A novel fuzzy logic inspired edge detection technique for analysis of malaria infected microscopic thin blood images

S.D Bias, Reni, S. and Kale, I.

A Novel Fuzzy Logic Inspired Edge Detection Technique for Analysis of Malaria Infected Microscopic Thin Blood Images

S.D. Bias

Department of Engineering
University of Westminster
London, UK
sbias254@gmail.com

S.Kareem Reni, I. Kale

Department of Engineering, ADVRG
University of Westminster
London, UK

Abstract—This paper proposes a novel, efficient, low complexity algorithm for edge detection, specifically focusing on the analysis of malaria infected microscopic thin blood smears. The algorithm proposes a simple, dynamic thresholding technique that is computed via histogram analysis, designed to capture as much information about the blood cells with minimal computational effort, which is followed by a morphological filtering process to remove noise and artifacts. A binary edge tracking system inspired by the works in fuzzy logic is introduced, defined by a semi ambiguous rule system that can be efficiently implemented in hardware.

Keywords: Canny, Edge detection, Edge detection, Fuzzy logic, Giesma staining, Histogram, Malaria, Morphological Processing, Morphology, Platelets, Red blood cells, White blood cells

I. INTRODUCTION

Edge detection is fundamental to computer vision and computer based automation; it is a versatile tool with a huge array of uses spanning medicine to traffic (such as in [1] and [2]). An increase in research done into this field has led to the expansion of its application to a wider subject area. One such subject is that of automatic disease detection; many diseases can be diagnosed by the shape or size of patients' blood cells (such as sickle cell [3]), and parasitic disease (such as malaria) can be diagnosed in a similar fashion, be it the parasitic component that is the focus, not just the cells.

The primary goal of any edge detection algorithm, regardless of the specifics involved, is to accurately identify and mark the edge points of an image. This task is often harder than face value. as lighting conditions, capture device, and the subject matter all have a significant impact on the performance of any algorithm. Due to the volume of research being done in this field, new and improving methods are always emerging, but few algorithms prove as useful and well performing as the Canny algorithm, the benchmark this paper will use. The inspiration for the algorithm described in this paper comes from the use of fuzzy logic in edge detection, such as in literature [4] and [5]. While the system proposed is not true fuzzy logic, it bears a distinct resemblance.

Analysis of thin blood films is the standard for malaria diagnosis and should be considered the first step to automated diagnosis of the disease. Malaria is a global epidemic affecting hundreds of millions of people, and according to the latest WHO 2017 malaria factsheet, in 2015 it killed an estimated 429000 people, where a significant number of these (>70%) deaths were children under 5 [6]. The disease is caused by the Plasmodium parasite transmitted in the bites of infected female *Anopheles* mosquitos. Microscopic thin blood films show the presence of the disease indicated by the heavily giemsa stained parasitic components along with other artifacts and platelets (fig 2.a). This paper proposes a quick thresholding method utilizing manually trained peak detection, and a binary removal of all subjects that are not the focus of segmentation, followed by a fast and comprehensive edge detection method derived from the investigation of how edges are formed in a binary image of blood cells.

II. METHOD

The algorithm is as follows (also shown in fig.1.): Firstly, the threshold is computed via histogram analysis and applied to the image. This is followed by morphological processing to tackle noise, and remove unwanted artefacts and platelets. The resultant binary image is further processed to track the edges, which is used to segment the original colour image.
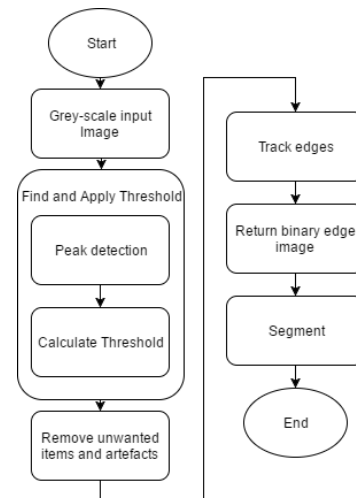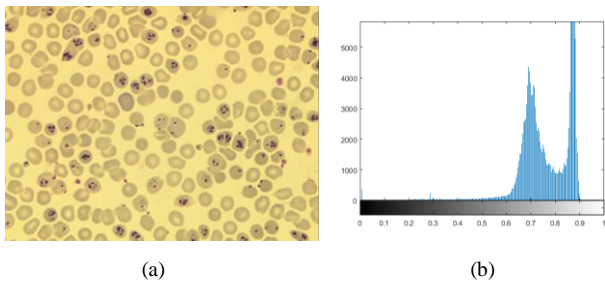

Fig 1. System level diagram

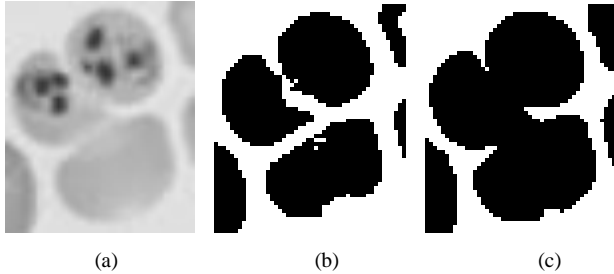Fig. 2. (a) Image of stained thin blood film, (b) its greyscale histogram

Fig. 3. (a) Original section. Post threshold: (b) using average value described, (c) using mid value between the peaks

A. Threshold aquisition

The histogram of a grey scale thin blood image contains two prominent peaks, one large peak, representing the background, and the smaller red blood cell (RBC) peak. Experiments showed that the mid-point between the two works well as a threshold value, however this cannot be completely relied upon as it leads to under segmentation. Hence a manually trained peak detection algorithm was implemented, and the boundary between the two peaks is taken as the lowest value between the two. The final threshold value is taken as the average value between the RBC peak and the boundary point between the two. This in turn reduces interference from close neighbouring cells as the values approaching the boundary of cells are likely to be between the RBC peak and the background peak. By bringing the threshold value towards the RBC peak, the interference issues can be avoided.

This method produces clear and accurate representations of the blood film in binary format, be it with many holes and spots that represent the middle bi-concave shape of the cell, platelets, and other artefacts, that should be marked for removal. Binary images are very simple and computationally faster, and often contain all the information needed for applications such as edge detection. This method eliminated the need to do further pre-processing in order to remove the dark components that many edge detectors would identify as separate entities inside the cells (fig.3. a.).

B. Platelet and artefact removal

When segmenting for a specific subject within an image, only the subject should be considered valid data. The case is no different when segmenting blood cells; platelets, debris, and other artefacts should be ignored. During thresholding however, keeping information about these components will be unavoidable, when it should be discarded and the edge points not marked. Other algorithms exist that focus on the reduction of noise, but these are non-specific and often very complex
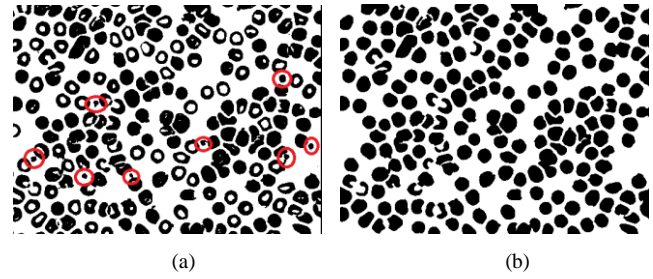

Fig. 4. (a) Binary image of blood film, (b) All unwanted items removed

such as the method shown in [7]. Background subtraction is a common method that would however be unsuitable for this application, as the "noise" being removed is in the foreground. The proposed algorithm makes uses a simple binary morphology. The images being worked with are of size 512 by 381, and the spots to be removed that represent platelets and other artefacts are approximately <150 connected pixels in area, the white center pallor are approximately <500 connected pixels in area (any sufficiently large number can be used as the background is the only other white component). The procedure for the removal of unwanted platelets and artefacts is as follows:

- Area of all the white connected components is calculated, remove any of 500 pixels or fewer

- Morphological closing is done, to separate out any small unwanted components that may be attached to the cells

- Area of all the black connected components is calculated, remove any of 150 pixels or fewer

Fig.4. a. is the result of thresholding the original greyscale image using the value calculated from the histogram. Highlighted are all the unwanted platelets and artefacts in the binary image that should be removed. Fig 4. b. shows the result of the simple removal procedure, as well as filling in the gaps in cells.

Conversely, this process can be done via morphology, but this leads to a loss of vital information on a scale that is not acceptable when dealing with the delicate nature of blood images. In order to remove the white center pallor, platelets, and other artefacts, morphological closing is done with a disk-shaped structuring element of radius 10. As seen in fig.5 (b) and (c), a significant number of the cells have been removed and as a result, will not be segmented. The same loss is observed regardless of whether this operation is done in greyscale or binary. The method described above maintains much more of the cell information, and consequently
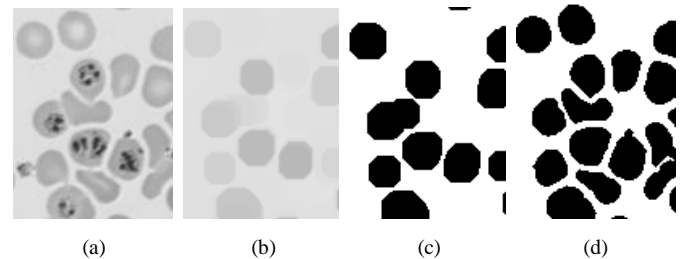

Fig.5. (a) section of blood sample, (b) after morphological closing (c) post-threshold of (c), (d) Post threshold using proposed algorithm

$$\begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 1 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

(a)

$$\begin{bmatrix} 0 & 0 & X \\ X & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} X & 0 & 0 \\ 1 & 1 & X \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & X & 0 \\ 1 & 1 & 0 \\ 1 & 1 & X \end{bmatrix} \begin{bmatrix} 1 & 1 & X \\ 1 & 1 & 0 \\ 1 & X & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 1 \\ X & 1 & 1 \\ 0 & 0 & X \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & X \\ X & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & X & 1 \\ 0 & 1 & 1 \\ X & 1 & 1 \end{bmatrix} \begin{bmatrix} X & 1 & 1 \\ 0 & 1 & 1 \\ 0 & X & 1 \end{bmatrix}$$

(b)

Fig. 6. (a) Full "crisp" logical rules, (b) ambiguous "fuzzy" rules
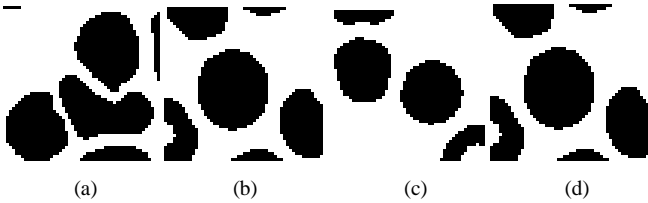

(a)　　　(b)　　　(c)　　　(d)

Fig. 7. Examples of post-threshold cells

maximizes accuracy, as only small targeted parts of the image are affected, rather than a global operation that impacts all cells.

Noise reduction can be done in a variety of ways depending on the type of noise being removed. Often the subject of noise, is random (such as AWGN), which is different to the small artefacts and platelets to be removed. One of the most common and effective methods is the recursive median filter (RMF), and proposed in [8] is the adaptive recursive median filter (ARMF) that shows significantly better results than the RMF in processing of greyscale image.

C. Edge tracking

The edge tracking is completed by comparing the output of a moving 3 by 3 window to a rule system shown in fig.7. The window starts in the top left moving towards the bottom right, propagating row wise. They were designed after a analyzing the most common formations of edges in binary images of blood, examples of which can be seen in fig.7. After a detailed image analysis, the system rules were established, that if match the output of the 3 by 3 window, mark the origin as an edge. The 0s in the rules represent the edge of a binarized cell, the 1s represent the white background, and the Xs are parts of the rule that can either be 1 or 0. The implementation of these can be
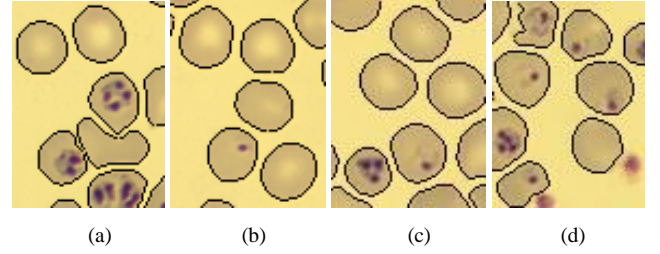

(a)　　　(b)　　　(c)　　　(d)

Fig.8. Examples of segmented cells

done via inspecting sections of the 3 by 3 window, where each rule has a unique set of criteria, as show in fig.6. (b), that must to be fulfilled in order for the point to be marked as an edge point. This operation is an entirely comparator based implementation and provides a comprehensive definition of edges in binary images.

D. Segmentation

Using the information in from the edge tracking, the corresponding pixels in the original colour image are marked as black, resulting in segmented cells with a black outline, as seen in fig.8.

Previously in literature [4], a method is described using fuzzy logic, and a fuzzy interference system (FIS). The authors propose using degrees of black and white to describe the intensity levels of pixels in images. The method uses a set of rules that references pre-defined degrees of "black" and "white". The image is first fuzzified and then compared to the rules, a completely unnecessary step, hence using more memory and thereby more power consuming and computationally more complex – instead of fuzzification, a threshold could be applied at the boundary between "blacks" and "whites", providing the same outcome and reducing complexity. The algorithm in this paper is a novel idea, inspired by the above method, but computationally more efficient.

In fig.6. (a), the fully expanded, fully logical rules can be seen, that almost completely define how edges form in binary images. Implementing a system with these rules is lengthy and cumbersome, and lead to a more complex, less efficient method. By inspecting smaller sections of the window, outputs that are not edges can be identified much sooner, leading to a faster and more efficient approach. The 8 rules set out fig.6. (b) are designed to be ambiguous in that, in some configurations, certain pixels are irrelevant and can have any binary state.

The segmentation of the cells is the most crucial part of the algorithm, as further processing on the resultant image can be used to determine patient prognosis. The algorithm limits the use of shape altering morphological procedures in order to retain as much of the information about the cells as possible. While the process of marking the edges in the original image using the output of the edge tracking isn't a difficult task, the process remains possibly the most important, as the result will be used for medical diagnostic purposes.

III. EXPERIMENTAL RESULTS

To better understand the performance of the proposed algorithm, a statistical analysis was carried out, documenting its sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV) [9-11]. These values are

calculated by identifying a number of parameters shown in table 1. Another important measure is accuracy, and is defined with the same parameters.

TABLE I.

	Item in sample is a blood cell	Item in sample is not a blood cell	
Item in sample is segmented	True positive (TP)	False positive (FP)	Total Positives (TP+ FP)
Item in sample is not segmented	False negative (FN)	True negative (TN)	Total Negative (FN+TN)
	Total blood cells (TP+FN)	Total items that are not cells (FP+TN)	

$$Sensitivity = {^{TP}}\!/_{TP + FN} \qquad (1)$$

$$Specificity = {^{TN}}\!/_{TN + FP} \qquad (2)$$

$$PPV = {^{TP}}\!/_{TP + FP} \qquad (3)$$

$$NPV = {^{TN}}\!/_{TN + FN} \qquad (4)$$

$$Accuracy = {^{TP + TN}}\!/_{TP + FP + FN + FN} \qquad (5)$$

Both the proposed algorithm and the Canny algorithm were tested on thin blood film samples containing 2946 cells, and 270 platelets and other artefacts, totaling 3216 objects.

TABLE II

Proposed algorithm				
Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
90.53	84.93	98.49	45.29	90.06

TABLE III

Canny				
Sensitivity (%)	Specificity (%)	PPV (%)	NPV (%)	Accuracy (%)
89.41	9.29	91.52	7.42	82.71

TABLE IV

	TP	TN	FP	FN
Proposed	2667	231	41	279
Canny	2634	25	244	312

As the results in table II and III show, the proposed algorithm outperforms the Canny algorithm in all measures. Specificity is one of the more important measures to take note of when handling blood images. As the eventual goal is a fully autonomous system, if items are segmented that aren't cells, further processing will become very difficult. The low specificity measure for the Canny algorithm is unsurprising, this is because by default, it doesn't implement any method of filtering out items that don't correspond to the subject matter of the image under scrutiny. Even when the fact that the Canny algorithm is non-specific, observing the results in table IV shows that it has more true positives and less false negatives, meaning if the image only contained that of the subject matter, the proposed algorithm would still be better suited.

IV. CONCLUSION

The algorithm described simplifies the process of edge detection by using non-conventional methodology. The thresholding via histogram analysis is akin to that of Otsu's method [12], but is a simpler and faster approach. Due to the discrete nature of computational data and the way images are stored, the novel edge tracking system and its corresponding fuzzy logic technique proposed offers a simple comparator based approach, and given a much more in-depth study, could entirely define the way edges are formed in images and eliminate the need for more complex processes that are often involved in edge detection.

ACKNOWLEDGMENT

REFERENCES

[1] Shun, S, S. Yan, Y. Wang, Y. Li, "Medical CT edge detection algorithm based on improved fuzzy clustering analysis", Intelligent Systems and Angineering Applications, 2014

[2] Hugn, N.V., N.T.T. Hien, P.T. Vinh, N.T. Thao, N.T. Dzung, "An utilization of edge detection in a modified bicubic interpolation used for frame enhancement in a camera-based traffic monitoring", Information and Communications, 2017

[3] American Society of Hematology, "Sickle cell anemia", http://www.hematology.org/Patients/Anemia/Sickle-Cell.aspx, Last accessed 20th August 2017 via Chrome

[4] Alshennawy, A.A., and A.A. Aly, "Edge Detection in Digital Images Using Fuzzy Logic Technique", World Academy of Science, Engineering, and Technology, 52, 2009

[5] Essa, A., "Edge detectent techniques using fuzzy logic", Signal Processing and Intergrated Networks, 2016

[6] WHO, "Malaria", http://www.who.int/mediacentre/factsheets/fs094/en/, last accessed 20th August 2017 via Chrome

[7] Karami, A., R. Heylen, P. Scheunders, "Band-specific shearlet-based hyperspectal image noise reduction", IEEE Transactions on Geoscience and Remote Sensing, vol.53, issue 9, 2015

[8] Tafti, A.D., E. Mirsadeghi, "A novel adaptive recursive median filter in image noise reduction base don using the entropy", Control System, Computing and Engineering, 2012

[9] Newberg, L., "Some useful statistics definitions", Renssalaer Polytechnic Institute, department of computer science, http://www.cs.rpi.edu/~leen/misc-publications/SomeStatDefs.html, last accessed 20th August 2017 via Chrome

[10] Parikh, R., A. Mathai, S. Parikh, G.C. Sekhar, and T. Thomas, "Understanding and using sensitivity, specificity, and predictive values", India Journal of Opthalmology, vol.56, no.1, January-February, 2008

[11] Akobeng, A.K., "Understanding diagnostic tests 1: sensitivity, specificity, and predicitive values", Acra Pædiatrica, vol.96, issue 3, March 2007

[12] Otsu, N., "A Threshold selection method from gray-level histograms", IEEE Transactions on Systems Man, and Cybernetics, vol.9, no.1, January 1979