Comparative Analysis of Thresholding Methods in Cancer Cells Image Processing

Daniel Martomanggolo Wonohadidjojo
Universitas Ciputra, Surabaya – Indonesia.
daniel.m.w@ciputra.ac.id

***Abstract*—Analysis of cancer cell images in medical field needs to be assisted using digital image processing. This paper presents the comparative analysis of image thresholding using two algorithms, Backtracking Search Algorithm and Particle Swarm Optimization. Two experimental designs were implemented. In the first design the images were thresholded and the performance was compared. In the second design, the images were enhanced before the thresholding was performed. In the second, the original and processed image histograms were presented and compared. In both designs, performance metrics were calculated to validate the comparative analysis. In the first experimental design, where BSA and PSO are implemented to threshold lung epidermoid carcinoma and chronic lymphocytic leukaemia cell images, the values of MSE, PSNR, MSSIM, FSIM and IEM show the superiority of BSA over PSO. In the second one where the thresholding method is implemented after image enhancement process, the histogram entropy and variance show that the thresholding method using BSA outperforms the one using PSO. These results show that in both designs the BSA outperforms PSO. Therefore, the thresholding method using BSA is more suitable for cancer cells image thresholding in processing the image samples for further analysis. This will provide a more reliable solution and effective way for assisting analysis of cancer cells where it minimize the difficulties arises in the conventional way of manual observation of microscopic images.**

***Index Terms*—Comparative Analysis; Image Thresholding; Backtracking Search Algorithm; Particle Swarm Optimization.**

I. INTRODUCTION

Quantitative analysis on cell image is used in biology and medicine. Research regarding cell shape, development, and behavior is one of the most important aspects in studying cell embryology, wound healing, defense mechanisms of the body, and others. Identification of cells which is undertaken conventionally is sometimes less accurate when it is performed with direct observation without digital image captured.

With increasingly advances in image processing techniques, the identification of cells will be more accurate by utilizing digital image processing method, such as computerized segmentation method automatically. Computer program is able to identify image processing quickly. On the resulted image, a group of uniform objects or almost uniform, there is a characteristic in every object. That particular characteristic is used to detect the read image.

In the study of image enhancement technique, a number of local image enhancement algorithms have been acquainted. A lot of images such as medical images are in poor contrast condition and suffer from noise. The ultimate purpose of enhancement in image is the attributes modification in the image, so that it fits a given task and presents a better detail of the image [1].

In processing the images taken from real-world, the entropic thresholding method is considered effective. However, there is a drawback in this method. The notable one is it only utilizes the distribution of gray level of an image. As a result, different images with the same histograms may have the same threshold, which is not acceptable [2].

In order to perform further analysis of microscopic images of cancer cells, ordinary image processing technique is not sufficient. For multiple thresholds segmentation, the technique needs to find the optimum values of threshold by exploring all the possible combination of trails for the number of thresholds. The computational complexity and the requirement of accurate measure in the case of multiple thresholds motivated the use of an efficient search algorithm [3]. At the thresholding stage of the process, the optimum threshold values should be determined. Therefore an optimization algorithm is required to find these values and processing the thresholding stage.

One of the advantages of using the Evolutionary Algorithms (EA) over classical optimization algorithms is its flexibility to find the solution over various kinds of problems. The abilities of EA are global exploration and local exploitation. The ability of the algorithm to use the entire space is referred to as global exploration, and its ability to search the best solution that is close to the last discovered solution refers to the local exploitation. In the first iteration, the global exploration ability is used by EA to find new solutions that are required, so that the local minima can be avoided. During the iterations, the solutions that are generated depend on its local exploitation ability [4].

In this study, two EAs are used to optimize the image thresholding process of microscopic sample cell image: Backtracking Search Algorithm (BSA) and Particle Swarm Optimization (PSO). The purpose of the study is to undertake a comparative analysis of the performance of Backtracking Search Algorithm and Particle Swarm Optimization to find the optimum threshold values in the thresholding stage of cancer cell images.

The contribution of this study is providing an automatic image enhancement and thresholding method of cancer cells which will provide a more reliable solution and effective way for assisting further analysis of cancer cells in image samples. Despite there were a number of studies on image thresholding by implementing PSO [5][6][7] and performance comparison of EA in image processing [8][9], however there were no previous studies that are conducting a comparative analysis of BSA and PSO in image thresholding method in cancer cell images. This study will increase and accelerate cancer research conducted by using established cancer cells line as an in vitro model. An example of further analysis is automatic

morphology classification of cancer cells that will enable correct detection and labeling of different cell lines. The morphology of cancer cells can infer invasive tumor cells and hence metastatic ability [10]. This will significantly improve the ability of analysts to identify the various types of cell line without the need to observe each one of microscopic images manually.

The rest of this paper is organized as follows: in Section II the method of this study is explained. In Section III, the experiments of this study are described. Then in Section IV the results of this study are presented and discussed. Finally, in Section V the study is concluded.

II. METHOD

A. Unsharp Masking Filter

When a highlighted and scaled image is augmented to its original version, then the principle of unsharp masking filter is applied. Even the given image is monochrome, the human visual system has an ability to perceive an image object based on its relative intensity with respect to its surrounding. Therefore, a color image is first converted to a gray image [11].

Figure 1 shows the filtering process.

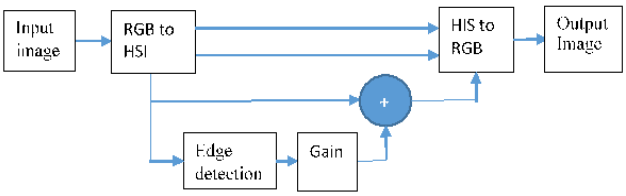

Figure 1: Unsharp masking filtering process [11]

B. Global Thresholding

A global thresholding technique is utilized when there is sufficient distinction between the distribution of intensity of the objects and its background. In global thresholding technique, the entire image uses a threshold value. When the pixel values of the components and that of background are fairly consistent in their respective values over the entire image, global thresholding could be used [12]. Global thresholding means to choose threshold value T that separates object from background. This can be characterized by (1)[12]. If g(x, y) is a threshold version of f(x, y) at some global threshold T,

$$g(x,y) = \begin{cases} 1 & if\ f(x,y) \geq T \\ 0 & otherwise \end{cases} \quad (1)$$

There are a number of global thresholding techniques, in this study, the optimal thresholding is implemented.

C. Thresholding using BSA

This process is applied to convert gray image into binary or black-and-white image so that it can be known which region includes the object and background of the image clearly. Threshold image results are then used further for the object recognition process as well as feature extraction.

BSA is an iterative global minimizer algorithm based on population. Its function is characterized by five processes: initialization, selection-I, mutation, crossover and selection-II [4]:

a. Initialization

The initialization of population P in BSA is defined with Equation (2).

$$P_{i,j} \sim U(low_j, up_j) \quad (2)$$

where $i=1,2,3,\ldots,N$, $j=1,2,3,\ldots,D$, where N is the size of population and D is the dimension of problem, U is the uniform distribution and P_i is the individual target in P.

b. Selection-I

In this stage the historical population oldP is determined to calculate the search direction. Its initialization is defined using Equation (3).

$$oldP_{i,j} \sim U(low_j, up_j) \quad (3)$$

At the beginning of each iteration, there is an option to redefine oldP using the 'if-then' rule as defined in Equation (4).

$$if\ a < b\ then\ oldP := P | a, b \sim U(0,1) \quad (4)$$

where:= is the update operation. Equation (4) is used to ensure the design of population is selected from the previous generation randomly, which is the historical population, and to remember it until it is updated. This is because BSA has a memory. When the oldP has been determined, Equation (5) is applied to change the order of individuals in oldP randomly.

$$oldP := permuting(oldP) \quad (5)$$

The function defined in Equation (5) is a random shuffling function.

c. Mutation

In this stage, the initial form of the trial population is generated using Equation (6).

$$Mutant = P + F.(oldP - P) \quad (6)$$

The Equation (6) defines that the search direction matrix (oldP-P) is controlled by F. A trial population is generated by BSA by using its experiences that is previously generated. It is the historical population that is used to calculate the search-direction matrix.

d. Crossover

Trial individuals are employed at this stage to obtain the better fitness values by evolving the target population. The binary integer-valued matrix (map) is calculated to determine which individual Mutant to be manipulated by the relevant individuals of P. This stage is defined in Equation (7).

$$Mutant_{i,j} = \begin{cases} P_{i,j}, & map_{i,j} = 1 \\ Mutant_{i,j}, & otherwise \end{cases} \quad (7)$$

e. Selection-II

In the next generation, the population is selected and updated. This is performed by a mechanism that is referred to as greedy selection. This is defined in Equation (8).

$$P_i^{next} = \begin{cases} Mutant_i, f(Mutant_i) \le f(P_i) \\ P_i \quad , \quad otherwise \end{cases} \tag{8}$$

D. Thresholding using PSO

The Particle Swarm Optimization (PSO) algorithm starts with a random population (called swarm) of candidate solutions (called particles), each one having the parameters to be optimized. Each particle adjusts its velocity vector at every iteration, and the best-known position of the swarm is updated according to the fitness function. The algorithm stores and progressively replace the best parameters of each particle, as well as the particle that best fits the parameters. The process continues until a predefined number of iterations is performed [13].

In order to find optimal threshold, PSO searches the solution so that the optimal segmentation is obtained. At the initialization stage, the PSO algorithm generates a swarm of m particles randomly, where each of them owns its k thresholds. At the next stage, the particles move to maintain the target partition on a search space according to the determined fitness function. The PSO algorithm uses the gray levels {gmin, ...,gmax} as a search space, where gmin and gmax are the minimum and maximum gray levels in a given image respectively [14].

E. Histogram Entropy

The notion of information entropy has been used as a measure of evaluation of image quality. In that idea, the image foreground and background is considered as two distinct signals of image. The entropy of each signal is calculated and summed. When the sum reaches its maximum level, the threshold is said to be optimal. The probability distribution of the gray levels on the black part of the image is defined in (9) [15].

$$\frac{p_0}{P_B}, \frac{p_1}{P_B}, \dots, \frac{p_s}{P_B} \tag{9}$$

The probability distribution of the white section of the image is:

$$\frac{(p_{s+1})}{1-P_B}, \frac{(p_{s+2})}{1-P_B}, \dots, \frac{(p_{n-1})}{1-P_B} \tag{10}$$

where *s* is the threshold, p_i is the probability of pixels with gray level *i* and P_B is the probability of gray level less than or equal to the threshold.

$$P_B = \sum_{i=0}^{s} p_i \tag{11}$$

The entropy of the image object is

$$H_B = -\sum_{i=0}^{s} p_i / P_B \log(\frac{p_i}{P_B}) \tag{12}$$

The entropy of the background is

$$H_w = -\sum_{i=s+1}^{n-1} p_i / (1-P_B) \log(\frac{p_i}{1-P_B}) \tag{13}$$

The threshold *s* is selected such that the total entropy, $H_B + H_W$ is maximized.

There are mostly gray-level values of the pixels in a bright image. On the contrary, the darker image has the gray-level values that close to black. The entropy of a grayscale image is a statistical measure of randomness that can be used to characterize the texture of the input image which can be defined as:

$$-\sum(p * log_2(p)) \tag{14}$$

where *p* is equal to the count of pixels for a particular gray level divided by the total number of pixels.

Therefore, the entropy of a single grayscale image is 0, and the entropy of an image with a uniform gradient with all values from 0 to 255 equally populated in the histogram is 1.

F. Morphological Operation

This process is focused on the shape of a segment or region of interest in the image. It is applied to the binary image obtained from the thresholding process. First a filling operation is undertaken on the input image in the form of an image of the boundary / contour, so that the object segment is obtained. The next process is creating morphological structuring element. This is further continued with erode operation, where the size of the object is minimized by scraping around the object and later the dilation operation is performed to increase the size of the object segment by adding a layer around the object.

III. EXPERIMENT

In order to validate the comparative analysis, 2 experimental designs were implemented. The first is comparing the maximum entropy based image thresholding methods using BSA and PSO. In this first design, 2 benchmark images from [16] and [17] were selected as experimental data.

In the first experimental design 2 benchmark images from [16] were used as data. The first image is the microscopic image of lung epidermoid carcinoma that was taken at 20 x magnification. The second is microscopic image of cells of chronic lymphocytic leukemia taken at 40 x magnification. All the images are JPEG formats with the same sizes (700×504 pixels). For each test image, independent runs were performed.

The second experimental design was undertaken by first enhancing the images before performing the thresholding. The same maximum entropy based image thresholding algorithms using BSA and PSO were applied. In this design, 3 benchmark images from [17] were selected as experimental data. The images are microscopic images of 3 types of breast cancer cell lines that were taken at 40 x magnification. The images were taken using Olympus CKX41 inverted microscope and Olympus DP72 camera. All the images are JPEG formats with the same sizes (4140×3096 pixels).

In order to perform quantitative analysis on the experiment, four measurement criteria are used in the first design and five in the second one. Moreover, in the second experiment, the histogram of each image was also observed and its variance and entropy were calculated.

In the first experimental design, the Mean Square Error (MSE), Peak Signal to Noise Ratio (PSNR), Feature Similarity Index (FSIM) and Mean Structure Similarity Index Map (MSSIM) are computed. In the second, in addition to the four criteria, Image Enhancement Metric (IEM) [18] was also calculated. The PSNR evaluates the similarity of the segmented image and the original image based on the mean

square error (MSE), FSIM is computed using phase congruency (PC) and image gradient magnitude (GM), and MSSIM is computed to evaluate the overall quality measurement of the entire images.

IV. Results and Discussion

In the first experimental design, Figure 2 and Figure 5 show the original images of lung epidermoid carcinoma and chronic lymphocytic leukemia respectively. Figure 3 and Figure 4 depict the results of processing the image in Figure 2 using BSA and PSO respectively. Figure 6 and Figure 7 show the results of thresholding the image in Figure 5 using BSA and PSO respectively. The performance metrics of the experiment on Figure 2 and Figure 5 are presented in Table 1 and Table 2 respectively. The Tables show that the proposed method is able to perform thresholding on the images effectively.

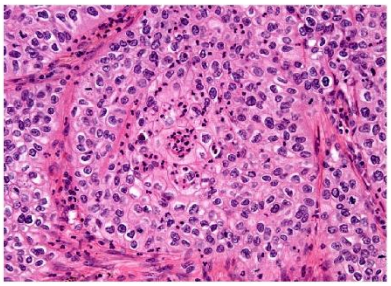

Figure 2: Lung epidermoid carcinoma

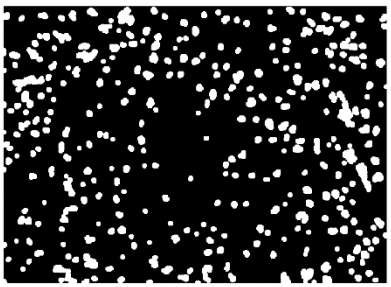

Figure 3: Thresholded image using BSA

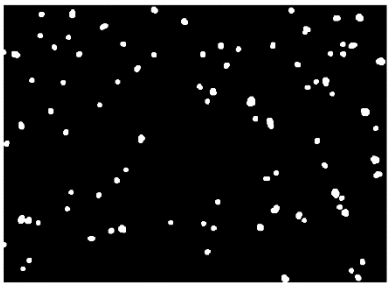

Figure 4: Thresholded image using PSO

Table 1
Performance metrics comparison

Metrics	Backtracking Search Algorithm	Particle Swarm Optimization
Mean Square Error	491.351452	702.973742
Peak Signal to Noise Ratio	21.219126	19.769527
MSSIM	0.989276	0.985899
FSIM	0.976117	0.954531

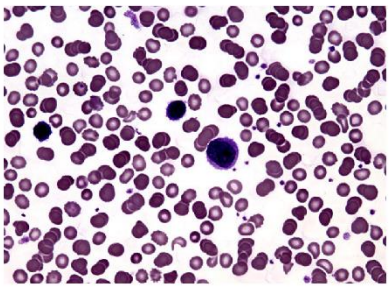

Figure 5: Chronic Lymphocytic Leukemia

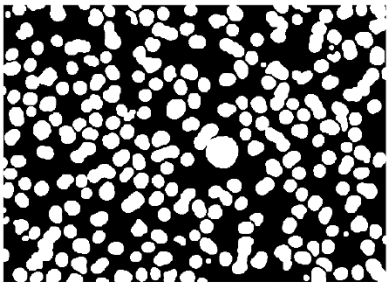

Figure 6: Chronic Lymphocytic Leukemia using BSA

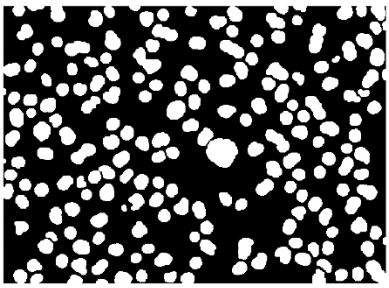

Figure 7: Chronic Lymphocytic Leukemia using PSO

Table 2
Performance metrics comparison

Metrics	Backtracking Search Algorithm	Particle Swarm Optimization
Mean Square Error	426.856725	7190.530694
Peak Signal to Noise Ratio	21.915008	9.585910
MSSIM	0.991251	0.914776
FSIM	0.924388	0.778111

In the second experimental design, Figure 8, 11 and 14 show the original images of 3 breast cancer cell lines. Figure 9, 12 and 15 depict the results of image enhancement and image thresholding using BSA as described in Section III. Figure 10, 13, and 16 show the results of the image processing using PSO.

The performance metrics of the experiment on Figure 8, 11 and 14 are presented in Table 3, 4 and 5 respectively. The Tables show that the proposed method is able to perform enhancement and thresholding on each of the image. Furthermore, the Tables show that the performance metrics of the image that were enhanced and then thresholded using BSA are better than the ones using PSO. The exception is only for IEM in Table 4 and FSIM in Table 5, where the metrics of PSO are better than BSA.

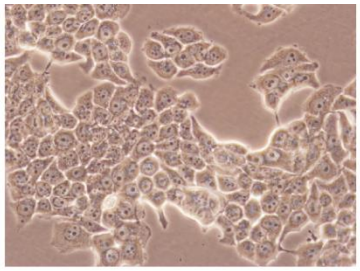

Figure 8: Breast cancer T47D

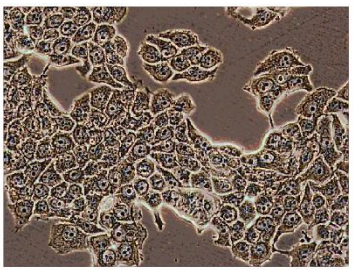

Figure 9: Breast cancer T47D thresholded using BSA

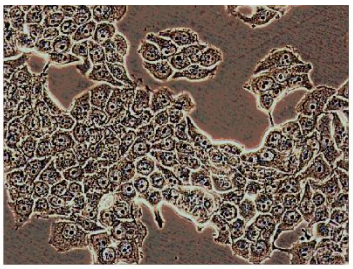

Figure 10: Breast cancer T47D thresholded using PSO

Table 3
Performance metrics comparison

Metrics	Backtracking Search Algorithm	Particle Swarm Optimization
Mean Square Error	413.155501	676.442837
Peak Signal to Noise Ratio	22.136872	19.922091
MSSIM	0.989868	0.983710
FSIM	0.962142	0.940422
IEM	0.3120	0.3047

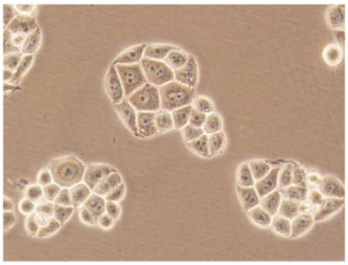

Figure 11: Breast cancer CamaI

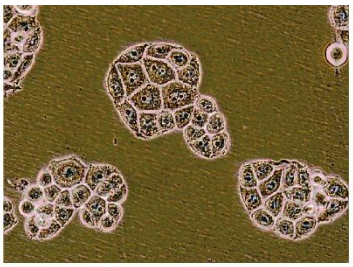

Figure 12: Breast cancer CamaI thresholded using BSA

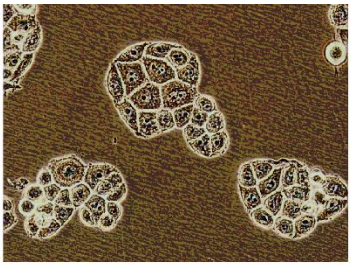

Figure 13: Breast cancer CamaI thresholded using PSO

Table 4
Performance metrics comparison

Metrics	Backtracking Search Algorithm	Particle Swarm Optimization
Mean Square Error	802.759937	1054.897501
Peak Signal to Noise Ratio	19.751658	17.902311
MSSIM	0.976256	0.970314
FSIM	0.893293	0.853484
IEM	0.2213	0.2475

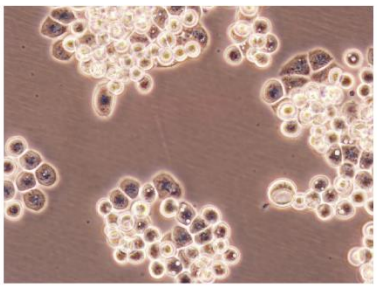

Figure 14: Breast cancer MDA-MB-453

Figure 15: Breast cancer MDA-MB-453 thresholded using BSA

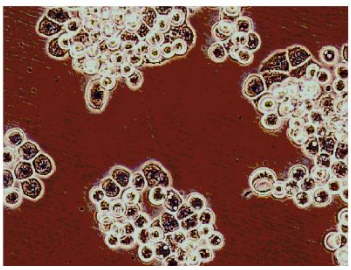

Figure 16: Breast cancer MDA-MB-453thresholded using PSO

Table 5
Performance metrics comparison

Metrics	Backtracking Search Algorithm	Particle Swarm Optimization
Mean Square Error	722.057271	1529.220368
Peak Signal to Noise Ratio	19.799552	16.577139
MSSIM	0.982678	0.955498
FSIM	0.823996	0.835903
IEM	0.1951	0.1733

The histograms of the original image are shown in Figure 17, 18 and 19. The histograms of the enhanced and

thresholded images using BSA are shown in Figure 20, 21 and 22 and using PSO are shown in Figure 23, 24 and 25. All the histograms of the enhanced and thresholded image depict better distribution in red, green and blue channel.

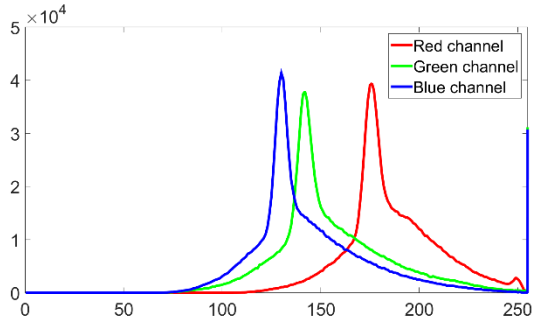

Figure 17: Histogram of original breast cancer image T47D

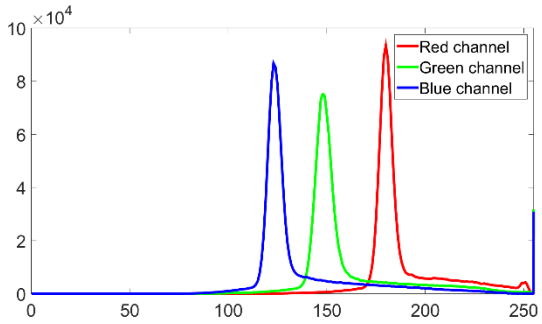

Figure 18: Histogram of breast cancer image CamaI

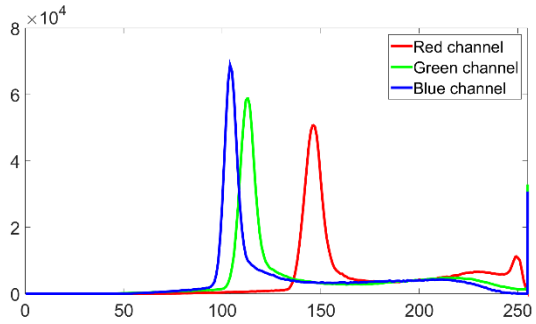

Figure 19: Histogram of breast cancer image MDA-MB-453

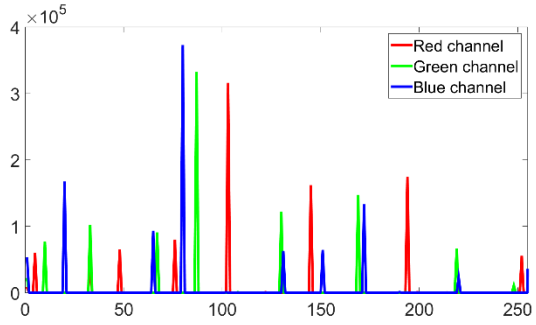

Figure 20. Histogram of breast cancer image T47D thresholded using BSA in Figure 8

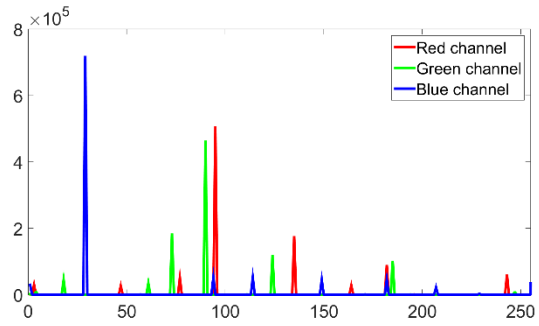

Figure 21: Histogram of breast cancer image CamaI thresholded using BSA in Figure 11

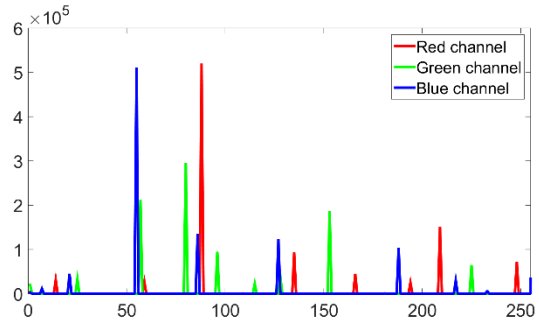

Figure 22: Histogram of breast cancer image MDA-MB-453 thresholded using BSA in Figure 14

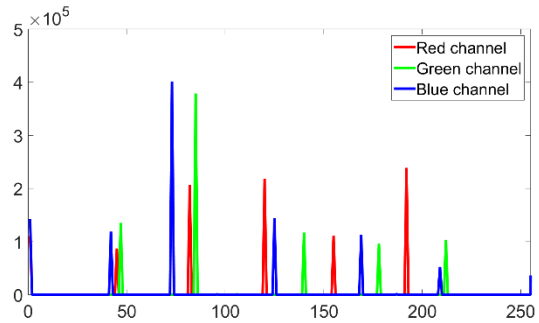

Figure 23: Histogram of Breast cancer image T47D thresholded using PSO Figure 9

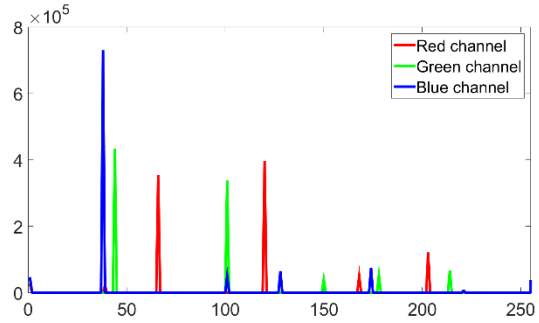

Figure 24: Histogram of breast cancer image CamaI thresholded using PSO in Figure 12

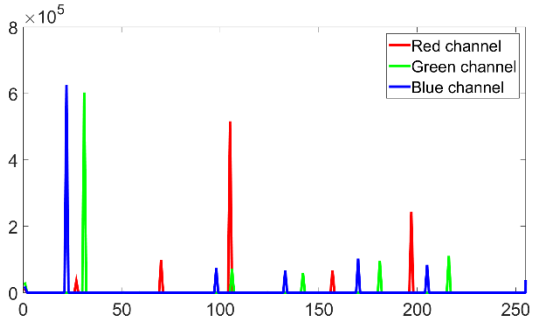

Figure 25: Histogram of breast cancer image MDA-MB-453 thresholded using PSO in Figure 15

Table 6, 7 and 8 show the comparison of variance and entropy of the enhanced and thresholded images using BSA and PSO. All the variance and entropy in that Tables show that the BSA outperforms PSO in the thresholding results.

Table 6
Variance and entropy comparison of T47D

Metrics	Backtracking Search Algorithm	Particle Swarm Optimization
Variance	7.3060e+03	7.6388e+03
Entropy	HR = 9.7136 HG = 9.7265 HB = 9.7211	HR = 9.6702 HG = 9.6648 HB = 9.6648

Table 7
Performance metrics comparison of CamaI

Metrics	Backtracking Search Algorithm	Particle Swarm Optimization
Variance	8.0261e+03	8.4496e+03
Entropy	HR = 9.7211 HG = 9.6756 HB = 9.6794	HR = 9.6702 HG = 9.6702 HB = 9.6447

Table 8
Performance metrics comparison of MDA-MB-453

Metrics	Backtracking Search Algorithm	Particle Swarm Optimization
Variance	7.0179e+03	8.6693e+03
Entropy	HR = 9.7211 HG = 9.7411 HB = 9.6756	HR = 9.6502 HG = 9.6756 HB = 9.6702

V. CONCLUSION

In this paper, a comparative study on optimized thresholding method using two metaheuristic algorithms was presented in order to assist analysis of cancer images. In both experimental designs, the performance metrics were calculated from the thresholded images and compared. Overall results indicate that both optimized thresholding methods are able to achieve acceptable threshold performance.

The metrics show that the optimized method using BSA outperforms the one using PSO. The method using BSA achieves higher performance metrics than the one using PSO. In the first experimental design, where BSA and PSO are implemented to threshold lung epidermoid carcinoma and chronic lymphocytic leukemia cell images, the values of MSE, PSNR, MSSIM, FSIM and IEM show the superiority of BSA over PSO. In the second experimental design where the thresholding method is implemented after image enhancement process, the histogram entropy and variance show that the thresholding method using BSA outperforms the one using PSO.

Therefore, the thresholding method using BSA is more suitable for cancer cells image thresholding in processing the image samples for further analysis. This will provide a more reliable solution and effective way for assisting analysis of cancer cells where it minimizes the difficulties arises in the conventional way of manual observation of microscopic images.

ACKNOWLEDGMENT

This research is funded by The Ministry of Research, Technology and Higher Education of the Republic of Indonesia.

REFERENCES

[1] S. B. Rana, "A Review of Medical Image Enhancement Techniques for Image Processing," vol. 5, no. 2, pp. 1282–1286, 2015.
[2] X. Zheng, H. Ye, and Y. Tang, "Image Bi-Level Thresholding Based on Gray Level-Local Variance Histogram," *Entropy*, vol. 19, no. 5, p. 191, 2017.
[3] K. Vennila, "Implementation of Multilevel Thresholding on Image using Firefly Algorithm," vol. 8, no. 3, 2017.
[4] P. Civicioglu, "Backtracking Search Optimization Algorithm for numerical optimization problems," *Appl. Math. Comput.*, vol. 219, no. 15, pp. 8121–8144, 2013.
[5] G. Chonghui and L. Hong, "Multilevel Thresholding Method for Image Segmentation Based on an Adaptive Particle Swarm Optimization Algorithm," pp. 0–5, 2017.
[6] B. H. Latha and C. Engineering, "An Efficient Implementation of Hybrid segmentation Based On Otsus And Particle Swarm Optimization Technique," no. April, pp. 12–16, 2017.
[7] S. P. Duraisamy and R. Kayalvizhi, "A New Multilevel Thresholding Method Using Swarm Intelligence Algorithm for Image Segmentation," *J. Intell. Learn. Syst. Appl.*, vol. 2, no. 3, pp. 126–138, 2010.
[8] P. Civicioglu, U. H. Atasever, C. Ozkan, E. Besdok, a. E. Karkinli, and A. Kesikoglu, "Performance Comparison Of Evolutionary Algorithms For Image Clustering," *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XL-7, no. October, pp. 71–74, 2014.
[9] P. Civicioglu and E. Besdok, *A conceptual comparison of the Cuckoo-search, particle swarm optimization, differential evolution and artificial bee colony algorithms*, vol. 39, no. 4. 2013.
[10] F. Keskin, A. Suhre, K. Kose, T. Ersahin, A. E. Cetin, and R. Cetin-Atalay, "Image Classification of Human Carcinoma Cells Using Complex Wavelet-Based Covariance Descriptors," *PLoS One*, vol. 8, no. 1, pp. 1–10, 2013.
[11] S. C. F. Lin *et al.*, "Intensity and edge based adaptive unsharp masking filter for color image enhancement," *Optik (Stuttg).*, vol. 127, no. 1, pp. 407–414, 2016.
[12] K. Bhargavi, "A Survey on Threshold Based Segmentation Technique in Image Processing," *Int. J. Innov. Res. Dev.*, vol. 3, no. 12, pp. 234–239, 2014.
[13] F. Balabanian, E. Sant'Ana da Silva, and H. Pedrini, "Image Thresholding Improved by Global Optimization Methods," *Appl. Artif. Intell.*, vol. 9514, no. April, pp. 1–12, 2017.
[14] C. Qi, "Maximum Entropy for Image Segmentation based on an Adaptive Particle Swarm Optimization," *Appl. Math. Inf. Sci.*, vol. 8, no. 6, pp. 3129–3135, 2014.
[15] D. M. Carabias, "Analysis of Image Thresholding Methods for their Application to Augmented Reality Environments," 2012.
[16] MicroscopyU, "Lung Epidermoid Carcinoma at 20x Magnification," 2017. [Online]. Available: https://www.microscopyu.com/gallery-images/lung-epidermoid-carcinoma-at-20x-magnification. [Accessed: 01-Jan-2017].
[17] MicroscopyU, "Chronic Lymphocytic Leukemia at 40x Magnification," 2017. [Online]. Available: https://www.microscopyu.com/gallery-images/chronic-lymphocytic-leukemia-at-40x-magnification. [Accessed: 01-Jan-2017].
[18] V. L. Jaya and R. Gopikakumari, "IEM : A New Image Enhancement Metric for Contrast and Sharpness Measurements," *Int. J. Comput. Appl.*, vol. 79, no. 9, pp. 1–9, 2013.