Automatic Blood Cancer Detection Using Image Processing

Ms. Chandni Yadav¹, Ms. Shrutika Zele², Ms. Tejasree Patil³, 
Ms. Vishakha Bombadi⁴, Mr. Tushar Chaudhari⁵

¹, ², ³, ⁴ Student, Department of Electronics & Telecommunication, SJCEM, Palghar, Mumbai, India
⁵Asst. Prof., Department of Electronics & Telecommunication, SJCEM, Palghar, Mumbai, India

Abstract - Microscopic pictures are reviewed visually by hematologists and the procedure is tedious and time taking which causes late detection. Therefore, automatic image handling framework is required that can overcome related limitations in visual investigation which provide early detection of disease and also type of cancer. The proposed strategy is effectively connected to many numbers of pictures, demonstrating accurate results for changing image standard. Distinctive picture handling calculations, for example, Image enhancement, Clustering, Mathematical process and Labelling are executed utilizing MATLAB.

Keywords - digital image processing acute Leukemia, normal white blood cell, enhancement.

1. INTRODUCTION

Cancer detection has always been a major issue for the hematologists for diagnosis and treatment planning. The manual identification of cancer from microscopic images is subjective in nature and may vary from expert to expert depending on their knowledge and other factors which include lack of specific and accurate measures to classify the microscopic images as normal or cancerous one. The automated identification of cancerous cells from microscopic biopsy images helps in alleviating the above-mentioned problems and provides better results if the biologically interpretable and clinically significant feature-based approaches are used for the identification of disease. About 32% population of India gets cancer at some point during their life time. Cancer is one of the most common diseases in India which has responsibility to maximum mortality with about 0.3 million deaths per year. Medical image processing has become one of the most important visualization and interpretation methods in biology as well as medicine over the past decade. This time has witnessed a tremendous development of new, powerful instruments for detecting, storing, transmitting, analyzing, and displaying medical images. This has led to a large growth in the application of digital image processing techniques for solving medical problems. The most challenging aspect of medical imaging lies in the development of integrated systems for the use of the clinical sector. Designing, implementing and validation of complex medical systems require a high collaboration between physicians and engineers. Main objective of analyzing through images is to gather information, detection of various diseases, diagnosis diseases, control and therapy, monitoring, evaluation and result. In blood diseases, the most dangerous disease is blood cancer i.e. leukemia. This study focused on detecting leukemia because, it needs fast and accurate detection. Leukemia is a fast-growing disease so it requires quick detection, regular test & monitoring for control it and to evaluate the treatment properly. Leukemia is detected by hematologist under microscope. This manual examination method has some drawbacks, it is time consuming and costly. Result may be affected by the factors such as hematologist experience and tiredness. The automation process overcome the above problems, images are cheap and do not require expensive testing and lab equipment.

In leukemia, bone marrow produces abnormal white blood cell. Compared with normal cells, abnormal white blood cell does not die when they should thus the number of abnormal white blood
cell become numerous and interfere normal white blood cells to carry out their duties. This leads to an imbalance of blood system in human body. Classification of Leukemia can be grouped on the basis of how quickly this disease develops and become severe. Leukemia is either Chronic or Acute. Chronic Leukemia: In this type, leukemic cells are works as a normal white blood cell at earlier stage, but after that gradually they will become severe chronic leukemia. Acute Leukemia: In this type, leukemia cells in blood will grow very fast and become severe at a short time. Generally, leukemia can be divided into 4 types that are:

- **Acute Lymphocytic Leukemia (ALL)**: Usually occurs in children aged 2-10 years. This type of leukemia is most common. It also always occurs in adults.
- **Acute Myeloid Leukemia (AML)**: This type of leukemia is common in children under the age of 1 year. It is extremely rare in teenagers. Even so it is mostly in adults aged 40 years.
- **Chronic Lymphocytic Leukemia (CLL)**: This type of leukemia often happened to older patients. It is extremely rare in patients under the age of 40.
- **Chronic Myeloid Leukemia (CML)**: This type of leukemia can occur in all but the most common is for adults age after 45 years.

For the detection and diagnosis of cancer from microscopic biopsy images, the histopathology normally looks at the specific features in the cells and tissue structures. The features used for finding cancer from the microscopic images include shape and size of cells, cell nuclei, and distribution of the cells. The brief descriptions of these features are given as follows.

- **Size and shape of the Cells**: It has been observed that the overall shape and size of cells in the tissues are mostly normal. The cellular structures of the cancerous cells might be either larger or smaller than normal cells. The normal cells have even shapes and functionality.

**II. BACKGROUND AND LITERATURE REVIEW**

**Methodology**: After studying the literature, we review following general methodology for the automatic blood cancer detection.

1. **Microscopic images**: Cancer infected blood cell images is collected from the authorized laboratory or from any government hospital in order to carry out the further processing.
2. **Enhancement**: Images may contain some artifacts initially, so there is a need to enhance the taken images. Mostly images contain some sort of noise, so before proceeding further these artifacts should be removed with the help of image enhancement techniques like to remove noise use various kinds of operators i.e. prewitt and sobel, canny etc.
3. **Segmentation**: Segmentation is a process of partitioning an image into sub parts, so that proper each and every area is scanned properly. Microscopic images consist of red blood cells, white blood cells and platelets. But in order to detect the presence of blood cancer we only required to count the number of white blood cells. So, with the help of segmentation process we will separated the white blood cells from red blood cells and platelets. Various techniques used for segmentation are region-based segmentation, k-means Zack algorithm, morphological operation, gradient magnitude and watershed transform etc.
4. **Feature Selection**: In this phase we try to extract some of the features from the processed image. Feature extraction is the process of converting the image into data so that we can check these values with the standard values and finally we can differentiate between the cancerous and noncancerous data. Some of the features which are necessary to be calculated are listed below.
   - **Color Features** – The mean color values of the grey images are acquired.
   - **Geometric Features** – The perimeter, radius, area, rectangularity, compactness, convexity, concavity, symmetry, elongation, eccentricity, solidity is obtained.
• **Texture Features** – The entropy, energy, homogeneity, correlation is obtained.

**Statistical Features** – The skewness, mean, variance and gradient matrix are obtained.

Radius – measured by averaging the length of the radial line segments defined by the centroid and border points.

Perimeter - the total distance between consecutive points of the border

Area – the number of pixels on the interior of the cell, defined separately for the nuclei and for the whole cell; as the features we assume the area of the nucleus and the ratio of the areas of the nucleus and the whole cell.

Compactness – given by the formula: perimeter²/area

Concavity – the severity of concavities in a cell

Concavity points – the number of concavities, irrespective of their amplitudes

Symmetry – the difference between lines perpendicular to the major axis to the cell boundary in both directions

Major and minor axis lengths.

5. **Classifier:** In this final phase, the extracted features are used to provide the final answer. All feature extracted are listed into the different columns with their values. When we give any image as an input to the proposed system then we first calculate the feature values. The values of the test image features are checked with the previously calculated values. Based on the values of the input image the classifier classifies that test image into either infected or not infected class.

III. **PROPOSED SYSTEM**

**A. BLOCK DIAGRAM**

![Block diagram of automatic blood cancer detection](image)

Above fig shows the block diagram of automatic blood cancer detecting using image processing. Microscopic image is taken which goes through an enhancement process further the image is segmented into parts to get the proper result using k-mean algorithm. The parameters such as area, perimeter, eccentricity are accounted and classification are done.

3.3.1 **Enhancements**

The main purpose of the pre-processing is to remove an image degradation such as noise and contrast enhancement of region of interests. The biopsy images acquired from microscope may be defective and deficient in some respect such as poor contrast and uneven staining, and they need to be improved through process of image enhancement which increases the contrast between the foreground (objects of interest) and background [7].
3.3.2 Segmentation: -
The k-means clustering based segmentation algorithms are used because of the preservation of the desired information. From the obtained results it is observed that the k-means clustering based segmentation method performs better in most of the cases as compared to other segmentation approaches under consideration for microscopic biopsy image segmentation.

3.4.3 Feature Extraction: -
Feature extraction is one of the vital steps in the analysis of microscopic images. The features are extracted at cell level and tissue level of microscopic images for better predictions. To get the shape information, we use region-based methods to extract ant circularity, area irregularity, and contour irregularity of nuclei as the three shape features to reflect the irregularity of nuclei in biopsy images. Based on these characteristics, some important shape and morphological based features are explained as follows [4].

Calculation of an area: Number of pixels in one cell makes its area. So, the number of pixels having same label constitute the area of the labelled cells.

Calculation of the perimeter: Any pixel whose four neighborhoods are white is surely not a boundary pixel as it lies interior of the cell. So, we get number of those pixels whose four of the neighborhoods are white. And if we subtract this value from the total area of the Image then this will give area outside the cell along the perimeter of the cell. Hence the form factor can be calculated easily using eqn. (1).

Form factor= \(4 \times \pi \times \frac{\text{area}}{\left(\text{perimeter} \times \text{perimeter}\right)}\) (1). There are sets of features used for finding the feature vector of microscopic images explained as follows.

- **Morphology and Shape Feature**: Describe the shape and morphology features. The considered shape and morphological features in this paper are area, major axis length, equivalent diameter, perimeter, orientation, convex area, solidity, filled area, and eccentricity.
- **Identification Process**: From the regions obtained in the segmentation process, by analyzing their shape, color, and special relation with respect to other regions to determine whether and analyzed region is a nucleus or a leukocyte.
- The features that were used to recognize cellular elements are: circularity to measure the perimeter complexity of a circular object (circularity = perimeter² / (4πarea)), eccentricity to find out how much the object deviates from being circular (eccentricity = dist.(center, focus)), color to determine if a region is darker than other, and containment proportion to establish whether a region contains or is contained by another region.

### 3.4.4 Classification

The classification of microscopic biopsy images is the most challenging task for automatic detection of cancer from microscopic biopsy images. For classification purposes, many classifiers have been used. Some commonly used method of classifications is artificial neural networks (ANN), Bayesian classification, K-nearest neighbor classifiers, support vector machine (SVM), and random forest (RF). Supervised machine learning techniques are used for the classification of microscopic images. There are various steps involved in the supervised learning approaches. First step is to prepare the data (feature set), the second step is to choose an appropriate algorithm, the third step is to fit a model, the fourth step is to train the fitted model, and then the final step is to use fitted model for prediction.

**FLOW CHART**

1. Take the microscopic image
2. Do the preprocessing on the image with the help of enhancement technique
3. Do the segmentation of the image
4. Feature selection is done on the parameters such as area parameter and various shapes.
5. Finally, the classification is done on the data obtained whether the image is cancerous or noncancerous.

*Fig 6. Flowchart of system architecture*
IV. EXPERIMENTAL RESULTS

The proposed algorithm is tested on the blood cell images as shown below. The total number of blood cells present in the image is processed using proposed algorithm (implemented using MATLAB).

Stage-1 Image enhancement and canny edge detection
In this Fig 4.1 used canny algorithm we convert original image into grey scale image by detecting the edge of the cell.

![Image enhancement and canny edge detection](image1.png)

*Fig 4. Result of Image enhancement and canny edge detection.*

Stage-3 Boundary extraction of purple cell
In this Fig 4.3 for boundary extraction of purple cell we converted original image into binary using erosion and binary image technique we extracted boundary from original image and extracted boundary of purple cell and original cell using subtraction of binary image and erosion.

![Boundary extraction of purple cell](image2.png)

*Fig 5 Result of boundary extraction of purple cell.*

Stage-4 clustering
In this Fig 4.4 for colour segmentation we used k-mean algorithm to segment the cell. The algorithm uses a similarity metric to assign all documents to one of k clusters. The clusters are represented as an average of all documents contained within the cluster. This average can be thought of as the centroid of the cluster.

![Clustering results](image3.png)

*Fig 6 Result of clustering cell.*
## V. CONCLUSION

Incident rates of blood cancer have been rising since last two decades. So early fast and effective detection of blood cancer is paramount important. If detected at early stage, blood has one of the highest cure rates. At earlier stage, blood cancer is easy to treat, but at later stage it becomes very difficult, cancerous usually result in near fatal consequences and extremely high costs associated with necessary treatments. We performed blood cancer detection and feature extraction through histogram equalization, k-mean algorithm. Using all these the final output given by the system will tell whether the cancer is recoverable or not. The future scope of the blood cancer detection system could be implemented as a standalone application and can be accurate and efficient.

## REFERENCES


II. S.Jagadeesh,” IMAGE PROCESSING BASED APPROACH TO CANCER CELL PREDICTION IN BLOOD SAMPLES”International Journal of Technology and Engineering Sciences Vol.1 (1) 2013, ISSN: 2320-8007


