Classification of Acute Lymphoblastic Leukemia through the Fusion of Local Descriptors

Shakhawan Hares Wady[1,2]

[1]Department of Applied Computer, College of Medicals and Applied Sciences, Charmo University, Chamchamal, Sulaimani, KRG, Iraq, [2]Department of Information Technology, University College of Goizha, Sulaimani, KRG, Iraq

ABSTRACT

Leukemia is characterized by an abnormal proliferation of leukocytes in the bone marrow and blood, which is usually detected by pathologists using a microscope to examine a blood smear. Leukemia identification and diagnosis in advance are a trending topic in medical applications for decreasing the death toll of individuals with Acute Lymphoblastic Leukemia (ALL). It is critical to analyze the white blood cells for the identification of ALL for which the blood smear images are utilized. This paper discusses and presents a micro-pattern descriptor, called Local Directional Number Pattern along with Multi-scale Weber Local Descriptor for feature extraction task to determine cancerous and noncancerous blood cells. A balanced dataset with 260 blood smear images from the ALL-IDB2 dataset was used as training data. Consequently, a proposed model was constructed by applying different individual and combined feature extraction methods, and fed into the machine learning classifiers (Decision Tree, Ensemble, K-Nearest Neighbors, Naïve Bayes, and Random Forest) to determine cancerous and noncancerous blood cells. Experimental results indicate that the developed feature fusion technique assured a reasonable performance compared to other existing works with a testing average accuracy of 97.69 ± 1.83% using Ensemble classifier.

Index Terms: Leukemia Diagnosis, Blood Smear, Feature Extraction, Machine Learning

1. INTRODUCTION

Hematology is the branch of medicine concerned with the study of blood and blood-forming organs, including the diagnosis, treatment, and prevention of illnesses of the blood, bone marrow cells, an immune-mediated disease, hemostatic, and vascular systems. Hematologists conduct and analyze a wide range of laboratory tests to assist clinicians in the diagnosis and treatment of disease. They deal with the blood and bone marrow to provide patients with immediate clinical care [1]. The amount and type of cells produced at any given time is controlled by your body's requirements. In some cases, estimating how bone marrow cells contribute to a clinical disease may be more essential than determining the patient's hematologic status. Hematologists typically receive blood smear samples and study them for abnormality; if they identify the presence of diseases, they perform a bone marrow biopsy and provide a diagnosis in a short period of time.

The bone marrow's primary function is to produce red blood cells (RBCs or erythrocytes), platelets (or thrombocytes), and white blood cells (WBCs or leukocytes) [2], as shown in Fig. 1. Complete Blood Count (CBC) is a laboratory hematology analyzer medical test that provides information that can be used to diagnose a disease. The CBC measures the production of all the cellular components, identifies the patient's oxygen-carrying capacity by evaluating RBC counts, and allows for the estimation of the immune system by evaluating WBC counts with differential. This test supports in the diagnosis of anemia, certain cancers, infections, and a variety of other
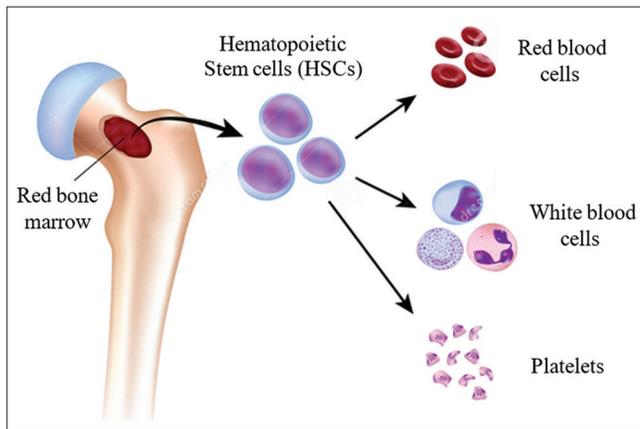
Fig. 1. Blood smear components (©Alila Medical Media, used with permission).

conditions, as well as monitoring the side effects of certain medications [3].

The absolute or relative count of WBCs, as well as their appearance on a blood smear, can be influenced by a variety of disorders. This is the case with parasite infections that cause an increase in WBCs, while the most severe cases are certainly leukemias [1]. Leukemia also known as Acute Lymphoblastic Leukemia (ALL), is a form of cancer of the blood or bone marrow that causes the body to produce cancerous WBCs called lymphocytes. These uncontrolled blood cells divisions harm the blood, lymphatic system, and bone marrow, throwing the immune system at risk [4]. They can also inhibit bone marrow's ability to produce RBCs and platelets. In addition, these cancerous WBCs can enter the bloodstream and cause harm to other regions of the human body, including as the liver, kidney, spleen, brain, and other organs, leading to the development of other deadly cancers.

In terms of how quickly it develops or gets worse, leukemia is characterized as either acute (which appears suddenly and grows quickly) or chronic (which develops more slowly). Acute leukemia can also be classified as lymphoblastic (ALL) or myelogenous depending on the cell type that is affected. Chronic lymphocytic leukemia (CLL) and chronic myeloid leukemia are two types of chronic leukemia [5-7]. ALL is the primary target of this research since it is predicted to have a superior survival rate than other categories. Only a suspicious and meticulous microscopic investigation of stained blood smears can accurately identify leukemia. Manual examination, due to the complicated nature of WBCs, results in inconsistencies in slide processing, resulting in non-standardized, unreliable, and subjective observations. Consequently, a cost-effective and robust automated system

is crucial to satisfy the demand for accurate diagnosis and classification without being impacted by experts. For this reason, several Computer-Aided Diagnosis (CAD) systems have been developed for identifying blast cells from microscopic blood images.

This paper proposes a new technique for detecting ALL-leukemia cells from microscopic blood images using a publicly available and widely used dataset (ALL-IDB). The rest of the paper is organized as follows. Section 2 puts forward a literature survey. Section 3 presents a complete CAD system for the detection of ALL, including sections such as an overview of system architecture, ALL-IDB dataset description, data preprocessing, feature extraction, feature fusion, classification, and performance metrics. Section 4 discusses the experimental results achieved after applying different feature extractors and comparing them with the existing approaches. Finally, Section 5 provides the conclusion of the work.

2. LITERATURE SURVEY

In biomedical image analysis and processing, machine learning and image processing methods have produced outstanding consequences, particularly in the field of ALL [8]. These methods are frequently performed in the classification on microscopic blood smear for ALL detection. For automatic classification of these disease-causing infections, a number of approaches have been effectively applied. Various procedures include convolutional neural networks (CNNs), ensemble learning, feature extraction, and feature selection, among others [9]. Recently, a number of works were executed with the help of numerous machines learning based procedures to detect and classify of leukemia microscopic blood smear images. Several transfer learning approaches, different system designs, and ensemble solutions were optimized to increase system performance in classifying acute leukemias, normal, and other disorders of the bone marrow and blood cells. A short review of some substantial contributions from the existing literature is provided.

Based on peripheral blood smear images, Al-Tahhan [10] intended to develop an improved classification model that is capable to categorize the ALL subtypes. In that article, the cytoplasmic vacuoles and the regularity of the nucleus membrane of ALL cells are the only geometric features that this approach relied on. Support Vector Machine (SVM) with multiple kernels, K-Nearest Neighbors (KNN) with various metric functions, and Artificial Neural Network (ANN)

were chosen and fine-tuned to identify automatically the subtypes of ALL utilizing ALL-IDB2 dataset. The authors in Shafique and Tehsin [4] developed a deep CNN (DCNN) model, trained on ALL-IDB dataset augmented with 50 private images, for the classification of ALL and its subtypes using pre-trained AlexNet and fine-tuning. Sharif *et al.* [11] recommended a YOLOv2-Nucleus-cytoplasm based scheme for WBCs localization using blood smear images. In that article, the Bag-of-Features were extracted from blood smear images of WBCs and optimized by using Particle Swarm Optimization for the classification task. On two challenging datasets, Leukocyte-Images for-Segmentation-Classification (LISC) and ALL-IDB, the classification results were computed. The experimental results reveal that the Optimized Naïve Bayes (O-NB) classifier outperformed the Optimized Discriminant Analysis (O-DA) classifier on ALL-IDB1, and ALL-IDB2 datasets. On the LISC dataset, however, the O-DA classifier outperformed the O-NB classifier.

Jha and Dutta [12] introduced the Sine Cosine Algorithm–based actor-critic neural network technique for leukemia classification. In that article, the developed entropy-based hybrid approach was used to segment the blood smear images, and the image-level features and statistical features were extracted from the segments. The relevant features were then fed into the designed classifier, which diagnoses leukemia using the ALL-IDB2 database. In another work [13], the authors proposed a new approach for detecting leukemia in blood images using transfer learning in CNNs and SVM for classification task. The procedure used pre-trained CNN methods (AlexNet, CaffeNet, and VGG-F) to extract images features directly without any prior preprocessing using three heterogeneous datasets. The authors of Di Ruberto *et al.* [1] developed a novel approach for leukocyte counting and classification of WBCs as healthy or affected by leukemia from microscopic blood images. The System, Man and Cybernetics-Image Data Base, Iran University of Medical Science-Image Data Base, and ALL-IDB public datasets for leukemia identification were utilized to evaluate the proposed approach. In Muthumayil *et al.* [14], the authors addressed a computer-based application technique based on Enhanced Virtual Neural Network classification for identifying and classifying CLL utilizing microscopic images of WBCs. The proposed technique attained the optimum accuracy in terms of detecting and classifying leukemia using WBCs images. In terms of sensitivity, specificity, accuracy, and misclassification error, the presented approach reached 97.8%, 89.9%, 76.6%, and 2.2%, respectively.

The authors in Singhal and Singh [15] proposed a new method for distinguishing ALL lymphoblast cells from healthy lymphocytes. Leukocytes were isolated from other blood cells first, and then lymphocytes were extracted. In this context, a novel CAD based on Gray Level Co-occurrence Matrices (GLCM) and shape-based features was developed for the diagnosis of hematological disorders such as leukemia (blood cancer). The auto SVM binary classifier was used to classify the retrieved features to determine the existence of leukemic cells. A new technique for segmenting and classifying ALL using the input blood smear images was proposed in Praveena and Singh [16]. In that article, the DCNN was trained using an optimization technique called Grey wolf-based Jaya Optimization Algorithm, which was built using the Grey Wolf Optimizer and Jaya Optimization Algorithm. Using the ALL-IDB2 dataset, the investigation was conducted by means of performance measures. The maximum accuracy, sensitivity, and specificity of the recommended approach were 0.9350, 0.9528, and 0.9389, respectively. The authors of Singhal and Singh [15] addressed an automatic leukemia detection system that used two texture descriptors, namely, Local Binary pattern (LBP) and GLCM generated from the nucleus image to detect ALL. The ALL-IDB2 dataset with 260 blood smear images (130 normal and 130 blast) was used as training data for a two-class classification. The experimental results from this paper demonstrated that LBP performed better than GLCM texture features with the classification accuracies of 93.84% and 87.30%, respectively.

The study Umamaheswari and Geetha [17] developed a scheme for optimized identification and detection of ALL using a novel customized-KNN classification model. During the preprocessing stage of this work, the medical image was prepared for segmentation through changing its size, brightness, and contrast. The nucleus portion was segmented using mathematical operators and Otsu's thresholding during the segmentation phase. Afterward, in the post-processing phase, mathematical morphological operators were used to make the nucleus portion appropriate for feature extraction task. Finally, using the developed customized KNN classification technique, the segmented portions were categorized into ALL affected and normal cells. This work was tested on over 80 images from the ALL-IDB2 dataset and achieved an average accuracy of 96.25%, sensitivity of 95%, and specificity of 97%. The authors in Tuba and Tuba [18] introduced Generative Adversarial Optimization (GAO)-based scheme for detecting acute lymphocytic leukemia as normal cells or blasts in microscopic digital blood images. In this study, the ALL-IDB2 image dataset was employed, and each blood image was characterized by 11 features

for classification tasks, along with 5 shape and 6 texture features. As a classifier, the SVM was used, and its criteria were tuned using an innovative optimization technique: The GAO technique.

Hariprasath *et al*. proposed a methodology for detection of acute lymphocytic leukemia from largest publicly accessible standard ALL-IDB dataset using statistical features. To differentiate between benign and leukemic cells, morphological and statistical features of blast cells were examined. SVM-R, SVM-L, and KNN classifiers were utilized and compared for classification process [19]. A unique strategy to automatically identify and classify ALL from peripheral blood smear images based on traditional image processing and machine learning approaches was proposed in Bodzas *et al*. [20]. In this work, to achieve the best segmentation results, substantial pre-processing and a three-phase filtration procedure were performed. Furthermore, sixteen robust features were extracted from microscopic images to discriminate between cancerous and noncancerous blood cells. Two popular machine learning classifiers, the ANN and the SVM, were computed to perform the classification task. A dataset of 31 peripheral blood smear images from a local dataset was utilized to test this approach, and they achieved a sensitivity of 100% and an average accuracy of 97.52%. The research paper [21] addresses a strategy for automatically detecting WBCs in complicated blood smear images based on Watershed Transform and circle fitting method. For the separation of overlapped WBCs, the presented technique employed segmentation and edge map extraction in preprocessing steps, in addition to parametric circle approximation, which identified both separated and overlapped WBCs. A dataset of 384 WBC images, including some overlapping cases, from ALL-IDB and ASH image bank was utilized to test this method.

This paper discusses and presents a micro-pattern descriptor, called Local Directional Number Pattern (LDNP) along with MWDT for feature extraction task. A balanced dataset with 260 blood smear images from the ALL-IDB2 dataset was used as training data. Consequently, a proposed model was constructed by applying different individual and combined feature extraction methods to overcome the most challenging parts of the detection of ALL-leukemia cells from microscopic blood images.

3. PROPOSED METHODOLOGY

3.1. System Architecture
To classify ALL-leukemia cells, the proposed approach used microscopic blood images as input data. To begin,

this method transformed RGB color images to gray scale and defined the areas of interest of healthy and blast cells by removing unwanted areas. Moreover, the system studied two feature descriptors: LDNP and MDWT. First, a feature vector was extracted from microscopic blood images of the ALL IDB2 database using the LDNP descriptor. Then the MDWT descriptor was performed to extract another feature vector from the same microscopic blood images. Various individual and combined feature extraction methods were constructed and fed into the classification model as input data. The number of features extracted by one method was insufficient to accurately classify ALL-leukemia cells. Nevertheless, using two distinct strategies to extract features could lead to a large number of features for classification task. In this circumstance, fusion was considered as a combination of the distinct feature vectors. Finally, the fused features were then classified microscopic blood images to determine cancerous and noncancerous blood cells using five well-known classifiers (Decision Tree [DT], Ensemble, KNN, Naïve Bayes [NB], and Random Forest [RF]). The significant stages of the proposed system design are shown in Fig. 2.

3.2. Dataset Description
The proposed model was trained as well as tested on a the ALL-IDB dataset containing images of leukemic blood smears along with images of non-leukemic blood smears. ALL-IDB [22] is a public image dataset which divided into two distinct versions: ALL-IDB1 and ALL-IDB2, and its images are in JPG format with a color depth of 24 bits. ALL-IDB was rated more accurate since professional oncologists classified/positioned ALL lymphoblasts for each image in the dataset. ALL-IDB1 is composed of 108 original RGB images taken with an optical laboratory microscope and either an Olympus Optical C2500L camera or a Canon PowerShot G5 camera. The first 33 images of the ALL-IDB1 are 1368 × 1712 pixels in size, while the rest of the images are 1944 × 2592 pixels in size; additionally, the images of these two image sets were collected under various magnifications, brightness, and hue staining. ALL-IDB1 depicts entire images comprising both cells or agglomerates; thus, it can be used to analyze the segmentation performance of techniques, image pre-processing methodologies, or classification approaches. ALL-IDB2 is a collection of cropped areas of interest derived from the ALL-IDB1 dataset of healthy and blast cells; furthermore, the gray level characteristics of ALL-IDB2 images are similar to those of ALL-IDB1 images. This dataset contains 260 microscopic images, 50% of these represent non-cancerous lymphocyte cells, and 50% cancerous (lymphoblast) cells. In this study, as a benchmark for evaluating the intelligence system's performance, the second version was considered.

In Fig. 3, representative example images from ALL-IDB2 dataset: Images of healthy lymphocyte and probable blast cells are given, respectively.

3.3. Preprocessing

Image preprocessing is a substantial step to improve the visualization of leukocytes in a blood image and accurate classification by removing noisy or unwanted pixels from each image. To eliminate superfluous text and machine annotation around images, the area of interest was localized (Fig. 4), cropped and resized images to a standard size 128 × 128 pixels after the raw images were converted from RGB to gray scale (Fig. 5). The area of interest on the microscopic blood images was determined by an area covering mostly WBC nuclei region to attain valuable information. As the raw images were captured in real life with a large variance

in exposure and contrast, image enhancement was required for superior classification performance. Consequently, the contrast enhancement of gray scale images was employed using Contrast Limited Adaptive Histogram Equalization technique and the median filter for a proper brightness and enhancement. Before importing the input images into a feature extraction stage, image adjustment was applied to improve the quality of microscopic blood smear images as shown in Fig. 6.

3.4. Feature Extraction

Feature extraction is the process of converting the raw pixel values from an image into a set of features; normally, this step comprises obtaining important features extracted from input patterns that can be used in the classification tasks [23]. In this study, two groups of features, including
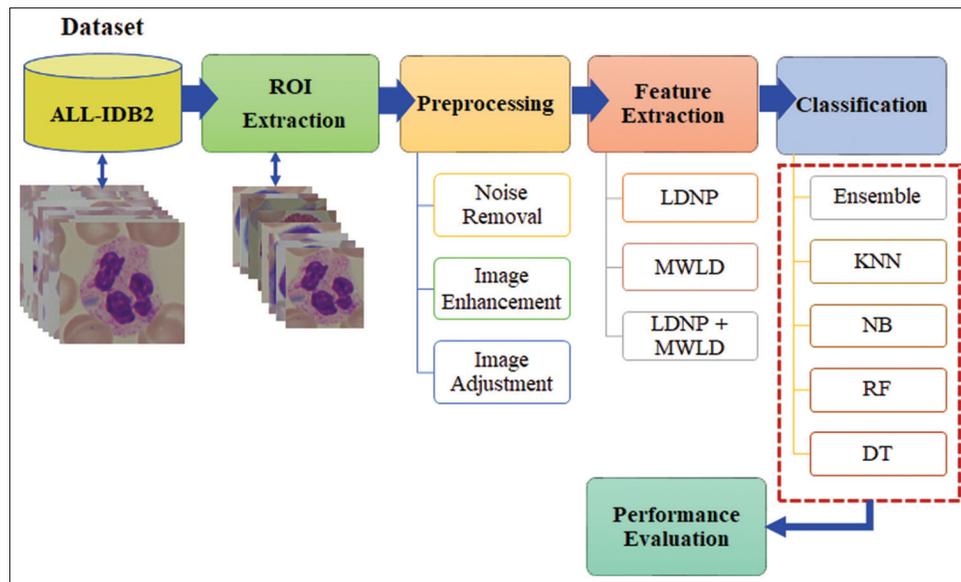

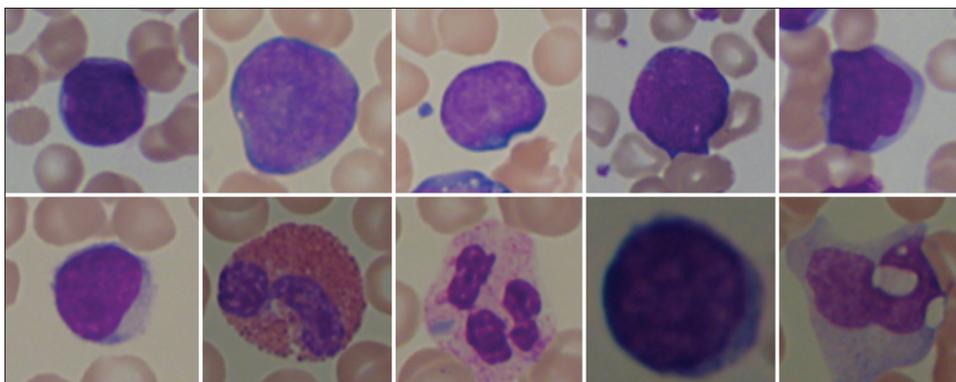**Fig. 2.** Workflow of proposed system.

Fig. 3. Example images contained in the ALL-IDB2: healthy lymphocyte (first row) and lymphoblast cells (second row).
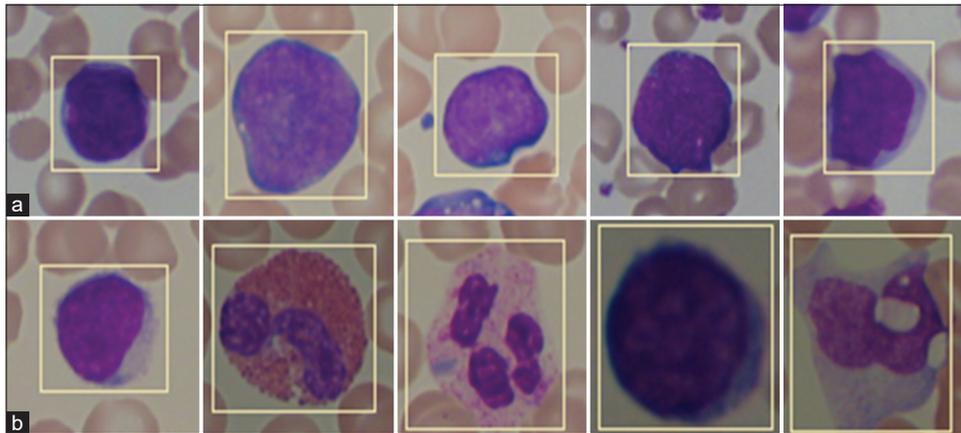
Fig. 4. Samples of blood smear images dataset used for proposed scheme; (a) localized ROI area of non-leukemic blood smears (first row), and (b) localized ROI area of leukemic blood smear cases (second row).
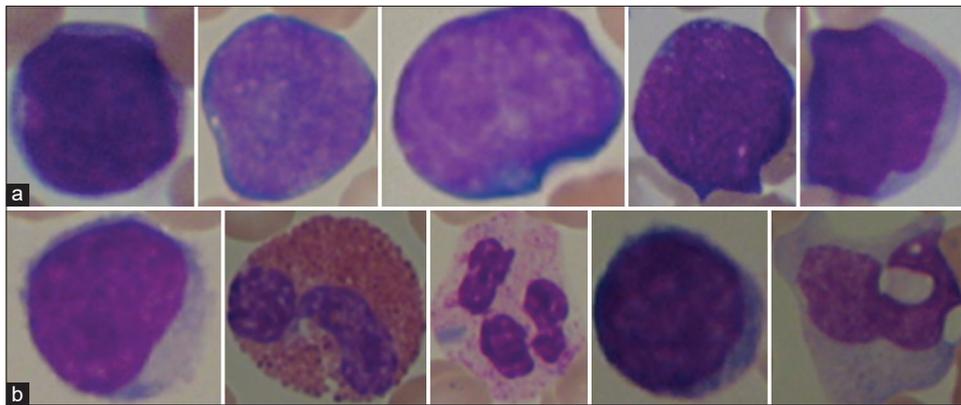

Fig. 5. Samples of blood smear images dataset used for proposed scheme; (a) cropped ROI area of healthy individuals (first row), and (b) cropped ROI area of leukemic blood smear cases (second row).
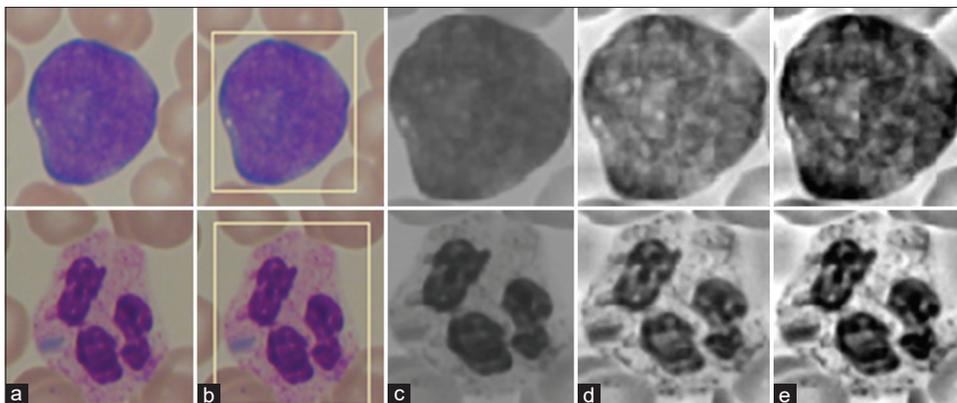

Fig. 6. Samples of blood smear images dataset used for proposed scheme; (a) original blood smear images, (b) localized ROI area, (c) denoised blood smear images, (d) contrast enhancement of blood smear images, and (e) adjusted blood smear images.

LDNP and MDWT are designed and proposed to distinguish noncancerous and cancerous blood cells from microscopic blood smear images.

3.4.1. LDNP

LDNP descriptor, introduced by Rivera *et al.* [24], is a six-bit binary code assigned to each pixel of an input image to

encode the structure information and intensity variations of a local texture pattern. The patterns are constructed using a compass mask to obtain the edge response value of the neighborhoods, as well as the direction of the top most positive and top most negative responses [25,26]. The positive and negative responses are critical in identifying the gradient direction of the neighborhood's prominent bright and dark areas. LDNP decreases the bit size per pixel from eight to a 6-bit binary code, where the first three bits represent the top positive directional number and the next three bits represent the top negative directional number. To generate the LDNP code, Kirsch compass mask (M_0, M_1., M_7) was experimented to calculate the edge response values and produce eight directional edge images or eight directional numbers for each pixel in microscopic blood smear images. Fig. 7 explains the working process of calculating bit for each pixel, and the Kirsch compass masks for eight directions [27] are revealed in Fig. 8.

To determine the meaningful descriptor for each region, LDNP is encoded using the maximum positive and maximum negative position values. The sign information is being used to determine the prominent values: The most significant bit is allocated to positive direction and the least significant bit is allocated to negative direction. Therefore, for the 3×3 neighborhood pixel window, the decimal LDNP coding is defined as:

$$LDN(x, y) = 8i_{(x, y)} + j_{(x, y)} \qquad (1)$$

Where (x, y) is the central pixel value of the region, $i_{(x,y)}$ is the direction index number of the maximum edge response, and $j_{(x,y)}$ is the top most negative direction. These directional numbers can be defined by equations 2 and 3, respectively.

$$i_{(x, y) = arg_i \max} \left\{ {}^i(x, y) \| 0 \le i \le 7 \right\} \qquad (2)$$

$$j_{(x, y) = arg_j \max} \left\{ {}^j(x, y) \| 0 \le j \le 7 \right\} \qquad (3)$$

Where is the convolution of the original image, I and the ith mask, M^i, characterized by:

$$^i = I * M^i \qquad (4)$$

In this work, LDNP descriptor was performed to extract a set of 56 features for each image was extracted from the dataset of microscopic blood images.

3.4.2. Multi-scale weber local descriptors
WLD is a powerful local descriptor that is based on Weber's Law and consists of two differential excitation and

orientation component [29,30]. WLD has a number of unique features, including edge detection and resistance to changes in illumination and noise. The differential excitation component is being used to extract the most important information from an image. For each pixel, the gradient orientation component of WLD is calculated. The differential excitation D (p_c) of a current pixel p_c is computed as:

$$D(p_c) = \arctan\left[\sum_{i=0}^{N-1} \left(\frac{p_i - p_c}{p_c} \right) \right] \qquad (5)$$

Where p_c represents the center pixel value, p_i denotes the value of ith pixel in the neighborhood, N is the total number of pixels in the neighborhood. If D (p_c) is positive, then center pixel is darker respect to the neighbor pixels and if D (p_c) is negative, then current pixel is lighter respect to the neighbor pixels. The orientation component of WLD is the gradient orientation that determines the directional property of the pixels. For pixel *pc*, it is calculated as follows: in Equation 6.

$$\varnothing(p_c) = \arctan\left(\frac{k_s^{11}}{k_s^{10}} \right) \qquad (6)$$

Where, K_s^{11} and K_s^{10} are the outputs of the filters $f11$ and $f10$ (Fig. 9), respectively.

After calculating differential excitation and gradient orientation, a concatenated two-dimensional WLD histogram using D and $\varnothing$ is created to produce the multi-WLD that forms the representation of the image. In this work, multi-WLD was computed to extract a set of 32 features for each image was extracted from the dataset of microscopic blood images.

3.5. Feature Fusion and Classification
Data fusion has been applied in several applications for machine learning and computer vision. Feature fusion strategy, in particular, can concatenate multiple feature vectors. The multi-feature fusion can improve the robustness of the model predictions [32]. This work proposed a fusion of feature vectors attained by a combination of LDNP (1×56) and MWLD (1×32) approaches. Equations 7 and 8 represent features extracted by LDNP and MWLD, respectively. The extracted feature vectors were combined by concatenation and represented by Equation 9.

$$F_{LDNP_{1 \times n}} = \left\{ LDNP_{1 \times 1} + LDNP_{1 \times 2} + LDNP_{1 \times 3} \ldots\ldots\ldots\ LDNP_{1 \times n} \right\} \qquad (7)$$

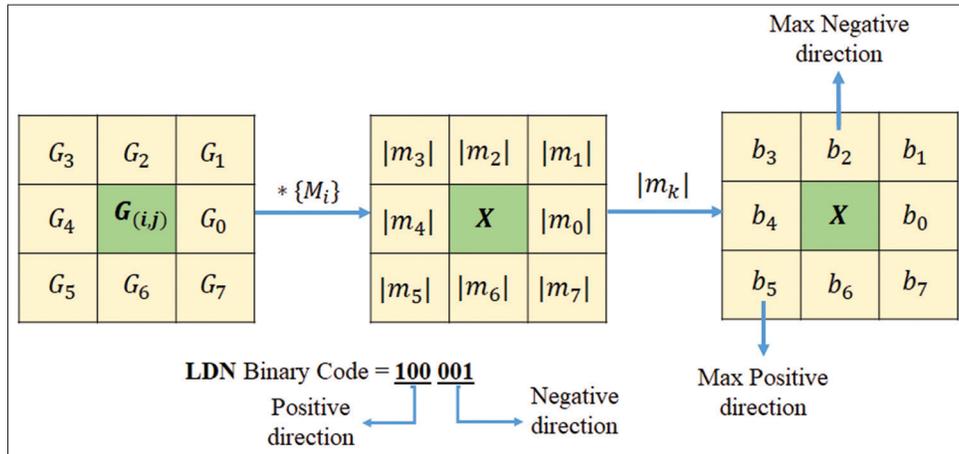$$F_{MWLD_{1 \times m}} = \left\{ MWLD_{1 \times 1} + MWLD_{1 \times 2} + MWLD_{1 \times 3} \ldots\ldots\ldots\ MWLD_{1 \times m} \right\} \qquad (8)$$
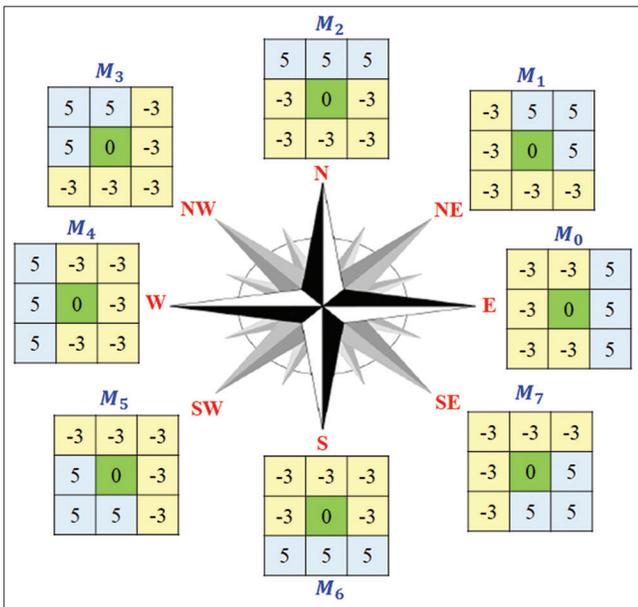
Fig. 7. The procedure of calculating LDNP code.

Fig. 8. Kirsch compass masks for eight directions [28].

Fig. 9. Filters used in simple WLD calculation [31].

$$Fused\ (Features\ vector)_{1 \times p}^{cat} = \{F_{\text{LDNP}_{1 \times n}}, F_{\text{MWLD}_{1 \times m}}\} \quad (9)$$

Then the features extracted by LDNP and MWLD were fused with 88 features. This fusion vector, which considered as the final input for the training and testing dataset, was fed to the classifiers to validate the proposed approach and identify blast cells from microscopic blood images. In the proposed workflow, machine learning models were used to identify patients affected by leukemia. To achieve the objective of identifying leukemia patients among normal healthy individuals,' five pattern recognition classifiers, namely, DT, Ensemble, KNN, NB, and RF classifiers were separately executed.

3.6. Model Evaluation

Model evaluation is used to estimate the parameter space and feature extraction results from various models. For the classification of ALL images, five widely used performance metrics were applied to assess the proposed model's performance: accuracy, recall, precision, F-measure, and MCC metrics. To calculate the metrics specified by Equations 10-14, four distinct performance parameters were used: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$Recall = \frac{TN}{TN + FP} \quad (11)$$

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

$$F - measure = \frac{2 * Precision * Recall}{Precision + Recall} \quad (13)$$

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \quad (14)$$

4. RESULTS AND DISCUSSION

The performance of the proposed system was evaluated based on extracted features derived from LDNP and

MWLD coefficients to identify automatically the class of the ALL images. The evaluation was performed using a dataset of 260 blood smear images, comprising 130 blasts and 130 non-blast cells. All experiments were conducted in MATLAB (R2021b) environment using microscopic blood images described in Section 3.2. Different scenarios using different features individually and combination of (LDNP + MWLD) features have been suggested. For each scenario, these extracted features were classified using five pattern recognition classifiers (i.e., DT, Ensemble, KNN, NB, and RF) to see which scenario could perform better classification performance. Furthermore, the entire dataset was divided into two groups: About 80% for training the model and 20% for evaluation of the classification performance using holdout cross-validation method. Performances of the suggested scenarios were analyzed through a number of different measures including accuracy, recall, precision, F-measure, MCC, and misclassification error rate computed from confusion matrix.

From the experiments performed on ALL-IDB2 dataset, the detailed category/class wise analysis of each scenario was estimated in terms of accuracy and average accuracy as (mean ± SD) with DT (Table 1), Ensemble (Table 2), KNN (Table 3), NB (Table 4), and RF (Table 5) classifiers respectively. According to the results in Tables 1-5, it is concluded that combining the LDNP and MWLD features

(scenario 3) together attain the highest average accuracy of 97.69 ± 1.83% followed by LDNP features (scenario 1) with 97.11 ± 1.44% along with Ensemble classifier, while the classification average accuracy of features derived from MWDT method had the lowest scoring (79.42 ± 4.8%) with NB classifier.

Based on the experimental outcomes for all three scenarios depicted in Fig. 10, it can be concluded that the fusion of features extracted from LDNP and MWLD methods has a positive impact on the performance and outperformed the other two scenarios with all classifiers. The consequences revealed that the fusion LDNP and MWLD techniques attained the highest average accuracy of 97.69 ± 1.83%, 96.92 ± 1.15%, 95.19 ± 1.63%, 93.84 ± 3.11%, and 89.03 ±

TABLE 1: Performance analysis on average accuracy with Decision Tree classifier

Descriptors	Features	Per Class Accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	82.3±8.54	83.46±7.03	82.88±5.62
MWDT	32	89.61±6.02	85.38±8.06	87.50±5.96
LDNP+ MWDT	88	91.92±5.57	86.15±6.58	**89.03±3.63**

The highlighted accuracy in bold indicates the best classification result. LDNP: Local directional number pattern

TABLE 2: Performance analysis on average accuracy with Ensemble classifier

Descriptors	Features	Per Class Accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	97.3±1.07	96.92±1.03	97.11±1.44
MWDT	32	94.99±2.07	95.76±2.13	95.38±2.59
LDNP+ MWDT	88	99.23±0.02	96.15±1.73	**97.69±1.83**

The highlighted accuracy in bold indicates the best classification result. LDNP: Local directional number pattern

TABLE 3: Performance analysis on average accuracy with KNN classifier

Descriptors	Features	Per class accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	99.23±0.12	92.69±3.27	95.96±1.78
MWDT	32	96.15±1.14	84.99±4.39	90.57±3.89
LDNP+ MWDT	88	96.15±1.24	97.69±1.34	96.92±1.15

The highlighted accuracy in bold indicates the best classification result. KNN: K-Nearest Neighbors, LDNP: Local directional number pattern

TABLE 4: Performance analysis on average accuracy with NB classifier

Descriptors	Features	Per class accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	91.15±3.45	93.07±4.22	92.11±4.29
MWDT	32	75.38±7.51	83.46±5.74	79.42±4.80
LDNP+ MWDT	88	94.23±2.88	93.46±4.02	93.84±3.11

The highlighted accuracy in bold indicates the best classification result. NB: Naïve Bayes, LDNP: Local directional number pattern

TABLE 5: Performance analysis on average accuracy with RF classifier

Descriptors	Features	Per Class Accuracy (%)		Average accuracy (%)
		Normal cells	Abnormal cells	
LDNP	56	96.92±1.05	90.38±6.34	93.65±3.39
MWDT	32	93.84±3.19	92.69±2.23	93.26±2.60
LDNP+ MWDT	88	97.69±1.68	92.69±2.36	95.19±1.63

The highlighted accuracy in bold indicates the best classification result. RF: Random forest, LDNP: Local directional number pattern

3.63% using Ensemble, KNN, RF, NB, and DT classifiers, respectively. On the other hand, the fusion of fusion LDNP and MWLD techniques with Ensemble classifier was sufficient to record maximal average accuracy performance of 97.69 ± 1.83% among the remaining classifiers for all scenarios.

The same fact has been concluded by examining other performance measures (precision, recall, F-measure, and MCC) to estimate the proposed framework. For all classifiers, the best precision rate was achieved with a set of features fusion LDNP and MWLD methods (scenario 3) and outperformed the other scenarios. The results of scenario 3 provided that all five classifiers reached the highest precision of 98.05 ± 1.74%, 97.77 ± 1.09%, 93.85 ± 5.21%, 93.15 ± 2.90%, and 87.25 ± 5.33% using Ensemble, KNN, NB, RF, and DT classifiers, respectively; however, the lowest precision rate of 82.14 ± 5.39% was recorded using MWDT method with NB classifier. Furthermore, the experiment verified maximum precision performance of 98.05 ± 1.74% with Ensemble classifier among the remaining classifiers for all scenarios. Comparison results of precision rates of all system scenarios with all classifiers are displayed in Table 6.

Furthermore, the fusion of features from LDNP and MWDT methods also performed the best in terms of recall rates as 97.69 ± 1.68%, 97.30 ± 1.14%, 96.15 ± 1.14%, and 94.23 ± 1.88% was attained using RF, Ensemble, KNN, and NB classifiers, respectively (Table 7); furthermore,

Fig. 10. Comparison of average accuracies for modeling the various scenarios using different classification methods.

MWDT performed the best recall rate of 93.46 ± 2.81% with DT classifier; conversely, the lowest recall rate of 75.38 ± 7.51% has been recorded when MWDT method was utilized with NB classifier. Based on the experimental results for all five classifiers depicted in Table 7, it can be verified that the value of recall rate with RF classifier was overall superior and outperformed the other classifiers. With regard to the F-measure rates, the results presented in Table 8 demonstrate the superiority of the fusion of LDNP and MWDT scenario and it was undoubtedly yielded excellent results which defiantly go beyond the other scenarios. The best performance with F-measure rate of 97.63 ± 1.62% was achieved using fusion of LDNP and MWDT features with Ensemble classifier, which surpassed other classifiers with F-measure rates of 96.89 ± 1.86%, 95.31 ± 1.56%, and 93.88 ± 3.11% for KNN, RF, and NB classifiers, respectively; however, the F-measure rate of MWDT method using NB classifier had the lowermost recording of 78.45 ± 5.43%.

With respects to the MCC rates, the outcomes depicted in Table 9 prove the superiority of the fusion of LDNP and MWDT scenario and it was unquestionably returned outstanding outcomes which defiantly go beyond the other scenarios. The finest performance with F-measure rate of 95.46 ± 2.85% was achieved using fusion of LDNP and MWDT features with Ensemble classifier, followed by F-measure rate of 93.97 ± 3.67% along with KNN classifier, while the classification F-measure rate of features derived from MWDT method had the lowest scoring (59.26 ± 9.59%) with NB classifier.

The experimentations from Fig. 11 besides obviously confirmed that the features derived by fusion of LDNP and MWDT scenario outperformed other scenarios and recorded the highest precision, recall, F-measure, and MCC rates with Ensemble classifier. Considering the obtained outcomes, the highest precision, recall, F-measure, and MCC scores of the features extracted using fused LDNP and MWDT methods were 98.05 %, 97.30%, 97.63%, and 95.46% respectively, and were achieved using 88 effective features. While, the lowest precision, recall, F-measure, and MCC rates were achieved
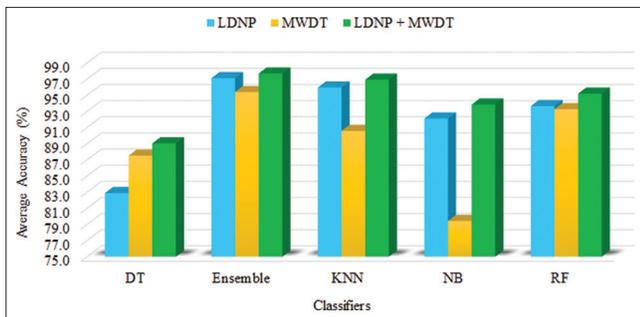
TABLE 6: The average classification precision in % over 10 runs for each scenario scenarios using different classification algorithms

Methods	DT	Ensemble	KNN	NB	RF
LDNP	83.51±5.90	97.01±1.89	93.33±4.41	93.25±5.96	91.24±5.02
MWDT	87.03±6.30	96.24±1.99	86.30±5.93	82.14±5.39	92.96±3.71
LDNP+MWDT	**87.25±5.33**	**98.05±1.74**	**97.77±1.09**	**93.85±5.21**	**93.15±2.90**

The best result per row is highlighted in bold. DT: Decision tree, KNN: K-Nearest Neighbors, NB: Naïve Bayes, RF: Random forest, LDNP: Local directional number pattern

TABLE 7: The average classification recall in % over 10 runs for each scenario scenarios using different classification algorithms

Methods	DT	Ensemble	KNN	NB	RF
LDNP	82.30±8.54	97.20±1.07	96.13±1.02	91.15±5.45	96.92±2.03
MWDT	**93.46±2.81**	96.53±2.36	96.15±1.44	75.38±7.51	93.84±3.19
LDNP+MWDT	91.92±5.57	**97.30±1.14**	**96.15±1.14**	**94.23±1.88**	**97.69±1.68**

The best result per row is highlighted in bold. DT: Decision tree, KNN: K-Nearest Neighbors, NB: Naïve Bayes, RF: Random forest, LDNP: Local directional number pattern

TABLE 8: The average classification MCC in % over 10 runs for each scenario scenarios using different classification algorithms

Methods	DT	Ensemble	KNN	NB	RF
LDNP	66.11±11.17	94.34±4.78	92.22±5.21	84.49±8.61	87.67±6.37
MWDT	**79.62±10.19**	92.78±4.57	81.27±7.68	59.26±9.59	86.78±5.17
LDNP+MWDT	78.55±7.27	**95.46±2.85**	**93.97±3.67**	**88.00±6.04**	**90.60±3.18**

The best result per row is highlighted in bold. DT: Decision tree, KNN: K-Nearest Neighbors, NB: Naïve Bayes, RF: Random forest, LDNP: Local directional number pattern

TABLE 9: The average classification MCC in % over 10 runs for each scenario scenarios using different classification algorithms

Methods	DT	Ensemble	KNN	NB	RF
LDNP	66.11±11.17	94.34±4.78	92.22±5.21	84.49±8.61	87.67±6.37
MWDT	**79.62±10.19**	92.78±4.57	81.27±7.68	59.26±9.59	86.78±5.17
LDNP+MWDT	78.55±7.27	**95.46±2.85**	**93.97±3.67**	**88.00±6.04**	**90.60±3.18**

The best result per row is highlighted in bold. DT: Decision tree, KNN: K-Nearest Neighbors, NB: Naïve Bayes, RF: Random forest, LDNP: Local directional number pattern
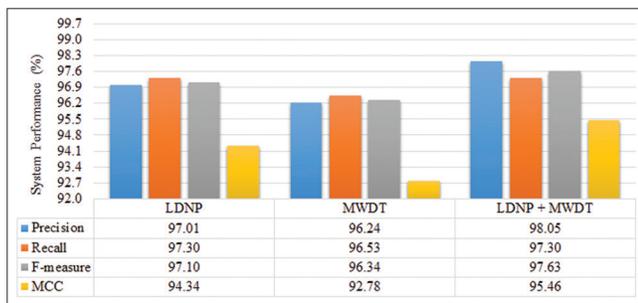

	LDNP	MWDT	LDNP + MWDT
Precision	97.01	96.24	98.05
Recall	97.30	96.53	97.30
F-measure	97.10	96.34	97.63
MCC	94.34	92.78	95.46

Fig. 11. Comparison of system performance for a different scenario using Ensemble classifier.

using MWDT method with a score of 96.24 %, 96.53%, 96.34%, and 92.78%, respectively, and were achieved with 32 extracted features.

In this work, the performance of the proposed scenarios was also evaluated through misclassification error rate metric used the same dataset and computing environment. As confirmed by Fig. 12, the misclassification error rates for the proposed scenarios were measured. The results prove that the fusion of LDNP and MWDT scenario with Ensemble

classifier outcomes in a lower misclassification error of 2.31% rate which confirms that the proposed scenario performs considerably much better than other suggesting scenarios. Thus, this scenario was chosen as a proposed technique for the classification of ALL images.

Finally, the performance of the proposed fusion scheme was also compared with some existing state-of-the-art methods as revealed in Table 10. The proposed scheme provides a promising outcome particularly in terms of average classification accuracy when comparing with the existing approaches. This is due to the combination carried out between LDNP and MWDT approaches which led to gaining their advantages. Nevertheless, the other researchers used some huge number of features, whereas in the proposed scheme, 88 features were utilized with the best performance results attained.

From the above experimental outcomes, it is obviously noticeable that the proposed system can effectively applied discriminates the normal cell cases from blasts in microscopic blood images more precisely, which might support the clinicians to made up a clear diagnosis based on their clinical specialists as well as the proposed tool as a second opinion.

TABLE 10: Comparison of proposed classification accuracy with recent techniques

Author	ALL-IDB2 datasets		Accuracy (%)
	Method	Classifier	
Umamaheswari and Geetha [17]	GLCM+Statistical+Geometrical Features	C-KNN	96.25
Tuba and Tuba [18]	Shape+Texture features	GAO-based methods	93.84
Singhal and Singh [15]	LBP	-------	93.84
	GLCM		87.30
Praveena and Singh [16]	GreyJOA	Deep CNN	93.50
Proposed work	LDNP+MWDT	Ensemble	**97.69**

LBP: Local binary pattern, GLCM: Gray level co-occurrence matrices, C-KNN: C-K-Nearest Neighbors, GAO: Generative adversarial optimization, LDNP: Local directional number pattern, CNN: Convolutional neural network, GreyJOA: Grey wolf-based Jaya optimization algorithm. The best result per row is highlighted in bold.
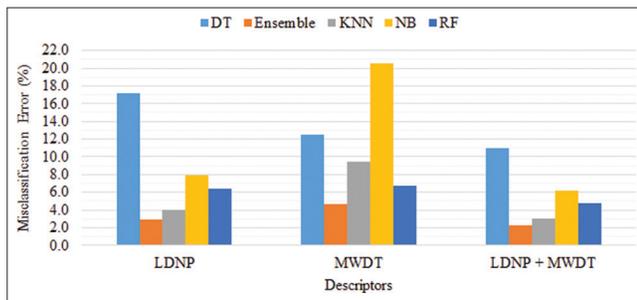

Fig. 12. Comparison of misclassification error rate for a different scenario.

5. CONCLUSION

Early diagnosis of leukemia especially acute leukemia in patients, gives the chance to cure cancer with the right treatment. Image processing-based technologies can be used to automatically and effectively analyze microscopic smear images in order to detect the incidence of leukemia. The main objective of this paper is to use feature fusion and a machine learning model to perform leukemia classification using input blood smear images. Each trained model was evaluated using benchmark performance metrics, for example, accuracy, precision, recall, F-measure, MCC, and misclassification error rate under three different scenarios concerned with balanced learning and classification method. The proposed method was tested on images of microscopic thin blood smears obtained from the publicly available leukemia benchmark dataset, that is, ALL-IDB2. The proposed feature fusion pipeline presented a higher average classification accuracy ($97.69 \pm 1.83\%$) compared to the accuracies attained using features achieved by individual feature extraction approaches, such as LDNP and MWDT. Furthermore, experimental outcomes revealed that the proposed model is more effective than previous works carried out for the classification of ALL.

REFERENCES

[1] C. Di Ruberto, A. Loddo and G. Puglisi G. "Blob detection and deep learning for leukemic blood image analysis". *Applied Sciences*, vol. 10, no. 3, p. 1176, 2020.

[2] G. Drałus, D. Mazur and A. Czmil. "Automatic detection and counting of blood cells in smear images using retinanet". *Entropy*, vol. 23, no. 11, p. 1522, 2021.

[3] B. George-Gay and K. Parker. "Understanding the complete blood count with differential". *Journal of PeriAnesthesia Nursing*, vol. 18, pp. 96-117, 2003.

[4] S. Shafique and S. Tehsin. "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks". *Technology in Cancer Research and Treatment*, vol. 17, pp. 1-7, 2018.

[5] D. A. Arber, A. Orazzi, R. Hasserjian, J. Thiele, M. J. Borowitz, M. Le Beau, C. D. Bloomfield, M. Cazzola and J. W. Vardiman. "The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia". *Blood*, vol. 127, pp. 2391-2405, 2016.

[6] F. Huang, P. Guang, F. Li, X. Liu, W. Zhang and W. Huang. "AML, ALL, and CML classification and diagnosis based on bone marrow cell morphology combined with convolutional neural network: A STARD compliant diagnosis research". *Medicine* (*Baltimore*), vol. 99, p. e23154, 2020.

[7] Y. Dong, O. Shi, Q. Zeng, X. Lu, W. Wang, Y. Li and Q. Wang. "Leukemia incidence trends at the global, regional, and national level between 1990 and 2017. *Experimental Hematology and Oncology*, vol. 9, pp. 1-11, 2020.

[8] M. Ghaderzadeh, F. Asadi, A. Hosseini, D. Bashash, H. Abolghasemi and A. Roshanpour. "Machine learning in detection and classification of leukemia using smear blood images: A systematic review". *Scientific Programming*, vol. 2021, p. 9933481, 2021.

[9] M. Kim, K. Chae, S. Lee, H. J. Jang and S. Kim. "Automated classification of online sources for infectious disease occurrences using machine-learning-based natural language processing approaches". *International Journal of Environmental Research and Public Health*, vol. 17, pp. 1-13, 2020.

[10] F. E. Al-Tahhan, M. E. Fares, A. A. Sakr and D. A. Aladle. "Accurate automatic detection of acute lymphatic leukemia using a refined simple classification". *Microscopy Research and Technique*, vol. 83, pp. 1178-1189, 2020.

[11] M. Sharif, J. Amin, A. Siddiqa, H. U. Khan, M. S. A. Malik, M. A. Anjum and S. Kadry. "Recognition of different types of leukocytes using YOLoV2 and optimized bag-of-features". *IEEE Access*, vol. 8, pp. 167448-167459, 2020.

[12] K. K. Jha and H. S. Dutta. "Nucleus and cytoplasm–based segmentation and actor-critic neural network for acute lymphocytic leukaemia detection in single cell blood smear images". *Medical and Biological Engineering and Computing*, vol. 58, pp. 171-186, 2020.

[13] L. H. S. Vogado, R. M. S. Veras, F. H. D. Araujo, R. R. V. Silva and K. R. T. Aires. "Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification". *Engineering Applications of Artificial Intelligence*, vol. 72, pp. 415-422, 2018.

[14] K. Muthumayil, S. Manikandan, S. Srinivasan, J. Escorcia-Gutierrez, M. Gamarra and R. F. Mansour. "Diagnosis of leukemia disease based on enhanced virtual neural network". *Computers, Materials and Continua*, vol. 69, pp. 2031-2044, 2021.

[15] V. Singhal and P. Singh. "Texture Features for the Detection of Acute Lymphoblastic Leukemia." 2016.

[16] S. Praveena and S. P. Singh. "Sparse-FCM and deep convolutional neural network for the segmentation and classification of acute lymphoblastic leukaemia". *Biomed Technologies*, vol. 65, pp. 759-773, 2020.

[17] D. Umamaheswari and S. Geetha. "A framework for efficient recognition and classification of acute lymphoblastic leukemia with a novel customized-KNN classifier". *The Journal of Computer Engineering and Information Technology*, vol. 26, pp. 131-140, 2018.

[18] M. Tuba and E. Tuba. "Generative adversarial optimization (GOA) for acute lymphocytic leukemia detection". *Studies in Informatics and Control*, vol. 28, pp. 245-254, 2019.

[19] S. Hariprasath, T. Dharani, S. Mohammad and N. Bilal. "Automated Detection of Acute Lymphocytic Leukemia Using Blast Cell Morphological Features". 2nd International Conference on Advances in Science and Technology (ICAST) 2019 on 8th, 9th April 2019 by K J Somaiya Institute of Engineering and Information Technology, Mumbai, India, 2019.

[20] A. Bodzas, P. Kodytek and J. Zidek. "Automated detection of acute lymphoblastic leukemia from microscopic images based on human visual perception". *Frontiers in Bioengineering and Biotechnology*, vol. 8, pp. 1-13, 2020.

[21] K. N. Sukhia, M. M. Riaz, A. Ghafoor and N. Iltaf. "Overlapping white blood cells detection based on watershed transform and circle fitting". *Radioengineering*, vol. 26, pp. 1177-1181, 2017.

[22] Conference II, Processing I. *ALL-IDB : The Acute Lymphoblastic Leukemia Image Database for Image Processing Ruggero Donida Labati*. Vincenzo Piuri, Fabio Scotti Università degli Studi di Milano, Department of Information Technology, IEEE International Conference on Image Processing, pp. 2089-2092, 2011.

[23] F. H. Ahmad and S. H. Wady. "COVID-19 infection detection from chest X-ray images using feature fusion and machine learning". *Scientific Journal*, Vol. 5, pp. 10-30, 2021.

[24] A. R. Rivera, J. R. Castillo and O. Chae. "Local directional texture pattern image descriptor". *Pattern Recognition Letters*, vol. 51, pp. 94-100, 2015.

[25] R. R. Rose, A. Suruliandi and K. Meena. "Local texture description framework-based modified local directional number pattern: A new descriptor for face recognition". *International Journal of Biometeorology*, vol. 7, pp. 147-169, 2015.

[26] S. P. Ramalingam. "Paturu Venkata Subbu Sita Rama CM. Dimensionality reduced local directional number pattern for face recognition". *Journal of Ambient Intelligence and Humanized Computing*, vol. 9, pp. 95-103, 2018.

[27] R. A. Kirsch. "Computer determination of the constituent structure of biological images". *Computers and Biomedical Research*, vol. 4, pp. 315-328, 1971.

[28] A. El Idrissi, Y. El Merabet and Y. Ruichek. "Palmprint recognition using state-of-the-art local texture descriptors: A comparative study". *IET Biometrics*, vol. 9, pp. 143-153, 2020.

[29] Sachinkumar and S. Raga. "Breast cancer detection technique based on multi-subspace randomization and collaboration". *International Journal of Advanced Research in Engineering and Technology*, vol. 11, pp. 687-701, 2020.

[30] A. Banerjee, N. Das and K. C. Santosh. "Weber local descriptor for image analysis and recognition: A survey". *The Visual Computer*, vol. 38, pp. 321-243, 2020.

[31] M. Hussain, S. Qasem, G. Bebis, G. Muhammad, H. Aboalsamh and H. Mathkour. "Evaluation of image forgery detection using multi-scale weber local descriptors". *International Journal on Artificial Intelligence Tools*, vol. 24, pp. 416-424, 2015.

[32] S. Ahmed. "Ethnicity identification based on fusion strategy of local and global features extraction". *International Journal of Multidisciplinary and Current Research*, vol. 4, pp. 200-205, 2016.