How to Cite:
Sailaja, M., & Sreenivasulu, G. (2022). An image processing approach for enumeration of leukemia infected cells for diagnosis. *International Journal of Health Sciences*, *6*(S9), 4062–4071. https://doi.org/10.53730/ijhs.v6nS9.13591

An image processing approach for enumeration of leukemia infected cells for diagnosis

M. Sailaja
Department of Electronics and Communication Engineering, Govt. Polytechnic for Women, Kadapa, AP, India-516 001.
Corresponding author email: sailaja172013@gmail.com

G. Sreenivasulu
Department of Electronics and Communication Engineering, S. V. University, Tirupati, AP-517502.

***Abstract*---**Image processing techniques are helped to improve the diagnosis of various diseases using medical images. Leukemia, a blood cancer, is one of the commonest malignancies affecting both adults and children. As per the reports, leukemia was the fifth and sixth cause of death in men (7%) and women (5.8%) with cancer, and it was the first cause of death in children with cancer between 1–4 and 5–14 years old, with 48.5% and 52.2% of deceases, respectively. Hematologists microscopically examine the blood under a light microscope for diagnosis. This process is very tedious, time-consuming, and not suitable for analyzing a large number of cells. The enumeration of infected cells for the diagnosis of leukemia will assist to overcome these drawbacks. Enumeration of blood cells plays a very important role in the health sector. In this paper, we have proposed the enumeration of leukemia-infected cell parasites in microscopic blood images. The infected parasites are detected by the proposed threshold method by maximizing the between-class variance and entropy of black and white pixels. The enumeration is carried out by connected component labeling technique from binary decision image obtained in maximizing entropy. The proposed yields more accurate results when compared to the traditional methods.

***Keywords*---**Leukemia, entropy, computerized diagnosis, threshold, connected component labeling, and microscopic images.

1. Introduction

Cancer is a major public health problem worldwide. **Leukemia is a blood cancer caused by a rise in the number of white blood cells in your body**. It

was the 11th leading cause of cancer-related mortality worldwide in 2021. Leukemia accounted for approximately 3.4% of all new cancer cases and 3.8% of all cancer deaths in 2021 according to the Surveillance and Epidemiology. With the advent of new treatments, such as mutational targeted inhibitors, proapoptotic agents, chimeric antigen receptor (CAR) T-cell therapy, and immunotherapy, mortality due to leukemia has descended recently, but leukemia is still a highly prevalent disease that leads to considerable disability and increased economic costs [1]. It not only results in a major personal burden, but also affects families and the economic structures of countries. A systematic analysis will help quantify health loss due to leukemia and guide policy-making by healthcare providers aimed at improving health systems and decreasing the burden of leukemia over time [2].

Microscopic blood tests are considered as the main procedure for the diagnosis of leukemia. Analysis of blood smears is the most common way of discovering leukemia, but not the only one technique. Interventional radiology is an alternative technique for the diagnosis of leukemia. However, radiological techniques, such as percutaneous aspiration, biopsy, and catheter drainage, suffer from inheriting limitations of imaging modality sensitivity and resolution of the radio images. Moreover, other techniques, such as Molecular Cytogenetics, Long Distance Inverse Polymerase Chain Reaction (LDI-PCR), and Array-based Comparative Genomic Hybridization (aCGH), need extensive work and time to identify leukemia types [3]. Due to time and cost requirements of these techniques, microscopic blood tests and bone marrow are the most common methods for identification of leukemia subtypes [4].

However, the manual method of diagnosis is tedious, requires expert technicians, and is prone to human erroneous. Hence, it affected the accuracy of the diagnosis by pathologists. Considering these problems, there is need for an automated system for diagnosis of leukemia [5]. Hence, we propose a entropy based technique by which we can extract the infected cells and consequently can enumerate these cells. The connected component labeling technique is carried out by using the detection of leukemia parasite images for enumerating number of infected cells. The leukemia infected sample microscopic blood images are shown in Figure-1.

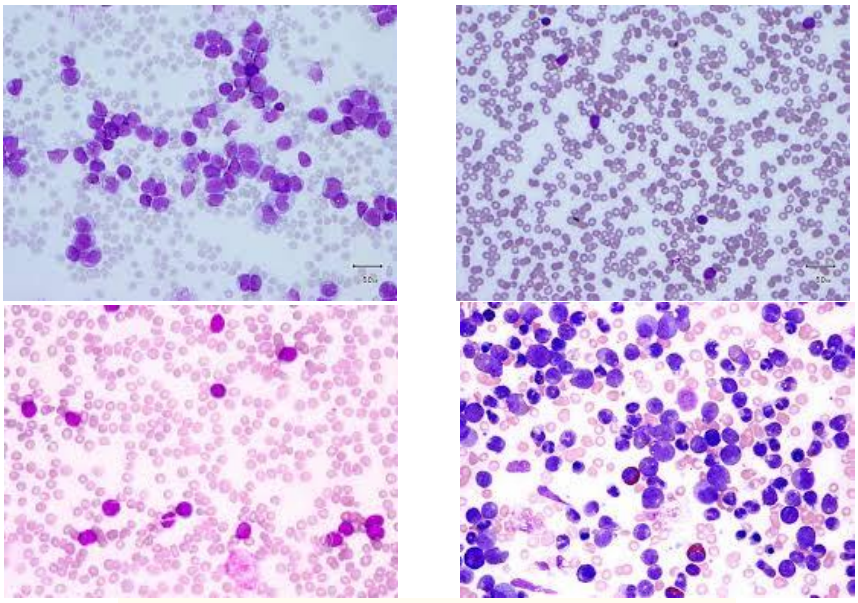


Fig.1: Sample microscopic images of Leukemia.

The rest of the article is organized as follows. The methodology for leukemia parasite enumeration is described in section 2. Section 3 presents the experimental results followed by conclusions are drawn in section 4.

2. Methodology

The enumeration of leukemia infected cells through microscopic blood images consists of the following steps sequentially. (i) Color space transformation, (ii) proposed threshold technique, (iii) Parasite detection,(iv) Connected component labeling. The following sub sections comprises in details of the steps involved.

The input color blood image be represented by

$$[f(x,y)]_{m\times n} = \begin{bmatrix} f_R(x,y) \\ f_G(x,y) \\ f_B(x,y) \end{bmatrix}_{m\times n} \tag{1}$$

Where $R(x, y)$, $G(x, y)$, and $B(x, y)$ denote the red, green, and blue color values of a pixel located at (x, y) with x, y are pixel coordinates in the width and height dimensions, respectively such that $1 \leq x \leq m$ and $1 \leq y \leq n$

2.1. Color space transformation

The original RGB image $f(x,y)$ is converted into gray scale image for single channel processing by using Equation (2).

$$\begin{aligned} \mathrm{h}(x,y) &= [w_1 \quad w_2 \quad w_3] \begin{bmatrix} f_R(x,y) \\ f_G(x,y) \\ f_B(x,y) \end{bmatrix} \\ &= w_1 f_R(x,y) + w_2 f_G(x,y) + w_3 f_B(x,y) \end{aligned} \tag{2}$$

where w_i $(for\ i = 1,2,3)$ are weights of three color channels (Red, Green ,Blue), are given by 0.29893,0.58704, 0.14020, respectively.

2.2. Leukemia infected cells detection

Image threshold is one of the widely used methods for image segmentation with a number of applications, such as image analysis, scene interpretation, and object recognition. It is useful in differentiating the foreground from the background. By selecting a threshold value T, the gray level image can be converted into a binary image. Suppose T is a threshold value and {a,b} is a pair of binary values with a=0 and b=1. The result of threshold image $g(x,y)$ at a threshold value T is a binary image $f(x,y)$ such that

$$f(x,y) = \begin{cases} b & ,\text{if } g(x,y) \geq T \\ a & ,\text{if } g(x,y) < T \end{cases} \tag{3}$$

The bi-level thresholding of an image $f(x,y)$ is defined in Equation (2). Herein, an image can be represented by L gray levels:

$$\left.\begin{aligned} G_1 &= \{f(x,y) \in I: 0 \leq f(x,y) \leq T-1\} \\ G_2 &= \{f(x,y) \in I: T \leq f(x,y) \leq L-1\} \end{aligned}\right\} \tag{4}$$

The multi-level thresholding uses more than one threshold value and creates an output image with multiple groups. The mathematical expressions for multi-level thresholding with $(n-1)$ number of thresholds are shown in Equation (5).

$$\left.\begin{aligned} G_1 &= \{f(x,y) \in I: \ 0 \leq f(x,y) \leq T_1 - 1\} \\ G_2 &= \{f(x,y) \in I: \ T_1 \leq f(x,y) \leq T_2 - 1\} \\ G_3 &= \{f(x,y) \in I: \ T_2 \leq f(x,y) \leq T_3 - 1\} \\ G_i &= \{f(x,y) \in I: \ T_{i-1} \leq f(x,y) \leq T_i - 1\} \\ G_n &= \{f(x,y) \in I: T_{n-1} \leq f(x,y) \leq L-1\} \end{aligned}\right\} \tag{5}$$

It is difficult to fix a threshold value for parasite detection because the medical images are heterogeneous. Since the medical images are heterogeneous and it is difficult to fix threshold value for parasite detection. So far no threshold method exists to find a threshold value for parasite detection using microscopic images. We attempt to solve the problem of finding a threshold value for parasite detection. However, it is not an easy task [6-8]. Therefore, the proposed method is carried out in two stages and the detailed procedure is shown as follows.

The **Heaviside function** $U(t)$ of a single variable t is defined by the following equation.

$$U(t) = \begin{cases} 1 & ,\text{if } t \geq 0 \\ 0 & ,\text{elsewhere} \end{cases} \tag{6}$$

The major mathematical equation of the stage-one threshold image corresponding to the gray scale original image $h(x,y)$ of size $m \times n$, is provided as follows.

$$f(x,y) = h(x,y)\left[U\left(\lambda - h(x,y)\right)\right] + \alpha[U(h(x,y) - \lambda)] \tag{7}$$

where $U(.)$ is a Heaviside function as defined in the Equation (6), and λ, α are the parameters. The value of the parameter α is set to be $(L-1) \times 0.7$. The selection α will be discussed in the experimental section. The selection of λ is

calculated as follows. Define the between-class variance $V_b(t)$ by using the within-class variance $V_w(t)$ as shown below [8].

$$\begin{aligned} V_b(t) &= V(t) - V_w(t) \\ &= V(t) - [W_1(t)V_1(t) + W_2(t)V_2(t)] \\ &= W_1(t)W_2(t)[\mu_1(t) - \mu_2(t)]^2 \end{aligned} \tag{8}$$

where $W_1(t)$, $V_1(t)$ and $\mu_1(t)$ indicate total amount, variance and the mean of inside class, respectively.
Analogous definition can be provided for the corresponding values $W_2(t)$, $V_2(t)$ and $\mu_2(t)$ of the between-class variance [24]. Thus the value of the parameter λ can be calculated by using the Equation (12):

$$(L-1)\underbrace{\text{Arg max}}_{0\le t\le N-1}\{W_1(t)W_2(t)[\mu_1(t) - \mu_2(t)]^2\} \tag{9}$$

Let h(i) be value of a normalized histogram. The value of 'i' takes integer values from 0 to 255.

$$\sum_{i=0}^{i_{max}} h(i)=1$$

Entropy of white pixels:

$$H_B(t)=-\sum_{i=0}^{t}\frac{h(i)}{\sum_{j=0}^{t}h(j)}\log\frac{h(i)}{\sum_{j=0}^{t}h(j)}$$

Entropy of black pixels:

$$H_W(t)=-\sum_{i=t+1}^{i_{max}}\frac{h(i)}{\sum_{j=t+1}^{i_{max}}h(j)}\log\frac{h(i)}{\sum_{j=t+1}^{i_{max}}h(j)}$$

Optimal threshold can be selected by maximizing the entropy of black and white pixels:

$$T=\underset{t=0\ldots i_{max}}{Arg\,Max}\,H_B(t)+H_W(t)$$

The detection of leukemia parasites is carried out for each element of the image , $h(x,y)$ using the optimal threshold value T_{opt} obtained in stage-two threshold. The leukemia parasite detection image is obtained as follows.

$$g(x,y) = W\big(h(x,y)-T_{opt}\big) \tag{11}$$

$$\text{s.t.,}\ \ W(k) = \begin{cases} 0, & k < 0 \\ 1, & k \geq 0 \end{cases} \tag{12}$$

Each value of an image $g(x,y)$ has either 0 or 1. The elements of the image $g(x,y)$ with value 1(white) is belongs to the signal (infected one) otherwise normal.

2.3. Connected component labeling

The enumeration of leukemia parasites carried out by using connected component labeling (CCL) algorithm [17-19]. It is the process of identifying the connected components in a binary image by assigning each component with a different label. All pixels of the each connected component have similar values. Pixel connectivity is used to identify which pixels are connected to other pixels in the surrounding neighborhoods. The most commonly used two connectivity's for connected component labeling in image analysis are 4-connectivity (N_4) and 8-connectivity (N_8) mathematical expressions are given in Equations (13) and (14) of $f(x,y)$. We have chosen 8-connected neighborhood and yields more accurate results [10].

$$N_4\big(f(x,y)\big) = \{f(x+i, y+j) : (i,j) = (1,0), (-1,0), (0,1), (0,-1))\} \tag{13}$$

$$N_8\big(f(x,y)\big) = N_4(f(x,y)) \cup \{f(x+i, y+j) : (i,j) = (1,1), (1,-1), (-1,1), (-1,-1)\} \tag{14}$$

In principle, CCL algorithm consists of the following steps.
Step 1 – Scan pixel by pixel in a binary mask of infected RBC image.
Step 2 – If pixel is not back ground then check neighbors.
Step 3 – If neighbors already labeled then assign neighbors parent label to main label.
Step 4 – If none of neighbors is labeled then assign new label to pixel.

The results obtained by applying CCL algorithm to the sample binary matrix and parasite detected images are shown in Figure-2. Besides, we calculated number of pixels (pels) infected with leukemia parasites in a microscopic images which assist pathologists for classification of degree (low, medium and high) of parasitemia for accurate treatment. The result of enumeration of infected pels and % of parasitemia for binary matrix as shown in Figure-3.

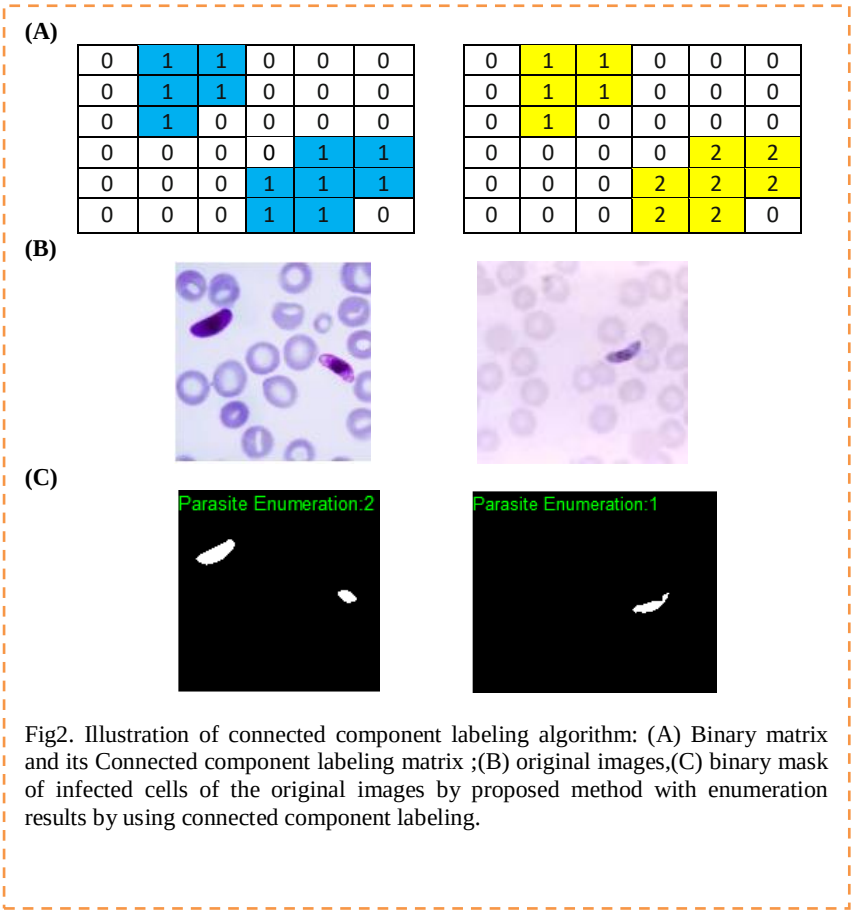


Fig2. Illustration of connected component labeling algorithm: (A) Binary matrix and its Connected component labeling matrix ;(B) original images,(C) binary mask of infected cells of the original images by proposed method with enumeration results by using connected component labeling.

3. Experimental results

The leukemia parasite enumeration method was performed using MATLAB 7.10.0 (R2010a) Programming software on a personal computer with an AMD Phenom II N830 triple-core processor 2.10 GHz, 3 GB system memory and 64-bit windows-7 operating system. The average computational time of one image was approximately 7 seconds.

0	1	1	0	1	0
0	1	1	0	0	0
0	1	0	0	0	0
0	0	0	0	1	1
0	0	0	1	1	1
0	0	0	1	1	0

(a)

0	1	2	0	3	0
0	4	5	0	0	0
0	6	0	0	0	0
0	0	0	0	7	8
0	0	0	9	10	11
0	0	0	12	13	0

(b)

Infected parasitemia(%)=36.11

(c)

Fig.3. illustration for enumeration of infected pels for parasite density: (a) Binary matrix, (b) enumeration of infected pels, (c) infected parasitemia (%) in blood sample. [% of parasitemia= (number of infected pels / total number of pels)*100]

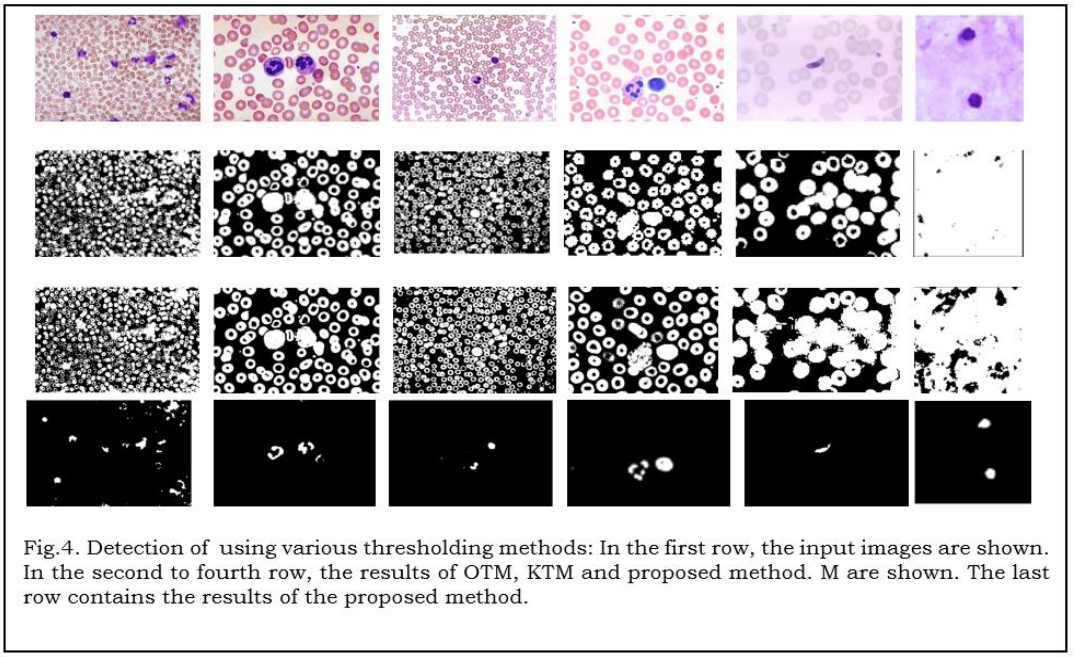
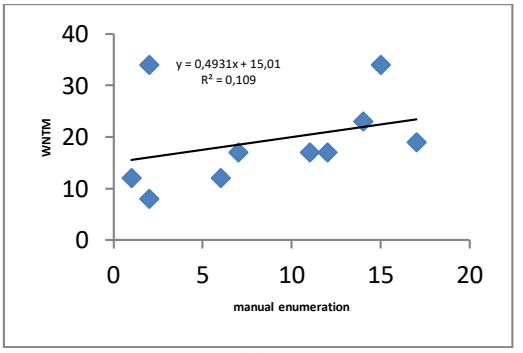
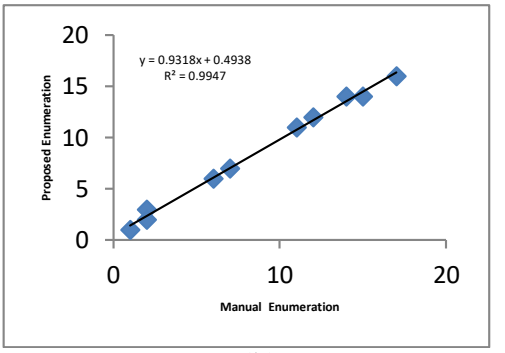


Fig.4. Detection of using various thresholding methods: In the first row, the input images are shown. In the second to fourth row, the results of OTM, KTM and proposed method. M are shown. The last row contains the results of the proposed method.

The proposed method will be compared with some of other our implementation thresholding methods, which includes Otsu (OTM) [24], Kittler (KTM) [26]. The leukemia infected microscopic images, together with the detection results based on OTM, KTM, and proposed method are shown in Figure 4. The linear regression to establish the correlation between the enumerations of various thresholding methods with proposed one using the manual enumeration as shown in Figure 6. The OTM and KTM linear regression of enumeration as shown in Fig.5. The proposed method yields highest correlation with manual infected cells when compared to KTM and OTM.



(a)



(b)

Fig.5. Linear regression of various threshold techniques enumeration with proposed technique on manual enumeration of ten sample images. (a)OTM, (b)KTM.

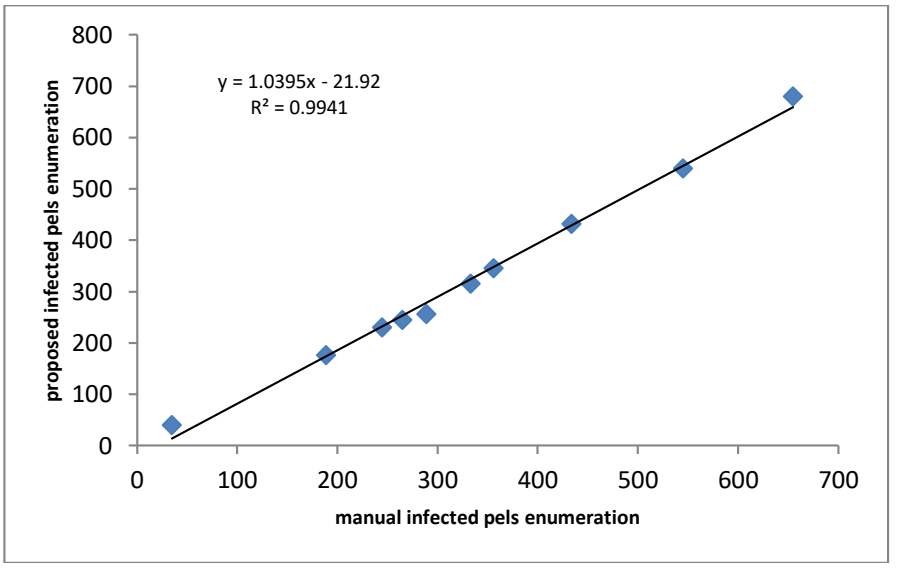


Fig.6. Linear regression of manual and proposed infected pels enumeration of ten samples

4. Conclusion

In this paper, we have proposed a technique to enumerate the infected cells with leukemia parasites in microscopic blood images for assisting diagnosis. First, the color image is converted in to gray scale image for single channel processing. The infected parasites are detected by the proposed threshold method that comprises the maximizing the between-class variance of an original image and maximizing entropy for optimal threshold value. The enumeration infected cells and infected pels (for % of parasitemia) using connected component labelling. The proposed method outperforms well when compared to the other traditional threshold methods. There is approximately perfect correlation between manual and proposed enumeration results. The work is useful in telepathology applications for computer aided diagnosis (CAD).

References

[1] Abdul-Nasir, A.S., Mashor, M.Y. & Mohamed, Z., Colour image segmentation approach for detection of leukemia parasites using various colour models and k-means clustering. WSEAS Trans. Biol. Biomed. 10, 41–55, 2013.

[2] Ahirwar, N., Pattnaik, S., Acharya, B. Advanced image analysis based system for automatic detection and classification leukemial parasite in blood images. Int. J. Inf. Technol. Knowl. Manag. 5, 59–64,2012.

[3] Arco, J., Go´rriz, J., Ram´ırez, J., A´ lvarez, I. & Puntonet, C. Digital image analysis for automatic enumeration of leukemia parasites using morphological operations. Expert Syst. Appl. 42, 3041–3047,2015.

[4] Boray Tek F, Dempster AG, Kale I. Computer vision for microscopy diagnosis of leukemia. Leukemia J 2009;8:153.
[5] D.K. Das, R. Mukherjee,C. Chakraborty, Computational microscopic imaging for leukemia parasite detection: a systematic review, Journal ofMicroscopy, Vol. 260, Issue 1 2015, pp. 1–19.
[6] Frean JA. Reliable enumeration of leukemia parasites in thick blood films using digital image analysis. Leukemia J 2009:8.
[7] Junaid Ahmed, Padmalaya Nayak, A survey on leukemia detection using image processing, E3S web of conferences, 2021 (ICMED 2021).
[8] K. Jain, Fundamentals of digital image processing. Englewood Cliffs, NJ, Prentice-Hall, 1989.
[9] Kaewkamnerd S, Uthaipibull C, Intarapanich A, Pannarut M, Chaotheing S. An automatic device for detection and classification of leukemia parasite species in thick blood film. BMC Bioinform 2012;13.
[10] Lim HN, Mashor MY, Hassan R. White blood cell segmentation for acute leukemia bone marrow images. IEEE International Conference on Biomedical Engineering (ICoBE) 2012; 357–361.
[11] Lim, Huey Nee, Mohd Yusoff Mashor, and Rosline Hassan. "White blood cell segmentation for acute leukemia bone marrow images." In Biomedical Engineering (ICoBE), 2012 International Conference on, pp. 357-361. IEEE, 2012.
[12] Linder, N., Turkki, R., Walliander, M., et al., A leukemia diagnostic tool based on computer vision screening and visualization of Plasmodium falciparum candidate areas in digitized blood smears. PLoS One 9, e104855, 2014.
[13] Makkapati, V.V. & Pathangay, V. Adaptive color illumination for microscopes. In Proceedings of National Conference on Communications, pp. 1–5. Bangalore, India,2011.
[14] NEGI—National Institute of Statistics, Geography, and Informatics. Statistics with respect to the World Day against Cancer.
[15] Pan C, Park DS, Yoon S, Yang JC. Leukocyte image segmentation using simulated visual attention. Expert Systems with Applications. 2012; 39(8): 7479–7494.
[16] Prasad Keerthana, Winter Jan, Bhat Udayakrishna M, Acharya Raviraja V, Prabhu Gopalakrishna K. Image analysis approach for development of a decision support system for detection of leukemia parasites in thin blood smear images. J Digit Imag 2012:542–9.
[17] R. C. Gonzalez, R. E. Woods, Digital image processing. 2nd ed. Reading, MA. Addison-Wesley, 1992, pp. 85-103.
[18] Ross NE, Pritchard CJ, Rubin DM, Duse AG. Automated image processing method for the diagnosis and classification of leukemia on thin blood smears. Med Biol Eng Comput 2006;44: 427–36.
[19] Suryawanshi, M.S. & Dixit, V. Improved technique for detection of leukemia parasites within the blood cell images. Int. J. Sci. Eng. Res. 4, 373–375, 2013.