Contents lists available at GrowingScience

International Journal of Data and Network Science

homepage: www.GrowingScience.com/ijds

Revolutionizing classification: A novel gray level co-occurrence matrix and statistical feature-based segmentation approach

Abdelwahed Motwakel[a*]

[a]*Department of management information systems, College of Business Administration in Hawtat Bani Tamim, Prince Sattam bin Abdulaziz University, Al-Kharj, Saudi Arabia*

C H R O N I C L E

Article history:
Received: July 11, 2024
Received in revised format: August 2, 2024
Accepted: August 18, 2024
Available online: August 18, 2024

Keywords:
Acute Lymphoblastic Leukemia
Gray Level Co-occurrence Matrix
Euclidean distance
Manhattan distance
Canberra distance
Chebyshev distance

A B S T R A C T

The accurate and efficient classification of leukemia images is crucial for early diagnosis and effective treatment planning. Traditional methods often face challenges in handling the complexity and variability of medical images. To address these challenges, we propose a novel approach that leverages the Gray Level Co-occurrence Matrix (GLCM) and statistical feature-based segmentation techniques. in this paper, we present a comprehensive framework for the automated classification of leukemia images using advanced image processing techniques. The methodology involves six key stages: input of leukemia images, preprocessing to enhance image quality, segmentation to isolate relevant features, feature extraction using texture analysis, classification using multiple distance metrics Euclidean, Manhattan, Canberra, and Chebyshev, and performance evaluation. Our results demonstrate significant improvements in classification accuracy, sensitivity, specificity, and error rates across various metrics and feature sets. For instance, using the Chebyshev distance, we achieved an average accuracy of 82.69%, sensitivity of 85.95%, and specificity of 82.77%. The Canberra distance provided optimal performance with 65 features, yielding an accuracy of 85.18%, sensitivity of 86.39%, and specificity of 86.31%. These findings underscore the efficacy of our approach in distinguishing between healthy and leukemic cells, thereby contributing to early diagnosis and effective treatment planning for leukemia.

1. Introduction

Leukemia is a type of cancer that affects the blood and bone marrow, characterized by an overproduction of abnormal white blood cells. It is a complex disease with various subtypes, each presenting unique challenges in diagnosis and treatment. Among these subtypes, Acute Lymphoblastic Leukemia (ALL) is particularly significant due to its prevalence, especially in children, and its rapid progression. Numerous researchers have conducted studies to discover and classify leukemia images (Terwilliger & Abdul-Hay, 2017). Muntasa et al. (2022, 2023) proposed a commutative convolutional neural network model for leukemia image classification using hypercomplex modeling A[+1, -1] and A[-1, +1]. They augmented the image datasets through rotation, zooming, and flipping, and evaluated their method using the ALL-IDB2 database. The proposed model achieved average accuracies of 96.43% for A[+1, -1] and 97.05% for A[-1, +1], with maximum accuracies reaching 100% for both. This method outperformed several other classifiers, including k-nearest neighbor, various support vector machines, naïve bayes, decision tree, and color hybrid modeling.

* Corresponding author.
E-mail address am.ismaeil@psau.edu.sa (A. Motwakel)

ISSN 2561-8156 (Online) - ISSN 2561-8148 (Print)

doi: 10.5267/j.ijdns.2024.8.017

Muntasa and Yusuf (2021) proposed a method to classify Acute Lymphoblastic Leukemia (ALL) using multi-distance models of the Gray Level Co-occurrence Matrix (GLCM), extracting 192 features from image enhancement results across three channels. The features were classified using Canberra and Chebyshev techniques, achieving 96.97% accuracy with 2.27% false positives and 0.75% false negatives. Our method was evaluated using the ALL Image database and compared to various approaches like SVM, Naïve Bayes, and Fuzzy-based detection, outperforming them all.

Muntasa and Yusuf (2019) proposed a model to detect Acute Lymphoblastic Leukemia (ALL) using principal features of the object, particularly through Color Orthonormal Basis Entropy (COBE) and Distribution of the Pixel Intensity (DoPI). Key innovations include combining three channel features for improved accuracy and developing a contrast enhancement filter to automatically recognize and remove noise. The approach involves four main stages: image enhancement, segmentation using the Otsu algorithm, feature extraction using COBE and DoPI, and performance measurement with the Manhattan method. Evaluated using the 2nd ALL-IDB, the method achieved a high accuracy of 91.67% and a small standard deviation of 0.022. Muntasa et al. (2020) proposed a novel convolutional neural network (CNN) architecture with pyramid-like kernel increases. The architecture utilizes the final convolution for fully connected layers, followed by the SoftMax function for image classification. Evaluated on the ALL-IDB2 database, the model achieved an accuracy of 99.17%, precision of 99.33%, and recall of 99%. This approach outperformed other models, including shape features, GLCM, CNN and SVM, AlexNet, and hypercomplex-valued CNNs.

Alagu and Bagan (2019) proposed a method to reduce diagnostic time and improve accuracy in classifying leukemia from microscopic blood smear images. The images were preprocessed and segmented into three clusters using the k-means clustering algorithm based on shape, color, and texture. Texture features were extracted using the grey level co-occurrence matrix (GLCM) and local binary pattern (LBP), and classification was performed using a support vector machine (SVM) with a Gaussian radial basis function (RBF) kernel. Testing on 367 images from the ALL-IDB database achieved accuracies of 90.5% for ALL-IDB1 and 95.3% for ALL-IDB2, outperforming other classifiers like Linear Discriminant (LD), Ensemble (Bagged trees), and KNN.

Sukhia et al. (2019) proposed a scheme for classifying Acute Lymphoblastic Leukemia (ALL) based on several key steps: preprocessing and segmenting white blood cell nuclei with the expectation maximization algorithm, extracting features, selecting features using principal component analysis, and classifying using sparse representation. This approach demonstrated significantly higher accuracy compared to existing methods for ALL classification. Muntasa and Yusuf (2019) proposed model extracted seven features—Energy (EN), Entropy (EP), Shannon Entropy, Log Energy Entropy (EE), Mean (ME), Variance (VA), and Correlation (CO) were proposed for object characterization. These features were measured using Euclidean Distance, Manhattan, Canberra, and Chebyshev methods, achieving maximum accuracies of 81.54%, 81.54%, 76.92%, and 82.31%, respectively. The method's accuracy was further validated using a confusion matrix. The proposed approach was evaluated on the Acute Lymphoblastic Leukemia-Image Database (ALL-IDB).

Shafique and Tehsin (2018) used CNN to utilize a pretrained AlexNet for automated detection and classification of acute lymphoblastic leukemia (ALL) and its subtypes L1, L2, L3, and Normal—previously overlooked in literature. Unlike traditional training methods, we fine-tuned the network on our dataset, replacing the last layers for classification into four classes. Data augmentation techniques were employed to prevent overtraining, and various color models were compared for performance. Our approach achieved notable results: 100% sensitivity, 98.11% specificity, and 99.50% accuracy for ALL detection, and 96.74% sensitivity, 99.03% specificity, and 96.06% accuracy for subtype classification, surpassing standard methods without requiring microscopic image segmentation. Talaat and Gamel (2024) proposed a model for blood microscopic images that distinguishes between leukemia-free and leukemia-affected samples. The method involves three primary steps: Image_Preprocessing, Feature Extraction, and Classification, employing an optimized CNN (OCNN). Fuzzy logic is utilized to optimize CNN hyperparameters, significantly enhancing classification performance. Results demonstrate that OCNN achieves exceptional accuracy of 99.99% on the C-NMC_Leukemia dataset, underscoring its effectiveness in improving CNN performance through fuzzy optimization. Pranav, and Rekha Sugandhi's proposed model involves extensive pre-processing to isolate WBCs using morphological techniques. Textural, geometrical, and statistical properties are then extracted from segmented regions. Four machine learning techniques—random forest (RF), support vector machine (SVM), naive Bayes classifier (NB), and K nearest neighbor (KNN)—are evaluated. SVM emerges as highly effective in identifying leukemia-inducing cells, outperforming other classifiers due to its robust performance across diverse blood smear images. The study proposes EMC-SVM as a viable method for accurate classification of leukocytes, demonstrating successful differentiation and categorization within sample images (More & Sugandhi, 2023).

2. The proposed method

The methodology for acute lymphoblastic leukemia (ALL) image classification involves six stages: inputting leukemia images, preprocessing to enhance quality, segmenting white blood cells, extracting textural features, classifying, and evaluating performance with metrics.

3. ALL-IDB2 as experimental datasets

We use the dataset of acute lymphoblastic Leukemia image database type 2 (ALL-IDB2) to evaluate our proposed method. It consists of 260 images, of which 130 images are categorized as healthy images, and the remaining are Leukemia images (Labati et al., 2011) . We obtain the ALL-IDB2 datasets from the department of computer science - università degli studi di Milano (in acronym "UniMi"). Expert oncologists have provided the classification and location of ALL lymphoblasts for every image in the dataset, as shown in the image samples in Figure. 4. The ALL-IDB2 image datasets are crop results from the ALL-IDB1 datasets with a 2592 × 1944 resolution size of 24-bit color depth.

4. Proposed Model Flowchart

Fig. 1 represents the research methodology for this study, which encompasses several critical stages aimed at the effective classification of leukemia images. Initially, the input consists of leukemia images, which serve as the primary data for analysis.

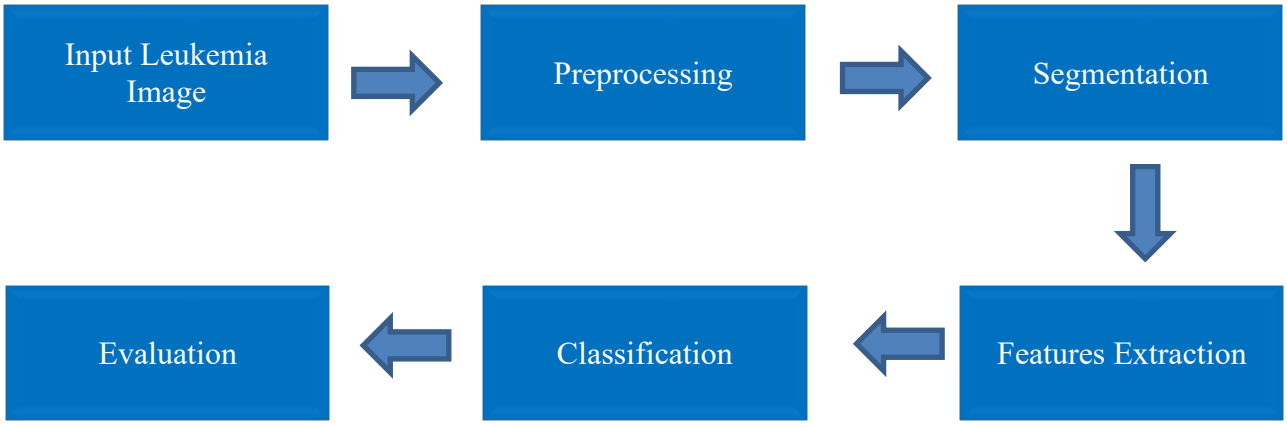

Fig. 1. Proposed Model flowchart

4.1 Preprocessing

The Preprocessing stage is carried out to prepare the images for feature extraction and is crucial in ensuring the integrity and quality of the data used in subsequent analysis. This stage involves sophisticated techniques designed to enhance the dataset images and make it suitable for accurate feature extraction. normalization is performed to scale the pixel values to a standard range, ensuring uniformity across all dataset images. The preprocessing stage plays a pivotal role in enhancing the reliability and accuracy of the feature extraction process, setting a solid foundation for the entire analysis workflow.

4.2 Segmentation

The segmentation stage is a critical process. This process involves several key steps, each contributing to the accurate extraction and identification of features within the image, as shown in Fig. 2. First, the input image is decomposed into its three primary color channels: green, red, and blue. A color image (Fr,g,b) can be represented by its red (Fr), green (Fg), and blue (Fb) channels, as shown in Eq. (1).

$$F_{r,g,b} = (F_r, F_g, F_b) \tag{1}$$

Eq. (1) can be broken down into three separate matrix equations, as demonstrated in Eqs. (2-4).

$$F_r = \begin{pmatrix} F_r(1,1) & \cdots & F_r(1,n) \\ \vdots & \ddots & \vdots \\ F_r(m,1) & \cdots & F_r(m,n) \end{pmatrix} \tag{2}$$

$$F_g = \begin{pmatrix} F_g(1,1) & \cdots & F_g(1,n) \\ \vdots & \ddots & \vdots \\ F_g(m,1) & \cdots & F_g(m,n) \end{pmatrix} \tag{3}$$

$$F_b = \begin{pmatrix} F_b(1,1) & \cdots & F_b(1,n) \\ \vdots & \ddots & \vdots \\ F_b(m,1) & \cdots & F_b(m,n) \end{pmatrix} \tag{4}$$

Moreover, each image channel is enhanced to compute the histogram and automatically adjust the limit values. In this instance, we used the standard gamma correction with $\boldsymbol{\gamma}$=1. If the enhanced color image is represented by the variable G, then the improved result matrix is a combination of Gr, Gg, and Gb, as described in Eq. (5).

$$G_{r,g,b} = (G_r, G_{g,} G_b) \tag{5}$$

Additionally, Eq. (5) is transformed into the HSV (Hue, Saturation, and Value) model, with the hue channel being selected for further processing.

$$Hue = \begin{cases} 60\,x\left(4+\frac{Red-Blue}{Saturation\,x\,Value}\right) if\ Value=Blue \\ 60\,x\left(2+\frac{Blue-Red}{Saturation\,x\,Value}\right) if\ Value=Green \\ 60\,x\left(\frac{Green-Blue}{Saturation\,x\,Value}\right) if\ Value=Red \\ 0 \qquad if\ Saturation=0 \end{cases} \tag{6}$$

$$Saturation = \begin{cases} 0 & if\ Value=0 \\ 1-\dfrac{\min(Red, Green, Blue)}{value} & if\ Value>0 \end{cases} \tag{7}$$

$$Value = \max(Red, Green, Blue) \tag{8}$$

Fig 3. shows the implementation of the Eqs. (1-8).

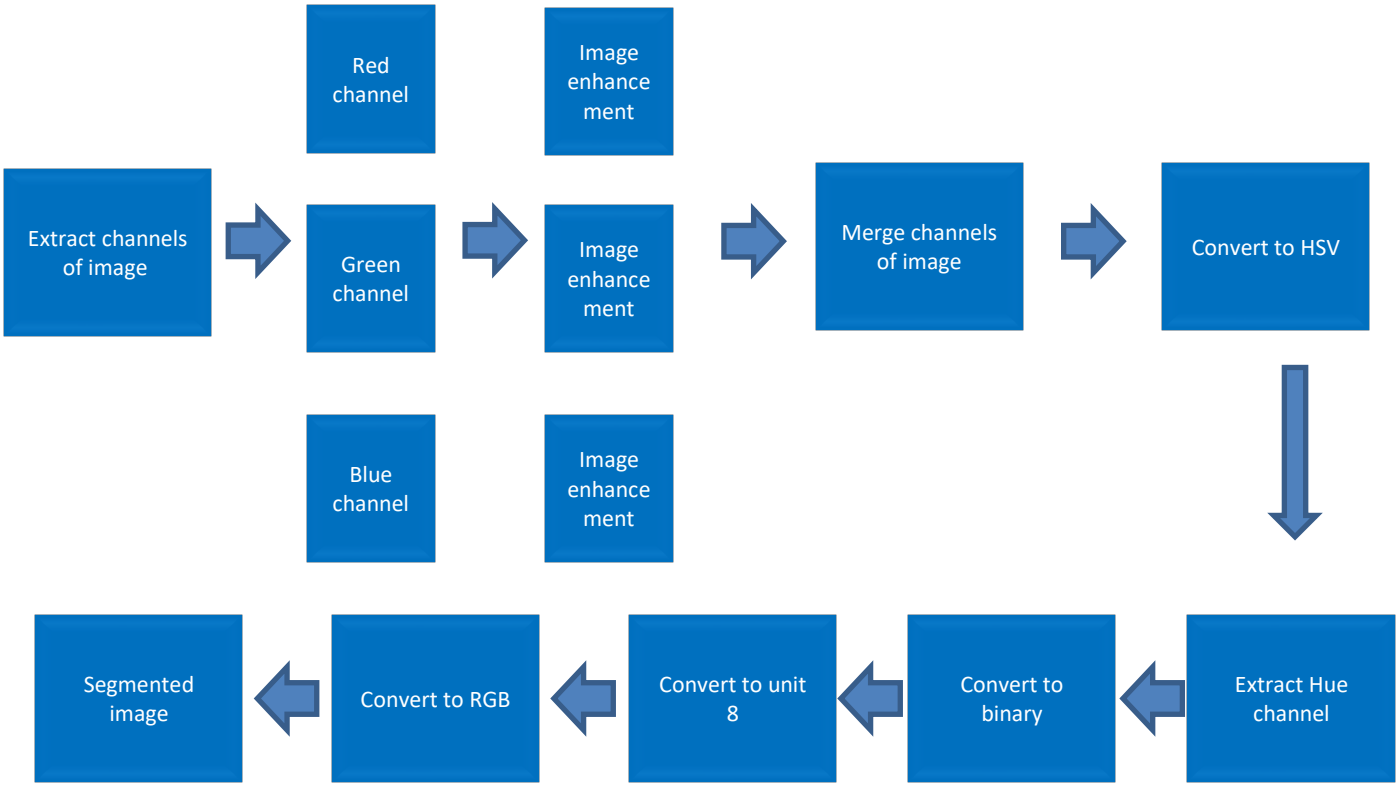

Fig. 2. The Segmentation stages

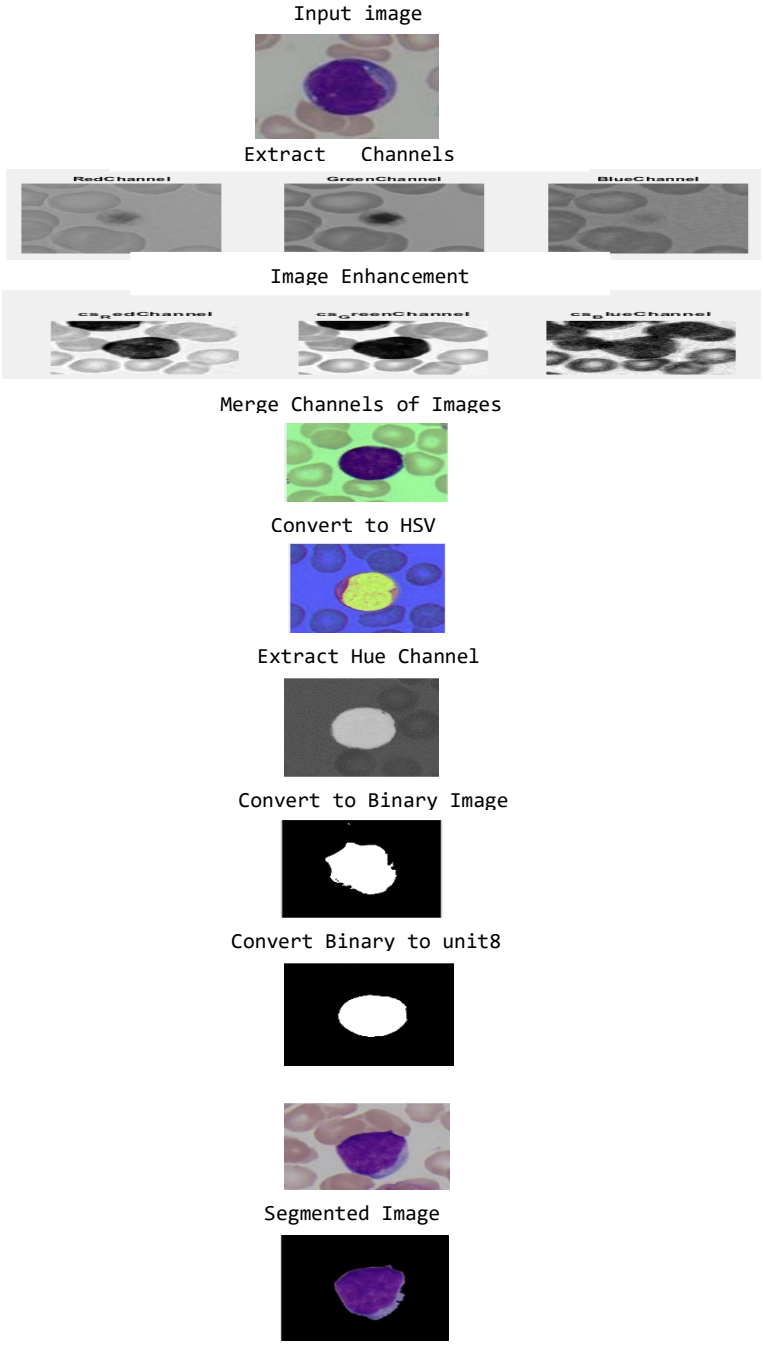

Fig. 3. The Implementation of Segmentation Stages

We proceeded by isolating the hue channel image into two components: lymphocytes and the background, with the lymphocyte section comprising both the cytoplasm and nucleus. To achieve this, we employed a mask matrix, producing a filtered image. This filtered image was then transformed into a binary image through the local entropy threshold method. However, the resulting binary image contained substantial noise, necessitating its removal. We established a noise area criterion of less than 125 pixels, removing objects smaller than this threshold. As a result, the image retained only two distinct objects: the background and the lymphocytes.

4.3 Features Extraction

In the classification and detection of Acute Lymphoblastic Leukemia (ALL) from microscopic images, texture analysis plays a critical role. The Gray Level Co-occurrence Matrix (GLCM) is a powerful statistical tool used for extracting texture features from images. This methodology focuses on extracting key features such as Contrast, Correlation, Energy, Homogeneity, Mean, and Variance using GLCM to improve the accuracy of ALL detection and classification. Spatial relationships (offsets) are defined for

GLCM computation. For each offset, a GLCM is constructed by counting the frequency of pixel pairs with specific gray level values occurring at the defined spatial relationship. From each GLCM, several texture features are extracted (PS & Vs, 2016).

Contrast: which measures the intensity contrast between a pixel and its neighbor over the whole image.

$$Contrast = \sum_{i,j} (i-j)^2 P(i,j), \tag{9}$$

where higher contrast values indicate a greater disparity in intensity between neighboring pixels, which may signify the presence of abnormal cell structures.

Correlation: which measures how correlated a pixel is to its neighbor over the whole image.

$$Correlation = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)P(i,j)}{\sigma_i \sigma_j}, \tag{10}$$

where high correlation values indicate a linear relationship between the pixel values, often reflecting more homogeneous cell regions.

Energy: which measures the sum of squared elements in the GLCM.

$$Energy = \sum_{i,j} P(i,j)^2 \tag{11}$$

where higher energy values suggest less texture complexity and more uniform cell regions.

Homogeneity: which measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.

$$Homogeneity = \sum_{i,j} \frac{P(i,j)}{1 + |i-j|}, \tag{12}$$

where higher homogeneity values indicate that pixels are more similar in intensity to their neighbors, characteristic of uniform tissue structures.

Mean: is the average value of the pixel intensities in the image.

$$Mean = \sum_{i,j} i.P(i,j), \tag{13}$$

where the mean provides the central tendency of the gray levels, which can help distinguish between different tissue types.

Variance: which measures the dispersion of pixel intensities around the mean.

$$Variance = \sum_{i,j} (i - \mu)^2 P(i,j) \tag{14}$$

where higher variance values indicate greater variability in pixel intensities, often seen in regions with mixed cell types.

4.4. Classification

The classification and detection of Acute Lymphoblastic Leukemia (ALL) in microscopic images is crucial for early diagnosis and treatment. In this research, we explore various distance metrics Manhattan, Euclidean, Canberra, and Chebyshev to enhance the accuracy of image classification algorithms in detecting ALL. These metrics are applied to feature vectors extracted from microscopic images, allowing us to compare their effectiveness in distinguishing between healthy and leukemic cells.

4.4.1. Manhattan Distance

The Manhattan distance calculates the distance between two points in a grid-based path by summing the absolute differences of their coordinates.

$$d_{Manhattan}(x,y) = \sum_{i=1}^{n} |x_i - y_i|. \tag{15}$$

4.4.2 Euclidean Distance

The Euclidean distance is the straight-line distance between two points in Euclidean space.

$$d_{Euclidean}(x,y) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}. \tag{16}$$

4.4.3. Canberra Distance

The Canberra distance is a weighted version of the Manhattan distance, giving more importance to differences in dimensions with smaller values.

$$d_{Canberra}(x,y) = \sum_{i=1}^{n}\frac{|x_i - y_i|}{|x_i| + |y_i|}. \tag{17}$$

4.4.4. Chebyshev Distance

The Chebyshev distance, also known as the L∞ distance, measures the greatest difference between corresponding coordinates of two points.

$$d_{Chebyshev}(x,y) = max_i|x_i - y_i| \tag{18}$$

4.5. Evaluation

The evaluation of the classification model is a crucial stage that ensures the reliability and effectiveness of the model. For this purpose, we utilize several performance metrics: Matthews Correlation Coefficient (MCC), sensitivity, specificity, accuracy, error rate, recall, and precision. Each metric provides unique insights into the performance of the model, helping us understand its strengths and weaknesses (Chicco & Jurman, 2023). Matthews Correlation Coefficient (MCC) which measures the quality of binary classifications. It considers true and false positives and negatives and is regarded as a balanced measure even for imbalanced datasets as the following Formula.

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}. \tag{19}$$

Sensitivity measures the proportion of actual positives that are correctly identified by the model as the following formula:

$$Sensitivity = \frac{TP}{TP - FN}. \tag{20}$$

Specificity measures the proportion of actual negatives that are correctly identified by the model as the following formula:

$$Specificity = \frac{TN}{TN - FP}. \tag{21}$$

Accuracy measures the overall correctness of the model, considering both true positives and true negatives as the following formula.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{22}$$

Error rate measures the proportion of incorrect predictions made by the model as the following formula.

$$Error\ Rate = \frac{FP + FN}{TP + TN + FP + FN}. \tag{23}$$

5. Experimental Results and Discussion

In this paper, we evaluated our proposed method using a dataset of 260 images, labeled 0 for healthy individuals and 1 for leukemia patients. Among these, 130 images are of healthy individuals, and 130 are of leukemia patients. We utilized random sampling for

training set selection and conducted experiments using four distance measures: Manhattan, Euclidean, Canberra, and Chebyshev. Fig. 4. Shows the positive and negative of the ALL-IDB samples.

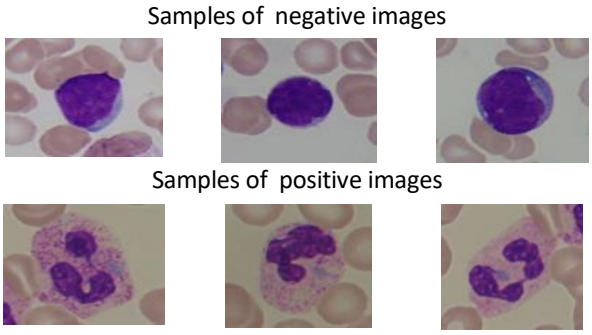

Fig. 4. Samples of ALL Images

Table 1 demonstrates that we developed 15 scenario models for each distance measure, performing 30 random samplings per scenario, resulting in a total of 450 experiments (15 scenarios × 30 samplings). We applied k-fold cross-validation with k=15, where each experimental scenario was evaluated using random sampling. The sample indexes for both training and testing sets were computer-generated, with the training set size consistently doubled, indicating equal numbers of acute lymphoblastic leukemia patient images and healthy individual images. For example, in scenario No. 1, the training set comprises 112 images (56 of leukemia patients and 56 of healthy individuals), which is similarly reflected in the third column of our data.

Table 1
Scenarios models

No	Training	Testing	No	Training	Testing
1	56 × 2=112	74 × 2=148	9	64 × 2= 128	66 × 2=132
2	57 × 2=114	73 × 2=146	10	65 × 2= 130	65 × 2=130
3	58 × 2= 116	72 × 2=144	11	66 × 2= 132	64 × 2=128
4	59 × 2= 118	71 × 2=142	12	67 × 2= 134	63 × 2=126
5	60 × 2=120	70 × 2=140	13	68 × 2= 136	62 × 2=124
6	61 × 2= 122	69 × 2=138	14	69 × 2= 138	61 × 2=122
7	62 × 2= 124	68 × 2=136	15	70 × 2= 140	60 × 2=120
8	63 × 2= 126	67 × 2=134			

We evaluated our proposed method using four measurement techniques: Euclidean Distance, Manhattan, Canberra, and Chebyshev.

5.1. Experimental Results and Discussion using Euclidean Distance

Fig. 5. present the MCC using Euclidean distance, where the average performance improves steadily from 73.04 at 56 features to a peak of 84.45 at 65 features. Minimum performance also increases, reaching 79.37 at 65 features. Maximum performance values peak at 90.48 with 62 features, while standard deviation remains low, indicating consistent results. Overall, the best performance and stability are observed with around 65 features.

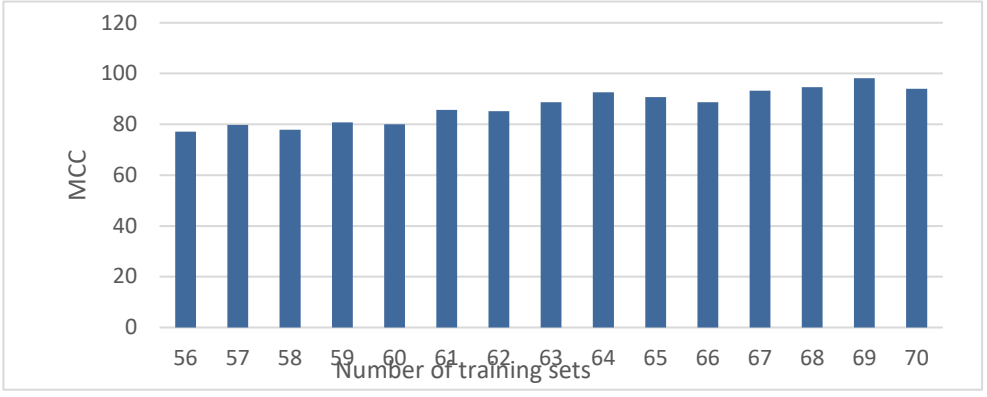

Fig. 5. The MCC for Euclidean distance

Fig. 6 presents the accuracy metrics using Euclidean Distance. The average accuracy improves, peaking at 65 features with 84.51. Minimum and maximum accuracy values also show improvements, with peaks at 60 and 61 features, respectively. The standard deviation remains low, indicating consistent performance, especially stable at 70 features. Overall, 65 features offer the best balance of high average accuracy and stability.

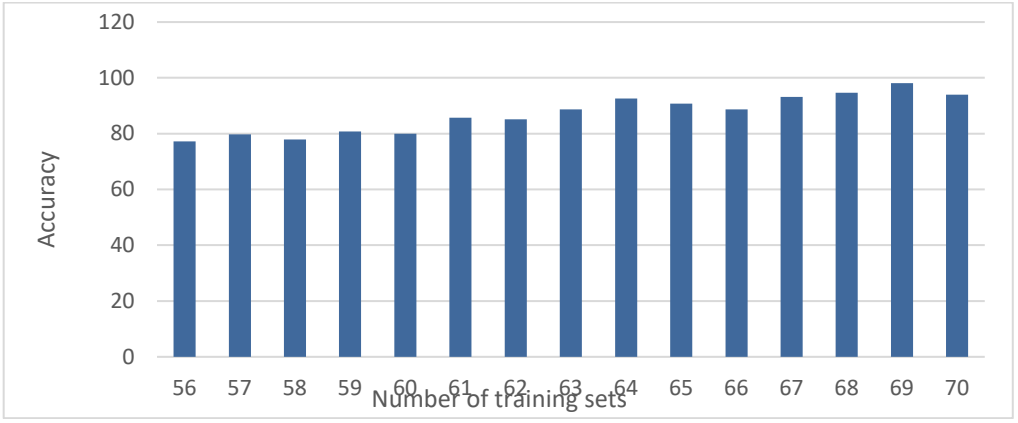

Fig. 6. The Accuracy for Euclidean distance

Fig. 7 presents error rates using Euclidean Distance. The average error rate decreases as features increase, reaching a low of 15.49 at 65 features before rising again. Minimum error rates show significant improvement, with the lowest at 62 features (8.82). Maximum error rates vary but show a notable reduction around 64 and 65 features. The standard deviation remains relatively low, indicating consistent performance, with the lowest at 70 features (0.024). Overall, the optimal performance is observed around 64 to 65 features.

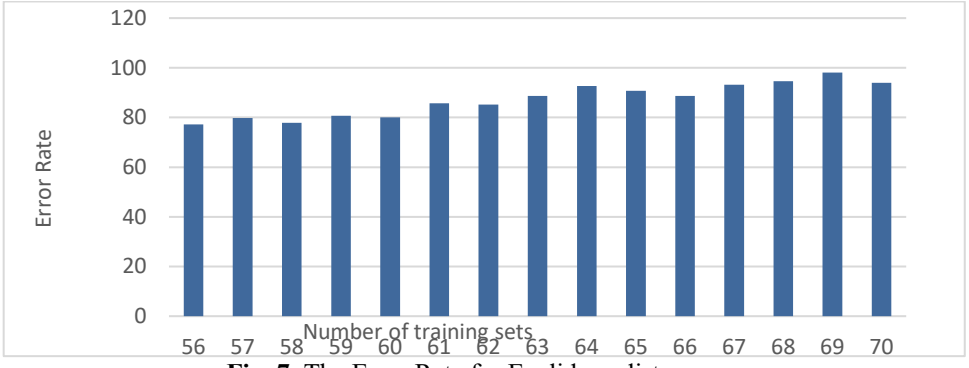

Fig. 7. The Error Rate for Euclidean distance

Fig. 8 presents the sensitivity using Euclidean Distance as a performance metric. The highest average sensitivity (85.48) is achieved with 56 features, while sensitivity generally declines as features increase, dropping to 75.29 at 70 features. The minimum sensitivity values show variability, with the lowest at 66.10 for 59 features. Maximum sensitivity remains relatively high, peaking at 96.83 for 63 features. Standard deviation values indicate consistent performance, with the lowest variability at 64 features (0.040).

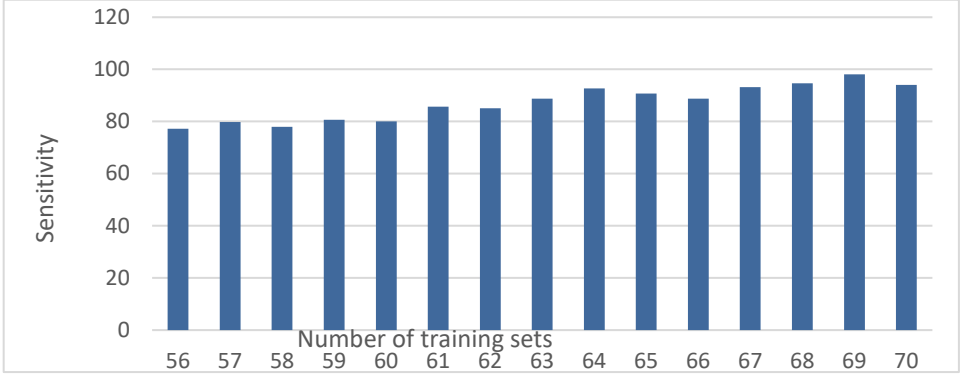

Fig. 8. The Sensitivity for Euclidean distance

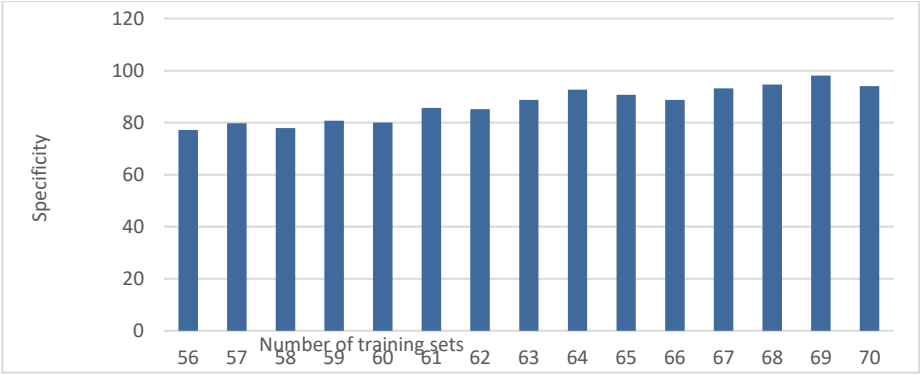

Fig. 9. The specificity for Euclidean distance

5.2. Experimental Results and Discussion using Manhattan distance

Fig. 10 presents MCC performance metrics using Manhattan distance. The average MCC improves steadily, peaking at 85.25 with 65 features. Minimum MCC values show improvement, particularly from 65 to 70 features. The maximum MCC value reaches 92.31 at 66 features. Standard deviation remains low, indicating consistent performance across feature sets. Overall, optimal performance and stability are achieved around 65 to 66 features.

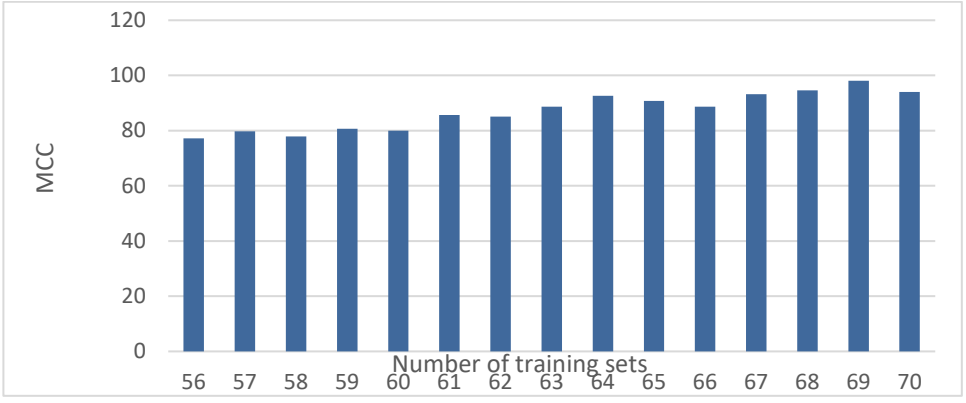

Fig. 10. The MCC for Manhattan distance

Fig. 11 presents the accuracy performance metrics for Manhattan distance which it shows an increasing trend in average accuracy, peaking at 85.1 with 65 features. Minimum accuracy improves significantly from 69.6 to 79.2, reflecting enhanced worst-case performance. Maximum accuracy reaches its highest at 92.2 with 66 features, indicating strong best-case scenarios. The standard deviation remains low, suggesting consistent performance, especially stable at 70 features. Overall, using around 65-66 features yields the best classification accuracy and reliability.

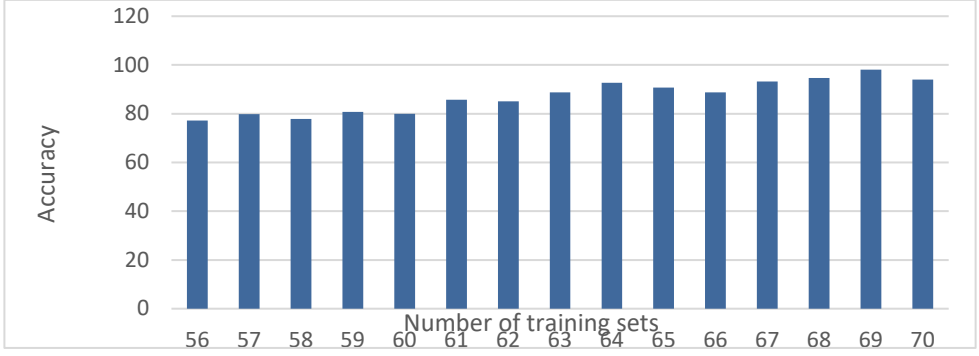

Fig. 11. The Accuracy for Manhattan distance

Fig. 12 present the average error rate metrics using Manhattan distance where it decreases steadily from 23.63 to 14.90 as the number of features increases, reaching the lowest at 65 features. Minimum error rates show significant improvement, dropping

from 18.24 to 10.77. Maximum error rates also decrease, with the best performance around 70 features at 25. The standard deviation values remain low, indicating stable error rates across feature sets. Overall, increasing features improves classification accuracy, with optimal performance around 65 features.

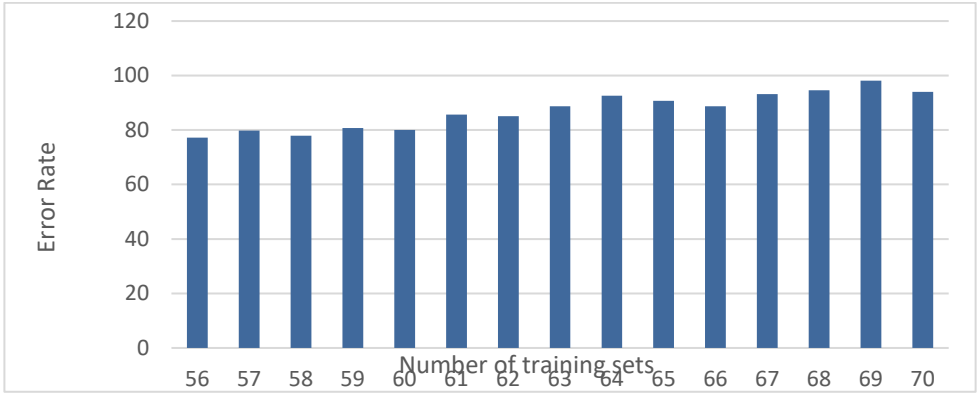

Fig. 12. The Error Rate for Manhattan distance

Fig. 13 presents the sensitivity for Manhattan distance as a performance metric for leukemia image classification. The average sensitivity shows a peak at 86.10 with 65 features, while minimum values improve up to 66 features. Maximum sensitivity is highest at 60 features with 96.88. Standard deviation values indicate consistent performance, particularly from 65 to 70 features. Overall, the sensitivity peaks and stabilizes around 65 features, highlighting optimal performance.

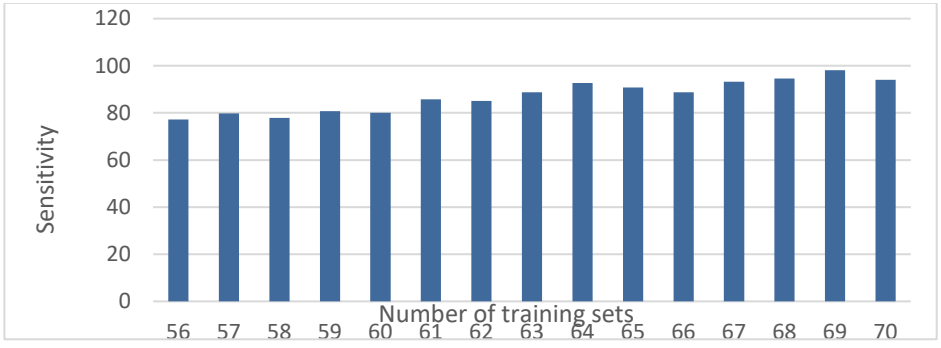

Fig. 13. The Sensitivity for Manhattan distance

Fig. 14 presents the sensitivity for Manhattan distance as a performance metric for leukemia image classification. Average sensitivity improves consistently, peaking at 85.97 with 69 features. Minimum values show variability but generally increase, with a low of 58.43 at 57 features and a high of 74 at 70 features. Maximum sensitivity values peak at 96.92 with 65 features. The standard deviation remains low, indicating stable performance across feature sets. Overall, sensitivity is optimized around 65 to 70 features.

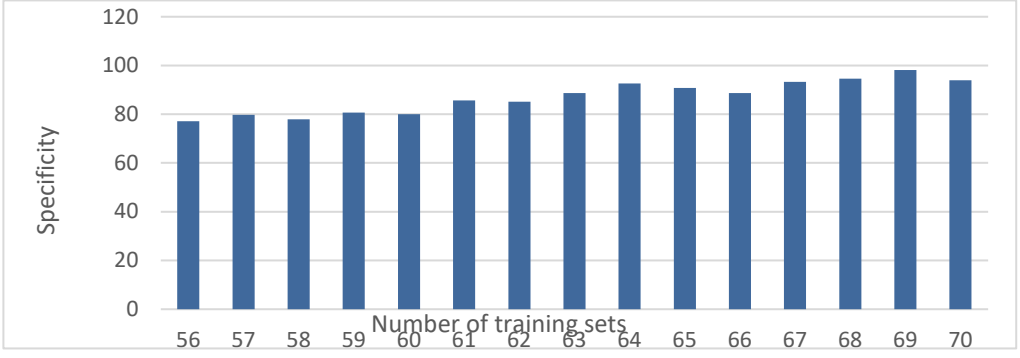

Fig. 15. The Specificity for Manhattan distance

5.3. Experimental Results and Discussion using Canberra distance

Fig. 15 presents the MCC performance metrics for leukemia using Canberra distance. Average MCC values show an upward trend, peaking at 85.25 with 65 features. Minimum and maximum MCC values also increase, with notable peaks at 65 features (minimum 80.92, maximum 91.47). Standard deviation values are low, indicating consistent performance, particularly stable at 70 features. Overall, 65 features provide optimal performance and stability.

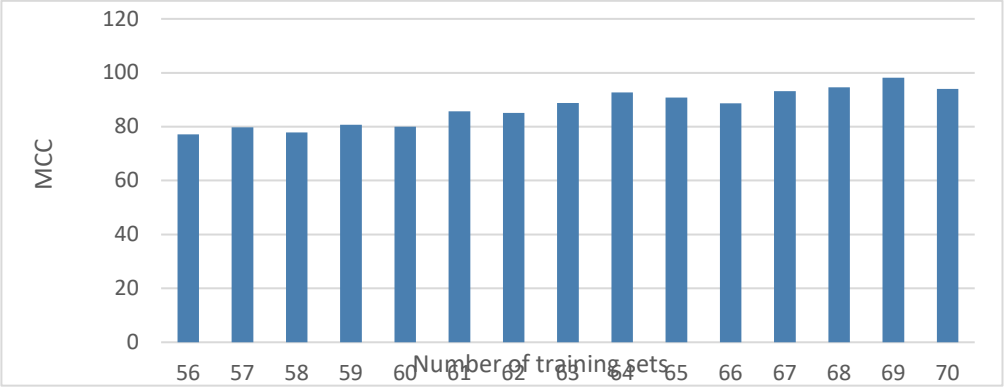

Fig. 15. The Specificity for Canberra distance

Fig. 16 presents accuracy metrics using the Canberra distance. Average accuracy peaks at 85.18 with 65 features, indicating the highest performance. Minimum accuracy improves steadily, peaking at 80.77 for 65 features. Maximum accuracy also peaks at 91.54 with 65 features. Standard deviation remains low, suggesting consistent performance across feature sets. Overall, 65 features achieve the best accuracy and stability.

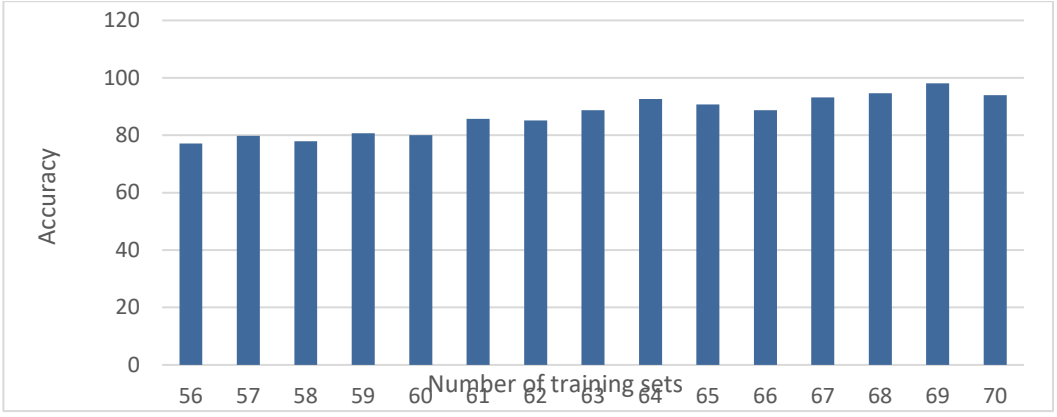

Fig. 16. The Accuracy for Canberra distance

Fig. 17 presents the error rate for Canberra distance as a performance metric. The average error rate decreases from 23.90 to a low of 14.82, with the minimum error rate also improving from 18.24 to 8.46. Maximum error rates show variability, peaking at 30.41 and later stabilizing around 25.83. The standard deviation remains relatively stable, indicating consistent performance. Overall, lower error rates are achieved with increased features, improving classification accuracy.

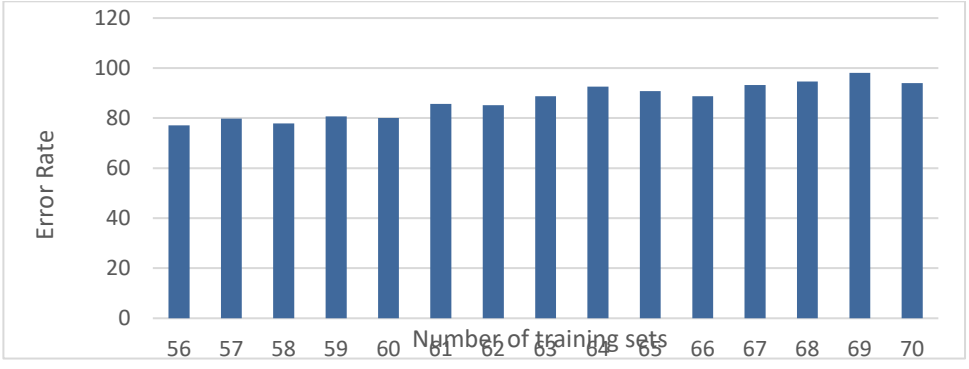

Fig. 17. The Error Rate for Canberra distance

Fig. 18 presents sensitivity metrics using Canberra distance. Average sensitivity peaks at 86.39 with 64 features and declines to 73.76 with 70 features. Minimum sensitivity improves until 64 features and fluctuates thereafter. Maximum sensitivity remains high, peaking at 96.43 with 56 features. Standard deviation indicates relatively stable performance, with the lowest variability at 63 features (0.035). Overall, optimal performance is observed around 64 features.

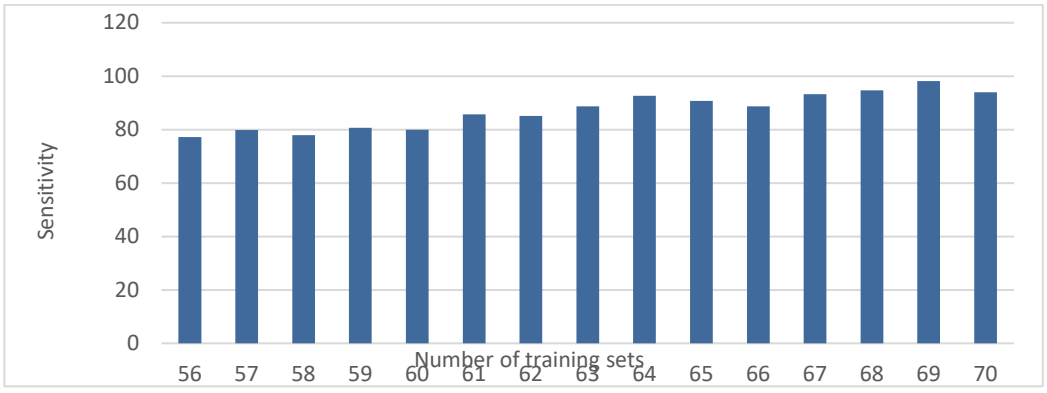

Fig. 18. The Sensitivity for Canberra distance

Fig. 19 presents specificity metrics using Canberra distance. The average specificity improves from 70.65 to a peak of 86.31 at 68 features. Minimum specificity shows variability, with a notable dip to 57.35 at 64 features. Maximum specificity peaks at 94 between 67 to 70 features. Standard deviation remains moderate, indicating relatively consistent performance. Overall, the optimal specificity performance occurs around 68 features.

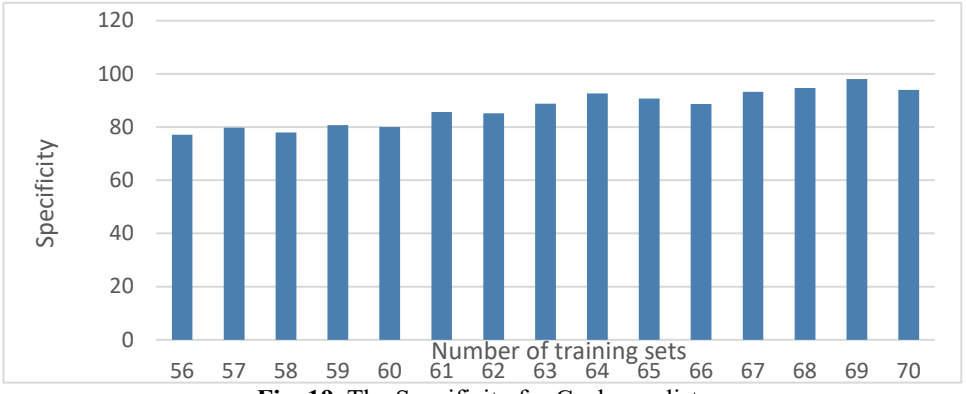

Fig. 19. The Specificity for Canberra distance

5.4. Experimental Results and Discussion using Chebyshev distance

Fig. 20 presents performance metrics for MCC using Chebyshev distance. Average MCC values increase steadily from 72.11 to 83.14 as the number of features ranges from 10 to 65. Minimum MCC values also show a gradual increase from 64.52 to 77.61, indicating improved worst-case performance. Maximum MCC values peak at 90.23, suggesting strong predictive capability at optimal feature counts. Standard deviation remains relatively low, indicating consistent model performance across feature sets.

Fig. 20. The MCC for Chebyshev distance

Fig. 21 presents performance metrics for accuracy using Chebyshev distance. The average accuracy across different experiments ranges from 74.89% to 82.69%, with a peak at 82.69% observed with 10 features. Minimum accuracy levels range from 67.14% to 76.92%, showing variability across feature sets. Maximum accuracy reaches up to 90.15%, indicating the potential for high classification performance. Standard deviations are generally low, indicating consistent performance evaluations.

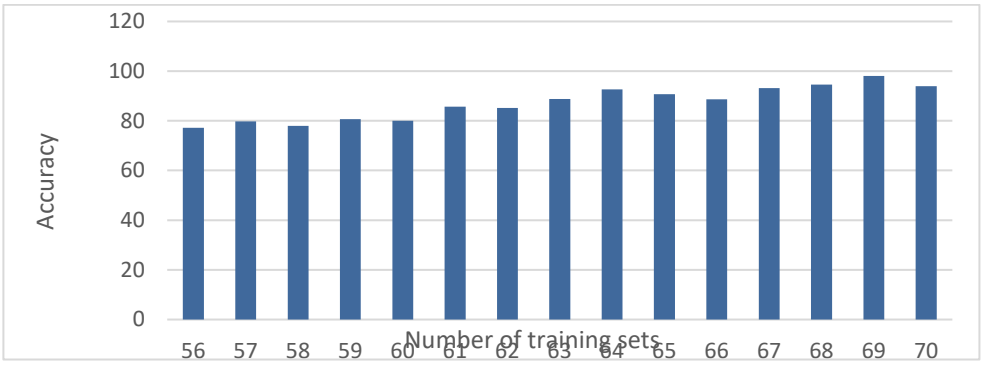

Fig. 21. The accuracy for Chebyshev distance

Fig. 22 presents performance metrics for error rate using Chebyshev distance. The figure shows varying performance across different measurements. The average error rate ranges from 17.31% to 25.11%, with fluctuations observed across different feature sets. Minimum error rates are notably lower, ranging from 9.85% to 19.86%, indicating effective classification in some scenarios. Maximum error rates range from 23.07% to 32.86%, suggesting variability in classification accuracy. Standard deviations are relatively low, indicating consistent performance measures across the dataset.

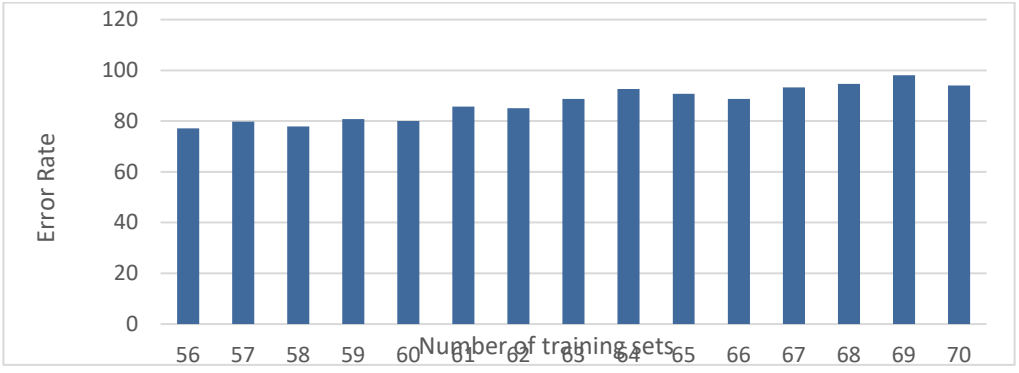

Fig. 22. The Error rate for Chebyshev distance

Fig. 23 presents performance metrics for Sensitivity using Chebyshev distance. The sensitivity results for Chebyshev distance as a performance metric indicate varying performance across different experiments. The average sensitivity ranges from 76.29% to 85.95%, with the highest average achieved at 85.95%. Minimum sensitivity values range from 67.14% to 78.46%, while maximum sensitivity values vary from 84.29% to 94.64%. Standard deviations across experiments are relatively low, ranging from 0.04 to 0.06, suggesting consistent sensitivity measurements. These findings highlight the robustness of Chebyshev distance in effectively classifying leukemia images with high sensitivity and reliability.

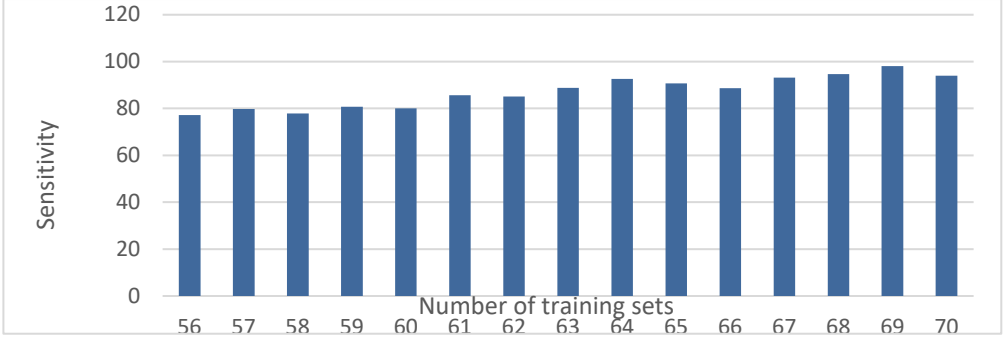

Fig. 23. The Sensitivity for Chebyshev distance

Fig. 24 presents specificity metrics for Chebyshev distance across different experiments. Specificity averages range from 68.15% to 82.77%, showing an increasing trend with varying feature sets. Minimum values range from 56.52% to 72%, indicating variability in worst-case performance across experiments. Maximum values span from 77.17% to 98.11%, demonstrating the potential for high performance. Standard deviations, ranging from 0.0369 to 0.0663, suggest consistent performance variability across experiments.

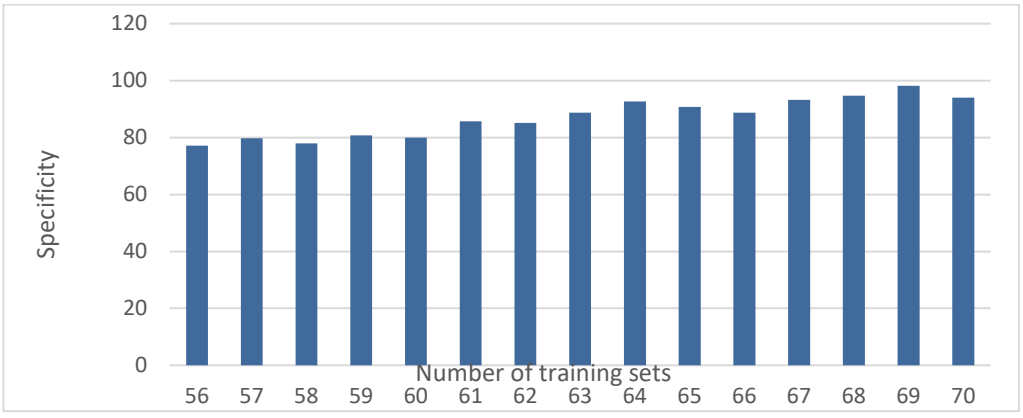

Fig. 24. The Specificity for Chebyshev distance

6. Conclusion

This research establishes a robust and reliable framework for the automated classification of leukemia images, leveraging state-of-the-art image processing and machine learning techniques. By incorporating multiple distance metrics Euclidean, Manhattan, Canberra, and Chebyshev we achieved high classification performance, with the Chebyshev distance showing an average accuracy of 82.69% and sensitivity of 85.95%. The optimal performance was observed with the Canberra distance using 65 features, which resulted in an accuracy of 85.18%, sensitivity of 86.39%, and specificity of 86.31%. These results validate the effectiveness of our proposed method in accurately classifying leukemia cells, thereby enhancing the potential for early diagnosis and targeted treatment strategies. Future work will focus on refining feature extraction techniques and exploring additional distance metrics to further improve classification accuracy and robustness.

Acknowledgement

The authors extend their appreciation to Prince Sattam bin Abdulaziz University for funding this research work through the project number (PSAU/2023/01/ 25906).

Author's note

The authors declare that there is no conflict of interest regarding the publication of this article. The authors confirmed that the paper was free of plagiarism.

References

Alagu, S., & Bagan, K. B. (2019, April). Acute Lymphoblastic Leukemia diagnosis in microscopic blood smear images using Texture features and SVM classifier. In *Alliance International Conference on Artificial Intelligence and Machine Learning (AICAAM)* (pp. 175-186).

Chicco, D., & Jurman, G. (2023). The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining*, *16*(1), 4.

Labati, R. D., Piuri, V., & Scotti, F. (2011, September). All-IDB: The acute lymphoblastic leukemia image database for image processing. In *2011 18th IEEE international conference on image processing* (pp. 2045-2048). IEEE. doi: 10.1109/ICIP.2011.6115881.

More, P., & Sugandhi, R. (2023). Automated and enhanced leucocyte detection and classification for leukemia detection using multi-class SVM classifier. *Engineering Proceedings*, *37*(1), 36.

Muntasa, A., & Yusuf, M. (2019). Modeling of the acute lymphoblastic leukemia detection based on the principal object characteristics of the color image. *Procedia Computer Science*, *157*, 87-98.

Muntasa, A., & Yusuf, M. (2020). A Novel Approach to Detect the Acute Lymphoblastic Leukemia Based on the Color Orthonormal Basis Entropy (COBE) and the Distribution of the Pixel Intensity (DoPI). *International Journal of Intelligent Engineering & Systems*, *13*(1).

Muntasa, A., & Yusuf, M. (2021). Multi Distance and Angle Models of the Gray Level Co-occurrence Matrix (GLCM) to Extract the Acute Lymphoblastic Leukemia (ALL) Images. *International Journal of Intelligent Engineering & Systems*, *14*(6).

Muntasa, A., Wahyuningrum, R. T., & Nafisah, D. (2023). A New Model: Commutative Hypercomplex-Convolutional Neural Network to Classify Acute Lymphoblastic Leukemia Images. *International Journal of Intelligent Engineering & Systems*, *16*(5).

Muntasa, A., Wahyuningrum, R. T., Tuzzahra, Z., Motwakel, A., Yusuf, M., & Mahmudi, W. F. (2022). A Pyramid Model of Convolutional Neural Network to Classify Acute Lymphoblastic Leukemia Images. *International Journal of Intelligent Engineering & Systems*, *15*(6).

PS, S. K., & Vs, D. (2016). Extraction of texture features using GLCM and shape features using connected regions. *International journal of engineering and technology*, *8*(6), 2926-2930.

Shafique, S., & Tehsin, S. (2018). Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks. *Technology in cancer research & treatment*, *17*, 1533033818802789.

Sukhia, K. N., Ghafoor, A., Riaz, M. M., & Iltaf, N. (2019). Automated acute lymphoblastic leukaemia detection system using microscopic images. *IET Image Processing*, *13*(13), 2548-2553.

Talaat, F. M., & Gamel, S. A. (2024). Machine learning in detection and classification of leukemia using C-NMC_Leukemia. *Multimedia Tools and Applications*, *83*(3), 8063-8076.

Terwilliger, T., & Abdul-Hay, M. J. B. C. J. (2017). Acute lymphoblastic leukemia: a comprehensive review and 2017 update. *Blood Cancer Journal*, *7*(6), e577-e577.