

Detection of Leukemia & its Staging Using Image Processing; Artificial Intelligence

SANNIHITHA PEETALA, DR B LEELA KUMARI

Department of Electronics and Communication, University college of Engineering, Jawaharlal Nehru technological University Kakinada Andhra Pradesh, INDIA.

Abstract—Leukemia or blood cancer is a life-threatening ailment of blood. It originates in bone marrow, & causes the formation of large number of abnormal cells. Leukemia is classified as Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML), Acute Myeloid Leukemia (ALL), Chronic Lymphatic Leukemia (CLL) & Chronic Myeloid Leukemia (CML). This Dissertation will aim at automated detection & staging of leukemia using combination of image processing & artificial intelligence technique. Also, detection of white blood in blood stain image, accurately, is crucial for prediction of leukemia with accuracy. This dissertation also aims of improving the accuracy of WBC detection using combination of contrast enhancement morphological area operation & Hough transform to find circles. WBC segmented by both the above methods. The outputs were then processed by Levenberg–Marquardt algorithm, which is pre trained with number of samples using artificial neural network to classify the blood stain as leukemia or non-leukemia and provide staging.

Keywords—White Blood Cells, Leukemia, Microscopic images, Morphological area, Hough transform, Artificial neural networks

Received: June 26, 2021. Revised: June 21, 2022. Accepted: August 13, 2022. Published: September 11, 2022.

1. Introduction

Over the last decade, medical imaging has grown to be one of the most essential means of visualizing and interpreting biology and medicine. During this period, there has been a great increase in the creation of new, powerful medical imaging tools. As a result, the use of digital image processing techniques to solve medical problems has grown significantly. The development of clinically useful integrated systems for medical imaging is the most difficult part of the field's work. In order to successfully design, develop, and validate sophisticated medical systems, physicians and engineers must work together closely. The primary goal of image analysis is to gather information, detect disease, diagnose disease, regulate and treat, monitor and evaluate. and to obtain information.

Leukemia, a form of blood cancer, is the most lethal of all blood illnesses. Because of the urgency of early and precise detection of leukemia, this research focuses on this topic. Because leukemia spreads rapidly, it necessitates early discovery, frequent testing, and close monitoring in order to keep it under control and appropriately assess treatment options.

A hematologist uses a microscope to diagnose leukemia in a patient. Manual testing has some limitations, such as being time-consuming and expensive. Hematologist experience and fatigue may have an impact on the outcome. Images are inexpensive and don't require expensive testing or lab equipment because of the automated process.

The bone marrow creates aberrant white blood cells in leukemia. When compared to regular white blood cells,

aberrant ones do not die as quickly. As a result, the number of aberrant white blood cells increases, preventing regular white blood cells from carrying out their duties. As a result, the human body's blood system becomes unbalanced.

2. Existing Techniques

Deepika kumaret. All (2021) A CNN-based early detection method for acute lymphoblastic leukemia was proposed and implemented in this research using a mix of center-cropping, image augmentations, and class rebalancing. The CNNs were proven to be more accurate in detecting all images when the images were cropped in the middle, rather than when the complete image was used. The ensemble model for image classification outperforms the single candidate model with significantly lower training sample counts. As a result, the aggregate results of the ensemble model can be improved by weighting models according to their performance. Despite the promising results of the microscopic cell picture categorization, the results, particularly for the Hem class, require further refinement. To address imbalance issues in the future, the adversarial network can generate synthetic data. As DL models are increasingly used in medical image interpretation scenarios, future research will examine the influence of data imbalance and properly accounting for subject information. [1]

Revella E. A. Armya et. All (2021) Machine learning and classification algorithms will be used to increase the accuracy of leukemia diagnosis. The microarray Leukemia dataset was classified using the J48, KNN, SVM, Random Forest, and

Nave Bayes methods. According to the results, the Nave Bayes classifier has the highest accuracy (98.61%), whereas the J48 classifier has the lowest (83.33 percent). SVM and Random Forest are the final two classifiers with accuracy ratio of 87.5 percent and 95.83 percent for SVM and 88.5 and 95.4 percentage points for Random Forest. Future work in this area will lead to the assessment of the efficacy of treatment provided to leukemia patients through the effective application of appropriate machine learning classification algorithms for all varieties of leukemia, which can be carried out in parallel for greater reaction time and accuracy. [2]

Prasidhi G. Fal Desai et. All (2018) Leukemia cells can be detected and classified using an automated method described in this article. Pre-processing, segmentation, feature extraction, and classification are the four important processes in the proposed approach. This technique was put to the test using data from the Goa Medical College (GMC) and an online leukemia dataset. For improved segmentation, images are transformed to the HSI color model. The K-means algorithm can produce accurate segmentation results. It splits all four types of cells correctly. To get better results, further testing might be done with different algorithms. More study can automatically determine the number of clusters. [3]

Jan-Niklas Eckardt et. All (2020) In cancer research, machine learning (ML) is becoming increasingly popular. To better understand cancer, machine learning algorithms that can deal with big medical datasets are needed. Acute myeloid leukemia (AML) has been employed as a model disease in a number of recent research because of ML's ability to incorporate data from many diagnostic modalities and functions to predict prognosis and offer therapy methods. Rapid diagnosis, precise risk classification, and effective treatment can be achieved by integrating these machine learning algorithms into various AML management platforms. The limitations of these techniques necessitate a strict regulatory system to ensure that machine learning is used properly. Current advances in machine learning approaches in AML management as a model illness for hematologic neoplasms are highlighted and examined in this comprehensive review to allow academics and clinicians to objectively assess this potentially game-changing technology. [4]

Supun S. Chandrananda et. All (2019) Each patient was evaluated for cell count by taking five photos. Testing revealed an accuracy rate of 86.7%, with a processing time of under 8 seconds. The classifier's trained sample to test sample ratio is higher since the training data contains 130 photos and the test data contains 250 photos. A high number of photos is required in the training sample in order to improve accuracy. Using caching for the classifier saves 95.32 percent of your time, proving it to be an excellent strategy (30.5 seconds). Apply image processing techniques to identify ALL and CLL from digital microscope images, then extract morphological features. The nucleus of the cell was measured for dimensions including diameter, radius, area, and perimeter to determine the type of leukocyte. On the image data set tested, the system had an accuracy rate of 91.76 percent. Upgrading to a fully automated system for cancer hospitals is likely to be easier with the help of Sri Lankan medical experts. [5]

Nizar Ahmed et. All (2019) White blood cells and bone marrow are attacked by leukemia, which weakens the

immune system. A common diagnostic approach is the examination of blood smears (small blood cells). For the first time, a new method for identifying leukemia from microscopic blood images utilizing a CNN architecture has been described. Our model was able to deal with a limited number of image samples because to the use of data augmentation techniques. To show that CNN surpasses earlier machine learning algorithms and tested it on one leukemia type (ALL samples) and found it to be 88% accurate, while it was 81% accurate for all other leukemia subtypes. Our tests also included cross-validation. It's true that classifying medical images takes time, but it's crucial to check whether the model under consideration remains stable during the classification process. [6]

Viswanathan P (2015) Leukemia is frequently missed in routine checkups because to the difficulties in interpreting blood smear images and the appearance of comparable symptoms in other diseases. In addition to being more time demanding, the diagnostic procedure is also more prone to errors. As a result, this research proposes a fuzzy C means cluster optimization for leukemia detection using morphological contour segmentation. This study's innovative approach to leukemia diagnosis includes nuclei contrast enhancement, morphological contour segmentation, and fuzzy C means. Basic addition and subtraction are used to raise the contrast required for nuclei separation. This technique uses morphological contour segmentation to eliminate the normal white blood cells from a microscopic blood picture so that nuclei's edges may be clearly seen (MCS). As a strategy to speed up the difficult and time-consuming procedure of finding leukemia, an entire slide image of leukemia WBCs has been proposed. To improve WBC segmentation from blood microscopic images, morphological contour processing of nuclei and leukemia has been improved. In order to produce an accurate leukemia rating, we trained a fuzzy C mean classification model using extracted features row vectors such as perimeters, densities, and percentages from segmented nuclear images. Leukemia will be clearly separated from normal blood in the suggested method because of this. By using the method proposed, researchers established a large-scale automated system for detecting and classifying normal and aberrant WBCs. [7]

Ashwini Rejintal et. All (2016) The visual examination of microscopic pictures by hematopathologists is a time-consuming and complex process that leads in delayed detection. An autonomous image handling framework is needed to overcome these limitations in visual inquiry, enabling for the early diagnosis of disease and cancer kinds. The proposed method has been successfully applied to a large number of photos, resulting in accurate findings even when the image standards have changed. MATLAB is used to perform a variety of image processing computations, including picture augmentation, clustering, mathematical method, and labelling. Using some of the most powerful image processing methods, can identify and slice disease cells. The segmentation is useful in identifying the cancer cell's exact size and shape, as well as its immediate surroundings. The picture's contrast and pixel values were first improved using image improvement methods. Using cell segmentation and feature extraction and hope to identify cancer cells. Angular second moment (energy), contrast, auto correlation, Entropy (variation), disparity (difference),

homogeneity (homogeneity), cluster prominence (prominence of clusters), and the Inverse Difference Moment are all factors to consider for accurate identification. The k mean approach yields the best segmentation results, as shown by the data. It is also easier to obtain a completely separated core since MATLAB-based calculations are less sensitive to variations in the image. [8]

Dr. P V Rama Raju et. Al (2017) An expert examines microscopic images in order to identify leukemia manually. This is a labor-intensive process that relies on the individual's knowledge and skill and cannot be generalized. This proposed method for leukemia diagnosis uses Image j as a tool to understand the entire detection process. Preprocessing and segmentation are also included in this tool. You may use this program to convert RGB images to grayscale images and to directly convert grayscale images to binary files. This tool's leukemia identification procedure is made easier by the use of macros installed at each phase. Java, Python, MATLAB and other programming languages all support the creation of macros. Leukemia can be diagnosed and classified as ALL or CLL based on a cell count and other factors including area and perimeter (or) In this way, nope. [9]

SatvikDasariraju et. Al (2020) In an effort to overcome the limits of manual diagnosis for AML, automated leukocyte recognition and categorization was created. The algorithm was able to appropriately detect immature leukocytes based on its 93 percent accuracy and 0.98 AUC-ROC. This new model was able to successfully categorize each of the four immature leukocyte categories during multiclass classification, which is a considerable improvement over earlier research. For classification purposes, the expected color characteristics of the nucleus in the B channel of LAB color space are crucial, whereas the N:C ratio was found to be relevant for both detection and categorization. This study has two purposes. The proposed approach can be utilized to assist doctors in the diagnosis of AML, saving both time and money. In order to expedite the selection of potentially malignant cells for further examination by a professional, the model's high binary classification accuracy may be utilized. To save lives in impoverished countries where diagnosis can take weeks, the proposed technique identifies immature leukocytes. When discovered early, those who have AML have a better chance of responding to treatment. Immature leukocyte classification may also help with therapy and prognosis options, which can differ depending on the type of malignant cell. This study can be utilized to improve classification performance in future investigations by utilizing the most relevant traits that were calculated and provided as features. [10]

3. Methodology

3.1 ANN Flow Chart

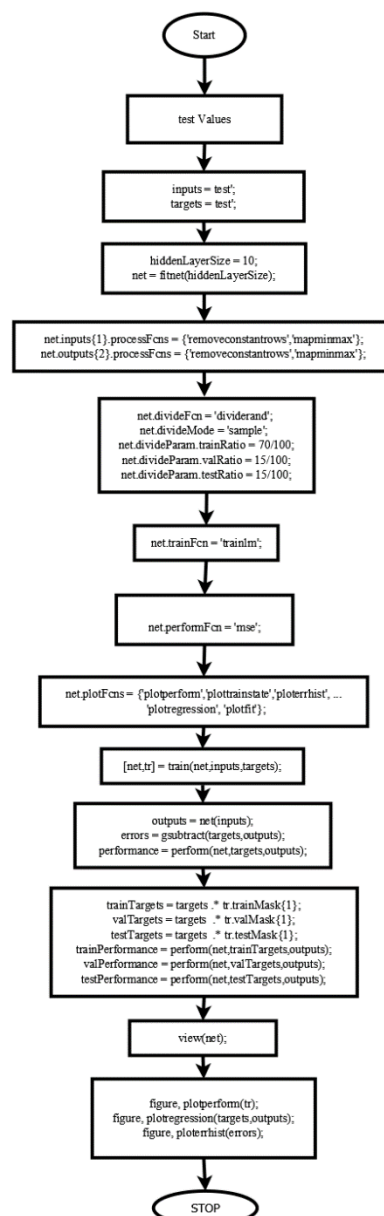


Figure: 3.1 ANN Flow Chart

The term "neural networks" refers to a class of adaptive systems that use a layered structure resembling a human brain to learn by connecting nodes or neurons (also known as artificial neural networks). Pattern recognition, classification of data, and forecasting of future events can all be taught to neural networks.

An artificial neural network processes the data and creates a series based on the input. To learn to recognize patterns in voice or images, for example, it may be taught to use many samples like a human brain does. How the system's components are linked together and the weights attached to each link determine its behavior.

3.2 Proposed Methodology Block Diagram

In the system block diagram, It can be seen 1st then input blood sample image then input image is processing then by blob analysis method WBC count 1 and by Hough transform RBC count 1 then by imfind circle method total number of cell count. And then RBC count 2 and bicolour differentiation & area condition WBC count 2.

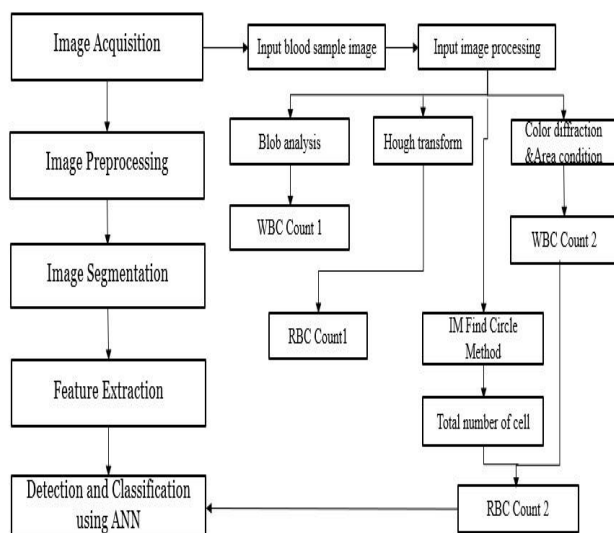


Figure: 3.2 Proposed Block Diagram

3.3 Flowchart of WBC Count Detection Using Blob Count Method

The WBC count detection operation is start with the input image. So 1st take a input image as Blood Cell Image and this image is stored as BO Img. After this process this input image is converted into the BC Img. Grayscale Image and this image is stored as BG Img. Then apply Linear contrast stretching on Grayscale Image and store as BL Img. After this, apply Histogram Equalization Process on BL Img. And stored as BH Img. Then by applying background Brightening on BH Img. And stored. Then obtain highlighted components by subtracting BH Img. From BB Img. And It stored as the BHL Img. Then remove other components of the image by using image addition of BB Img. And this is stored as BR Img. Then applying 2-D order statistic filtering on this BR Img. And stored as BF Img. Then compliment BBW Img. And Stored BCBW Img. Then by applying morphological operation on BCBW Img. And stored this as BMC Img. Then applying area condition operation on WBC and remove ourcts having are less then and this stored as BMC img. Then all connected label is computed and compute number of WBC by number of observed connected components.

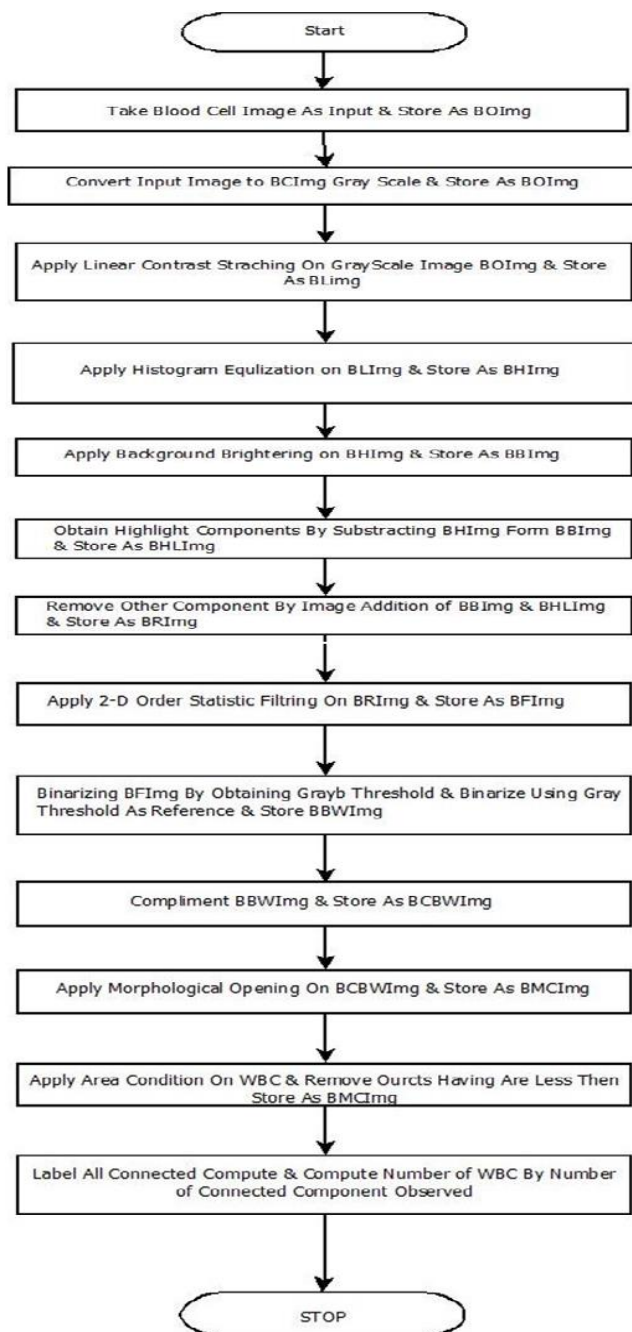


Figure: 3.3 Flowchart of WBC Count Detection

3.4 Flowchart of RBC Count Detection:

1st Extract Green Channel from Original Image. Then Apply Adoptive Histogram Equalization. Then Apply Circular Hough Transform BOGCHE Img. Then Get NO. Circles Radius Then Assigns RBC Count. Then it assigns as RBC Count.

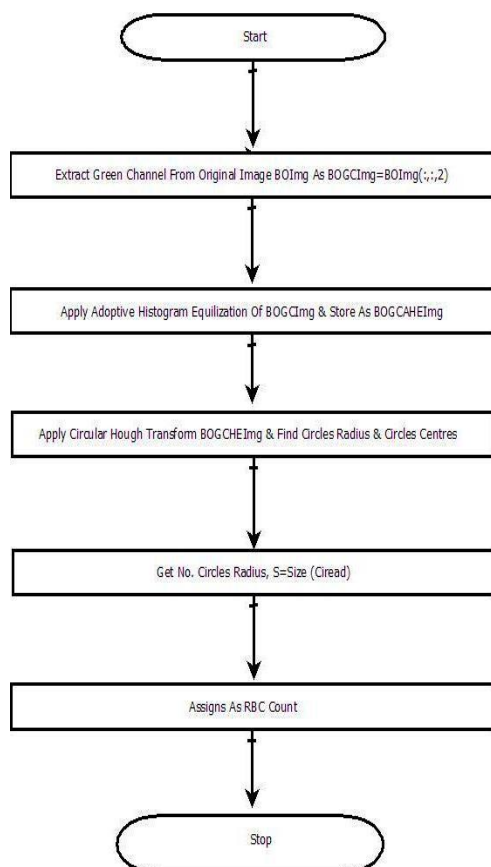


Figure:3.4 Flowchart of RBC Count Detection

3.5 WBC Detection Using Color Area differentiation Method

First, it is seen a grayscale image of a stained blood cell slide in this flowchart. Then use the Im adjust function to apply linear contrast stretching to a grayscale image. Finally, using image addition, combine the Im adjust and histogram equalized images to achieve background brightening. Then, to normalize the image's histogram, then subtracted the background to brighten the highlight component. Then, using image addition, brighten the image by removing non-interesting elements. Then apply 2-D order statistic filtering technique using 1st order and domain a unity matrix of size 3X3. Then by using Otsu's method binaries filtered image. Then obtain complement of the binary image using image compliment technique. Then apply morphological opening on the image with structural element disk of radius 9Pixels. Then use connect component analysis to count number of objects in binary compliment image and remove objects from list whose area is less than 5000 pixels. Then display detected WBC Image.

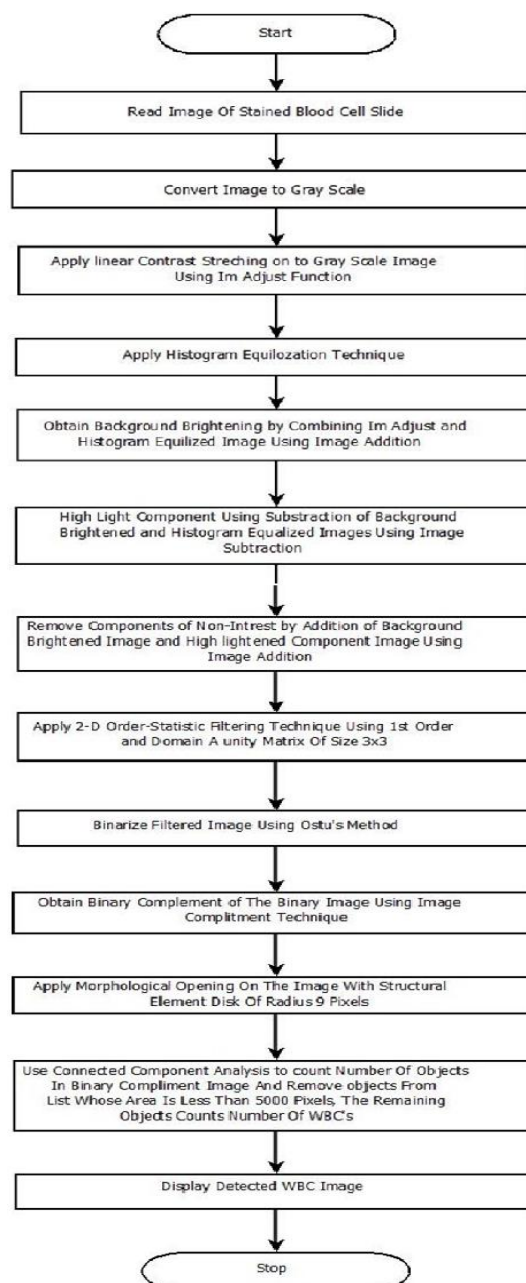


Figure: 3.5 WBC Detection Using Color Area differentiation Method

4. Results

The first step of the system is the acquisition of images. Retrieving photographs from sources that automatically capture input images is part of this process. The raw image is first turned into a grayscale image during pre-processing, which is the following step after image acquisition. Afterward, the noise is eliminated by employing a median filter to remove it. There are a number of ways to improve the process's outcomes:

- “New initialization of the network parameters or multiple trainings”.
- “Successive increasing of the number of hidden neurons. This concern is of prime importance, when defining the network structure, and will be explained later in more detail”.
- “To train the network with different learning algorithms”:

Learning functions such as Levenberg-Marquardt and back propagation were also employed to train the network during this stage (BP). When it comes to fitting issues with a ANN trained data, LM theoretically provides the results.

The three methods outlined above are used to train the network until it was able to generalise to the greatest extent possible. Using the early-stopping-approach, the network is generalised. We obtain data for training, validation, and testing in our network. Weight adaptation is the method to train the Artificial neural network. To avoid overfitting, this error rises and the weights are calculated based on this minimum error.

When the number of LM-algorithm hidden neurons is increased to 20, the mean square deviation drops. It was decided to use a network having 10 hidden neurons with Algorithm to keep the network's complexity less and prevent the risk of overfitting the network. The diagram depicts the network's structure.

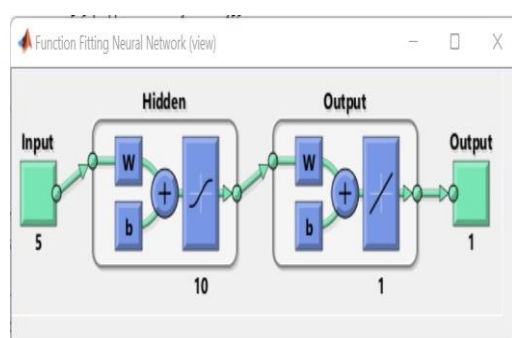


Figure 4.1 Structure of the chosen network

Few attempts were made to get the performance curves of training, validation, and test data to follow a nearly identical path. " There is a correlation between the three curves if the network's response to learning data, validation and testing data is similar.

Overfitting is therefore less likely, but it is not completely eliminated. It is shown in Figure 4.2 that the network with 10 hidden neurons trained in LM algorithm.

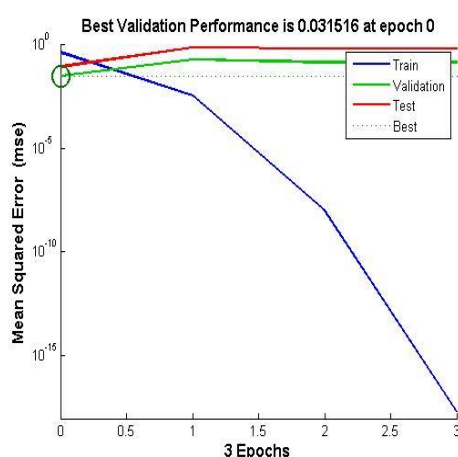


Figure 4.2 Performance-plot of the chosen network

The process to evaluate the generalization of a network is a regression analysis between the output and the targets, which typically is represented by a regression factor (R). With a perfect network performance R should be three but Figure 4.3 shows the regression plot with $R=2.1$.

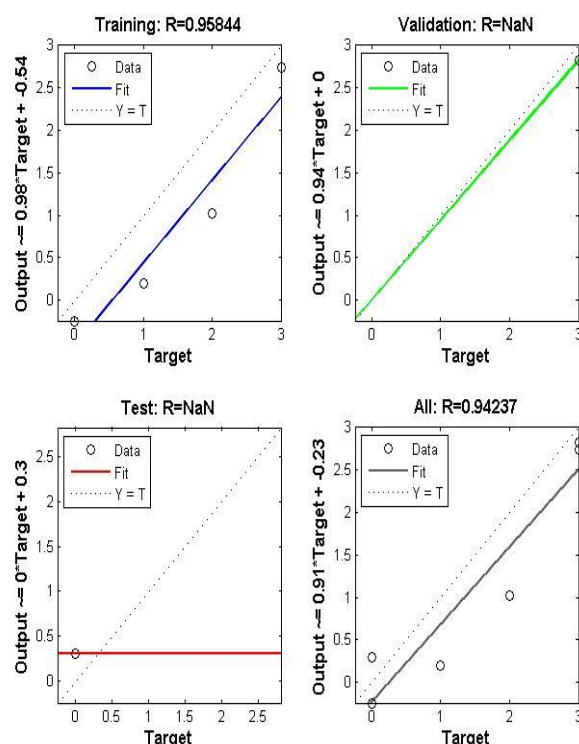


Figure 4.3 Regression plots of the chosen network

The quality of network is trained using error histogram which indicates the distribution of the residuals between target and output. This histogram is able to indicate outliers. It checks if the quality of the learning data is bad or if these outliers are different from the test data.

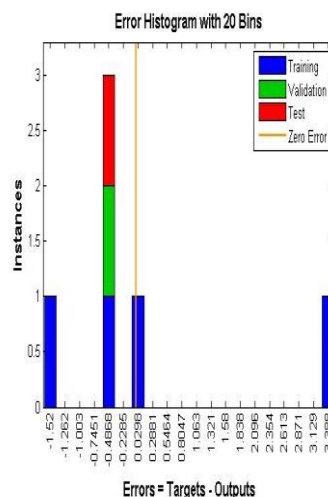


Figure 4.4 Error histogram of the chosen network

Obtained count of RBC and WBC using Morphological operation, Linear contrast stretching. Circular transform for RBC is carried and by subtracting the RBC from Total number of cells WBC is obtained, with these parameters which are used for train the neural network hence it is estimated whether Person exist leukemia or not.

GUI Outputs for leukemia detection and staging:

Here using color leukemia is classified representing stage of leukemia, if the red color is displayed it indicate third stage

of leukemia, yellow indicate that patient has early stage of leukemia and green indicate non leukemia.

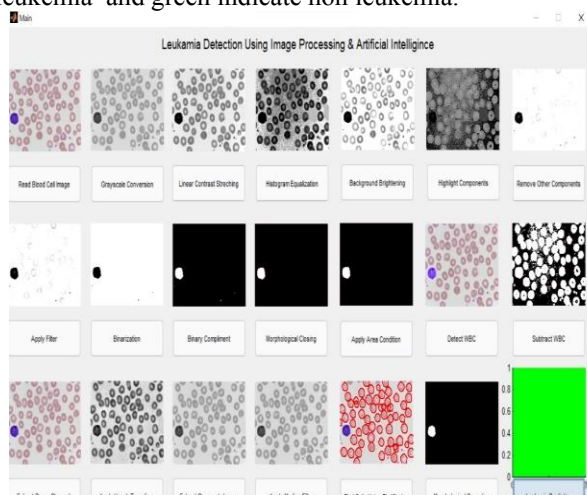


Figure: 4.5 Non leukemia detection

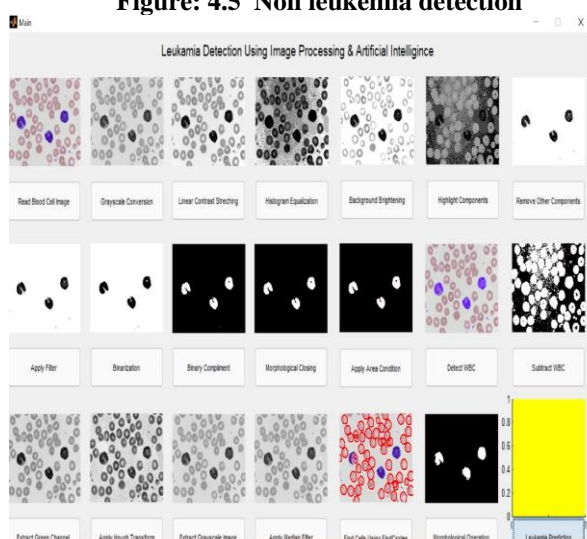


Figure: 4.6 Early stage detection

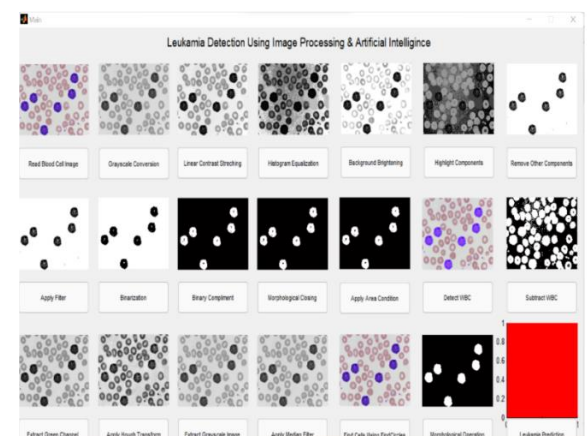


Figure: 4.7 Third stage leukemia Detection

V. Conclusion

Computed blood cell count, which measures the density of various blood cells in a person's body, is an essential tool for diagnosing anemia, infection, and leukemia, among other conditions. By automating this process, we can improve diagnosis times while significantly lowering treatment costs. For the total blood cell count on blood smear images, we use

a artificial neural network in this paper. In this paper, we presented a low cost detection of leukemia & its staging using image processing and artificial intelligence.

Comparison With Base Paper

S. No.	Parameters	Base Paper	Proposed Work
1.	Title	Automated Complete Blood Cell Count and Malaria pathogen detection using convolution neural network	Leukemia detection using image processing & artificial intelligence
2.	Techniques Used	Convolution neural network to perform this complete blood cell count on blood smear image. Convolution neural network detection of malarial pathogens.	Dual method WBC count using color differentiation with area condition & blob analysis. Dual method RBC count using Hough transform & Im find circles. Levenberg Marquardt function fitting neural network for leukemia prediction.
3.	WBC Count Method	Common convolution neural network for complete blood cell count & malarial parasite detection.	<ul style="list-style-type: none"> Color differentiation with area condition. Blob detection & parameter analysis
4.	RBC Count Method	Color differentiation with area condition. Blob detection & parameter analysis.	<ul style="list-style-type: none"> Hough Transform Imfind Circles
5.	Disease Detected	Malaria parasite	Leukemia
6.	Staging or Severity	No Staging or severity information	Staging & severity predicted by Levenberg Marquardt ANN

Table: 1 Proposed Work Comparison with Base Paper

5. Future work

1. Detection of WBC Geometry such as shape size also apart from count value for enhancement of prediction accuracy.
2. Analysis of RBC population area coverage to estimate canvas area (reference area) & use WBC area occupied parameter, to enable WBC oversize/ undersize detection & further input to ANN to improvise prediction accuracy.
3. Simulation of various treatment choices such as radiotherapy, chemotherapy, naturopathy, electropathy & their variants & combination.
4. Automatic prediction of prognosis treatment choice, hospitalization time etc.

References

- [1.] Deepika Kumar, Nikithajain, Md. Ayushkumar, Swetha Mittal "Automatic Detection of White blood cancer bone marrow Microscopic Images Using Convolutional Neural Networks" IEEE Access 2021.
- [2.] Revella E. A. Armya, Adnan Mohsin Abdulazeez, Amira Bibi Sallow "Leukemia Diagnosis using Machine Learning Classifiers Based on Correlation Attribute Eval Feature Selection" Asian Journal of Research in Computer Science 2021.
- [3.] Prasidhi G. Fal Desai, Geeta Shet "Detection of Leukemia Using Image Processing" International Journal of Advance Research in Science and Engineering 2018.
- [4.] Jan-Niklas Eckardt, Martin Bornhauser, Karsten Wendt "Application of machine learning in the management of acute myeloid leukemia: current practice and future prospects" Blood Advances 2020.
- [5.] Supun S. Chandrananda, Sachethana H. Pathirage, Shavini R. Marapana "Detection of Leukemia using Image Processing and Machine Learning" Research Gate 2019.
- [6.] Nizar Ahmed, Altug Yigit, Zerrin Isik and Adil Alpkokak "Identification of Leukemia Subtypes from Microscopic Images Using Convolutional Neural Network" Diagnostics 2019.
- [7.] Viswanathan P "Fuzzy C means Detection of Leukemia based on Morphological Contour Segmentation" Procedia Computer Science 2015.
- [8.] Ashwini Rejintal, Aswini N "Image Processing based Leukemia Cancer Cell Detection" International Journal of Engineering Research & Technology 2016.
- [9.] Dr. P. V. Rama Raju, Ramesh Babu, Naga Raju "Detection of Leukemia Using Image Processing" JETIR 2017.
- [10.] Satvik Dasariraju, Marc Huo and Serena McCalla "Detection and Classification of Immature Leukocytes for Diagnosis of Acute Myeloid Leukemia Using Random Forest Algorithm" Bioengineering 2020.
- [11.] Kakara Deepika, B. Leela Kumari "Assessment of Breast Cancer Detection and Implementation" International Journal of Innovative Technology and Exploring Engineering (IJITEE) 2019.