
- [127] J. Laosai and K. Chamnongthai, "Acute leukemia classification by using SVM and K-means clustering," in *Proc. Int. Electr. Eng. Congr. (iEECON)*, Mar. 2014, pp. 1–4.
- [128] N. Chatap and S. Shibu, "Analysis of blood samples for counting leukemia cells using support vector machine and nearest neighbour," *IOSR J. Comput. Eng.*, vol. 16, no. 5, pp. 79–87, 2014.
- [129] M. Y. Mashor, A. S. A. Nasir, and R. Hassan, "Performance comparison between multilayer perceptron and fuzzy ARTMAP networks for acute leukemia detection," *Science*, vol. 2, no. 5, pp. 1160–1166, 2011.
- [130] A. A. Nasir, M. Y. Mashor, and R. Hassan, "Classification of acute leukaemia cells using multilayer perceptron and simplified fuzzy ARTMAP neural networks," *Int. Arab J. Inf. Technol.*, vol. 10, no. 4, pp. 1–9, 2013.
- [131] P. M. Gumble and S. Rode, "Analysis & classification of acute lymphoblastic leukemia using KNN algorithm," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 5, no. 2, pp. 94–98, 2017.
- [132] C. Reta, L. Altamirano, J. A. Gonzalez, R. Diaz-Hernandez, H. Peregriana, I. Olmos, J. E. Alonso, and R. Lobato, "Segmentation and classification of bone marrow cells images using contextual information for medical diagnosis of acute leukemias," *PLoS ONE*, vol. 10, no. 6, Jun. 2015, Art. no. e0130805.
- [133] M. Saraswat and K. V. Arya, "Feature selection and classification of leukocytes using random forest," *Med. Biol. Eng. Comput.*, vol. 52, no. 12, pp. 1041–1052, Dec. 2014.
- [134] J. Prinyakupt and C. Pluempitwiriwajew, "Segmentation of white blood cells and comparison of cell morphology by linear and naïve Bayes classifiers," *Biomed. Eng. OnLine*, vol. 14, no. 1, pp. 1–19, Dec. 2015.
- [135] S. Selvaraj and B. Kanakaraj, "Naïve Bayesian classifier for acute lymphocytic leukemia detection," *ARNP J. Eng. Appl. Sci.*, vol. 10, no. 6, pp. 6888–6892, 2015.
- [136] R. S. Priya, "Analysis of blood samples using ANFIS classification," *Int. Res. J. Eng. Technol.*, vol. 4, no. 3, pp. 1460–1466, Mar. 2017.
- [137] M. Hosseini and M. Zekri, "Review of medical image classification using the adaptive neuro-fuzzy inference system," *J. Med. Signals Sensors*, vol. 2, no. 1, p. 49, 2012.
- [138] C. Mondal, M. K. Hasan, M. T. Jawad, A. Dutta, M. R. Islam, M. A. Awal, and M. Ahmad, "Acute lymphoblastic leukemia detection from microscopic images using weighted ensemble of convolutional neural networks," 2021, *arXiv:2105.03995*. [Online]. Available: <http://arxiv.org/abs/2105.03995>
- [139] B. W. Matthews, "Comparison of the predicted and observed secondary structure of T4 phage lysozyme," *Biochim. Biophys. Acta-Protein Struct.*, vol. 405, no. 2, pp. 442–451, Oct. 1975. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/0005279575901099>
- [140] M. Z. Hoque, A. Keskinarkaus, P. Nyberg, and T. Seppänen, "Retinex model based stain normalization technique for whole slide image analysis," *Comput. Med. Imag. Graph.*, vol. 90, Jun. 2021, Art. no. 101901.
- [141] S. S. S. Poon, R. K. Ward, and B. Palcic, "Automated image detection and segmentation in blood smears," *Cytometry*, vol. 13, no. 7, pp. 766–774, 1992.
- [142] J. Laosai and K. Chamnongthai, "Classification of acute leukemia using medical-knowledge-based morphology and CD marker," *Biomed. Signal Process. Control*, vol. 44, pp. 127–137, Jul. 2018.
- [143] A. S. Negm, O. A. Hassan, and A. H. Kandil, "A decision support system for acute leukaemia classification based on digital microscopic images," *Alexandria Eng. J.*, vol. 57, no. 4, pp. 2319–2332, Dec. 2018.
- [144] F. Scotti, "Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images," in *Proc. IEEE Int. Conf. Comput. Intell. Meas. Syst. Appl. (CIMSAS)*, Jul. 2005, pp. 96–101.
- [145] S. Kumar, S. Mishra, and P. Asthana, "Automated detection of acute leukemia using k-mean clustering algorithm," in *Advances in Computer and Computational Sciences* Singapore: Springer, 2018, pp. 655–670.
- [146] N. Z. Supardi, M. Y. Mashor, N. H. Harun, F. A. Bakri, and R. Hassan, "Classification of blasts in acute leukemia blood samples using k-nearest neighbour," in *Proc. IEEE 8th Int. Colloq. Signal Process. Appl.*, Mar. 2012, pp. 461–465.
- [147] M. Sajjad, S. Khan, Z. Jan, K. Muhammad, H. Moon, J. T. Kwak, S. Rho, S. W. Baik, and I. Mehmood, "Leukocytes classification and segmentation in microscopic blood smear: A resource-aware healthcare service in smart cities," *IEEE Access*, vol. 5, pp. 3475–3489, 2017.
- [148] A. Setiawan, A. Harjoko, T. Ratnaningsih, E. Suryani, and S. Palgunadi, "Classification of cell types in Acute Myeloid Leukemia (AML) of M4, M5 and M7 subtypes with support vector machine classifier," in *Proc. Int. Conf. Inf. Commun. Technol. (ICOIACT)*, Mar. 2018, pp. 45–49.
- [149] L. Faivdullah, F. Azahar, Z. Z. Htike, and W. Y. N. Naing, "Leukemia detection from blood smears," *J. Med. Bioeng.*, vol. 4, no. 6, pp. 488–491, 2015.
- [150] 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.
- [151] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "SSD: Single shot multibox detector," in *Proc. Eur. Conf. Comput. Vis. Cham, Switzerland: Springer*, 2016, pp. 21–37.
- [152] F. N. Iandola, S. Han, M. W. Moskewicz, K. Ashraf, W. J. Dally, and K. Keutzer, "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," 2016, *arXiv:1602.07360*. [Online]. Available: <http://arxiv.org/abs/1602.07360>
- [153] P. Tiwari, J. Qian, Q. Li, B. Wang, D. Gupta, A. Khanna, J. J. Rodrigues, and V. H. C. de Albuquerque, "Detection of subtype blood cells using deep learning," *Cognit. Syst. Res.*, vol. 52, pp. 1036–1044, Dec. 2018.
- [154] 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.
- [155] T. Tran, O.-H. Kwon, K.-R. Kwon, S.-H. Lee, and K.-W. Kang, "Blood cell images segmentation using deep learning semantic segmentation," in *Proc. IEEE Int. Conf. Electron. Commun. Eng. (ICECE)*, Dec. 2018, pp. 13–16.
- [156] A. Tareef, Y. Song, D. Feng, M. Chen, and W. Cai, "Automated multi-stage segmentation of white blood cells via optimizing color processing," in *Proc. IEEE 14th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2017, pp. 565–568.
- [157] W. Yu, J. Chang, C. Yang, L. Zhang, H. Shen, Y. Xia, and J. Sha, "Automatic classification of leukocytes using deep neural network," in *Proc. IEEE 12th Int. Conf. ASIC (ASICON)*, Oct. 2017, pp. 1041–1044.
- [158] T. T. P. Thanh, C. Vununu, S. Atoev, S.-H. Lee, and K.-R. Kwon, "Leukemia blood cell image classification using convolutional neural network," *Int. J. Comput. Theory Eng.*, vol. 10, no. 2, pp. 54–58, 2018.

- [159] J. Amin, M. Sharif, M. Raza, T. Saba, and M. A. Anjum, "Brain tumor detection using statistical and machine learning method," *Comput. Methods Programs Biomed.*, vol. 177, pp. 69–79, Aug. 2019.
- [160] J. Amin, M. Sharif, M. Raza, T. Saba, and A. Rehman, "Brain tumor classification: Feature fusion," in *Proc. Int. Conf. Comput. Inf. Sci. (ICIS)*, Apr. 2019, pp. 1–6.
- [161] L. Vogado, R. Veras, F. H. D. Araújo, R. R. V. Silva, and K. Aires, "Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification," *Eng. Appl. Artif. Intell.*, vol. 72, pp. 415–422, Jul. 2018.
- [162] J. Zhao, M. Zhang, Z. Zhou, J. Chu, and F. Cao, "Automatic detection and classification of leukocytes using convolutional neural networks," *Med. Biol. Eng. Comput.*, vol. 55, no. 8, pp. 1287–1301, Aug. 2017.
- [163] 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," in *Proc. 10th Int. Conf. Mach. Vis., Soc. Photo-Opt. Instrum. Eng. (ICMV)*, A. Verikas, P. Radeva, D. Nikolaev, and J. Zhou, Eds., vol. 10696, Apr. 2018, Art. no. 1069612.
- [164] L. Lin, W. Wang, and B. Chen, "Leukocyte recognition with convolutional neural network," *J. Algorithms Comput. Technol.*, vol. 13, Nov. 2018, Art. no. 1748301818813322.
- [165] J. L. Wang, A. Y. Li, M. Huang, A. K. Ibrahim, H. Zhuang, and A. M. Ali, "Classification of white blood cells with PatternNet-fused ensemble of convolutional neural networks (PECNN)," in *Proc. IEEE Int. Symp. Signal Process. Inf. Technol. (ISSPIT)*, Dec. 2018, pp. 325–330.
- [166] S. Agaian, M. Madhukar, and A. T. Chronopoulos, "Automated screening system for acute myelogenous leukemia detection in blood microscopic images," *IEEE Syst. J.*, vol. 8, no. 3, pp. 995–1004, Sep. 2014.
- [167] K. K. Anilkumar, V. J. Manoj, and T. M. Sagi, "Colour based image segmentation for automated detection of leukaemia: A comparison between CIELAB and CMYK colour spaces," in *Proc. Int. Conf. Circuits Syst. Digit. Enterprise Technol. (ICSDDET)*, Dec. 2018, pp. 1–6.
- [168] J. Rodellar, S. Alférez, A. Acevedo, A. Molina, and A. Merino, "Image processing and machine learning in the morphological analysis of blood cells," *Int. J. Lab. Hematol.*, vol. 40, pp. 46–53, May 2018.
- [169] P. Viswanathan, "Fuzzy C means detection of leukemia based on morphological contour segmentation," *Proc. Comput. Sci.*, vol. 58, pp. 84–90, Jan. 2015.
- [170] A. Rejintal and N. Aswini, "Image processing based leukemia cancer cell detection," in *Proc. IEEE Int. Conf. Recent Trends Electron., Inf. Commun. Technol. (RTEICT)*, May 2016, pp. 471–474.
- [171] A. R. J. Begum and T. A. Razak, "Diagnosing leukemia from microscopic images using image analysis and processing techniques," in *Proc. World Congr. Comput. Commun. Technol. (WCCCT)*, Feb. 2017, pp. 227–230.
- [172] R. J. A. Cabrera, C. A. P. Legaspi, E. J. G. Papa, R. D. Samonte, and D. D. Acula, "HeMatic: An automated leukemia detector with separation of overlapping blood cells through image processing and genetic algorithm," in *Proc. Int. Conf. Appl. Syst. Innov. (ICASI)*, May 2017, pp. 985–987.
- [173] P. Jagadev and H. G. Virani, "Detection of leukemia and its types using image processing and machine learning," in *Proc. Int. Conf. Trends Electron. Informat. (ICEI)*, May 2017, pp. 522–526.
- [174] Y. Wang and Y. Cao, "Quick leukocyte nucleus segmentation in leukocyte counting," *Comput. Math. Methods Med.*, vol. 2019, pp. 1–10, Jun. 2019.
- [175] P. Kumar and S. M. Udwadia, "Automatic detection of acute myeloid leukemia from microscopic blood smear image," in *Proc. Int. Conf. Adv. Comput., Commun. Informat. (ICACCI)*, Sep. 2017, pp. 1803–1807.
- [176] S. Alférez, A. Merino, L. E. Mujica, M. Ruiz, L. Bigorra, and J. Rodellar, "Automatic classification of atypical lymphoid b cells using digital blood image processing," *Int. J. Lab. Hematol.*, vol. 36, no. 4, pp. 472–480, Aug. 2014.
- [177] N. Patel and A. Mishra, "Automated leukaemia detection using microscopic images," *Proc. Comput. Sci.*, vol. 58, pp. 635–642, Jan. 2015.
- [178] S. Hariprasath, T. Dharani, and B. N. S. Mohammad, "Automated detection of acute lymphocytic leukemia using blast cell morphological features," in *Proc. 2nd Int. Conf. Adv. Sci. Technol. (ICAST)*, 2019, pp. 1–6.
- [179] E. Tuba, I. Strumberger, N. Bacanin, D. Zivkovic, and M. Tuba, "Acute lymphoblastic leukemia cell detection in microscopic digital images based on shape and texture features," in *Proc. Int. Conf. Swarm Intell. Cham, Switzerland: Springer*, 2019, pp. 142–151.
- [180] C. Zhang, S. Wu, Z. Lu, Y. Shen, J. Wang, P. Huang, J. Lou, C. Liu, L. Xing, J. Zhang, J. Xue, and D. Li, "Hybrid adversarial-discriminative network for leukocyte classification in leukemia," *Med. Phys.*, vol. 47, no. 8, pp. 3732–3744, Aug. 2020.
- [181] N. H. A. Halim, M. Y. Mashor, and R. Hassan, "Automatic blasts counting for acute leukemia based on blood samples," *Int. J. Res. Rev. Comput. Sci.*, vol. 2, no. 4, p. 971, 2011.
- [182] S. Biswas and D. Ghoshal, "Blood cell detection using thresholding estimation based watershed transformation with Sobel filter in frequency domain," *Proc. Comput. Sci.*, vol. 89, pp. 651–657, Jan. 2016.
- [183] B. C. Ko, J.-W. Gim, and J.-Y. Nam, "Automatic white blood cell segmentation using stepwise merging rules and gradient vector flow snake," *Micron*, vol. 42, no. 7, pp. 695–705, Oct. 2011.
- [184] F. I. Sholeh, "White blood cell segmentation for fresh blood smear images," in *Proc. Int. Conf. Adv. Comput. Sci. Inf. Syst. (ICACSIS)*, Sep. 2013, pp. 425–429.
- [185] Q. Li, Y. Wang, H. Liu, J. Wang, and F. Guo, "A combined spatial-spectral method for automated white blood cells segmentation," *Opt. Laser Technol.*, vol. 54, pp. 225–231, Dec. 2013.
- [186] S. Ravikumar and A. Shanmugam, "WBC image segmentation using modified fuzzy possibilistic C-means algorithm," *J. Theor. Appl. Inf. Technol.*, vol. 63, pp. 315–323, May 2014.
- [187] A. Jati, G. Singh, R. Mukherjee, M. Ghosh, A. Konar, C. Chakraborty, and A. K. Nagar, "Automatic leukocyte nucleus segmentation by intuitionistic fuzzy divergence based thresholding," *Micron*, vol. 58, pp. 55–65, Mar. 2014.
- [188] Y. Liu, F. Cao, J. Zhao, and J. Chu, "Segmentation of white blood cells image using adaptive location and iteration," *IEEE J. Biomed. Health Informat.*, vol. 21, no. 6, pp. 1644–1655, Nov. 2016.
- [189] V. P. Ananthi and P. Balasubramaniam, "A new thresholding technique based on fuzzy set as an application to leukocyte nucleus segmentation," *Comput. Methods Programs Biomed.*, vol. 134, pp. 165–177, Oct. 2016.
- [190] 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.
- [191] B. Mourtada, D. Xavier, C. M. Amine, and B. M. Lamine, "Multi features based approach for white blood cells segmentation and classification in peripheral blood and bone marrow images," *Int. J. Biomed. Eng. Technol.*, vol. 1, no. 1, p. 1, 2019.
- [192] S. Nazlibilek, D. Karacor, T. Ercan, M. H. Sazli, O. Kalender, and Y. Ege, "Automatic segmentation, counting, size determination and classification of white blood cells," *Measurement*, vol. 55, no. 3, pp. 58–65, 2014.
- [193] R. B. Hegde, K. Prasad, H. Hebbar, and B. M. K. Singh, "Comparison of traditional image processing and deep learning approaches for classification of white blood cells in peripheral blood smear images," *Biocybernetics Biomed. Eng.*, vol. 39, no. 2, pp. 382–392, Apr. 2019.
- [194] P. Ghosh, D. Bhattacharjee, and M. Nasipuri, "Blood smear analyzer for white blood cell counting: A hybrid microscopic image analyzing technique," *Appl. Soft Comput.*, vol. 46, pp. 629–638, Sep. 2016.
- [195] S. N. M. Safuan, M. R. M. Tomari, and W. N. W. Zakaria, "White blood cell (WBC) counting analysis in blood smear images using various color segmentation methods," *Measurement*, vol. 116, pp. 543–555, Feb. 2018.
- [196] M. M. Amin, S. Kermani, A. Talebi, and M. G. Oghli, "Recognition of acute lymphoblastic leukemia cells in microscopic images using K-means clustering and support vector machine classifier," *J. Med. Signals Sensors*, vol. 5, no. 1, pp. 49–58, 2015.
- [197] V. Shankar, M. M. Deshpande, N. Chaitra, and S. Aditi, "Automatic detection of acute lymphoblastic leukemia using image processing," in *Proc. IEEE Int. Conf. Adv. Comput. Appl. (ICACA)*, Oct. 2016, pp. 186–189.
- [198] M. MoradiAmin, A. Memari, N. Samadzadehghadam, S. Kermani, and A. Talebi, "Computer aided detection and classification of acute lymphoblastic leukemia cell subtypes based on microscopic image analysis," *Microsc. Res. Technique*, vol. 79, no. 10, pp. 908–916, Oct. 2016.
- [199] M. A. Khosrosereshki and M. B. Menhaj, "A fuzzy based classifier for diagnosis of acute lymphoblastic leukemia using blood smear image processing," in *Proc. 5th Iranian Joint Congr. Fuzzy Intell. Syst. (CFIS)*, Mar. 2017, pp. 13–18.
- [200] W. Srisukham, L. Zhang, S. C. Neoh, S. Todryk, and C. P. Lim, "Intelligent leukaemia diagnosis with bare-bones PSO based feature optimization," *Appl. Soft Comput.*, vol. 56, pp. 405–419, Jul. 2017.
- [201] H. J. Escalante, M. Montes-y-Gómez, J. A. González, P. Gómez-Gil, L. Altamirano, C. A. Reyes, C. Reta, and A. Rosales, "Acute leukemia classification by ensemble particle swarm model selection," *Artif. Intell. Med.*, vol. 55, no. 3, pp. 163–175, Jul. 2012.

- [202] S. C. Neoh, W. Srisukham, L. Zhang, S. Todryk, B. Greystoke, C. P. Lim, M. A. Hossain, and N. Aslam, "An intelligent decision support system for leukaemia diagnosis using microscopic blood images," *Sci. Rep.*, vol. 5, no. 1, pp. 1–14, Dec. 2015.
- [203] N. D. Khalilabad and H. Hassanpour, "Employing image processing techniques for cancer detection using microarray images," *Comput. Biol. Med.*, vol. 81, pp. 139–147, Feb. 2017.
- [204] S. Dumyan and A. Gupta, "An enhanced technique for lymphoblastic cancer detection using artificial neural network," *Int. J. Adv. Res. Comput. Sci. Electron. Eng.*, vol. 6, no. 4, pp. 38–42, 2017.
- [205] Z. Moshavash, H. Danyali, and M. S. Helfroush, "An automatic and robust decision support system for accurate acute leukemia diagnosis from blood microscopic images," *J. Digit. Imag.*, vol. 31, no. 5, pp. 702–717, Oct. 2018.
- [206] K. K. Jha and H. S. Dutta, "Mutual information based hybrid model and deep learning for acute lymphocytic leukemia detection in single cell blood smear images," *Comput. Methods Programs Biomed.*, vol. 179, Oct. 2019, Art. no. 104987.
- [207] T. Pansombut, S. Wikaisuksakul, K. Khongkrapan, and A. Phon-on, "Convolutional neural networks for recognition of lymphoblast cell images," *Comput. Intell. Neurosci.*, vol. 2019, pp. 1–12, Jun. 2019.
- [208] M. Makem and A. Tiedeu, "An efficient algorithm for detection of white blood cell nuclei using adaptive three stage PCA-based fusion," *Informat. Med. Unlocked*, vol. 20, 2020, Art. no. 100416.
- [209] J. Rawat, A. Singh, H. S. Bhadauria, and J. Virmani, "Computer aided diagnostic system for detection of leukemia using microscopic images," *Procedia Comput. Sci.*, vol. 70, pp. 748–756, 2015.
- [210] A. T. Sahlol, M. A. Abdeldaim, and A. E. Hassanien, "Automatic acute lymphoblastic leukemia classification model using social spider optimization algorithm," *Soft Comput.*, vol. 23, no. 15, pp. 6345–6360, 2019.
- [211] L. H. Nee, M. Y. Mashor, and R. Hassan, "White blood cell segmentation for acute leukemia bone marrow images," *J. Med. Imag. Health Inform.*, vol. 2, no. 3, pp. 278–284, 2012.
- [212] S. Skrobanski, S. Pavlidis, W. Ismail, R. Hassan, S. Counsell, and S. Swift, "Use of general purpose GPU programming to enhance the classification of leukaemia blast cells in blood smear images," in *Proc. Int. Symp. Intell. Data Anal.* Berlin, Germany: Springer, 2012, pp. 369–380.
- [213] F. Kazemi, T. Najafabadi, and B. Araabi, "Automatic recognition of acute myelogenous leukemia in blood microscopic images using K-means clustering and support vector machine," *J. Med. Signals Sensors*, vol. 6, no. 3, p. 183, 2016.
- [214] E. Suryani, S. Palgunadi, and T. N. Pradana, "Classification of acute myelogenous leukemia (AML M2 and AML M3) using momentum back propagation from watershed distance transform segmented images," *J. Phys., Conf. Ser.*, vol. 801, Jan. 2017, Art. no. 012044.
- [215] C. Matek, S. Schwarz, K. Spiekermann, and C. Marr, "Human-level recognition of blast cells in acute myeloid leukaemia with convolutional neural networks," *Nature Mach. Intell.*, vol. 1, no. 11, pp. 538–544, Nov. 2019.



**AFSHAN SHAH** received the B.E. (electronics) and M.E. degrees from Mehran University of Engineering and Technology, Jamshoro, in 2011 and 2013, respectively. She is currently pursuing the Ph.D. degree with the Department of Electrical Engineering, COMSATS University Islamabad. She is also working with the Department of Electrical Engineering, Sukkur IBA University, as a Lecturer (electronic engineering). Her current research interests include medical image analysis and computer aided diagnosis.



**SYED SAAD NAQVI** received the B.Sc. degree in computer engineering from the COMSATS Institute of Information Technology, Islamabad, Pakistan, in 2005, the M.Sc. degree in electronic engineering from The University of Sheffield, U.K., in 2007, and the Ph.D. degree from the School of Engineering and Computer Science, Victoria University of Wellington, New Zealand,

in 2016. He is currently working as an Assistant Professor with the COMSATS Institute of Information Technology. His research interests include saliency modeling, medical image analysis, scene understanding, and deep learning methods for image analysis.



**KHURAM NAVEED** received the Ph.D. degree from COMSATS University Islamabad (CUI), Islamabad, Pakistan, in 2020. He was a Visiting Researcher with Tandon School of Engineering, New York University (NYU), New York, USA, from January 2018 to July 2018. He is currently working as a Lecturer with the Department of Electrical and Computer Engineering, CUI. His research interests include signal and image processing using multiscale and multi-component signal decomposition methods. His recent research focuses on the design machine learning and deep learning techniques for medical diagnosis as well as other real world applications.



**NEMA SALEEM** received the B.Sc. degree (Hons.) from Alexandria University, Egypt, in 1987, the M.Sc. degree in mitral valve diagnosis from Alexandria University, Egypt, in 1990, and the Ph.D. degree in classification of breast tumors by acutance measure and shape factors from Calgary University, Canada, in 1996. She was the Chair of the Department of Electronics and Communication Engineering at Effat for one academic year. She has been an Assistant Professor with Alexandria University, since 1987, Effat University, since 2008, Alexandria Higher Institute of Engineering & Technology for five years, and Hadhramout University for one year. She supervised many capstone projects, master and Ph.D. theses, and has published many papers in international conferences and journals. She participated at NCAAA and ABET accreditation. She has received many awards and appreciations from different places. She taught gifted students for six years and organized many events as the Science and Engineering Fair and the 5th GCC Robotics Challenge.



**MOHAMMAD A. U. KHAN** received the B.S. degree from UET Peshawar, Pakistan, in 1992, the M.Sc. degree from Loughborough University of Technology, U.K., in 1994, and the Ph.D. degree from Georgia Tech, in 1999. Since 1999, he has been actively involved in image processing, pattern recognition, and machine learning. He has published extensively in residual vector quantization, directional filter bank, retinal vessel extraction, and signature verification.



**KHURRAM S. ALIMGEER** was born in 1977. He received the bachelor's degree in IT, in 2002, the M.S. degree in electronic engineering, in 2006, and the Ph.D. degree in electrical engineering with specialization in antenna design, in 2014. He is currently an Assistant Professor with CUI and working as a Researcher with RF-Laboratory, COMSATS University Islamabad (CUI), Islamabad Campus, Pakistan. He has published more than 80 research articles at reputed journals and conferences in the fields of antenna design, wireless communications, image processing, and mathematical modeling. He received the Gold Medal for his M.S. degree. He has been serving as an editor and a reviewer for few reputed impact factor journals, since 2008.

...