ISSN: 2249-7196
IJMRR/March-May 2024/ Volume 14/Issue 3/175-185
B. Y. Namratha Sri / International Journal of Management Research & Review

Blood Cell Counting Using Watershed Algorithm

Dr. E. Govinda[1], N. Nagamani[2], B. Y. Namratha Sri[3], G. Venu Madhav[4], P. Lavanya[5]

[1]Associate Professor, Department of Electronics and Communication Engineering Avanthi Institute of Engineering & Technology, Tamaram, Makavarapalem, Anakapalle – 531113, Andhra Pradesh, India.

[2]Assistant Professor, Department of Electronics and Communication Engineering Avanthi Institute of Engineering & Technology, Tamaram, Makavarapalem, Anakapalle – 531113, Andhra Pradesh, India.

[3,4,5]UG Student, Department of Electronics and Communication Engineering Avanthi Institute of Engineering & Technology, Tamaram, Makavarapalem, Anakapalle – 531113, Andhra Pradesh, India.

Abstract

Blood cell counting is traditionally carried out either manually with a haemocytometer or through automated analyzers, and both approaches have drawbacks. Manual counting is slow and laborious, and errors creep in because cells overlap and visual inspection is not consistent from one observer to the next. Automated analyzers are expensive and cannot detect variations in cell shape or other irregularities. This paper proposes a digital image processing method that improves accuracy while reducing the time and cost of blood cell analysis. The process begins with acquisition of a stained blood sample image using a microscope and camera. The image is then pre-processed to remove noise and to extract the colour planes corresponding to red blood cells and white blood cells. Cells are separated using the watershed algorithm, which treats pixel intensity as elevation and floods the image from marker points so that touching and overlapping cells are split into distinct regions. Counting is then performed on the segmented regions using shape and size criteria. The method achieved an accuracy of 94% for red blood cell counts and 92% for white blood cell counts when compared with manual counting. A MATLAB application module was also developed, allowing a user to upload an image and receive the cell counts immediately.

Index Terms*— Blood cell counting, Canny edge detection, digital image processing, morphological operations, red blood cells, segmentation, watershed algorithm, white blood cells.*

I. Introduction

Digital image processing (DIP) is a developing area of computer science engineering with branches in almost every field, and one of the fastest-growing of these is medicine. DIP uses computer algorithms to process digital images, and its impact on modern practice has been considerable. At present, blood samples are taken to a laboratory, processed with various substrates, and the results are produced after some delay. The work described here is interdisciplinary, combining biomedical and computer science methods: a stain is applied and absorbed by the blood sample, an image of the stained sample is captured, and that image is processed by software so that the result is displayed immediately.

The haemocytometer is the device normally used in laboratories to count blood cells. It is a microscope glass slide with a rectangular indentation that forms a chamber, into which a grid of perpendicular lines is etched. Because the

volume of the chamber is known, the number of cells in a given volume of fluid can be counted and the concentration calculated. A physician views the haemocytometer through a microscope and counts the cells with a hand counter.

The manual method has three well-known drawbacks. Counting cells by hand is time consuming and laborious. Overlapping cells are a considerable overhead, because a clump has to be resolved by eye. Visual inspection is not consistent, so two observers, or the same observer on two occasions, may not produce the same count.

In automated blood cell counting a complete blood count is performed by an automated analyzer. The blood is mixed without shaking and placed on an analyzer rack. The instrument contains several components that analyze different elements of the blood, and a cell counting component determines the numbers and types of cells present. Results are printed out or sent to a computer for review. The drawbacks here are cost, since an automated analyzer is not cheap, and sensitivity, since variations in cell shape and other irregularities cannot be detected.

A. Related Work

Berge et al. [1] presented morphological iterative threshold techniques. Segmentation was performed on red blood cells including clumped cells, and boundary curvatures were used to construct a Delaunay triangulation. Real microscopy images prepared in the laboratory were used. The method could not tolerate a high degree of overlap between cells, and the iterative threshold could not detect faded red cells.

Khan et al. [2] proposed a method to count red blood cells, white blood cells and platelets. Several pre-processing steps convert the image to binary form, and an optimal threshold value calculated from the histogram plays the major role in segmentation and counting. An accuracy of 95.23% was reported against manual counting and a hematology analyzer, but overlapping cells could not be detected. Iterative thresholding therefore carries a high probability of losing valid information.

Nguyen et al. [3] introduced distance transformation to address the overlapping cell problem, concentrating on clumped cells at a higher degree of overlap. Chiu et al. [4] proposed a randomized Hough transform for detecting circles, which is fast but becomes less efficient on complex images because of its probabilistic sampling.

Mahmood and Mansor [5] analyzed ten samples of normal blood cells. The image was converted to HSV colour space and the saturation channel was selected for analysis. Morphological operators and thresholding were used for cell segmentation, and a circular Hough transform was applied to check for the circularity that indicates a red blood cell. Their method achieved approximately 96.54% accuracy compared with manual counting. Bhamare and Patil [6] presented a preliminary study of automatic blood cell counting using digital image processing.

II. Literature Review

A literature review contextualizes a research effort by giving an overview of the current state of knowledge on the chosen topic. Beyond summarizing individual sources, it identifies patterns and discrepancies across the literature, points to gaps that warrant further investigation, and allows the methodological approaches used in earlier studies to be evaluated for their strengths and limitations. The sources examined for this work are summarized below.

A. Improved Red Blood Cell Counting in Thin Blood Smears [1]

This paper addresses accurate counting of red blood cells in thin blood smears, a step that matters in both diagnostics and research. Using digital image processing, the authors present a method aimed at improving the precision and reliability of the count. The contribution lies in improving medical assessment in areas where precise red cell counts are needed for diagnosis and monitoring.

B. An Accurate and Cost-Effective Approach to Blood Cell Count [2]

The authors address accuracy and cost together. Accurate counts are needed for diagnosing a range of conditions, but conventional counting equipment can be prohibitively expensive. The proposed image processing method matches the accuracy of conventional methods at significantly lower cost, which is relevant in resource-constrained settings.

C. Cell Splitting with High Degree of Overlapping [3]

This work presents a technique for segmenting cells that overlap heavily in peripheral blood smear images. Overlap affects both counting and classification accuracy. The method uses image processing algorithms to delineate individual cells even under extensive overlap, which improves the reliability of the analysis that follows.

D. Fast Randomized Hough Transform for Circle Recognition [4]

An algorithm for rapid recognition of circles and circular arcs in digital images is introduced. Identifying circular shapes accurately matters in biomedical imaging, since red blood cells are approximately circular in a smear. The randomized Hough transform gives better performance than the conventional transform, though at the cost of some robustness on complex images.

E. Red Blood Cell Estimation Using Hough Transform [5]

The authors estimate the number of red blood cells in digital images using the Hough transform. The method provides a robust means of detecting and quantifying red cells and contributes to the refinement of automated analysis systems.

F. Automatic Blood Cell Analysis Using DIP [6]

This paper is a preliminary exploration of automating blood cell analysis with digital image processing. Manual methods are labour-intensive and subjective, which motivates automation. The study investigates the feasibility of analysing blood cell images automatically and lays groundwork for later development, although it is an initial step rather than a complete system.

G. Reference Texts and Later Work [7]–[10]

Sridhar [7] and Gonzalez and Woods [8] provide comprehensive treatments of digital image processing covering fundamental principles through to advanced methods, and were used as the theoretical basis for the techniques applied here. Thejashwini and Padma [9] present a technique for automating the counting of red and white blood cells using image processing, reducing human error and improving efficiency. Varun et al. [10] describe a blood cell count system using digital image processing, demonstrating large-scale automated analysis of blood cell populations.

III. Proposed Method

The proposed method uses the watershed algorithm, which segments complex images by detecting intensity gradients and delineating individual regions. Because manual counting is slow and error-prone, an automated algorithm offers

an accurate and efficient alternative. Customized pre-processing improves image quality and optimizes cell boundary detection before segmentation is applied. The complete flow is shown in Fig. 1.

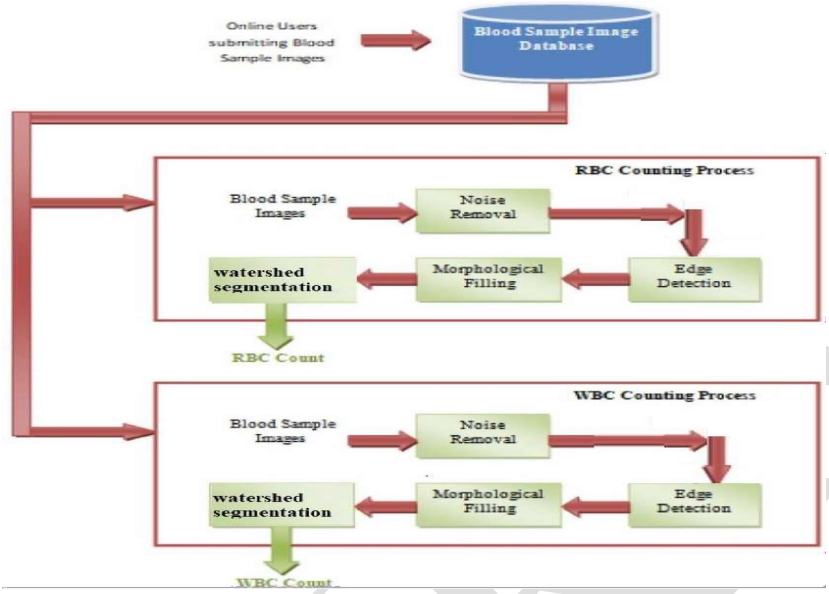

Fig. 1. Block diagram for blood cell counting using the watershed algorithm.

The method proceeds through the following stages.

Image acquisition. The blood cell image is acquired using a high-resolution microscope or other suitable imaging device.

Pre-processing. The acquired image is converted to grayscale. Noise reduction, using Gaussian smoothing or median filtering, improves image quality. Contrast is then enhanced, using techniques such as histogram equalization, so that cells separate more cleanly from the background.

Thresholding. A suitable thresholding technique segments the cells from the background. Adaptive thresholding is particularly useful where lighting varies across the field of view.

Morphological operations. Erosion and dilation remove small noise and fill gaps between cells, giving better segmentation results.

Distance transformation. The distance transform of the thresholded image is calculated, which helps determine the approximate location of each cell nucleus.

Watershed segmentation. The watershed algorithm is applied to the distance-transformed image to separate clustered cells. Markers are obtained from the local maxima of the distance transform, which represent potential cell centres.

Post-processing. Segmented regions that do not meet the size, shape or intensity criteria for a blood cell are removed, and the segmentation is refined with further morphological operations where necessary.

Counting. The segmented cells within the regions obtained from watershed segmentation are counted, giving the total number of cells in the image.

Quantification and analysis. The segmented cells are analysed for size, shape and distribution, and metrics such as cell density and average cell size are calculated.

Validation and optimization. Results are compared with manual counting, and the algorithm parameters are refined iteratively to improve accuracy and efficiency.

IV. Implementation Of The Processing Stages

A. Blood Sample Image Database

A blood sample image database is a repository of digital images of blood samples obtained through various imaging techniques, including brightfield microscopy, phase contrast microscopy, fluorescence microscopy and digital image analysis systems. Diversity in sample sources matters: images from individuals of different demographics, health statuses and medical conditions make the findings more generalizable. Annotations and metadata accompanying the images — staining protocols, imaging parameters and clinical context — add to their usefulness. Image quality and resolution directly affect the reliability of measurement, cell identification and morphological assessment, so a database is usually structured into training, validation and test subsets to support development and evaluation of algorithms.

B. Noise Removal

Noise removal suppresses background noise and artefacts without damaging the cell structures that must later be segmented. Smoothing filters and adaptive thresholding are applied to enhance contrast and reduce noise. A Gaussian smoothing filter is used here: it examines each pixel together with its neighbours, mixes their values, and so blurs out random variation while leaving the larger structures intact. The Gaussian kernel in discrete form is

$$G(x, y) = (1 / 2\pi\sigma^2) \cdot e^{-(x^2 + y^2) / 2\sigma^2} \quad (1)$$

where G(x, y) is the value of the kernel at position (x, y), σ is the standard deviation of the Gaussian distribution, and (x, y) are the spatial coordinates within the kernel. A larger σ gives stronger smoothing but also blurs genuine cell boundaries, so the value must be chosen with the cell size in mind.

C. Edge Detection

Edge detection is the basis for accurate segmentation of individual cells from the background and from neighbouring cells. Blood cell images often contain complex backgrounds and varying levels of noise, and edge detection algorithms delineate cell boundaries from abrupt changes in intensity or gradient. This provides the information needed to isolate individual cells and count them.

Edge detection also contributes to noise reduction, since prominent features are retained while low-intensity variation and background clutter are disregarded. Where cells are clustered or overlapping, the boundaries at which cells meet can be identified, which allows a clump to be partitioned into distinct entities so that each cell is counted once. The morphological features extracted at this stage — size, shape and contour — are also used later for distinguishing cell types and identifying abnormal cells.

D. Canny Edge Detector

The Canny edge detector is used to delineate cell boundaries precisely. Images are first pre-processed with Gaussian blurring to smooth pixel intensity variation and remove noise that is not relevant to cell edges. The detector then calculates the gradient magnitude and direction at each pixel. Gradient magnitude highlights areas of rapid intensity

change, which correspond to cell edges, while gradient direction gives the orientation of those edges and helps localize the boundary accurately.

Non-maximum suppression follows, retaining only the local maxima along each edge so that edges are thinned to a single pixel width and unnecessary clutter is removed. Thresholding then classifies pixels as strong, weak or non-edge according to gradient magnitude. Strong pixels are taken as definite cell boundaries and weak ones are retained for further analysis. Hysteresis thresholding connects weak edge pixels to strong ones, forming continuous edges and discarding isolated noise artefacts that were wrongly classified as edges.

The gradient magnitude and direction are computed as

$$G = \sqrt{(Gx^2 + Gy^2)} \quad (2)$$

$$\theta = tan^{-1} (Gy / Gx) \quad (3)$$

where Gx and Gy are the horizontal and vertical gradient components obtained with Sobel operators.

E. Morphological Operations

Morphological operations manipulate the shape and structure of objects in an image, drawing on mathematical morphology. The two primary operations are erosion and dilation.

Erosion shrinks or thins the boundaries of objects. A structuring element is scanned over the image and each pixel is replaced by the minimum value within the neighbourhood the element defines. On a binary image, the centre pixel is set to 1 only if the structuring element fits entirely within the object, and to 0 otherwise. This erodes away small protrusions, noise and fine detail while preserving larger structures, which makes it useful for separating touching objects.

Dilation expands or thickens boundaries. The structuring element is scanned over the image in the same way, but each pixel is replaced by the maximum value in the neighbourhood. This enlarges objects and fills gaps or holes, which is useful for connecting broken edges and making objects more robust against noise.

The two can be combined. Morphological opening is erosion followed by dilation, which removes small objects and smooths boundaries. Closing is dilation followed by erosion, which fills small gaps within objects. In this work erosion is the main operation used, invoked in MATLAB as

```
output_image = imerode(binary_image, se);
```

F. Watershed Segmentation

The watershed algorithm is used for segmentation where the objects of interest are not clearly separated. The name comes from the geographical watershed, where terrain determines the flow of water. The image is conceptualized as a topographic surface in which intensity values represent elevation: high intensity forms peaks and low intensity forms valleys.

The algorithm first identifies markers within the image, typically by detecting local minima in the gradient magnitude or intensity map. These markers are the starting points from which the image is flooded. As the water level rises from each marker, catchment basins form and grow until they meet, and the lines at which they meet become the boundaries between regions. Each basin corresponds to one blood cell.

Without precautions the algorithm over-segments, splitting a single cell into several regions wherever the gradient surface has spurious minima. Marker adjustment and the merging of neighbouring regions are therefore applied afterwards to refine the result. Careful marker placement, aligned with the centre of each cell, is what keeps false positives low. The strength of the method for this application is that it handles varying cell sizes, shapes and densities, which is exactly what a blood smear presents.

G. Software Environment

The method was implemented in MATLAB, whose name stands for Matrix Laboratory. MATLAB is a high-performance language for technical computing that integrates computation, visualization and programming in one environment. Its basic data element is an array that does not require dimensioning, which suits image processing, where an image is naturally a two- or three-dimensional array. The built-in routines and graphics commands make results immediately available for inspection, and toolboxes are available for signal processing, image processing and related fields. Processing steps were written as M-files, which are external files containing sequences of MATLAB statements that can be saved, edited and re-run.

V. Results And Discussion

The method was tested on microscope images of stained blood samples. The original test image, of size 182 × 277 pixels, is shown in Fig. 2.

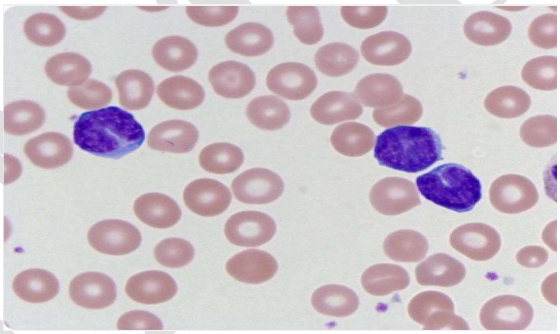

Fig. 2. Original image (182 × 277).

Fig. 3 shows the pre-processing stage, with the original image, the grayscale conversion and the adaptive histogram equalized image side by side. Converting to grayscale removes the colour information that is not needed for shape-based counting, and adaptive equalization brings out cells that were poorly separated from the background in the original.

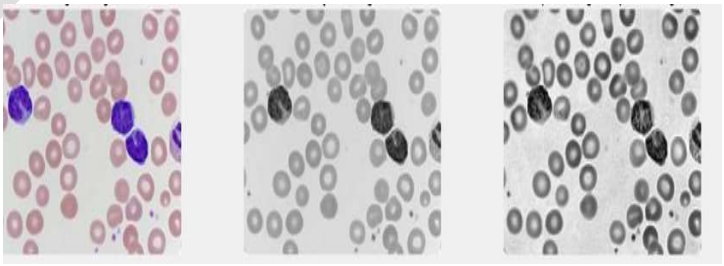

Fig. 3. Pre-processing: original, grayscale and adaptive histogram equalized image.

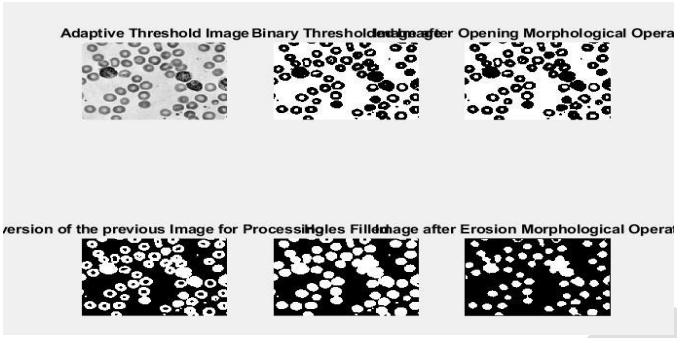

Fig. 4. Image enhancement.

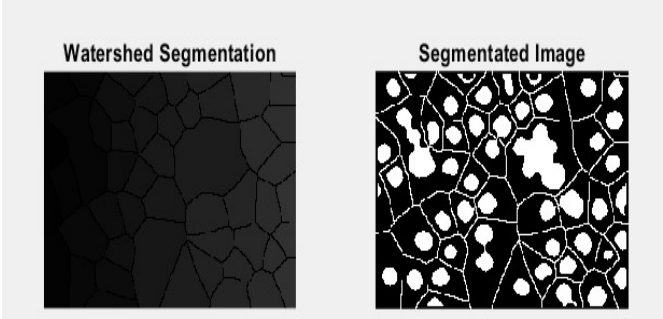

Fig. 5. Image segmentation.

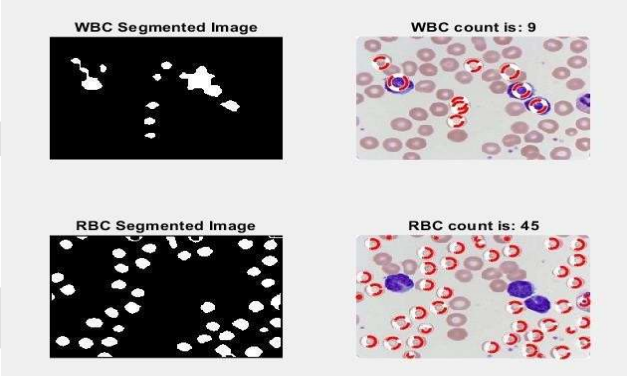

Fig. 6. Image post-processing.

Figs. 4 to 6 show enhancement, segmentation and post-processing. After enhancement the cell boundaries are distinct enough for the watershed transform to operate on. Segmentation separates the touching cells into individual regions, and post-processing removes the regions that fail the size and shape tests, which are mostly fragments at the image border and residual noise.

A. Counting Using the Application Module

An application module was developed in MATLAB so that the method can be used without writing code. The user uploads an image and the module carries out the digital image processing internally, returning the red and white blood cell counts directly. The stages performed by the module are shown in Figs. 7 to 11.

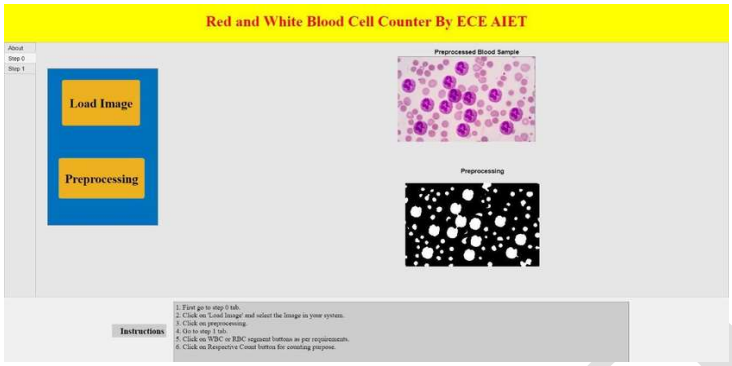

Fig. 7. Pre-processing in the application module.

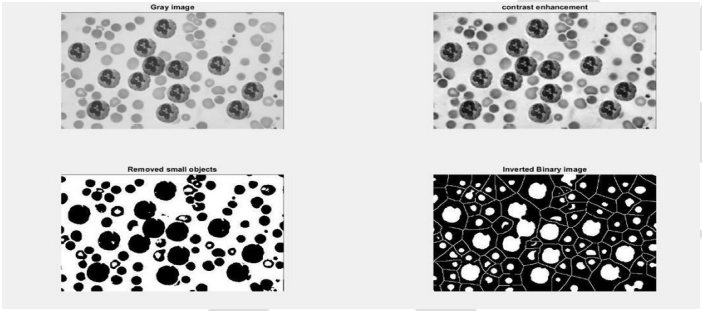

Fig. 8. White blood cell segmentation.

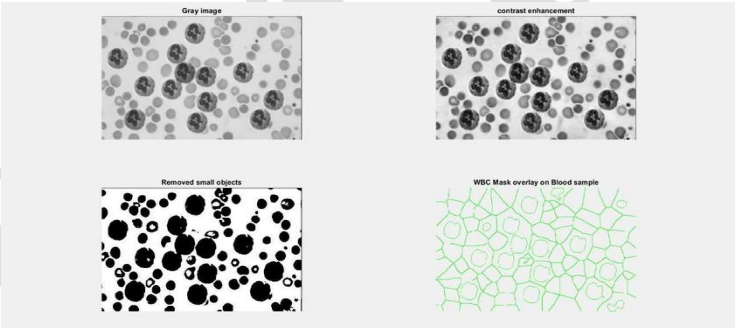

Fig. 9. Red blood cell segmentation.

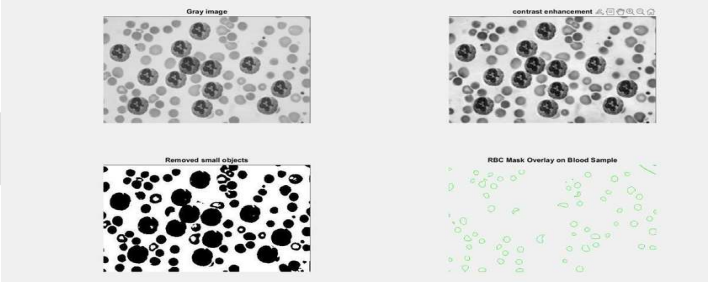

Fig. 10. Image enhancement in the application module.

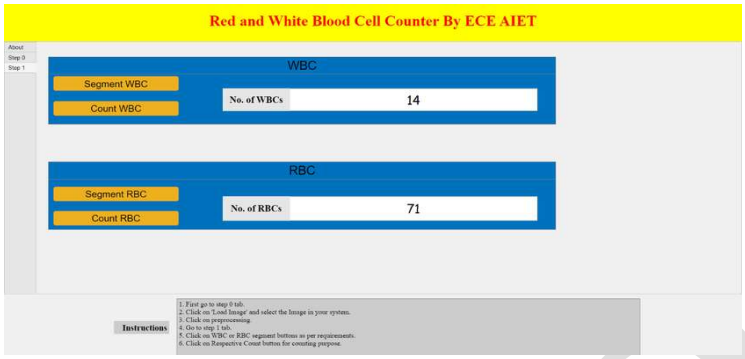

Fig. 11. Counting of blood cells using the application module.

White and red cells are segmented separately, because the two differ in size and in staining response, and treating them in one pass would force a single set of size thresholds onto two populations. Segmenting them separately allows the threshold and marker parameters to be set appropriately for each.

Compared with manual counting, the method achieved 94% accuracy for red blood cells and 92% for white blood cells. The residual error comes mainly from two sources. Cells that touch the image border are partially visible and cannot be counted reliably, and very heavy clumps occasionally survive post-processing as a single region. Red cell accuracy is higher than white cell accuracy because red cells are more uniform in size and shape, so the shape criteria used in post-processing separate them more cleanly.

Against the alternatives, the method is considerably cheaper than an automated analyzer, since it needs only a microscope, a camera and a computer. It is faster and more consistent than manual counting, and unlike simple thresholding methods it handles overlapping cells, which was the main limitation identified in the earlier work reviewed in Section II.

VI. Conclusion

The work demonstrates a cost-effective and efficient alternative to traditional blood cell counting. The watershed algorithm proved effective at segmenting images containing overlapping cells and irregularly shaped objects. By treating the image as a topographic surface in which pixel intensities represent elevation, and by placing markers at the centres of cells, the image can be partitioned into regions that each correspond to a single cell.

The ability of the algorithm to handle varying cell sizes, shapes and densities makes it versatile across different types of blood samples and imaging conditions, and its robustness against noise and artefacts supports its use in both clinical and research settings. Automation reduces the need for manual intervention, saving time and removing the inconsistency associated with counting by eye, and it makes large-scale analysis of blood samples practical.

The limitations should also be stated plainly. The algorithm is susceptible to over-segmentation in regions of low contrast or with complex backgrounds, and post-processing is needed to correct for this. Accuracy also depends on the quality of staining and of the captured image. Within these limits the results — 94% for red blood cells and 92% for white blood cells against manual counting — show that the approach is a workable tool for hematology.

VII. Future Scope

The technique can be extended to platelet counting and to the identification of the different types of white blood cell, which would take it from a cell count towards a fuller hematological assessment. Refining the methodology to raise accuracy further, particularly for white cells, is the immediate next step.

Integration into digital health platforms is a second direction. As healthcare systems become more digital, automated counting could feed directly into electronic health records, giving both real-time and historical data for patient monitoring. A third direction is research use: automated counting allows rapid analysis of large datasets, which supports investigation of disease mechanisms, treatment responses and prognostic indicators related to blood cell populations. Finally, adding a convolutional neural network for classifying the extracted features would allow cell types to be distinguished automatically rather than by size and shape rules alone.

References

[1] Berge, D. Taylor, S. Krishnan, and T. S. Douglas, "Improved red blood cell counting in thin blood smears," in Proc. 8th IEEE Int. Symp. Biomedical Imaging: From Nano to Macro (ISBI '11), Chicago, IL, USA, Apr. 2011, pp. 204–207.

[2] S. Khan, A. Khan, and A. Naseem, "An accurate and cost effective approach to blood cell count," International Journal of Computer Applications, vol. 50, no. 1, pp. 18–24, 2012.

[3] N. Nguyen, A. Duong, and H. Vu, "Cell splitting with high degree of overlapping in peripheral blood smear," International Journal of Computer Theory and Engineering, vol. 3, no. 3, pp. 473–478, 2011.

[4] S.-H. Chiu, J.-J. Liaw, and K.-H. Lin, "A fast randomized Hough transform for circle/circular arc recognition," International Journal of Pattern Recognition and Artificial Intelligence, vol. 24, no. 3, pp. 457–474, 2010.

[5] N. H. Mahmood and M. A. Mansor, "Red blood cells estimation using Hough transform technique," Signal & Image Processing, vol. 3, no. 2, pp. 53–64, 2012.

[6] M. G. Bhamare and D. S. Patil, "Automatic blood cell analysis by using digital image processing: A preliminary study," International Journal of Engineering Research & Technology (IJERT), vol. 2, no. 9, Sep. 2013.

[7] S. Sridhar, Digital Image Processing. Oxford University Press, 2011.

[8] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 3rd ed. Pearson Education, 2009.

[9] Thejashwini M and M. C. Padma, "Counting of RBC's and WBC's using image processing technique," International Journal on Recent and Innovation Trends in Computing and Communication, vol. 3, no. 5, pp. 2948–2953, 2015.

[10] D. Dyanesh Varun, K. Reddy, S. Lakshmi Priya, Vasavi, and S. Abirami, "Blood cell count using digital image processing," in Proc. IEEE Int. Conf. Current Trends Towards Converging Technologies, Coimbatore, India, 2018.