Automated Leukemia Detection using K-means Clustering for Feature Extraction

Srwa Ahmed Mustafa[1], Asst. Prof. Dr. Gullanar M Hadi[1]

[1]Department of Software Engineering and Informatics, College of Engineering, Salahaddin University- Erbil, Kurdistan Region, Iraq

***Abstract*— Leukemia is a kind of blood cancer that may cause significant damage to a person's general health. It is characterized by the production of an excessive number of white blood cells. To address leukemia quickly and effectively, it is important to have a diagnosis that is both correct and quick. Not too long ago, experts started using AI methods to help find cancer much earlier. One of the hardest parts of making a method to find leukemia is separating the nuclei from the rest of the picture. When medical staff use quick and accurate division methods, they can find patients faster and treat them more effectively. So far, hybrid clustering algorithms have been very helpful in the process of picture segmentation in the field of medical image processing. To find leukemias like chronic myeloid leukemia (CML) and chronic lymphocytic leukemia (CLL), this study looks into segmentation methods that use machine learning (ML) and deep learning (DL). The study looks at how many ML and DL algorithms can be used to automatically diagnose different types of leukemia. It is checked to see how well the ML and DL algorithms do at segmentation, pre-processing, feature extraction, selection, and total classification accuracy.**

***Index Terms*— Machine learning, CLL, k-means clustering, CML, Deep learning, Segmentation, Leukemia**

I. Introduction

Leukemia is a type of blood cancer that can have serious effects on the health of men and women of all ages. It is marked by the growth of white blood cells that is not normally occurring. In leukemia instances, a timely and accurate diagnosis is essential for both efficient treatment and timely action. Recently, researchers have been using artificial intelligence techniques to detect cancer. Nucleus segmentation is a crucial step in any strategy for leukemia detection. When precise and effective segmentation techniques are used, medical practitioners may identify illnesses more rapidly and provide more effective treatment. Physicians employ hybrid clustering algorithms widely as a potent tool for medical picture segmentation. This work investigates several segmentation strategies that use ML and DL to differentiate between chronic lymphocytic leukemia (CLL) and chronic myeloid leukemia (CML), among other leukemias, in order to automate the diagnosis of various leukemias. There is also a list of all the ML and DL methods that are used in this area. The study checks how well their methods work in several areas, such as feature extraction, leukemia picture segmentation, pre-processing, and making the images more readable. It also checks how accurate the ML (Vogado et al., 2018, Mishra et al., 2019, Inbarani et al., 2020, Kazemi et al., 2016, Amin et al., 2015, Dasariraju et al., 2020, Ranjitha and Duth 2021, Kumar et al., 2018, Fatma and Sharma 2014) and DL (Vogado et al., 2018, Das and Meher 2021) systems are at classifying things in general. Figure 1 shows four types of white blood cells that are linked to acute myeloid leukemia.

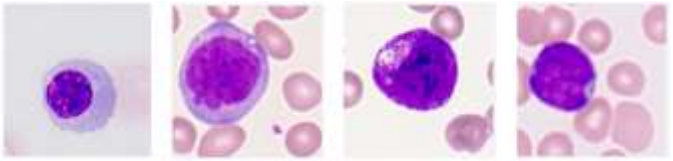

Figure 1 Acute myeloid leukemia (a) Erythroblast; (b) monoblast; (c) promyelocyte; (d) myeloblast (Dasariraju et al., 2020).

. (a) (b) (c) (d)

Gene expression profiles and medical images are two examples of the kinds of massive datasets made possible by recent advances in high-throughput technologies. They make use of the information available by utilizing machine learning techniques. This allows us to discover patterns that could be indicators of the presence of leukemia or the classification of its different types. The researchers presented a method, for detecting leukemia in blood images by employing neural networks (CNNs). They also evaluated how dependable pre trained CNNs are, at extracting features in comparison, to state of the art methods. This study discovered that photos containing a quantity of leukocytes necessitate a number of characteristics, for classification whereas images, with only one leukocyte require fewer attributes. In addition , the primary advantage of (Vogado et al., 2018), One of the reasons, for its performance compared to advanced methods is that it doesn't require a segmentation procedure. The model achieved a success rate of 99%. Outperformed nine techniques that were previously described in existing studies.

II. Related Work

Detecting leukemia relies heavily on machine learning (ML) algorithms. Research has indicated the existence of a method to differentiate between damaged blood cells, in microscopic images particularly for diagnosing acute lymphoblastic

DOI:http://doi.org/10.24086/cocos2024/paper.1529

leukemia (Mishra et al., 2019). We begin the recommended process by preparing the input photos. This involves utilizing the Y component of the image and applying a threshold using the triangle method. Afterward the texts properties are determined by utilizing the S transform (DOST) and simplicity is enhanced through linear discriminant analysis. The suggested Adaboost approach utilizes a forest, as the underlying classifier, the RF (ADBRF) classifier to process the gathered features. The effectiveness of this method was confirmed by using the ALL IDB1 dataset. With a 99% success rate, the results of five rounds of k-fold stratified cross-validation show that the suggested method performs better than the most advanced systems.

Lin et al. (Lin et al., 2013) suggested a microscopic blood image tumor prediction method. They distinguished six ALL subtypes using evolutionary algorithms for feature selection and silhouette statistics. They improved accuracy with microarray data and gene expression. CFS/SVM had 96% accuracy with more than 20 predicted genes, while 23 genes had 100% accuracy. A flow cytometer-based lymphoblast identification method by Zong et al. The categorization accuracy of (Zong et al., 2005) is 96.67%. Gene expression data might be expensively and tediously obtained from bone marrow samples. Ross et al. (Ross et al., 2003) offered a system for classifying pediatric ALL. Researchers analyzed 132 distinct blasts using higher-density nucleotide arrays. With 97% accuracy, it included recently discovered genes in class prediction algorithms.

Hybrid histogram-based soft covering rough k-means clustering (HSCRKM), a machine learning-based leukemia detection technique, was introduced by Inbarani et al. (Fatma and Sharma 2014) in their work. This approach combines the rough set, rough k-means clustering, and soft covering approaches. The segmented nucleus image was used to extract color, shape, and gray-level co-occurrence matrix (GLCM) features. The ML prediction techniques accurately categorized cells as either malignant or non-cancerous. The empirical results demonstrate that the HSCRKM clustering method, along with all ML algorithms (except naive Bayes), achieves a prediction accuracy of above 80%. Logistic regression and neural networks provide greater accuracy, surpassing other prediction algorithms with a rate over 90%.

Kazemi et al. (Kazemi et al., 2016) developed an automated technique to identify AML and its common subtypes, M2-M5, and photographed AML and normal blood smears under a microscope. In order to retrieve the whole nucleus from pictures including many nuclei, they use discriminative criteria such as cytoplasm-to-nucleus ratio, shape, color, texture, and Hausdorff dimension (HD) after image preprocessing. A technique, for identifying blood cells includes separating them based on their color. By utilizing a method known as support vector machine (SVM) classifiers in conjunction, with 10 cross validation it becomes feasible to distinguish between images that exhibit signs of cancer and those that do not present any indications. Using support vector machine (SVM) classifiers we can effectively differentiate between images that show signs of cancer and those that do not. This is achieved through a 10fold validation method. Moradi et al. (Amin et al., 2015) They demonstrate how computers can differentiate between cells by analyzing the imaging features of the nucleus. The SVM classifier assesses the sensitivity, specificity as well as accuracy of cells resulting in values of 98%, 95% as well as 97%. (Dasariraju et al., 2020) We used the RF approach to classify blood cells.The model effectively identified blood cells with a precision of 92.99% sorting them into four categories at an accuracy rate of 93.45%. According to Escalante et al. (Escalante et al., 2012), Researchers have successfully developed a method to identify leukemia

Study	Methods	Objective	Type of Leukemia	Dataset	Number of images in the dataset	Evaluation Metrics
(Salehi et al., 2020)	HSCRKM	An improved version of the soft-covering rough k-means clustering method that makes use of histogram-based approaches	Leukemia	ALL-IDB	368	Classification accuracy more than 80%
(Kazemi et al., 2016)	k-means clustering and SVM	identification and detection of AML	AML	AML image s	105	SVM=96% Multi-SVM=87%
(Amin et al., 2015)	k-means clustering and SVM	classification of cancerous and noncancerous cells	AML	Microscopi c image (ALL-IDB)	312	97%
(Dasariraju et al., 2020)	Random Forest	Identification and categorization of undeveloped white blood cells	immature leukocytes	Labelled images of leukocytes	100 AML patients	93.45%
(Fatma and Sharma 2014)	HSCRKM	clustering using hybrid histograms with smooth covering and rough k-means	leukemia	ALL IDB		90%
(Lin et al., 2013)	GA for feature selection and SVM	reliability via the use of gene expression and microarray data	Acute myeloid leukemia	ALL IDB	23	98%
(Zong et al., 2005)	ANN	Identification of different white blood cell categories	Acute myeloid leukemia	flow cytometer data	132	96.67%
(Escalante et al., 2012)	EPSMS	categorization of acute leukemia according to physical characteristics	ALL	Microscopi c image (ALL-IDB)		2-classes=97.68% Multi-Classes=94.21%
(Piuri and Scotti 2004)	Based on morphologi cal properties of a color image	separates the WBC from other blood cell components	leukocytes	Leukocytes images	134	92%
(Patel and Mishra 2015)	K-means clustering Zack algorithm, SVM	After identifying white blood cells using the Zack algorithm and the K-means clustering approach, researchers used the term "applied" to name the cells	Leukemia	Microscopi c image (ALL-IDB)	27	Classification accuracy 95.57%
(Setiawan et al., 2017)	GA-PSO-FCM and Clustering Based on SVM	This optimization method improves the clarity of diagnosing leukemia.	Acute lymphoblastic leukemia (ALL)	Microscopi c image (ALL-IDB)	180	Geometric mean 91 to 96%
(Xu et al., 2017)	K-means, Hidden Markov random field	The data was segmented using this method, extracted the traits, separated the blast cell nucleus from the background, and categorized the cells.	Acute myeloid leukemia	Microscopi c image	81	Segmentation accuracy 96 to 98% (average)
(Ahmed et al., 2019)	Convolutio nal neural network (CNN)	identified the subtypes of leukemia	Leukemia	Microscopi c image ASH Image Bank	903	Acc Leukemia = 88.25% Acc Healthy cell = 81.74%

Table 1 presents an overview of the literature on clustering methods that are based on ML and DL.

specifically focusing on categorizing its subtypes using particle swarm optimization. Researchers categorize acute leukemia subtypes using morphology-based ensemble building. In terms of multi-class categorization, the authors' accuracy was 94.21%, and in terms of two-class classification, it was 97.68%. The Foran et al. (Foran et al., 2000) framework distinguishes between leukemia and lymphoma. At 83%, they correctly diagnosed 19 individuals with leukemia and lymphoma. Although researchers have not confirmed the approach in ALL and on a larger dataset, they have successfully separated leukemia and lymphoma from one another. According to Scotti et al. (Piuri and Scotti 2004), examining color image morphology may help distinguish between leukocytes and WBCs. The recommended method first isolates WBC from other components of blood cells. This approach is suitable for identifying tumor distortion components in cell morphology.

DOI:http://doi.org/10.24086/cocos2024/paper.1529

The scientists collected 134 images of leukocytes. With a feature size of 23, Parallel FF-NN has a 92% accuracy rate in identifying distinct WBC types. Scotti proposed a further method (Scotti 2005) for distinguishing ALL from microscopic images. A classification error of 0.0133 was found in 150 images using feed-forward neural networks. These images show that morphological features are better for identifying lymphoblasts for ALL detection.

III. Methodology

The process of using ML and DL to diagnose leukemia is multifaceted and methodical; it combines sophisticated computer methods with the complexities of hematological diagnosis. The first phase involves the methodical gathering of a varied and well-annotated dataset, including gene expression profiles and medical imaging, from individuals with leukemia as well as healthy individuals. Ethical concerns guide the collection of this data to ensure patient privacy and regulatory compliance. Figure 2 presents a schematic of the method for detecting leukemia, and Figure 3 illustrates the pipeline for automated leukemia diagnosis, together with a detailed synopsis of noteworthy developments at every stage.

(Mishra et al., 2019) suggests preprocessing, sub imaging, feature extraction, feature reduction, and classification, much like earlier classification techniques. The typical preprocessing techniques that we used were background smoothing and noise reduction. Two important contributions are the PCA+LDA-based feature reduction and the DOST-based feature extraction. Then, we use a random forest (ADBRF) classifier based on AdaBoost to classify the relevant attributes. Like previous classification schemes, (Mishra et al., 2019) suggests categorization, feature extraction, feature reduction, sub-imagery, and preprocessing. The typical preprocessing techniques that we used were background smoothing and noise reduction. The feature reduction based on PCA+LDA and the feature extraction based on DOST are two significant contributions. All the important characteristics are sorted using the AdaBoost-based random forest (ADBRF) classifier. the third Separated the leukemia nucleus image using a mix of rough k-means clustering and a soft covering-based rough approximation. Leukemia nucleus picture segmentation using a unique HSCRKM method after preprocessing. Histogram peak values automatically determine the cluster number (K). (Kazemi et al., 2016) The suggested approach uses color picture segmentation to remove WBCs from the background and separate nuclei. Acute leukemia has little cytoplasm; thus, it is focused on the nucleus and isolated from its key properties. (Dasariraju et al., 2020) The approach included segmentation, feature extraction, classification, and feature significance calculation. Segmentation yielded cell and nucleus binary masks for each picture. A random forest system classified immature and adult cells using 16 criteria. Finally, random forest metrics determined each feature's relevance.

CNNs that have already been trained on ImageNet, a massive natural image dataset with 1000 categories (Vogado et al., 2018), can be utilized. Images of lymphocytes and lymphoblasts may be found in the database under these categories. Researchers use Vgg-f, CaffeNet, and AlexNet. While the design of both models is identical, the number of neurons in the fully connected layers and the size of the filters in the convolutional layers are different.UNITS

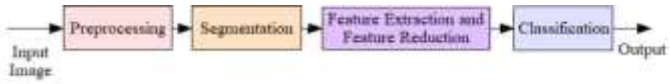

Figure 2 Steps of process for Leukemia detection based on ML and DL (Das et al., 2022)

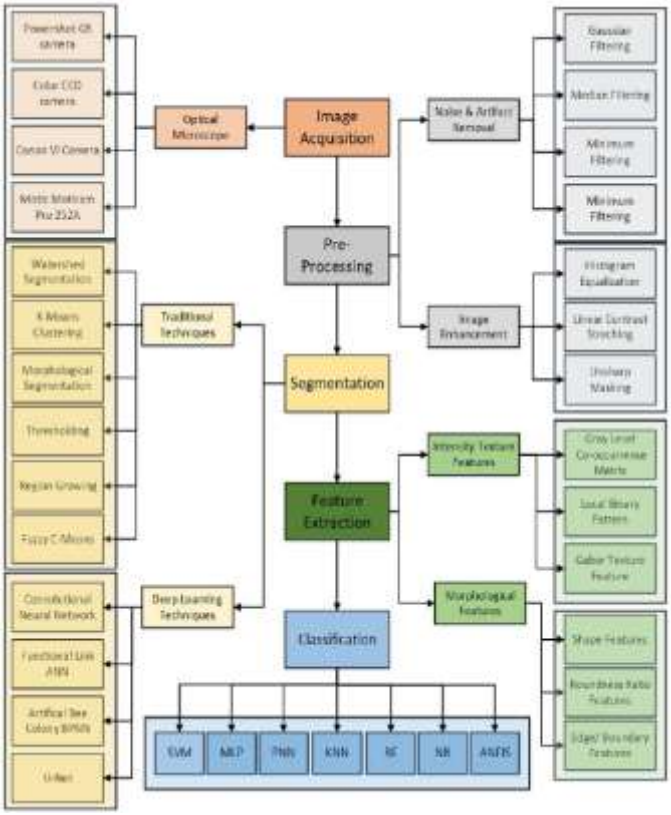

Figure 3 Diagram of the automated leukemia diagnostic pipeline and important breakthroughs (Shah et al., 2021)

A. Image Acquisition

Leukemia image acquisition often entails getting high-quality medical images to diagnose and analyze the features of leukemia. Healthcare practitioners use various imaging techniques to visually examine and study abnormalities in blood cells and tissues, recording different facets of the condition. After a comprehensive literature review, it is evident that approximately 17% of studies automating the analysis of blood smear images utilized publicly accessible datasets. ALL-IDB and ASH are two popular databases. Reference (Shah et al., 2021) provides further information on other available datasets, including Bloodseg, Cellavision, JTSC, SMC-IDB, and HUMS-IDB. (Vogado et al., 2018) Created three diverse image databases by using conventional databases described in existing research. The categorization of two of these databases is based on the leukocyte count in the photos. The hybrid-leukocyte database consists of four databases: ALL-IDB, ALL-IDB2, Leukocytes, and CellaVision. Each database contains just one leukocyte per image. For this particular experiment,

DOI:http://doi.org/10.24086/cocos2024/paper.1529

researchers particularly employed the datasets from the ALL-IDB (Inbarani et al., 2020). For this experimental investigation, we captured a total of 368 pictures, classifying 175 as benign and 193 as malignant. Due to their usual design, digital microscopes are not suitable for use in the RGB color space. In the preparatory phase, we convert the RGB source images into a LAB color space. The researchers used (Kazemi et al., 2016) to create a database of nine males and eight women, ages 16 to 69, who were Shariati Hospital AML patients. Clinical, blood, and bone marrow tests validated each case. Ten control participants and seventeen AML patients provided 27 bone marrow and peripheral blood smear slides to the pathology department at Shariati Hospital.

Using (Amin et al., 2015), 21 bone marrow slides and peripheral blood smears were prepared for this study from 14 ALL patients and 7 controls. Pathology specimens are prepared and subjected to giemsa staining in order to visualize cell components at the Al-Zahra and Omid hospitals in Isfahan. A Nikon 1 V1 high-resolution camera and a Nikon Eclipse 50i light microscope, equipped with a 100X oil immersion objective and a 1000X effective magnification, were utilized to acquire the images. In addition, photographs in JPEG format are obtained using the maximum resolution of the camera, which is 2592 × 3872 pixels in RGB color space. 18,365 leukocyte-centered images with ground truth labels identifying each one by sort were utilized in (Dasariraju et al., 2020). Medical examiners with training in cytomorphology generated ground truth annotations.

B. Preprocessing

Preprocessing is essential in leukemia identification, notably for gene expression profiles, medical imaging, and clinical data. The purpose is to improve data quality, decrease noise, and prepare it for analysis. A number of variables, including the camera, the microscope, the light source, the capturing camera angle, changes in illumination, and noise, may alter the quality of microscopic images. Stain normalization, which adjusts all stain slides to account for recording environmental variations, particularly light, is thus a crucial preprocessing step. It reduces light and color fluctuations in microscopic pictures from various labs, increasing segmentation and classification (Anghel et al., 2019). Histogram equalization or its modified variations, such as adaptive and Contrast Limited Adaptive Histogram Equalization, are simple solutions (Das et al., 2020). Gehlot et al. (Anghel et al., 2019) used two U-Net modules to create an effective coupled self-supervised framework. The first and second modules are used for stain normalization and identity transformation, respectively. The pre-processing steps for leukemia photos are shown in Figure 4. Preprocessing is done using the Gaussian laplacian (LoG) or LoG-based modified high-boosting (LoGMH) approach (Das et al., 2021). Edge enhancement and deblurring are accomplished using the Laplacian approach. However, noise has an effect on it. Before performing the Laplacian operation, LoG applies a Gaussian filter to decrease noise. To raise performance, LoGMH combines high-boosting operation with the benefits of LoG.

Figure 4 Procedure for pre-processing: (a) input image; (b) related Y component; (c) image equalized by histogram; (d) image thresholding; (e) after background removal operation.

IV. Image Segmentation

Image segmentation is a computer vision job that separates an image into semantically significant sections. to split an image into sections or areas that depict similar objects or locations. Medical imaging, autonomous automobiles, item recognition, and other fields all depend on image segmentation. Image segmentation facilitates the extraction of valuable information from visual data, opening up a wide range of applications. Researchers determine the segmentation strategy based on the task needs and image properties. The majority of the reviewed research here segmented the leukemia images using k-means clustering. For example, (Amin et al., 2015) We utilized a two step segmentation method, starting with k means clustering to create a group of nuclei. Afterwards we separate the core elements. Remove any unnecessary components, within this group.

V. Results and Discussion

This section focuses mainly on the procedure of extracting characteristics from images of leukemia. It also highlights how crucial it is to summarize discoveries, from the literature recognize trends and obtain an understanding of the current status of leukemia images.

A. Feature Extraction from leukemia images

Identifying malignancies is crucial especially when dealing with datasets such, as gene expression profiles or medical images with dimensions. Feature extraction plays a role, in this process. The objective is to identify the characteristics that accurately distinguish forms of leukemia or differentiate between leukemia samples and those, from individuals, without the condition (Vogado et al., 2018) utilized a gain ratio method to choose features. This approach involves utilizing decision trees to define the characteristics of vectors and assessing impurity using entropy to measure information gain. Due, to its enhanced precision, in measuring complexity it constructs trees that outperform information gain. While (Mishra et al., 2019) A novel approach was employed to distinguish between lymphocytes and diseased lymphoblasts. This involved extracting features, from the nucleus and cytoplasm areas, which were then used as input, for a classification algorithm. We have extracted a total of 21 characteristics from each image of leukemia. These features include attributes related to color and shape such, as elongation, area, perimeter and roundness (Inbarani et al., 2020).

According to reference (Kazemi et al., 2016), Selecting the features is a stage as it has a substantial influence, on the

DOI:http://doi.org/10.24086/cocos2024/paper.1529

classifiers effectiveness. Based on the guidance, from hematologists it has been noted that when developing a feature set it is important to consider characteristics that lead to categorization. Hematologists have identified traits when examining blood samples, including definition (HD) texture features, the ratio of nucleus to cytoplasm (N;C ratio) irregularities, shape features and color characteristics. We collected the extracted features from photographs, within our system (Amin et al., 2015) employed a two-step process: In the phase a set of attributes is formed by extracting information, from nuclei using a process called "feature generation." The subsequent step involves selecting the combination of traits that optimize the efficiency of recognition In (Dasariraju et al., 2020), The goal of feature extraction was to gather a set of characteristics that could be used to categorize types of blood cells. Each image generated 16 features, which were then categorized into four groups; nucleus size, nucleus shape, elliptical properties and color characteristics.

VI . K-MEANS CLUSTERING FOR LEUKEMIA FEATURES

Using K Means clustering a technique, for grouping data points can be applied within a data analysis pipeline to detect leukemia. This becomes particularly handy when dealing with gene expression records or other datasets that possess a number of dimensions.Initialization: Randomly initialize cluster centroids.

- Assignment: The gene expression profile should be allocated to the cluster whose centroid is closest to it.
- Update Centroids: To recalibrate the centroids we should make use of the gene expression profile, within each cluster.
- Repeat: Continue iterating through the assignment and centroid update phases until convergence is achieved.

VII. DISCUSSION

The suggested technique by (Vogado et al., 2018) The improvements could greatly impact the results of each classifier. Both SVM and KNN exhibit levels of accuracy demonstrating their performance as separate classifiers. The classifiers, namely MLP, RF, KNN, and SVM, achieved accuracies of 98.93%, 98.40%, 99.20%, and 99.20%, respectively. (Mishra et al., 2019) The classifier received 35 reduced characteristics and an output class level for classification. Adjusting the number of trees and selecting attributes for each node determines the optimal settings for random forest classification. The study revealed that using RF with 35 trees yields satisfactory results. The ADBRF classifier achieves a performance metric of 100% for 5-fold cross-validation.

The performance evaluation of the HSCRKM method in (Inbarani et al., 2020) achieved a 93% classification accuracy by utilizing the LR, NN, and DT algorithms. The NB, KNN, and RF accuracy scores are, in that order, 84%, 85%, and 86%. Furthermore, it is noteworthy to mention that the SVM attains the least accurate result, standing at 84%. In general, the HSCRKM algorithm achieves an accuracy rate of 82%. The accuracy rates attained by the proposed methodology are as follows: 13% for k-means clustering, 5% for FCM, and 4% for PSO-based clustering. For optimal performance, precise segmentation is required. The experimental findings demonstrate that the HSCRKM algorithm partitions the nucleus effectively. According to the analysis of the literature review, a considerable number of authors achieve an accuracy rate surpassing 90%. However, for the research, an extremely limited number of photographs are being utilized. In this study, we evaluate the efficacy of the proposed HSCRKM method through the examination of approximately 350 images.

The technique proposed by (Kazemi et al., 2016) attained a sensitivity of 95%, a specificity of 98%, and an accuracy of 96% for the binary SVM classifier. The accuracy achieved by the SVM classifier is 87%, which indicates a reliable and effective method, for categorizing AML and its various subtypes. By employing the classification method, we were able to attain accuracy rates of 98%, 95% and 97%, for sensitivity, specificity and overall accuracy respectively using a SVM classifier (Amin et al., 2015). The researchers employed an SVM classifier to differentiate among reactive, atypical, normal as well as L1 L3 cells as other cell kinds. The results obtained through this method were, in line, with what was anticipated. Once we finished training the forest model we enhanced its effectiveness by investigating the predefined performance indicators (Dasariraju et al., 2020). We employed the forest approach to assess how well the models can accurately distinguish among kinds of blood cells, in both the training as well as testing datasets, for binary classification. During the testing phase the machine accurately identified whole the images. Achieved an accuracy rate of 92.99% after training, on the dataset. The testing set indicated values of accuracy, recall as well as specificity that were, than 90%.

CONCLUSION

This investigation seeks to give an overview of the progress, in classifying as well as identifying Acute Lymphoblastic Leukemia (ALL) utilizing machine learning as well as other learning approaches. To ensure a diagnosis of Acute Lymphoblastic Leukemia (ALL) various patterns of segmentation methods, for extracting features and approaches, to classification were thoroughly examined. Unlike segmentation tasks, which often need the use of unsupervised machine learning technologies, classification difficulties are

DOI:http://doi.org/10.24086/cocos2024/paper.1529

more likely to favor supervised techniques, as shown by the findings of our study. On the other hand, deep learning, and more especially transfer learning, has emerged as the technique of choice for the automatic and more reliable classification of ALL conditions. Because it generates conclusions that are acceptable even with very little datasets, this is the reason why it is becoming more popular. The fact that we have previously explored the challenging issues and potential future directions for the study of this topic is another fascinating feature of the situation. Scholars will find this study to be helpful in evaluating the recent advancements in ALL detection, and the authors feel that it will encourage further research to be conducted in this respective field.

REFERENCES

1. Vogado, L.H.S., Veras, R.M.S., Araujo, Flavio.H.D., Silva, R.R.V. and Aires, K.R.T. (2018). Leukemia diagnosis in blood slides using transfer learning in CNNs and SVM for classification. Engineering Applications of Artificial Intelligence, 72, pp.415–422. doi:https://doi.org/10.1016/j.engappai.2018.04.024.
2. Mishra, S., Majhi, B. and Sa, P.K. (2019). Texture feature based classification on microscopic blood smear for acute lymphoblastic leukemia detection. Biomedical Signal Processing and Control, 47, pp.303–311. doi:https://doi.org/10.1016/j.bspc.2018.08.012.
3. Inbarani H., H., Azar, A.T. and G, J. (2020). Leukemia Image Segmentation Using a Hybrid Histogram-Based Soft Covering Rough K-Means Clustering Algorithm. Electronics, 9(1), p.188. doi:https://doi.org/10.3390/electronics9010188.
4. Kazemi, F., Najafabadi, T.A. and Araabi, B.N. (2016). Automatic recognition of acute myelogenous leukemia in blood microscopic images using k-means clustering and support vector machine. Journal of medical signals and sensors, 6(3), p.183.
5. Amin, M.M., Kermani, S., Talebi, A. and Oghli, M.G. (2015). Recognition of acute lymphoblastic leukemia cells in microscopic images using k-means clustering and support vector machine classifier. Journal of medical signals and sensors, 5(1), p.49.
6. Dasariraju, S., Huo, M. and McCalla, S. (2020). Detection and Classification of Immature Leukocytes for Diagnosis of Acute Myeloid Leukemia Using Random Forest Algorithm. Bioengineering, [online] 7(4), p.120. doi:https://doi.org/10.3390/bioengineering7040120.
7. Ranjitha, P. and Duth, S. (2021). Detection of blood cancer-leukemia using k-means algorithm. In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 838-842). IEEE.
8. Kumar, S., Mishra, S., Asthana, P. and Pragya. (2018). Automated detection of acute leukemia using k-mean clustering algorithm. In Advances in Computer and Computational Sciences: Proceedings of ICCCCS 2016, Volume 2 (pp. 655-670). Springer Singapore.
9. Fatma, M. and Sharma, J. (2014). Leukemia image segmentation using K-means clustering and HSI color image segmentation. Int. J. Comput. Appli, 94, pp.6-9.
10. Das, P.K. and Meher, S. (2021). An efficient deep convolutional neural network based detection and classification of acute lymphoblastic leukemia. Expert Systems with Applications, 183, p.115311.
11. Lin, T.C., Liu, R.S., Chao, Y.T. and Chen, S.Y. (2013). Classifying subtypes of acute lymphoblastic leukemia using silhouette statistics and genetic algorithms. Gene, 518(1), pp.159-163.
12. Zong, N., Adjouadi, M. and Ayala, M. (2005). Artificial neural networks approaches for multidimensional classification of acute lymphoblastic leukemia gene expression samples. In Proceedings of the 9th WSEAS International Conference on Computers (pp. 1-6).
13. Ross, M.E., Zhou, X., Song, G., Shurtleff, S.A., Girtman, K., Williams, W.K., Liu, H.C., Mahfouz, R., Raimondi, S.C., Lenny, N. and Patel, A. (2003). Classification of pediatric acute lymphoblastic leukemia by gene expression profiling. Blood, 102(8), pp.2951-2959.
14. Escalante, H.J., Montes-y-Gómez, M., González, J.A., Gómez-Gil, P., Altamirano, L., Reyes, C.A., Reta, C. and Rosales, A. (2012). Acute leukemia classification by ensemble particle swarm model selection. Artificial intelligence in medicine, 55(3), pp.163-175.
15. Foran, D.J., Comaniciu, D., Meer, P. and Goodell, L.A. (2000). Computer-assisted discrimination among malignant lymphomas and leukemia using immunophenotyping, intelligent image repositories, and telemicroscopy. IEEE Transactions on Information Technology in Biomedicine, 4(4), pp.265-273.
16. Piuri, V. and Scotti, F. (2004). Morphological classification of blood leucocytes by microscope images. In 2004 IEEE International Conference onComputational Intelligence for Measurement Systems and Applications, 2004. CIMSA. (pp. 103-108). IEEE.
17. Scotti, F. (2005). Automatic morphological analysis for acute leukemia identification in peripheral blood microscope images. In CIMSA. 2005 IEEE International Conference on Computational Intelligence for Measurement Systems and Applications, 2005. (pp. 96-101). IEEE.
18. Patel, N. and Mishra, A. (2015). Automated leukaemia detection using microscopic images. Procedia Computer Science, 58, pp.635-642.
19. Srisukkham, W., Zhang, L., Neoh, S.C., Todryk, S. and Lim, C.P. (2017). Intelligent leukaemia diagnosis with bare-bones PSO based feature optimization. Applied soft computing, 56, pp.405-419.
20. Su, J., Liu, S. and Song, J. (2017). A segmentation method based on HMRF for the aided diagnosis of acute myeloid leukemia. Computer methods and programs in biomedicine, 152, pp.115-123.
21. Ahmed, N., Yigit, A., Isik, Z. and Alpkocak, A. (2019). Identification of leukemia subtypes from microscopic images using convolutional neural network. Diagnostics, 9(3), p.104.
22. Das, P.K., Diya, V.A., Meher, S., Panda, R. and Abraham, A. (2022). A systematic review on recent advancements in deep and machine learning based detection and classification of acute lymphoblastic leukemia. IEEE access.
23. Shah, A., Naqvi, S.S., Naveed, K., Salem, N., Khan, M.A. and Alimgeer, K.S. (2021). Automated diagnosis of leukemia: a comprehensive review. IEEE Access, 9, pp.132097-132124.
24. Anghel, A., Stanisavljevic, M., Andani, S., Papandreou, N., Rüschoff, J.H., Wild, P., Gabrani, M. and Pozidis, H. (2019). A high-performance system for robust stain normalization of whole-slide images in histopathology. Frontiers in medicine, 6, p.193.
25. Das, P.K., Jadoun, P. and Meher, S. (2020). Detection and classification of acute lymphocytic leukemia. In 2020 IEEE-HYDCON (pp. 1-5). IEEE.
26. Das, P.K., Meher, S., Panda, R. and Abraham, A. (2021). An efficient blood-cell segmentation for the detection of hematological disorders. IEEE Transactions on Cybernetics, 52(10), pp.10615-10626.

DOI:http://doi.org/10.24086/cocos2024/paper.1529