AN EFFICIENT METHOD TO DETECT LEUKEMIA USING ARTIFICIAL INTELLIGENCE

Nikita Sharma, Shivali Sharma

Abstract

Leukemia is the blood cancer that initiates in the bone marrow and resulted in abnormal cells generation. Leukemia is broadly classified into ALL (Acute lymphoblastic Leukemia), AML (Acute myeloid Leukemia), CLL (chronic lymphocytic Leukemia) and CML (Chronic myeloid Leukemia). This research has make an effort for draw a methodology for the classification and detection of leukemia. Work done by number of authors has been analyzed and presented in this research. For the classification, fuzzy logic has been used and for the optimization, genetic algorithm has been used. Parameters, like FAR, FRR and accuracy has been used for the calculation of the performance.

Keywords: Leukemia detection, Genetic algorithm, Fuzzy logic, FAR, FRR and accuracy

I. Introduction

The significant part of human body is blood as it retains one alive. It performs a number of crucial functions like transferring the carbon dioxide, oxygen, mineral and so on for growing the body in order to sustain metabolism. Blood comprises of three major components, particularly, WBC (White blood cell), RBC (Red Blood Cell), and Platelets. Insufficient blood may infect the metabolism that may be very precarious if early treatment is not taken. One of the common blood disorders is Leukaemia. Leukaemia is the general type of cancer in children and sources are present in blood cells [1].

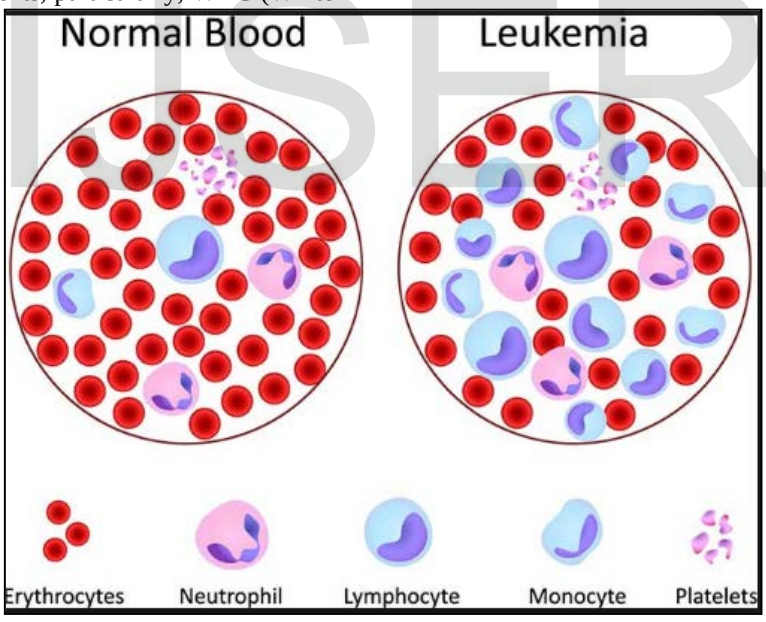

Figure 1: Leukaemia disease

Leukaemia is nothing but cancer of blood cells in which number of white cells is increasing and those are immature cells that destroy other cells [2]. Leukemia is broadly classified into four types:

i. Acute leukemia
ii. Chronic leukemia
iii. Lymphoid leukemia
iv. Myeloid leukemia

The description for the same is depicted in below table [3]:

Table 1: Leukemia types

Leukemia types	Description
Acute	• Speedy growth of immature blood cells • Occur usually in young adults and children • Requires on-time treatment
Chronic	• Build mature blood cells excessively • Occur generally in older patients • Examining before the treatment
Lymphoid	• Effects plasma as well as lymphocytes cells • It is lymphocytic leukemia
Myeloid	• Effects basophils, eosinophils and neutrophils • It is myelogenous leukemia

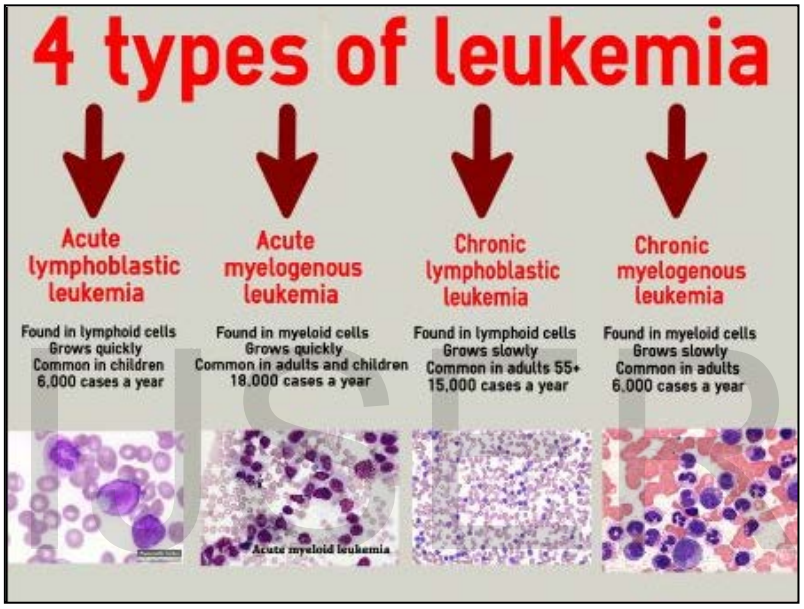

Figure 2: Leukemia Types

The satisfactory causes of leukemia are unidentified and in most case, it's unsettled why leukemia has developed [4]. Research into possible causes is going on all the time. Like other cancers, leukemia isn't transferable and can't be approved on to other people. There is integer of factors that may amplify a person's risk of budding leukemia [5]. Having a scrupulous hazard factor doesn't denote you will definitely get this category of disease and personnel lacking any recognized risk factors can still develop it. The recognized risk factor of generating this type of cancer i.e. leukemia are clarified here [6].

i. **Exposure to radiation:** People who exposed to high level of release, such as nuclear developed accidents, have the main risk of developing leukemia than people who have not been exposed. On the other hand, a small numeral of people in the UK will be uncovered to emission levels high adequate to augment their risk.

ii. **Smoking:** Smoking increases the risk of initial leukemia. This may be due to the intense levels of benzene in cigarette smoke.

iii. **Exposure to benzene:** In very unusual cases, leukemia may begin due to the long term contact to benzene (and possibly other solvents) used in industry.

iv. **Cancer treatments:** Now and then, a few anti-cancer treatments such as chemotherapy or radiotherapy can be a basis for leukemia to build up after some years of this behaviour. The risk increase when persuaded types of chemotherapy drugs are mutual with radiotherapy. While leukemia develops since of earlier anti-cancer treatment, this is called lower leukemia or treatment related leukemia.

v. **Blood disorders:** People with certain blood disorders, such as myelodysplasia or myeloproliferative disorders have a distended risk of initial AML.

vi. **Genetic disorders:** People with a certain hereditary disorder, excluding Down's syndrome and Franconia's anaemia, have an inflated risk of embryonic leukemia.

From all leukemia types as discussed above, the symptoms are generally caused by the normal blood cells lacking as compare to the abnormal white cells presence [7]. Because of more number of leukemia cells in the bone marrow, it become unable to occur more number of blood cells that are needed by the body. The symptoms of leukemia are [8]:

i. Poor blood clotting
ii. Affected immune system
iii. Anaemia
iv. Patients might also feel nausea, fever, chills, night sweats, flu-like symptoms, weight loss, bone pain, and tiredness. If the liver/spleen grows unnecessarily the patient might feel full and will eat less, following in weight loss.
v. Weight loss could also occur independently of hepatomegaly (enlarged liver) or splenomegaly (enlarged spleen). A headache is more popular among patients whose harmful cells have invaded the central nervous system (CNS).As all these signs could be due to other diseases, a diagnosis of leukemia could only be verified after medical tests are conducted out.

In order to improve patient diagnosis, various image processing software are developed to extract useful information from medical images [9]. An essential part of the diagnosis and treatment of leukaemia is the visual examination of the patient's peripheral blood smear under the microscope. Morphological changes in the white blood cells are commonly used to determine the nature of the malignant cells, namely blasts. Morphological analysis of blood slides are influenced by factors such as haematologists experience and tiredness, resulting in non-standardized reports [10]. So, there is always a need for a cost effective and robust automated system for leukaemia screening which can greatly improve the output without being influenced by operator fatigue. Earlier various methods have been introduced as studied in literature review but these techniques have not provided good accuracy with less error rate. So, to enhance the detection rate there is urgent need to develop that type of algorithm that will provide high accuracy with less error rate. So, this research work has presented an application for disease detection based on leukaemia cells. The work has used GA (Genetic algorithm) for feature extraction and the classification of cells is executed by using Fuzzy logic for the recognition as well as the normal cells differentiation from the blast cell. Parameters, namely, FAR (False Acceptance rate), FRR (False Rejection rate) and accuracy has been calculated for the detection of the leukemia [11].

II. A Glance of Existing techniques

Number of researchers has tried their hand for the detection of leukemia. This section has highlighted the eminent work covering the symptoms, types and the feature extraction algorithms with the classification method for the detection of the disease.

B. Rajwa et al. [2017] proposed a system that has been used to distinguished normal and abnormal samples along with the progress of disease on the basis of Flow cytometry (FC) bone marrow data. **S. Savvopoulos et al., 2016]** proposed a mathematical model presenting the B-cell dieses. For representing CLL five hypothetical cases have been taken. **Beatriz et.al [2016]** proposed a methodology for classifying DNA microarray. The proposed method performs a feature selection process based on a swarm intelligence algorithm to find a subset of genes that best describe a disease. After that, different ANN has been trained using the subset of genes. Finally, four different datasets have been used to validate the accuracy of the proposal and test the relevance of genes to correctly classify the samples of the disease. **Nimesh et.al [2015]** proposed automated approach of leukaemia detection. In a manual method of Leukaemia detection, experts check the microscopic images. This is lengthy and time taking process which depends on the person's skill and not having a standard accuracy. The automated Leukaemia detection system analyses the microscopic image and overcomes these drawbacks. It extracts the required parts of the images and applies some filtering techniques. K-mean clustering approach is used for white blood cells detection. The histogram equalization and Zack algorithm is applied for grouping white blood cells. Some of the features like mean, standard deviation, colour, area, perimeter, etc. are calculated for leukaemia detection. The SVM is used for classification. The proposed system is tested on image dataset and 93.57% accuracy is achieved. The proposed system is successfully implemented in MATLAB. **Himali et.al [2015]** discusses about methods for detection of leukemia. Various image processing techniques have been used for identification of RBC and immature white cells. A different disease like anemia, leukemia, malaria, deficiency of vitamin B12, etc. could be diagnosed accordingly. The main objective of the research is to detect the cells affected from Lukemia and count them. According to detection of immature blast cells, leukemia can be identified and also define that either it is chronic or acute. To detect immature cells, a number of

methods have been used like histogram equalization, linear contrast stretching, some morphological techniques like area opening, area closing, erosion, dilation. Watershed transform, K means, histogram equalization & linear contrast stretching, and shape based features are accurate 72.2%, 72%, 73.7 %, and 97.8% respectively. **Jyoti Rawat et.al [2015]** proposed a novel technique to differentiate ALL (acute lymphoblastic leukemia) lymphoblast cells from healthy lymphocytes. The technique first separate leukocytes from the other blood cells and then lymphocytes have been extracted. In this context, a novel computer aided diagnostic system (CAD) has been designed for detection of hematological disorders like leukemia (blood cancer) based on Gray level co–occurrence matrices (GLCM) and shape based features. The features thus extracted classified by the auto support vector machine (SVM) binary classifier to find the presence of lymphoblast cell (leukemic cells). GLCM texture feature with feature vector length 13 reveals, classification accuracies of 86.7% and 72.4% for cytoplasm and nucleus respectively while for shape based features illustrated, classification accuracies of 56.1% and 72.4% respectively for a feature vector length 11 in both regions of lymphocyte. The classification accuracy of the combined texture-shape feature achieved is 89.8% with feature vector length 37 which shows better results as compared to an individual. **Lokman et.al [2015]** proposed a computer-aided diagnosis system that can detect and classify leukemia from blood microscopic images. Satisfactory preliminary experimental results demonstrate the efficiency of the proposed system. **Tejashree et.al [2015]** proposed a method to detect leukemia which is a childhood cancer. Acute Lymphoblastic Leukemia (ALL) growth is faster so for the treatment of patient we need early detection. The traditional methods like Fluorescent in Situ Hybridization (FISH), Immunophenotyping, Cytogenetic analysis, and Cytochemistry are time-consuming. So for the fast detection, the proposed method is useful because the segmentation used in this method is an automated segmentation which is based on the Otsu‟s method. Thus the automated segmentation performs the important role for the detection of leukemia. The result of shape features also important for leukemia detection. The result of contour signature is used for feature extraction. Results obtained encourage the future work to develop automated segmentation system which is independent of stains used in blood smear image and which is helpful for segmenting overlapping cells.

III. Simulation model

This work has analysed the existed implemented features extraction as well as classification algorithms for the medical data processing. A new feature extpraction algorithm has been proposed in leukemia data set for obtaining features, termed as, average mean, correlation, intensity and maximum mean value. Fuzzy logic as a classification technique has been proposed for leukemia cells classification. For the optimization of the work, Genetic algorithm has been used. Parameters, like FAR, FRR and accuracy has utilized for performance calculation.

Genetic algorithm

$u \leftarrow 0$;

Init population [p(u)];

Eval population [p(u)];

While not termination do

P’ (u) ← Variation [P(u)];

Eval population [P’(u)];

P(u+1) ← Apply genetic operator [P’ (u)∪ T];\

u←u+1;

end while

The steps followed for the classification process are defined below:

Fuzzy logic

Identification of principal Inputm outpur and procedure tasks

Identification of linguistic variables being utilized and describe fuzzy set and membership function

Utilize the fuzzy sets and linguistic variable to produce the process rules

Verify the de-fuzzification method

Check the system and amend if required

The procedure via which the simulation has been carried out is explained below:

Step: 1 Upload dataset.
Step: 2 Training of uploaded dataset has been done on the basis of Genetic Algorithm.
Step: 3 Total data is stored in MATLAB database. From here, selection of data has been done for training purpose. Training is done on the basis fitness function

contained in genetic algorithm. Fitness function is applied to get the optimal solution from large pool of solutions. When training get completed then move towards classification process utilized by fuzzy logic.

Step: 4 Fuzzy Logic has been utilized for classification process

Step: 5 On the basis of training, Fuzzy Logic detects the disease.

Step: 6 In the end, parameter evaluation has been done using FAR (false acceptance rate), FRR (false rejection rate) and accuracy. FAR and FRR must be low as they are indirectly related with error rate of system. So error rate must be low in order to get high accuracy.

IV. Simulation Results

The results obtained after the simulation of the research are defined in this section. GA has been used for the optimization and fuzzy logic is being used for the classification of the diseases.

Table 2: Evaluation of performance parameters

Number of samples	FAR	FRR	Accuracy
1	0.0021521	0.00041795	99.6825
2	0.0020965	0.00041855	98.8232
3	0.0021352	0.00041698	99.6912
4	0.0021263	0.00040981	99.7455
5	0.0020925	0.00041795	98.8976
6	0.0022121	0.00042123	99.6925
7	0.0021181	0.00041882	99.7232
8	0.0021271	0.00041972	99.6836
9	0.0021093	0.00041792	99.6927
10	0.0021274	0.00041865	99.7521

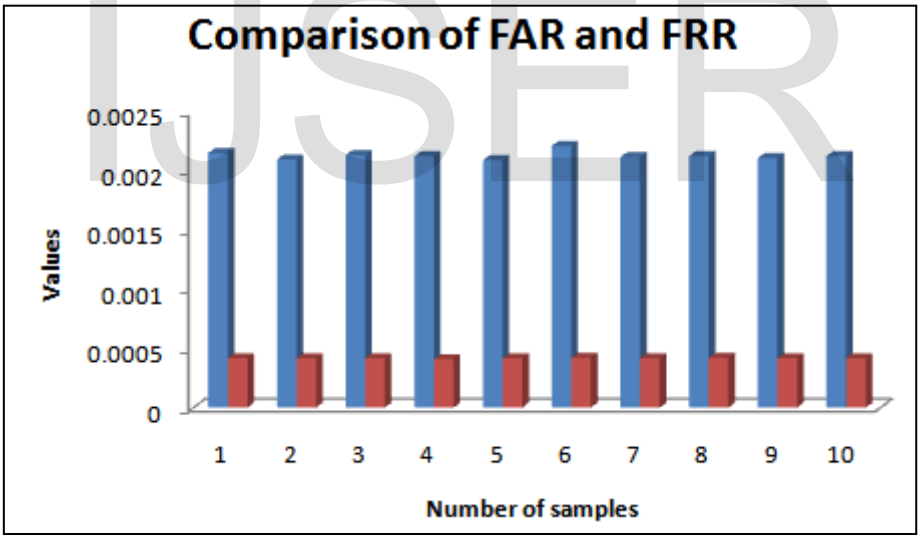

Figure 3: Comparison of FAR and FRR

Table 2 and figure 2 defines the values of FAR and FRR being obtained after the evaluation of the proposed work. As depicted, the X-axis defines the number of samples and Y-axis defines the obtained values. Red color bar is for FRR and blue bar defines the results of FAR. The average value of FAR 0.0021263 is 0.000041855 and is the average value of FRR.

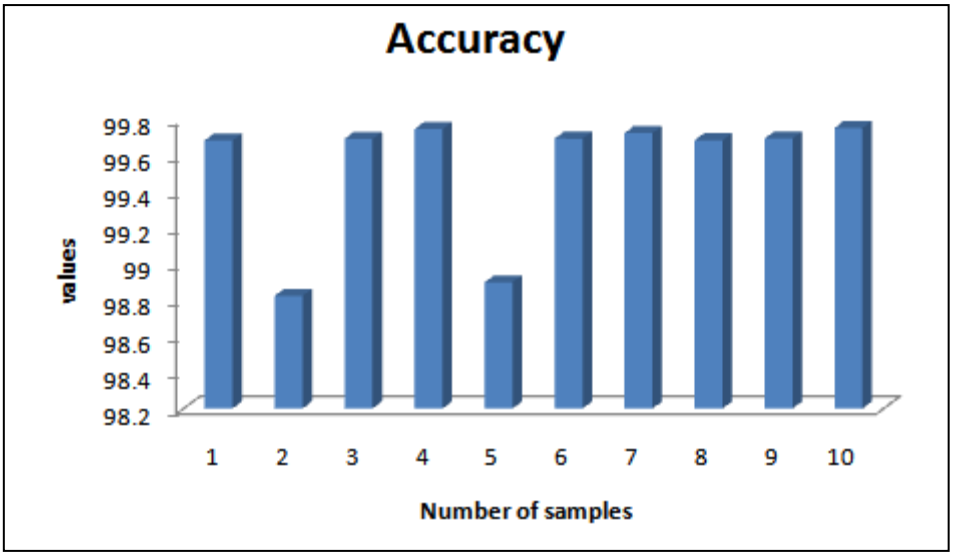

Figure 4: Obtained Accuracy

The results for the accuracy are shown in the above graph. The Number of samples are shown in the X-axis and obtained values of accuracy is shown in Y-axis. The average value for Accuracy came out to be 99.7455.

V. Conclusion

Leukemia is known as a blood disease with abnormal WBC (White blood cells)that occurred in large scale very quickly because of which WBC cannot execute its functions appropriately, therefore, it requires quick detection. Image analysis has a significant role in the detection and diagnosing of leukemia at initial stage. Genetic algorithm has been used for the optimization and Fuzzy logic for the optimization. Parameters, namely, FAR, FRR and accuracy parameters has been used for the calculating the performance for the detection of leukemia. The average value of FAR is 0.0021263, FRR is 0.00041855 and the average value of accuracy is 99.7455.

References

1. B. Rajwa, P. K. Wallace, E. A. Griffiths and M. Dundar, "Automated Assessment of Disease Progression in Acute Myeloid Leukemia by Probabilistic Analysis of Flow Cytometry Data," in IEEE Transactions on Biomedical Engineering, vol. 64, no. 5, pp. 1089-1098, May 2017.
2. S. Savvopoulos, R. Misener, N. Panoskaltsis, E. N. Pistikopoulos and A. Mantalaris, "A Personalized Framework for Dynamic Modeling of Disease Trajectories in Chronic Lymphocytic Leukemia," in IEEE Transactions on Biomedical Engineering, vol. 63, no. 11, pp. 2396-2404, Nov. 2016.
3. E. Pefani, N. Panoskaltsis, A. Mantalaris, M. C. Georgiadis and E. N. Pistikopoulos, "Chemotherapy Drug Scheduling for the Induction Treatment of Patients With Acute Myeloid Leukemia," in IEEE Transactions on Biomedical Engineering, vol. 61, no. 7, pp. 2049-2056, July 2014.
4. Beatriz A. Garro et.al, "Classification of DNA Microarrays Using Artificial Neural Networks And Abc Algorithm", Applied Soft Computing, vol.38, 2016.
5. Nimesh et.al, "Automated Leukaemia Detection Using Microscopic Images", Elsevier, Volume 58, 2015, Pages 635-642.
6. Himali. P, "Leukemia Detection using Digital Image Processing Techniques", International Journal of Applied Information Systems (IJAIS), vol. 10, no.1, November 2015.
7. Jyoti Rawat, "Computer Aided Diagnostic System for Detection of Leukemia Using Microscopic Images", Elsevier, vol. 70, pp. 748-756, 2015.
8. Lokman Faivdullah, Farid Azahar, Zaw Zaw Htike, and Wei Yan Nyein Naing, "Leukemia Detection from Blood Smears", vol. 4, no. 6, December 2015.
9. Tejashree et.al, "Blood Microscopic Image Segmentation & Acute Leukemia Detection", International Journal of Emerging Research in Management &Technology, vol.4, no.9, 2015.
10. El-Nasser, Ahmed Abd, Mohamed Shaheen, and Hesham El-Deeb, "Enhanced leukemia cancer classifier algorithm", Science and Information Conference (SAI), IEEE, 2014.

11. Zadeh, Hossein Ghayoumi, Siamak Janianpour, and Javad Haddadnia, “Recognition and Classification of the Cancer Cells by Using Image Processing and Lab View”, International Journal of Computer Theory and Engineering, 2013.
12. Xavier Thomas, “First contributors in the history of leukemia”, World J Hematol, vol.2, pp. 62-70, 2013.
13. Manisha Pokharel, “Leukemia: A Review Article”, IJARPB, Pokharel, IJARPB, vol.2, pp. 397-407, 2012.
14. L. H. Nee, M. Y. Mashor, R. Hassan, “White Blood Cell Segmentation for Acute Leukemia Bone Marrow Images”, International Conference on Biomedical Engineering (ICoBE), Penang, Malaysia, February 2012.
15. Kasmin, Fauziah, Anton Satria Prabuwono, and Azizi Abdullah, “Detection of Leukemia in Human Blood Sample Based on Microscopic Images: A Study”, Journal of Theoretical & Applied Information Technology, vol.46, 2012.
16. Nikumbh, Sarvesh, Shameek Ghosh, and Valadi K. Jayaraman, “Biogeography-based informative gene selection and cancer classification using SVM and random forests”, Evolutionary Computation (CEC), IEEE, 2012.
17. Xue-wen Chen and Michael McKee,” Finding expressed genes using genetic algorithms and support vector machines”, Department of Electrical and Computer Engineering, California State University, USA, 2003.
18. Deb, Kalyanmoy, and A. Raji Reddy. “Reliable classification of two-class cancer data using evolutionary algorithms”, Bio Systems, vol. 72, no.1,pp.111-129,2003.
19. Huiqing Liu and et.al,” A Comparative Study on Feature Selection and Classification Methods Using Gene Expression Profiles and Proteomic Patterns”, IEEE, 2002.
20. B. Rajwa, P. K. Wallace, E. A. Griffiths and M. Dundar, "Automated Assessment of Disease Progression in Acute Myeloid Leukemia by Probabilistic Analysis of Flow Cytometry Data," in IEEE Transactions on Biomedical Engineering, vol. 64, no. 5, pp. 1089-1098, May 2017.