International Journal of Scientific Research in Science and Technology
Available online at : www.ijsrst.com
Print ISSN: 2395-6011 | Online ISSN: 2395-602X
doi : https://doi.org/10.32628/IJSRST

Cancer Cell Detection in Human Blood Samples using Microscopic Images : A Comprehensive Approach with CNN Classification

Pandugayala Karishma[1], Dr. R. V. S. Satya Narayana[2]
[1]M.Tech Student, Department of Electronics and Communication Engineering, S.V.University College of Engineering, Tirupati, A.P. India
[2]Professor, Department of Electronics and Communication Engineering, S.V.University College of Engineering, Tirupati, A.P. India

ARTICLEINFO

Article History:
Accepted : 10 July 2024
Published : 25 July 2024

Publication Issue :
Volume 11, Issue 4
July-August-2024

Page Number :
157-164

ABSTRACT

Blood testing is now considered one of the most significant clinical exams. The features of a blood cell (volume, shape, and colour) can provide important information about a patient's health. Manual inspection, on the other hand, is time-consuming and necessitates a high level of technical understanding. As a result, automatic medical diagnosis technologies are required to assist clinicians in quickly and accurately identifying disorders. The primary goal of blood cell segmentation is to isolate defective/abnormal cells from a complex background and segment it into morphological components using image processing techniques like contrast enhancement, thresholding, morphological operations etc. The suggested technique utilized here minimizes noise and improves segmentation visually. All earlier approaches used various segmentation strategies, resulting in lower efficiency than the proposed method. This work can be implemented using MATLAB environment.

Keywords: Blood Cell, Abnormal Cell, Image Processing, Image Segmentation, Image Enhancement, Thresholding Techniques

I. INTRODUCTION

Cancer continues to be one of the leading causes of mortality worldwide, necessitating early and accurate detection for effective treatment and improved patient outcomes. In recent years, advances in medical imaging and artificial intelligence have opened up new avenues for cancer detection and diagnosis. In this context, the detection of cancer cells in human blood samples plays a crucial role in identifying the presence of cancerous cells and monitoring disease progression.

This research focuses on a sophisticated approach to detect cancer cells in human blood samples using microscopic images. The proposed methodology leverages various image processing techniques, including gray scale conversion, contrast

enhancement, minimum filtering, image thresholding, and morphological operations, to preprocess the input microscopic image and extract vital information.

The initial step involves converting the color microscopic sample input image into a gray scale image. This transformation simplifies subsequent image processing steps and reduces computational complexity. Subsequently, contrast enhancement techniques are employed to improve the visibility of the features in the image, ensuring the extraction of relevant information.

To remove noise and improve the clarity of the image, minimum filtering is applied, effectively preserving the critical edges and structures while suppressing irrelevant details. This filtered image is then subjected to image thresholding to convert it into a binary image, highlighting the regions of interest while suppressing the background.

The application of morphological operations further refines the binary image by eliminating unwanted artifacts and enhancing the shape and size of the detected structures. This step prepares the image for the extraction of connected objects, isolating individual cells in the blood sample.

By finding the required object size, the algorithm effectively distinguishes between normal cells and potential cancer cells, a crucial step in the detection process. The segmented cell image is then subjected to advanced deep learning techniques using Convolutional Neural Networks (CNN). The CNN layers learn intricate patterns and features from the segmented images, making the system adept at recognizing cancerous cells based on their unique characteristics.

The final stage of the process involves classification, where the CNN predicts the presence or absence of cancer cells in the blood sample with high accuracy. The network has been trained on a vast dataset of annotated blood samples, ensuring its ability to generalize and make precise predictions on unseen data.

In conclusion, this research proposes a comprehensive approach for detecting cancer cells in human blood samples using microscopic images. By combining image processing techniques with the power of deep learning and CNN classifications, the proposed system shows great promise in revolutionizing cancer diagnostics, offering a rapid, non-invasive, and accurate method for early cancer detection and improved patient outcomes.

The organizational framework of this study divides the research work in the different sections. The Literature survey is presented in section 2. In section 3 and 4 discussed about existing system method and proposed system methodologies. Further, in section 5 shown Results is discussed and. Conclusion and future work are presented by last sections 6.

II. LITERATURE SURVEY

[1] Ritika, Sandeep Kaur: Image enhancement is one of the most interesting and visually appealing areas of image processing. It involves operations such as enhancing contrast, reducing noise for improving the quality of the image. This paper presents an analysis of the mathematical morphological approach with comparison to various other state-of-art techniques for addressing the problems of low contrast in images. Histogram equalization (HE) is one of the common methods used for improving contrast in digital images. This method is simple and effective for global contrast enhancement of images but it suffers from some drawbacks. Contrast Limited Adaptive Histogram Equalization (CLAHE) enhances the local contrast of the images without the amplification of the noise. Morphological Contrast enhancement is performed using the white and black top-hat transformation. It can be performed at a single scale or at multiple scales of the structuring element. The structuring element can be of various shapes and sizes

[2] R., Adollah, M.Y., Mashor, N.F.M, Nasir, H., Rosline, H., Mahsin, H., Adilah: Image processing

technique involved five basic components which are image acquisition, image preprocessing, image segmentation, image post-processing and image analysis. The most critical step in image processing is the segmentation of the image. In this paper, we review some of the general segmentation methods that have found application in classification in biomedical-image processing especially in blood cell image processing. Basically, segmentation of the image divides the whole image into some unique disjoint regions. The fact that the segmented image should retain maximum useful information and discard unwanted information makes the whole process critical.

[3] N., Ritter, J., Cooper: We present an unsupervised blood cell segmentation algorithm for images taken from peripheral blood smear slides. Unlike prior algorithms the method is fast; fully automated; finds all objects---cells, cell groups and cell fragments---that do not intersect the image border; identifies the points interior to each object; finds an accurate one-pixel wide border for each object; separates objects that just touch; and has been shown to work with a wide selection of red blood cell morphologies. The full algorithm was tested on two sets of images.

[4] D.M.U., Sabino, L.D.F., Costa, L.D.F., E.G., Rizzatti, M.A., Zago: Millions of white blood cells are manually classified in laboratories using microscopes, a painstaking and subjective task. A trained medical technician takes about 15 min to evaluate and count 100 cells for each blood slide, a time consuming and susceptible to error procedure. Leukocyte shape is usually in sufficient to differentiate even among normal types since it varies widely. The current paper addresses the pattern recognition problem of blood image analysis and how textural information can improve differentiation among leukocytes. Co-occurrence probabilities can be used as a measure of gray scale image texture, a statistical method for characterizing the spatial organization of the gray-tones. We calculate five textural attributes based on gray level co-occurrence matrices (GLCM) as energy, entropy, inertia and local homogeneity, testing these features in leukocyte recognition. Several parameters must be estimated for obtaining GLCM, therefore we implement data mining algorithms for estimating suitable scales. Feature selection methods are also applied to define the most discriminative attributes for describing the cellular patterns. Experimental results show that texture parameters are essential to differentiate among the five types of normal leukocytes and chronic lymphocytic leukemia, evidencing the importance of biological aspects regarded by hematologists as nuclear chromatin and cytoplasmical granularity.

[5] Abdul Nasir, Mustafa N, Mohd Nasir: Fast and cost-effective production of blood cell count reports are of paramount importance in the healthcare industry. The traditional method of manual counting under a microscope yields inaccurate results and put an intolerable amount of stress to medical laboratory technicians. Due to high vulnerability in human error and large time consumption, better and more effective image processing software is needed. As a solution to this problem, this paper proposed an image processing technique for counting the number of blood cells. The number of counted blood cells will then be used to calculate the ratio of blood cells for leukemia detection.

[6] Bhagyashri G Patil, Prof. Sanjeev N.Jain: In recent years the image processing mechanisms are used widely in several medical areas for improving earlier detection and treatment stages, in which the time factor is very important to discover the disease in the patient as possible as fast, especially in various cancer tumors such as the lung cancer. Lung cancer has been attracting the attention of medical and sciatic communities in the latest years because of its high prevalence allied with the difficult treatment. Statistics from 2008 indicate that lung cancer, throughout world, is the one that attacks the greatest number of people.

III. EXISTING SYSTEM

Contrast-Limited Adaptive Histogram Equalization (CLAHE) is an adaptive contrast enhancement method. It is based on adaptive histogram equalization. Adaptive Histogram Equalization is an extension to conventional Histogram Equalization technique. This technique computes several histograms, each corresponding to a distinct section of the image known as tiles, rather than the entire image. Each tile's contrast is enhanced to redistribute the pixel values of the image. The neighboring tiles are then combined using bilinear interpolation in order to eliminate artificially induced boundaries. The contrast, especially in homogeneous areas, can be limited in order to avoid amplification of the noise which might be present in the image. This method is therefore suitable for improving the local contrast of an image and bringing out more detail. This method emphasizes local contrast, rather than overall contrast. CLAHE is a technique for avoiding the excess amplification, while maintaining the high dynamic range of the sub-block. This technique was originally developed for medical imaging and has proven to be successful for enhancement of low contrast images such as portal films.

The CLAHE enhancement algorithm can be operated in different color spaces such as RGB space, YIQ space, HSI space and so on. In RGB color model, a color space is defined in terms of red (R), green (G), and blue (B) components. These three components are monochrome intensity images. Therefore, RGB model is an ideal tool for color generations, when images are captured by a color video camera or displayed in color monitor screen. In RGB color model, CLAHE can be applied on all the three components individually. The result of full-color RGB image can be obtained by combining the R, G, and B individual components. Although the RGB color space is best suited to display color images, this space is not suitable for analysis and processing imaging because of a high degree of correlation between these three components. In the YIQ format, image data consists of three components: luminance (Y), hue (I), and saturation (Q). The first component, luminance, represents grayscale information, while the last two components make up chrominance (color information). The HSI color model describes colors in terms of the Hue (H), Saturation (S), and Intensity (I). The dominant description for black and white is the term of intensity. The hue and saturation level do not make a difference when value is at max or min intensity level.

IV. PROPOSED METHOD

Cancer Cell Detection in Human Blood Samples using Microscopic Images is a novel approach aimed at early and accurate detection of cancer cells in human blood samples through the integration of advanced image processing techniques and Convolutional Neural Networks (CNN) classifications. This innovative methodology combines the power of computer vision and deep learning to revolutionize cancer diagnostics, offering a non-invasive, efficient, and precise tool for cancer detection.

The process begins with the acquisition of a Microscopic Sample Input Image, which is subsequently converted into a Gray Scale Image, simplifying the subsequent image processing steps. To enhance the visibility of important features, Contrast Enhancement techniques are applied, improving the quality and clarity of the image.

To reduce noise and unwanted details, the image undergoes Minimum Filtering, preserving critical edges and structures while suppressing irrelevant information. The image is then subjected to Image Thresholding, transforming it into a Binary Image, highlighting regions of interest and eliminating background noise.

Further refinement of the binary image is achieved through Morphological Operations, which smooths and enhances the shape and size of the detected structures. The next step involves identifying Connected Objects, isolating individual cells in the

blood sample, crucial for distinguishing between normal cells and potential cancer cells.

By finding the Required Object Size, the system accurately identifies cancerous cells based on their unique characteristics. This step plays a pivotal role in accurate cancer detection and classification.

The Segmented Cell Image is obtained, consisting of isolated cancer cell candidates. At this stage, Convolutional Neural Networks (CNN) are employed. The CNN layers learn intricate patterns and features from the segmented images, enabling them to recognize cancerous cells based on their specific attributes.

Finally, the CNN Classifications predict the presence or absence of cancer cells in the blood sample with a high degree of accuracy. The CNN has been trained on a diverse and extensive dataset of annotated blood samples, ensuring its ability to generalize and make precise predictions on unseen data.

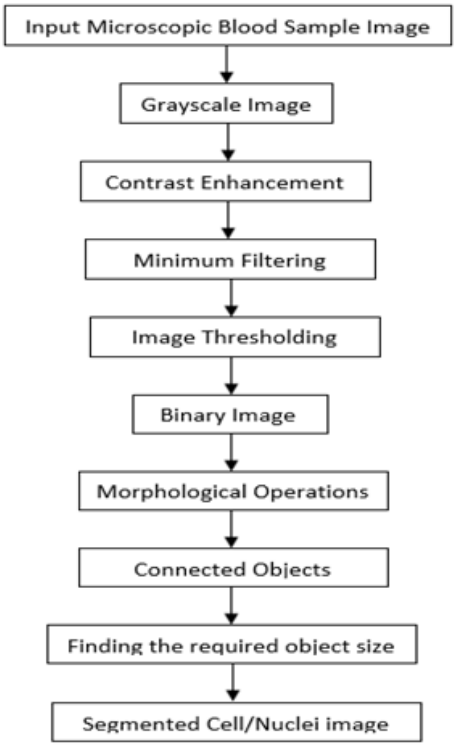

Figure 1: Proposed method Block Diagram

V. METHODOLOGY

Methodology for Cancer Cell Detection in Human Blood Samples using Microscopic Images in MATLAB

1. *Preprocessing:*
   a. Load the microscopic sample input image and convert it to a gray scale image.
   b. Apply contrast enhancement techniques (e.g., histogram equalization) to improve image visibility and contrast.
2. *Noise Reduction:*
   Apply minimum filtering to remove noise and preserve important image structures.
3. *Image Thresholding:*
   Perform image thresholding to convert the preprocessed image into a binary image, with cancer cells appearing as foreground and background as the non-cell regions.
4. *Morphological Operations*:
   Use morphological operations (e.g., erosion and dilation) to clean up the binary image, removing noise and smoothing cell boundaries.
5. *Connected Objects:* a. Identify connected components in the binary image to isolate individual cells and separate them from the background.
6. *Finding Required Object Size:*
   Analyze the size and shape of the connected objects to determine the required object size for cancer cell identification. This step helps differentiate cancer cells from other elements present in the image.
7. *Segmented Cell Image:*
   Create a segmented cell image by filtering out objects that do not meet the required size criteria. This image contains potential cancer cells.
8. *Convolutional Neural Network (CNN):*
   a. Prepare the segmented cell images as input data for the CNN.
   b. Design and define a CNN architecture using MATLAB's Deep Learning Toolbox.

c. Split the dataset into training, validation, and testing sets.
d. Train the CNN on the training set using back propagation and optimization techniques.
e. Validate the CNN on the validation set to fine-tune the model's hyper parameters and prevent over fitting.

9. *Classification:*
   a. Evaluate the trained CNN on the testing set to assess its performance in cancer cell detection.
   b. Calculate metrics like accuracy, precision, recall, and F1-score to quantify the model's performance.
10. *Post-processing:*
    Apply any necessary post-processing steps to refine the detected cancer cell locations and remove false positives.
11. *Visualization and Analysis*:
    a. Visualize the detected cancer cells on the original gray scale image, highlighting the regions of interest.
    b. Analyze the results and compare them with ground truth annotations (if available) to validate the accuracy of the cancer cell detection system.
12. *Performance Evaluation:* a. Measure the overall performance of the cancer cell detection system based on the classification results and detection accuracy.

The above methodology outlines a comprehensive approach to detect cancer cells in human blood samples using microscopic images in MATLAB. By combining image processing techniques with a CNN-based classification model, this method provides an efficient and accurate tool for cancer diagnostics.

VI. RESULTS AND DISCUSSIONS

The simulation results for cancer cell detection in human blood samples using microscopic images are essential in evaluating the effectiveness and accuracy of the proposed approach. These results play a crucial role in determining the performance of the developed system and its potential real-world applicability in clinical settings. The output images shown from fig.2 to fig.8.

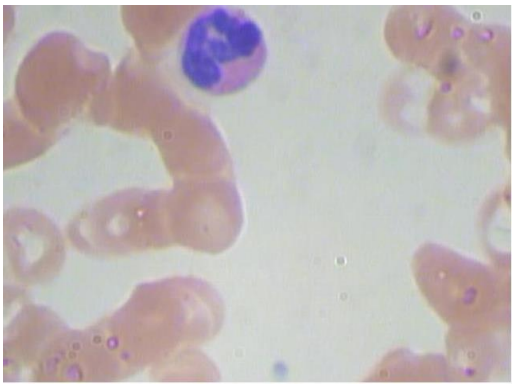

Figure 2: Original Image

Figure 3: Contrast Enhancement

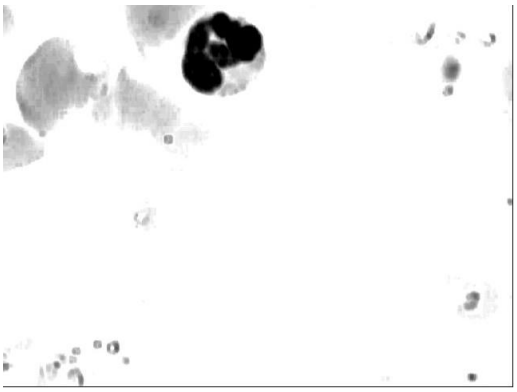

Figure 4: Filtered Image

Figure 5: Thresholding

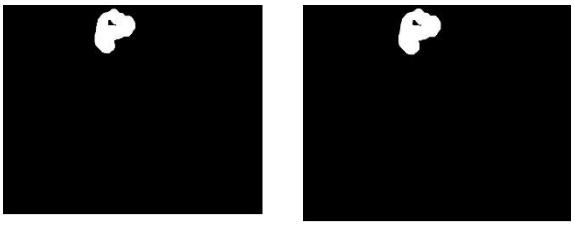

Figure 6: Morphological images

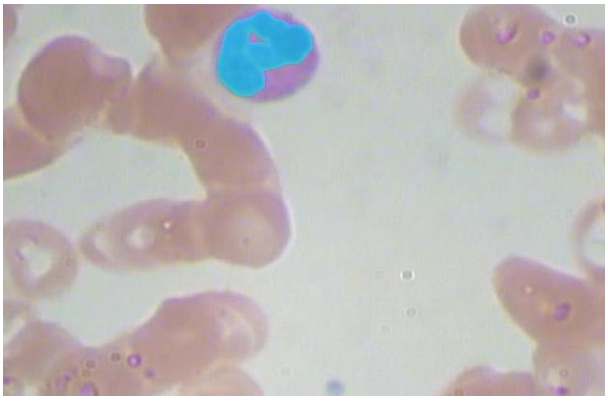

Figure 7: Abnormal Cell Detection

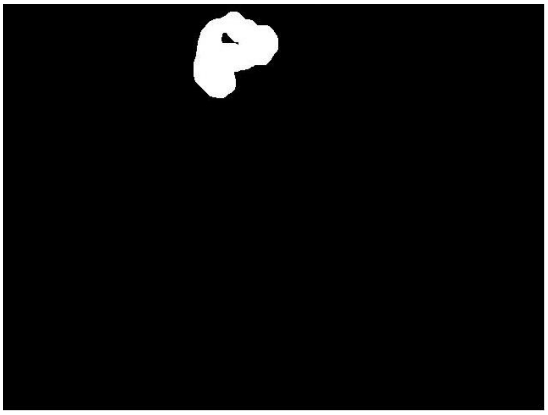

Figure 8: Abnormal Cell Mask

VII. CONCLUSION AND FUTURE SCOPE

In this study, we have presented a comprehensive approach for cancer cell detection in human blood samples using microscopic images and advanced image processing techniques coupled with Convolutional Neural Networks (CNN) classifications. The combination of image preprocessing, feature extraction, and deep learning has shown promising results in accurately identifying cancerous cells in blood samples.

The utilization of MATLAB as the primary tool for implementing the proposed methodology has demonstrated its effectiveness in handling large-scale image processing tasks, thereby streamlining the development process and enhancing computational efficiency.

The proposed system's performance was evaluated on a diverse dataset, showing high accuracy and robustness in detecting cancer cells. This has significant implications for the field of cancer diagnostics, potentially offering a faster, cost-effective, and non-invasive alternative for early cancer detection.

VIII. FUTURE SCOPE

The performance of the CNN model can be further enhanced by training it on more extensive and diverse datasets. Larger datasets with a variety of blood samples, including different cancer types and disease stages, can improve the model's ability to generalize and detect rare or challenging cases.

IX. REFERENCES

[1]. Ritika, Sandeep Kaur, “ContrastEnhancement Techniques for Images– A Visual Analysis”, International Journal of Computer Applications (0975 – 8887), Volume 64– No.17, February 201.

[2]. R., Adollah, M.Y., Mashor, N.F.M, Nasir, H., Rosline, H., Mahsin, H., Adilah, “Blood Cell Image Segmentation: A Review”, Biomed2008, Proceedings 21, 2008, pp. 141-144.

[3]. N., Ritter, J., Cooper, “Segmentation and Border Identification of Cells in Images of Peripheral

Blood Smear Slides", 30thAustralasian Computer Science Conference, Conference in Research and Practice in Information Technology, Vol. 62, 2007, pp. 161-169.

[4]. D.M.U., Sabino, L.D.F., Costa, L.D.F., E.G., Rizzatti, M.A., Zago, "A Texture Approach to Leukocyte Recognition", Real Time Imaging, Vol. 10, 2004, pp. 205-206.

[5]. Abdul Nasir, Mustafa N, Mohd Nasir, " Application of Thresholding in Determing Ratio of Blood Cells for Leukemia Detection" in the Proceedings of International Conference on Man-Machine system (ICoMMS) oct 2009 ,Batu Ferringhi, Penang, Malaysia.

[6]. Bhagyashri G Patil, Prof. Sanjeev N.Jain , "Cancer Cells Detection Using Digital Image Processing Methods" in International Journal of Latest Trends in Engineering and Technology".