Turkish Journal of Computer and Mathematics Education Vol.12 No.9 (2021), 2128– 2135

Research Article

CML: Chronic Myeloid Leukemia Detection Using Particle Swarm Optimization and Fuzzy C-Means Clustering

Dhanya M [a] and Dr. N. Kavitha [b]

[a] Research Scholar, Department of Computer Science, Nehru Arts and Science College, Coimbator
[b] Head of the Department and Associate Professor, Department of Computer Science, Nehru Arts and Science College, Coimbatore

Article History: Received: 10 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 20 April 2021

Abstract: Leukemia is a blood cancer caused by a rise in the number of white blood cells in your body. Those white blood cells crowd out the red blood cells and platelets that your body needs to be healthy. The extra white blood cells don't work right. However, the macroscopic examination of human blood necessitates various microscopic techniques, including color imaging, segmentation, and clustering, to distinguish patients with the disease. The white blood cell count can be higher than normal and prevent the immune system from functioning properly. The most reliable method to detect leukemia is with the assistance of a microscope. Automation of this form of Leukemia diagnosis is required because fewer and more expensive precise tests are required than the above. Blood slides don't reveal the findings because of hematologists' combined experience and tiredness, and those results can vary significantly from one PSO to another. As it is reasonably inexpensive and effective, blood-stained histopathology imaging to test for leukemia is a low-cost and innovative choice. The fuzzy-c segmentation coefficients of this paper are compared to those for image segmentation k. Using the GLCM to separate color characteristics from photographs performs better precision than using Fuzzy-classifications. Fuzzy's consistency level is 96%, while sharpness is 83% CNN is mainly a television outlet. The whole project has been done using the MATLAB toolkit.

Keywords*:* CML, PSO, CNN, Morphological, GLCM

Introduction

Image processing techniques are most widely used for detection of various medical diseases leukemia is one of the most interesting areas for researchers because it belongs to the category of blood cancer which can affect the persons of all ages starting from children to the old age people (Ratley, 2020). Leukemia often affects the bone marrow, the lymphatic node, or the blood/marrow system. This depletes blood cells and compromises the structure and functionality of tissues. There are four main varieties of blood cancer: lymphoma, Leukemia, myeloma, and chronic myelogenous leukemia. A lot of different Image processing techniques have been developed for Leukemia diagnosis. If accurate image segmentation is carried out, there may be problems with automatic haematology. In addition to hematological imaging, computer algorithms for classifying leukemia are often used in the areas of work in biological picture processing (Vengatesan, 2019). CML is a disease distinguished by the overproduction of white blood cells (leukocytes) in the bone marrow (the plural of leukocytes) and/amplification of the blood.

The human body's essential material is blood since it takes us around oxygen and nourishes the cells (Wang, 2013). Furthermore, it is involved in certain vital bodily processes, such as transporting oxygen, carbon dioxide, and nutrients to the entire body to sustain normal metabolic processes (Yurttaş, 2020). There are three primary contributors of blood: red blood cells, white blood cells, and platelets (Ananya, 2018) Without immediate medical care, the low blood flow could significantly hamper the metabolism, potentially life-threatening. Leukemia is one of the less common blood diseases. Most pediatric tumors are leukemia. As cancer begins in the body, it most often starts in body cells. Very commonly, as the body requires new cells, they will be shaped and mature. When a cell ages, it is replaced by a newer cell (Ekata, 2014). Often, these patterns fail to fit On presentation, patients' total age is between 45 to 55, but children are all diagnosed with CML. In chronic myeloid Leukemia (CML), the disease is present for an extended period (Jagadev, 2017). During the chronic process, CML usually follows a slow, steady course. Usually, the normal history of CML development from the chronic to the rapid and blast phases is three to five years (Kumar, 2017). The authors examined CML by using clinical characteristics, molecular pathology, care, the disease's pathophysiology, treatment options, and its development over the years, all in CML, as described in (Raje, 2013). The paper is organized as follows: Section II consists of a literature survey, section III describes a proposed method, section IV contains results and discussion, and section V is the conclusion (Shaikh, 2017; Mandal, 2019).

2. Background Study

(Ananya, 2018) chronic myeloid leukemia develops when granulocytes, white blood cells, are found in the human body but are not under strict supervision. It mutates the genetic sequences, which culminates in the development of different genetical mutations. It gives the wrong template, which results in CML. Nationally and

pathologically with hematologically, to study and understand the different stages of the disease, the authors were analyzed each patient in the different stages of the disease to determine how the disease will progress with (CML). When the authors were studying the combination divide and conquer algorithm and the miles-per-hour estimate, the authors found that they fit well together. Our approach suggested the most efficient method that needs less time and space complexity. They analyzed the reports on astride, from stage to stage, to find the dose.

(Chang, 2009) To gain a uniform control strategy for disease dynamics, the authors discussed a unified methodology for drug scheduling. To use in CML dynamics simulations, the authors have proved its ability and centered on the fact that an intuitive approach is used to improve the framework.

(Fatma, 2014) The technician has suggested a classification technique for classifying Leukemia pictures. An analysis of fifty Leukemia photographs includes both the statistical and color elements. A function dataset is extracted and used as an input to a neural network for Classification is built from it. In both usual and irregular smear experiments, the feature extraction feature was introduced. The extracted function values depend on the nature of the images they are pulled. The amount of training datasets governs the Classifier's classification accuracy. All turned out according to schedule. As the dataset grew, the findings improved to 92% of the Classification.

(Kumar, 2017) Things have become too sophisticated in the last twenty years; everything is automated. This thesis suggests a tool to classify AML in a blood smear automatically. images are made into three different levels: [additive, luminance, and subtractive]. The nucleus can be obtained, and they are further divided into these three types: [additive, light, luminance, and luminance darkening] in the segmentation phase.

(Raje, 2014) The primary objective of this research is to fractionate blood cells into their respective nuclei and features to detect leukemia. As nuclei form, for accuracy, attributes, such as field, perimeter, and so on, are considered. This study shows that mean and standard deviation-based image segmentation and Otsu's quantization performed well.

(Zhang, 2012) Dynamics of CML progression and mutations are part of our investigation. The authors have initially implemented a new mathematical calculation to help solve certain established mathematical models' Finally. The authors proved that our model was realistic and the calculations are feasible. The authors were also found that our model describes CML progression dynamics and the formation of mutations as accurately as possible. The authors were must stabilize our model, e.g., at the equilibrium. Let me put it to you this way: it's better before or after therapy, depending on your position.

3. System Model

The suggested approach prioritizes the input to improve the picture quality by applying it to preprocess. At this stage, WBC dissection is done to separate the cluster. For each segmented WBC image, statistical and color-based features are extracted for each element. This information is applied to the Classifier for categorizing Leukemia types. The technique is laid out in a numbered series, starting with number one.

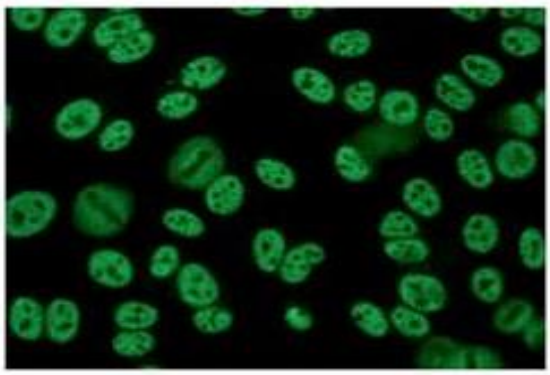

Figure 1. Input Color Image

Figure 1 illustrates the input blood cell image

3.1 Pre-processing

Leishmen stain, Wright's stain, and Giemsa stain are used to hide the blood smears. Digital microscopes typically produce pictures in the color space of RGB. This is moderately boring and very challenging to disaggregate. Finally, because of changing lighting, age stain, and different camera settings, the picture is subject to image bleaching and variation in darkness.

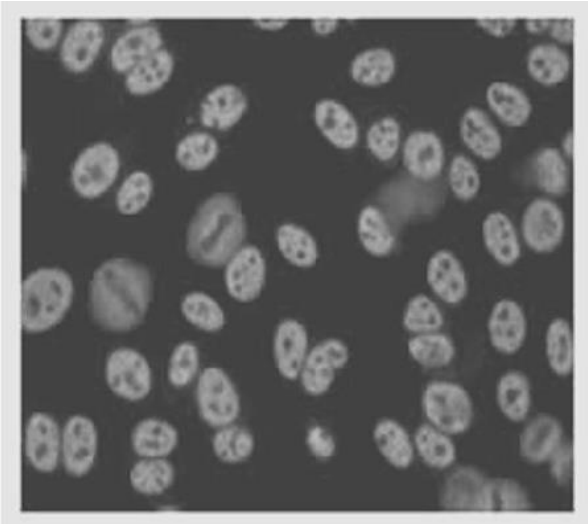

Figure 2. Grayscale Image Processing

Figure 2 illustrates the Grayscale image processing with reshaping and resizing, and sharpening with the filter value

3.2 Segmentation

To select the main information from the provided image, segmentation is needed. We found that image segmentation plays a significant role in its capabilities of feature extraction and Classification. A clustering algorithm, called K-means, has been hired. This data-retrieval action algorithm is particularly useful for broad and highly dimensional datasets. Nucleation of the white cells is extracted from images using color segmentation in this article.

3.3 Fuzzy C means for image segmentation

The FCM c-means algorithm generalizes and simultaneously claims a point to belong to several clusters. May yield a partition that is on the gentler side. The FCM clustering algorithm has been commonly used for image segmentation with magnification. Retaining knowledge is more advantageous in the FCM approach to K-Means segmentation than rough segmentation. Clustering images that contain large quantities of data offers notable advantages, especially for images with several pixels. FCM is specially assigned pixels to fuzzy clusters without a defined classification. Unlike the K-Means classification, which does not accommodate non-overlapping clusters, Fuzzy C-me uses degrees as the class assignment criteria. A lower- however, points near the outer edge of a cluster are less associated with the core of the distribution than points in the center. However, it has several major disadvantages, including numerical difficulty and decreases dramatically in performance when faced with noise. single data may belong to multiple clusters

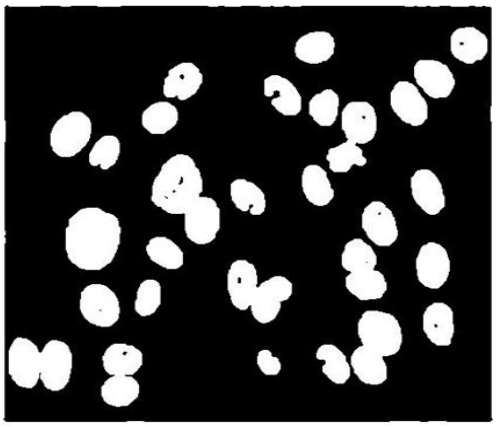

Figure 3. Morphological Operation

3.4 Feature Extraction

The pattern of features in a picture provides useful insight into what the image depicts. In other words, picture prioritization is finding the appropriate elements and discarding the extraneous ones. Preprocessing is useful for picture clarity but does not improve the detail that can be extracted from it. Various undesirable or undesired distortions and detail and features may be put into a picture to make it more suitable for different post-processing techniques. It is important to decide what features are needed and the appropriate ones to focus on until starting the feature extraction method. It is difficult to select the 'the best way' to feature extraction, but features should be extracted that communicate sufficiently valid information about the picture. It must be possible to incorporate these features in computing without much trouble. Computer models of human characteristics and preferences must be able to adapt well to human perceptions and can be applied to a broad range of images.

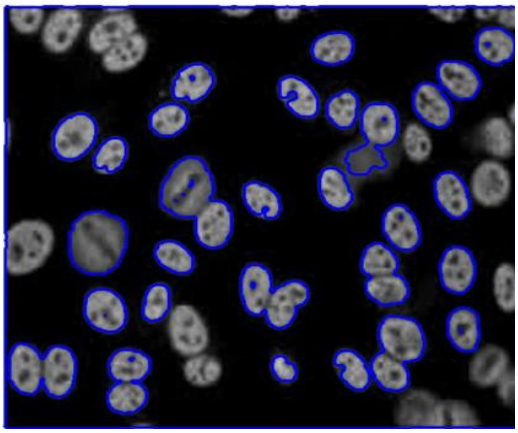

Figure 4. Identifying the Affected and Normal Blood Cell Separately

Figure 4 represents the classified affected and normal blood cells separately.

3.5 Neural Network and Supervised Learning

A neural network was originally devised to store and manipulate information, but it derives its internal structure from the nervous system mechanisms (as the brain does). An CNN is made up of many elements called axons, which are linked to each other in a highly interactive manner. A nerve cell is like a computer CPU, working like a single computing entity (Fatma, 2014).

Algorithm: 1: Preprocessing

Step 1: Image collection for the blood cell
Step 2: Preprocessing
Step 3: Color image to grayscale conversion
Step 4: Resizing the image
Step 5: Reshaping the image to the actual contour
Step 6: Sharpening the image
Step 7: Gaussian filter to make the blood cell more visible

Algorithm: 2: Binary Formation

Step 1: Structuring element of the image
Step 2: Morphological image processing
Step 3: Region max thresholding
Step 4: Cache thresholding
Step 5: Open/close image operation
Step 6: Erosion/dilation operation

Algorithm: 3: Clustering

Step 1: FCM (fuzzy c means) clustering with radial thresholding
Step 2: Boundary box calculation
Step 3: Contorur and cell part identification
Step 4: Optimization with **Particle Swarm Optimization (algorithm)**

Step 5: Finding the local and global best fitness data
Step 6: Error reduction of the data with the clustered region
Step 7: Optimized centroid is identification with the region co-occurrence level

Algorithm: 4: Feature Extraction

Step 1: GLCM (gray level co-occurrence)
Step 2: other such features are (area, mean, std, variance)
Step 3: kurtosis, skewness, perimeter

Algorithm: 5: Classification using CNN

Step 1: Different types of category (eg:'centromere', 'golgi', 'homogeneous', 'nucleolar', 'nuclear_membrane', 'speckled')
Step 2: which category of the data under Classification (neural networks)
Step 3: Resulting with good or affected

4. Results and Discussion

Performance in section III were effectively extracted. These images have unique characteristics that were taken into consideration when creating the datasets. A multilayer feed-forward neural network was trained and tested to identify leukemia according to its form, and that method was found to be more efficient. There are fifty different input Leukemia pictures,

Figure 5. Confusion Matrix

Figure 5 represents the confusion matrix about Training and Testing confusion matrix

For instruction, 38 images were used; while testing, we only used 12. The performance of the method was estimated from twenty-one experiments, on average. The study showed that the machine got much better at Classification with more results.

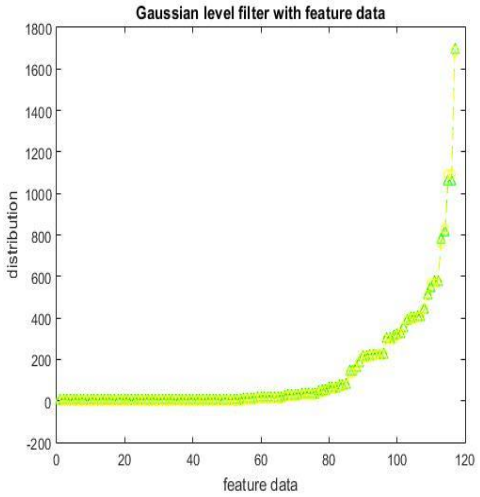

Figure 6. Performance Graph

Figure 6 represents the proposed performance graph; in X-axis denotes the feature data, and Y-axis denotes the distribution.

The Classifier's consistency is tested using an accuracy parameter. Four potential conclusions can be derived from a classifier analysis: Where there are cancerous images, the right ones are known as TN, when all the noncancerous ones are considered to be incorrect, and where all the noncancerous ones are accurately defined as noncancerous, this is called a true negative (TN) picture.

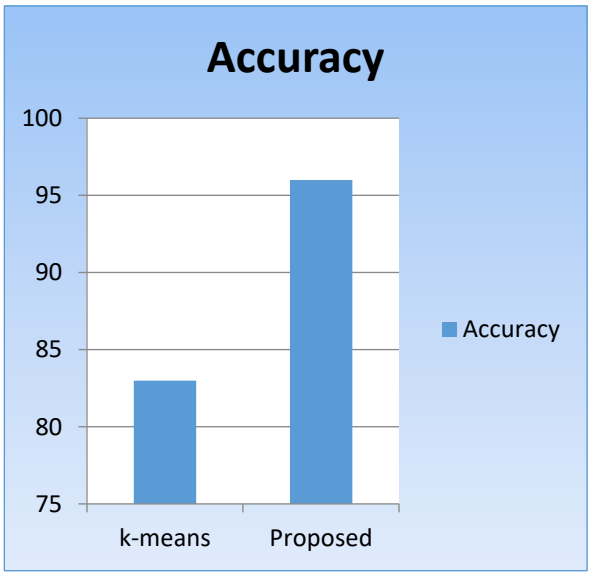

Figure 7. Accuracy Comparison Chart for K-means and Proposed Method

In figure 7 represents the accuracy comparison between the k-means algorithm and combining FCM, CNN.

Accuracy: The closeness of the output to the real value is defined as accuracy. The accuracy of the system running both CNN with the full feature set and CNN with a reduced feature set is 96%.

$$\text{Accuracy}=\frac{T_{p+T_N}}{T_p+T_{N+T_P+T_N}} \qquad (1)$$

Thus, using a reduced feature set to classify the image, the time taken is less than the existing systems, while the accuracy remains unchanged.

5. Conclusions

Since leukemia affects the blood cells, it is a difficult illness to treat. It has the potential to lead to cause death if left unchecked. As it is reasonably inexpensive and effective, blood-stained histopathology imaging to test for leukemia is a low-cost and innovative choice. Due to this article, segmentation and fuzzy-c are compared. The

name of the Fuzzy/PSO classification optimization approach states that features are optimized first extracted using the Fuzzy method, and then Classifier is used to extract color features. It's is paired with Blurry C math, so it has a mean. With the derived features in the Periphery, density, and percentage of segmented nuclei, the machine learning model accurately classifies the tissue samples. The proposed scheme helps ensure that Leukemia pictures are segmented correctly. Thus, a broader system was developed that can identify and classify the WBCs quickly with less computing time and lower error rate; full-scale automated systems may be used. It may be applied to additional forms of cancers in the future, perhaps. Fuzzy's consistency level is 96%, while sharpness is 83%. The results are divided by segmentation, and color and form features are found Once the values are gathered, they are included in the CNN classifier. Classifies the segmented picture according to its characteristics and outputs the classifications.

References

1. Ananya, B., Prabisha, A., & Kanjana, V. (2018). *Novel Approach to Find the Various Stages of Chronic Myeloid Leukemia Using Dynamic Short Distance Pattern Matching Algorithm. 2018 3rd International Conference for Convergence in Technology (I2CT), pp. 1-5.*
2. Bhasha, A.C., Balamurugan, K. End mill studies on Al6061 hybrid composite prepared by ultrasonic-assisted stir casting. Multiscale and Multidiscip. Model. Exp. and Des. (2020). https://doi.org/10.1007/s41939-020-00083-1
3. Chang, H., & Astolfi, A. (2009). *Enhancement of the immune response to chronic myeloid leukemia via controlled treatment scheduling. 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 3889-3892.*
4. ChinnamahammadBhasha, A., Balamurugan, K. Studies on Al6061nanohybrid Composites Reinforced with SiO_2/3x% of TiC -a Agro-Waste. Silicon (2020). https://doi.org/10.1007/s12633-020-00758-x
5. Ekata, Tyagi, P. K., & Singhal, H. (2014). *Neurofuzzy inference system for diagnosis of leukemia. 2014 Innovative Applications of Computational Intelligence on Power, Energy and Controls with Their Impact on Humanity (CIPECH).*
6. Fatma, M., & Sharma, J. (2014). *Identification and Classification of acute Leukemia using neural network. 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems(MedCom), pp. 142-145.*
7. Jagadev, P., & Virani, H. G. (2017). *Detection of Leukemia and its types using image processing and machine learning. 2017 International Conference on Trends in Electronics and Informatics (ICEI), pp. 522-526*
8. Kumar, P., & Udwadia, S. M. (2017). *Automatic detection of acute myeloid leukemia from microscopic blood smear image. 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pp. 1803-1807.*
9. Latchoumi, T.P. and Kannan, V.V., 2013. Synthetic Identity of Crime Detection. International Journal, 3(7), pp.124-129
10. Loganathan, J., Latchoumi, T.P., Janakiraman, S. and parthiban, L., 2016, August. A novel multi-criteria channel decision in co-operative cognitive radio network using E-TOPSIS. In *Proceedings of the International Conference on Informatics and Analytics* (pp. 1-6). https://doi.org/10.1145/2980258.2982107
11. Laurino, M., Stano, M., Betta, M., Pannocchia, G., & Landi, A. (2013). *Combining pharmacological therapy and vaccination in Chronic Myeloid Leukemia via model predictive control. 2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 3925-3928.*
12. Mandal, S., Daivajna, V., & V, R. (2019). *Machine Learning based System for Automatic Detection of Leukemia Cancer Cell. 2019 IEEE 16th India Council International Conference (INDICON), pp. 1-4.*
13. Raje, C., & Rangole, J. (2014). *Detection of Leukemia in microscopic images using*

image processing. 2014InternationalConferenceonCommunicationandSignal Processing, pp.255-259.
14. Ratley, A., Minj, J., & Patre, P. (2020). *Leukemia Disease Detection and Classification Using Machine Learning Approaches: A Review. 2020 First International Conference on Power, Control and Computing Technologies (ICPC2T), pp. 161-165.*
15. Shaikh, M. B. N., & Deshpande, S. (2017). *Computer aided leukemia detection using digital image processing techniques. 2017 2nd IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), pp. 344-348*
16. Vengatesan, K., Kumar, A., Karuppuchamy, V., Shaktivel, R., & Singhal, A. (2019). *Face Recognition of Identical Twins Based On Support Vector Machine Classifier. In 2019 Third International conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC), pp. 577–580.*
17. Wang, K., Cheng, J., Tan, Q., & Sun, D. (2013). *Probing mechanical behaviors of chronic myeloid leukemia cells in doxorubicin resistance by robotic manipulation with optical tweezers. 2013 13th IEEE International Conference on Nanotechnology (IEEE-NANO 2013),pp. 191-194*
18. Yurttaş, N. Ö., & Eşkazan, A. E. (2020). *Novel therapeutic approaches in chronic myeloid leukemia. Leukemia Research, 91, pp.106337.*
19. Zhang, B., Tanaka, G., Chen, L., & Aihara, K. (2012). *Dynamical modeling of chronic myeloid leukemia progression and the development of mutations. 2012 IEEE International Conference on Bioinformatics and Biomedicine Workshops, pp. 130-135.*