AUTOMATIC DETECTION OF WHITE BLOOD CELLS CANCER DISEASES

Sivaranjani.A[1], Sornalatha Janani.P[2], Shalini.M[3]

[1] *Student, Electronics and Communication Engineering, Prince Shri Venkateshwara Padmavathy Engineering College, Tamil Nadu, India*
[2] *Student, Electronics and Communication Engineering, Prince Shri Venkateshwara Padmavathy Engineering College, Tamil Nadu, India*
[3]*Assistant Professor, Electronics and Communication Engineering, Prince Shri Venkateshwara Padmavathy Engineering College, Tamil Nadu, India*

ABSTRACT

Blood smear images are efficient and reliable source of information in extraction of features in case of cancer detection. In assessment of cancer detection like leukemia cancer image feature extraction and reduction has to be carried out efficiently. It discusses a robust image feature extraction algorithm using scale invariant feature points for key point extraction and GLCM Shape Analysis. For matching the key points, k nearest network algorithm is developed with tumbling functions of model. This technique assists in comparing the scale invariant feature points and GLCM features. From the obtained outcomes, the capability of the obtained accuracy is estimated and it is wrapped up that the proposed approach is efficient in regards to SVM algorithm.

Keyword: - *MATLAB, Adaptive median filter, K means clustering, KNN classifier, Arduino Uno*

1. INTRODUCTION

A hybrid automated system to facilitate the diagnosis of different white blood cells cancer diseases. These diseases are Leukemia with its subtypes and Myeloma which are detected automatically. These diseases are divided into two categories, each category includes similar symptoms diseases. There are two approaches. Each approach is applied on one of the two diseases category by computing different features. The main scope of the project is to early detection of white blood cell cancer and reduce the misdiagnosis case in addition to improve system learning methodology. Image processing software is used for the detection of cancer cells and different types of cancer such as Acute Myeloid Leukemia (AML), Acute Lymphoblastic Leukemia (ALL), Chronic Lymphocytic Leukemia (CLL) and Chronic Myelocytic Leukemia (CML). The input images are send to the preprocessing stage where the RGB images are converted into grayscale images then the noise is removed by adaptive median filter. The preprocessed output is send to the segmentation process in which the K means clustering algorithm is used for clustering purpose. The various features used in feature extraction are geometrical, textural and colour. Then, KNN classifier algorithm is used for classifying the types of cancer diseases. The output is displayed in LCD.

2. PROPOSED SYSTEM

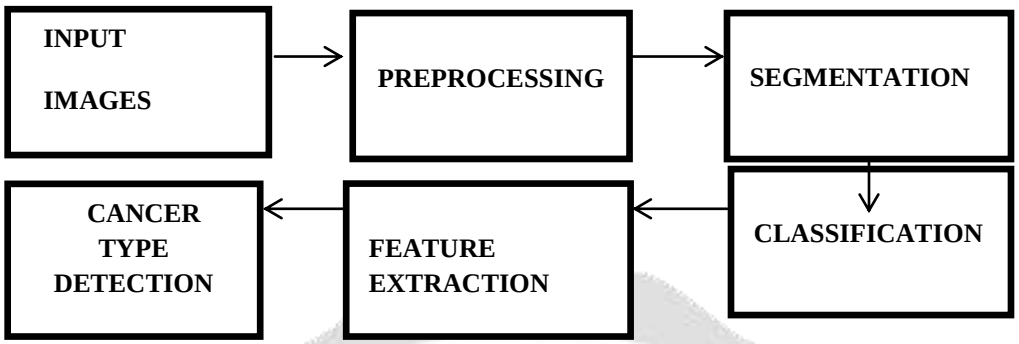


Fig-1:Block Diagram of Cancer Type Detection

The luminance of a pixel value of a gray scale image ranges from 0 to 255.The conversion of a color image into a gray scale image is converting the RGB values(24 bit) into gray scale value(8bit).The original image has been converted to gray scale in order to minimize the CPU time. A representation of additive color mixing. Projection of primary color lights on a screen shows secondary colors where two overlap the combination of all three of red, green, and blue in appropriate intensities makes white. The Fig-1 shows that the step of cancer type detection. This conversion is done using the luminosity method which relies on the contribution of each color of the three RGB colors. Using this method, the gray scale image is brighter since the colors are weighted according to their contribution in the RGB image not averagely. The module descriptions are preprocessing, segmentation, feature extraction and KNN classification.

2.1 Preprocessing

The purpose of the pre-processing stage is to remove unwanted effects such as noise from the image, and transform or adjust the image as necessary for further processing. The resolution of the image is reduced by a factor of four to 512·384 to speed up performance of the system. Also, the test images will be subjected to selective median filtering and unsharp masking to isolate noise which may have been accumulated during image acquisition and due to excessive staining.

2.2 Segmentation

The technique of partitioning the image into segment can be defined as image segmentation. Considering the similar property, segmentation is implemented. This similar property is cluster together our propounded approach implements Lloyd's clustering technique which aids in the segmentation of blood microscopic images on the basis of alike properties. This technique broadens the k-mean clustering algorithm by introducing repeated segmentation scheme which explores the centroid of each set in the segment and eventually re-segment the input based on the closest centroid. This technique aids in the extraction of important image characteristics, based on which information can be easily perceived. A simple thresholding approach is applied to give initial labels to pixels in the blood cell images. The algorithm is based on priority information about blood smear images. Then the labels are adjusted with a shape detection method based on large regional context information to produce meaningful results.

2.3 Feature extraction

In pattern recognition and in image processing, feature extraction is a special form of dimensionality reduction. When the input data to an algorithm is too large to be processed and it is suspected to be notoriously redundant, then the input data will be transformed into a reduced representation set of features. Transforming the input data into the set of features is called feature extraction. If the features extracted are carefully chosen it is expected that the features set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. When performing analysis of complex data one of the major problems stems from the number of variables involved. Analysis with a large number of variables generally requires a large amount of memory and computation power or a classification algorithm which over fits the training sample and generalizes poorly to new samples. Feature extraction is a general term for methods of constructing combinations of the variables to get around these problems while still describing the data with sufficient accuracy.

2.4 Geometrical Features

- 'Area' — Scalar; the actual number of pixels in the region. 'Centroid' – 1-by-Q vector that specifies the center of mass of the region. Note that the first element of Centroid is the horizontal coordinate (or *x*-coordinate) of the center of mass, and the second element is the vertical coordinate (or *y*-coordinate).
- 'Eccentricity' — Scalar that specifies the eccentricity of the ellipse that has the same second-moments as the region. The eccentricity is the ratio of the distance between the foci of the ellipse and its major axis length. The value is between 0 and 1.
- 'EquivDiameter' — Scalar that specifies the diameter of a circle with the same area as the region. Computed as sqrt(4*Area/pi).
- 'Extent' — Scalar that specifies the ratio of pixels in the region to pixels in the total bounding box. Computed as the Area divided by the area of the bounding box.
- 'MajorAxisLength' — Scalar specifying the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments as the region.
- 'MinorAxisLength' — Scalar; the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments as the region.
- 'Orientation' — Scalar; the angle (in degrees ranging from -90 to 90 degrees) between the *x*-axis and the major axis of the ellipse that has the same second-moments as the region.
- 'Perimeter' — Scalar; the distance around the boundary of the region. region props computes the perimeter by calculating the distance between each adjoining pair of pixels around the border of the region.

2.5 KNN Classification

The KNN binary (as two class) is given more accurate data classification which beneficial to select k as an odd number which avoids the irregular data. The KNN procedure is the technique in ML procedures: It is an object which classified through a mainstream selection of its neighbors, with the determination assigned occurrence for most mutual class amongst its *k* nearest neighbors (k is a positive integer, classically small). Classically Euclidean distance is used as the distance metric; however, this is only suitable for endless variables. KNN is a new process that deliveries all available cases and categorizes novel cases built on an evaluation quantity (e.g., distance functions).KNN procedure is identical simple. It works built on a minimum distance from the interrogation instance to the training samples to regulate the K-nearest neighbors. The information for KNN procedure contains numerous attribute which will be used to categorize. The information of KNN can be any dimension scale from insignificant, to measurable scale.

3. SIMULATED RESULTS

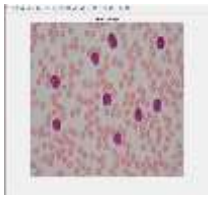


Fig-2:Input Image

The input image of blood sample is taken as normal image in fig-2

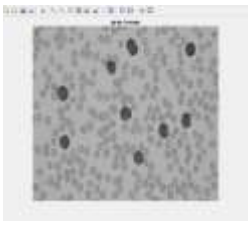


Fig-3:Grayscale Conversion

The RGB image is converted in to GRAY scale image(refer fig-3).

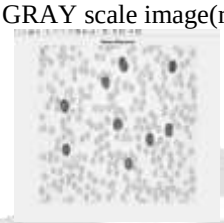
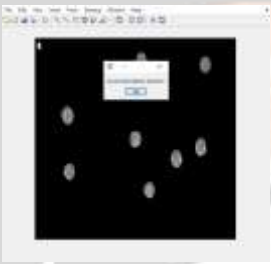


Fig-4:Noise removed image

The Fig-4 shows the noise removed image.In this image the median filter is used to remove the noise.



. **Fig**-5:Detection of ALL

In Fig-5 Acute Lymphoblastic Leukemia is a type of cancer in which the bone marrow makes too many immature lymphocytes.

Fig-6: Detection of AML

The fig-6 shows the cancer type of acute myeloid leukemia. It is a cancer of the myeloid line of blood cells, characterized by rapid growth of abnormal cells.

Fig-7: Detection of CLL

The fig-7 shows Chronic Lymphocytic Leukemia is a type of cancer. These cells do not function well and crowd out healthy blood cells.

Fig-8:Detection of CML

The fig-8 shows Choronic Myelocytic Leukemia is the type of cancer that starts in certain blood forming cells of bone marrow.

4. CONCLUSIONS

Image processing approach are used for the detection of cancer types. In this approach input images are taken then converted into gray scale images then automated the segmentation, feature extraction and classification of white blood cells using KNN, NN, and SVM classification algorithm. Several improvements were made to the SVM algorithm, including an initialization step to find 12-neighbor connected component. Additionally, the proposed model features an enhanced accuracy of selecting the correct circle from three candidate circles, the capability to detect irregular cells, the use of a dynamic number of iterations, and improved detection of overlapping cells. The proposed method performed the segmentation and classification of WBCs well when results were compared with the ground truth, which was determined.

5. REFERENCES

[1]. R. Bhattacharjee, L. M. Saini, Detection of acute lymphoblastic leukemia using watershed transformation technique, in: Signal Processing, Computing and Control (ISPCC), 2015 International Conference on, IEEE, 2015, pp. 383–386.

[2]. Detection of Leukemia in Microscopic White Blood Cell Images by Shubhangi Khobragade1, Dheeraj D Mor2,Dr. C.Y.Patil 3,2015.

[3]. Implementation of Back propagation Neural Network and Blood Cells Imagery Extraction for Acute Leukemia Classification,2017.

[4]. E. Sarhan, E. Khalifa, A. M. Nabil, Post classification using cellular automata for landsat images in developing countries, in: International Conference on Image Information Processing (ICIIP), Shimla, India, IEEE, 2011, pp. 1–4

[5]. Automated Detection Of White Blood Cells Cancer Diseases by Hend Mohamed,Rowan Omar,Nermeen Saeed,Ali Essam,Nada Ayman,Taraggy Mohiy Ashraf AbdelRaouf,IEEE,2018.