

RESEARCH

Open Access



An object localization optimization technique in medical images using plant growth simulation algorithm

Deblina Bhattacharjee¹, Anand Paul^{1*}, Jeong Hong Kim¹ and Mucbeol Kim²

*Correspondence:

anand@knu.ac.kr

¹ Department of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea

Full list of author information is available at the end of the article

Abstract

The analysis of leukocyte images has drawn interest from fields of both medicine and computer vision for quite some time where different techniques have been applied to automate the process of manual analysis and classification of such images. Manual analysis of blood samples to identify leukocytes is time-consuming and susceptible to error due to the different morphological features of the cells. In this article, the nature-inspired plant growth simulation algorithm has been applied to optimize the image processing technique of object localization of medical images of leukocytes. This paper presents a random bionic algorithm for the automated detection of white blood cells embedded in cluttered smear and stained images of blood samples that uses a fitness function that matches the resemblances of the generated candidate solution to an actual leukocyte. The set of candidate solutions evolves via successive iterations as the proposed algorithm proceeds, guaranteeing their fit with the actual leukocytes outlined in the edge map of the image. The higher precision and sensitivity of the proposed scheme from the existing methods is validated with the experimental results of blood cell images. The proposed method reduces the feasible sets of growth points in each iteration, thereby reducing the required run time of load flow, objective function evaluation, thus reaching the goal state in minimum time and within the desired constraints.

Keywords: Nature-inspired computing, Plant growth simulation algorithm, Medical image segmentation, Object recognition

Background

Various computing techniques inspired from nature has been extensively used in solving problems spanning from optimization, pattern recognition, machine learning, image detection to computer vision. There is hardly any field that is left uninfluenced by the nature-based computing techniques. Image Processing is one such field where recently biomimicry methods are being used invariably. One such application of nature-inspired computation is in the field of medical image processing, especially focused on object localization. In this paper, the problem of object localization in medical images has been solved with the application of the highly efficient plant growth simulation algorithm (PGSA) (Li and Wang 2008) applied to the analysis of white blood cell (WBC) images. WBCs, also known as leukocytes, play a very important role in the diagnosis of a myriad

of diseases. In most haematology labs, such cell differential analyses are performed using manual microscopy but this traditional process is not without its limitations. Such analysis require the availability of highly experienced personnel. Due to the substantial possibility of an inter and intra-observer variability in these manual examinations and due to the highly labour intensive routine procedures, new methods for cell analysis are being developed using digital image processing techniques to form better and reliable systems for disease diagnosis. However, high variability of cell shape, edge, localized features, the contrast between cell boundaries and background, cell size and positions have posed as challenges for the efficient object localization during the analysis of such smear images. In this paper, the process of detecting the white blood cells have been undertaken because as per haematology, the WBC tests are tougher and analysing them manually are more difficult as compared to red blood cells. The WBC detection problem has been solved in this article by viewing it as a circle detection problem because WBCs can be approximated as circular in shape. In medical imaging, detecting circular features holds huge significance (Karkavitsas and Rangoussi 2007). As per the existing conventional method of Hough Transform for circle detection in digital images (Muammar and Nixon 1989), an edge detector is used to first find the two necessary and sufficient parameters for circles i.e. the coordinates and the radius of the circle. Upon finding these, the image is averaged over its pixels, followed by filtration for detecting the image peaks. Finally, the image is transformed using a histogram. However, in order to cover all parameters (x, y, r) , a lot of memory is required by this method. It also implies a high computational time complexity decreasing its processing speed. Also, the method is not resistant to noise thereby resulting in even lower accuracy (Atherton and Kerbyson 1993). To overcome such a problem, some other approaches based on the Hough transform, for instance the probabilistic Hough transform (Fischler and Bolles 1981; Shaked et al. 1996), the randomized Hough transform (RHT) (Xu et al. 1990), the fuzzy Hough transform, Circular Hough transform with local maximization (Yadav et al. 2014), one-dimensional Hough Transform (Zhou et al. 2014), Hough Transform of curves (Campi et al. 2013) and recently scanline-based hybrid Hough Transform (Seo and Kim 2015) have been proposed with better time complexities but having average memory usage and no noise resistance. In order to overcome the drawbacks of the Hough Transforms, many optimization techniques have been applied to the circle detection problem. These methods have produced much higher accuracy, stability, computational speed and robustness as compared to the discussed Hough Transform as well as other methods like Otsu method based on circular histogram (Wu et al. 2006), WBC identification based on support vector machines (Wang and Chu 2009), and modified transformation methods as proposed in the scientific literatures (Becker et al. 2002; George et al. 2014). These optimization techniques are nature-inspired methods including genetic algorithms (GA) (Ayala-Ramirez et al. 2006), simulated annealing with differential evolution (DE) (Das et al. 2008), harmony search algorithm (Pan et al. 2010, 2011; Cuevas et al. 2012a, b, c), swarm intelligence methods like ant colony optimization based on ant regeneration and recombination to solve the circle detection problem (Chattopadhyay et al. 2008), adaptive bacterial foraging algorithm with adaptive chemotactic step size to facilitate faster convergence (Dasgupta et al. 2008), artificial bee colony algorithm for circle detection (Cuevas et al. 2012a, b, c), clonal selection algorithm for circle detection based on

artificial immune system (Isa et al. 2010) and fuzzy cellular neural network (Tong et al. 2005; Wang and Cheng 2007a, b), all of which has been discussed in context of solving the circle detection problem.

Hereunder each optimization method will be discussed one by one and their possible problems that were unaddressed in the cited literature before introducing our proposed scheme. In summary, the genetic algorithm is the most favored computational intelligence model for multi-circle detection so far and has been proven to be more suitable for multi-circle detection problem among other computational intelligence based methods. However, due to the nature of global optimization of genetic algorithm, multi-circle detection requires additional processing. The ideal case would be an algorithm with an inbuilt computational intelligence with a niche adaptability and robustness that only needs to run once like regular deterministic approaches. In Das et al. (2008), simulated annealing and differential evolution has been combined to perform circle detection. Although the method here is robust to noise, it fails to detect circle locations with considerable precision, under both clear and noisy conditions seen from their result samples. In ant colony algorithm (Chattopadhyay et al. 2008), the final circle detection criterion is to threshold the deviation error derived from the detected radius which is the distance between corresponding edge pixels and circle center. This method is essentially a closed loop tracking method, and its performance is questionable when circular shape edges are not enclosed (Chattopadhyay et al. 2008). In Cuevas et al. (2012a, b, c), where the artificial bee colony algorithm has been used, the potential problem is that a lot of memory space would be used if the iteration is set to a large number, but it saves rerun computations. For the fuzzy cellular neural network (Wang and Cheng 2007a, b), the basic limitation is that it takes single inputs where only one WBC is analysed. Moreover, with an exponential increase in the number of iterations the detected circle gets distorted covering the surrounding area, thereby giving more false positives and losing out on the true positives. Also for the adaptive bacterial foraging algorithm in Dasgupta et al. (2008), the method is not inherently capable of detecting multiple circles. In the clonal selection algorithm for circle detection (Isa et al. 2010), both the antigens and antibodies are designed as 10-by-10 images and representations are a binary string which makes it not very practical to process normal resolution images.

In this paper, the detection of WBC has been done by the PGSA. The PGSA is a stochastic evolutionary computation technique based on the natural growth process of a plant towards the global optimal solution—sunlight. Based on plant phototropism, the PGSA regards the feasible region of Integer programming as plant growth environment and evaluates the probability on different growth points according to the changes in the object function. It then grows towards the global optimal solution—light source. The plant grows a trunk from its roots; some branches will grow from the nodes on the branches. This repeats until a plant is formed. The plant branches out through a number of iterations (which can be considered as generations) towards the globally optimal solution, thereby forming an optimal configuration structure that can help it to absorb maximum sunlight for photosynthesis. In the literature (Wang and Cheng 2007a, b), PGSA is compared with other optimization algorithms where the results have shown that the optimal network given by PGSA is the best option as compared to the existing optimization techniques namely genetic algorithms, particle swarm optimization, gradient

descent and Tabu search, with a higher rate of accuracy and faster global optimization. As per the analyses, PGSA has the following advantages: (1) the objective function and constraint satisfaction are dealt separately, (2) it does not require any predefined error coefficient, rates of cross-over and mutation therefore resulting in stable solutions, (3) it has a search mechanism with ideal direction and randomness balancing property which is determined by the plant growth hormone (morphactin) concentration and thus finds the global optimal solution quickly. In literatures Luo and Yu (2008), Xu et al. (2012), Lu and Yu (2013), Kumar and Thanushkodi (2013), certain improvements have been carried out on PGSA by studying growth characteristics of plants. The algorithm uses the variable growth rate of the plant vertex to reduce the search time and uses the vertical growth characteristics of the early growth to reduce search space; hence, it is possible to obtain a more optimal solution in less time. Thus, PGSA gives minimum loss while showing greater convergence stability.

The PGSA based circle detector uses three edge points on the image which are randomly selected that represent candidate circles in the edge map of the blood sample image. First, a validation is done to check if these candidates are really present in the image edge map which is generated in the pre-processing stage that will be discussed later in the paper. This is done by calculating the fitness value of such candidates. The better a candidate circle approximates the actual edge circle, the better will be the fitness function value. Hence, the edge map should be accurate and precise enough. Further, the segmentation of the image, which will be mentioned in the pre-processing stage later in the paper, plays an important role to accurately measure the similarity of a candidate circle with an actual WBC. Guided by the values of the new objective function, the set of encoded candidate circles are evolved using the PGSA algorithm so that they can fit into the actual WBC on the image. The approach generates a subpixel detector which can effectively identify leukocytes in real images. PGSA is relatively new, having been introduced in the year 2005 and has never been applied to image processing techniques. This paper aims to apply this highly efficient evolutionary technique towards medical image processing, by proposing a new WBC detector algorithm that efficiently recognizes WBC under different complex conditions while considering the whole process as a circle detection problem.

Circle detection using PGSA

Plant growth simulation algorithm (PGSA)

The PGSA is a bionic random algorithm guided by plant phototropism (the ability of a plant to bend towards the light source). The light source is the global optimal solution and the PGSA simulates the mechanism of plant phototropism by assessing the morphactin concentration on the growth points of the plant. This morphactin concentration decides the growth of branches and leaves and is dependent on the intensity of light. PGSA regards the feasible region of Integer programming as plant growth environment and evaluates the probability on different growth points according to the changes in the light intensities taken as the corresponding objective function (Li and Wang 2008; Bhattacharjee and Paul 2016). The algorithm emphasises on a plant system's method of making decisions which are based on plant's growth rules and probability models. Biological experiments state the following plant growth laws: First, the node on the plant with a

higher morphactin concentration has a greater probability to grow into a branch. Second, the morphactin concentrations on these nodes vary according to the environmental information and the relative positions of these nodes on the plant. If a node has the highest morphactin concentration and hence, if it grows into a branch, the morphactin concentrations of all the remaining plant nodes will be freshly allotted as per the new environment and the just branched node will have a concentration equal to zero.

Mathematical model for plant growth

According to the probability model of PGSA, $f(i)$ is the fitness function for the node i on the plant to grow in the given environment. The biological laws of plant growth specifies that if the value of $f(i)$ is smaller, then node i has a better environment for growing into a new branch. The mathematical model of PGSA is as follows. Given a root R_0 , a trunk T grows from the root. Assuming that there are n nodes on the trunk T that might provide a more thriving growth environment than the root, i.e. the fitness function $f(R_{Ti}) < f(R_0)$ ($i = 1, 2, \dots, n$), the morphactin concentrations C_{Ti} of these nodes are shown in Eq. (1).

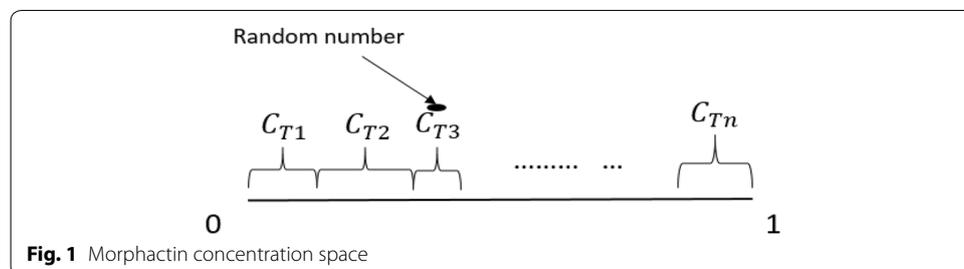
$$C_{Ti} = \frac{f(R_0) - f(R_{Ti})}{\sum_{i=1}^n (f(R_0) - f(R_{Ti}))} \quad (i = 1, 2, \dots, n) \tag{1}$$

The above equation shows that the morphactin concentration of a node in a plant is dependent on the growth environment of all the nodes on the plant. A change in the concentration of one node, therefore effects the rest. These concentrations can be imagined as a state space of the interval $[0,1]$ because the $\sum C_{Ti} = 1$. The state space can be shown as in Fig. 1.

Now, from the state space of concentrations a random number is obtained which returns a random concentration C_{TM} . The corresponding preferential node R_{TM} takes the priority in the next iteration to branch out. However, the node R_{TM} will grow into a branch only if the random number β satisfies the following two Eqs. (2) and (3).

$$0 \leq \beta \leq \sum_{i=1}^M C_{Ti} \quad (M = 1) \tag{2}$$

$$\sum_{i=1}^{M-1} C_{Ti} < \beta \leq \sum_{i=1}^M C_{Ti} \quad (M = 2, 3, \dots, n) \tag{3}$$



After the branch sprouts out, the morphactin concentration of the current node, R_{TM} , is set to zero and the morphactin concentration of the remaining nodes is reallocated as follows:

$$C_{Ti} = \frac{f(R_0) - f(R_{Ti})}{\sum_{i=1, i \neq M}^n (f(R_0) - f(R_{Ti})) + \sum_{j=1}^p (f(R_0) - f(R_{bj}))} \quad (4)$$

$$C_{Ti} = \frac{f(R_0) - f(R_{bj})}{\sum_{i=1, i \neq M}^n (f(R_0) - f(R_{Ti})) + \sum_{j=1}^p (f(R_0) - f(R_{bj}))} \quad (5)$$

After the reallocation of the concentrations to all the nodes on the plant except R_{TM} , the state space of concentrations is again formed with the same interval $[0, 1]$. Assuming, the newly grown branch b has p nodes, such that $f(R_{bi}) < f(R_0)$ ($i = 1, 2, \dots, p$), again a random number β is thrown in the state space and a new node branches out in the next iteration. The new state space has greater number of nodes now, i.e. the nodes previously present (n nodes) and the nodes on the new branch b (p nodes). This growth process stops in the bionic world when the plant has reached its maturity and cannot further branch out.

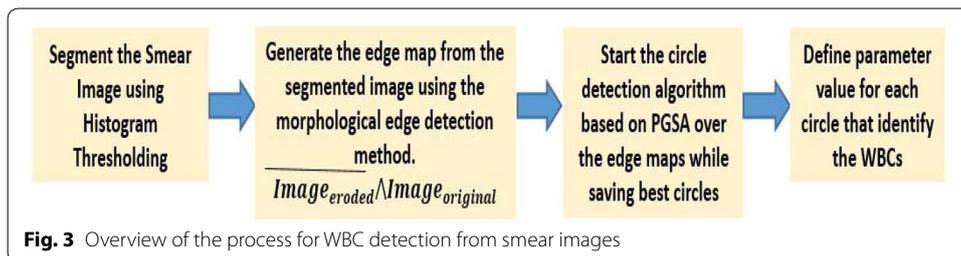
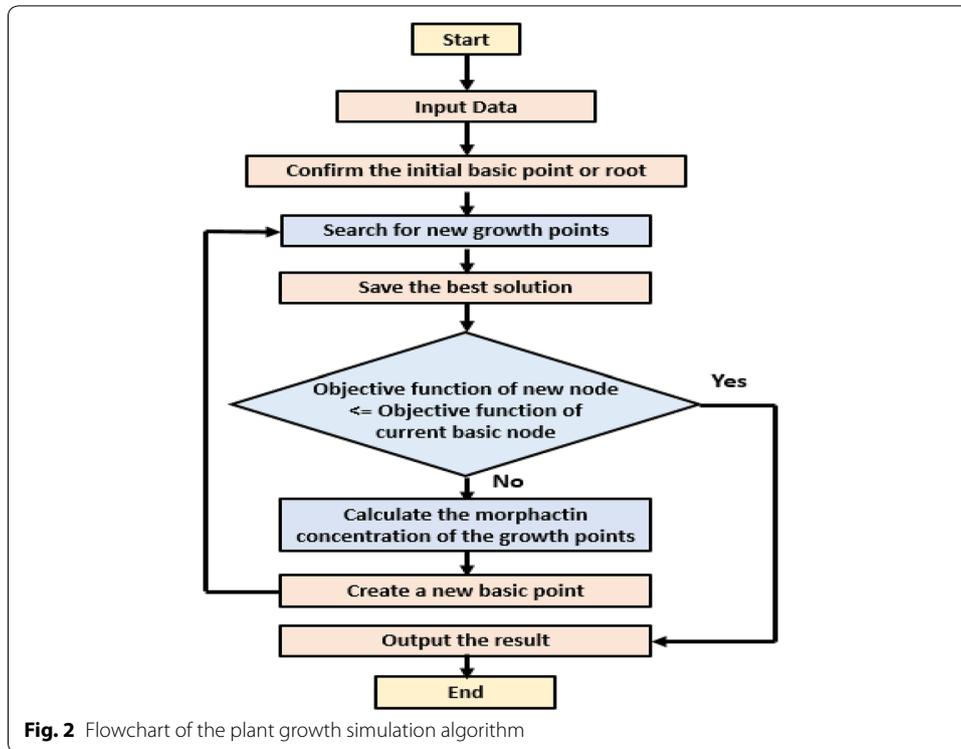
The PGSA has huge potential to be used in optimization problems. Here, the control parameters are the fitness function $[f(i)]$, the initial solution (root), search domain of candidate solutions (length of the trunk and branches) and candidate solutions (plant nodes). Further, it has a well-balanced exploration to exploitation ratio (Crepinsek et al. 2013). This method keeps exploring the entire search space with random node selection in the search interval $[0, 1]$. Although, candidate solutions (nodes) grow in each iteration the search space still remains in the interval $[0, 1]$. After the exploration, upon the selection of the best candidate solution (preferential node), the morphactin concentrations are reassigned by a neighbourhood like search that assesses the nodes in the vicinity of the just grown branch. The morphactin is not only assigned to the nodes on the new branch by exploitation but also to the previous nodes on the trunk by exploration in a given iteration. This is mainly because the objective function (growth environment) is dependent on the concentration of all the nodes on the plant. Thus PGSA has a well-balanced exploration to exploitation ratio which is necessary for any search optimization algorithm.

There have been many algorithms, both traditional and evolutionary, that have been used for circle detection. As has been stated in the background study, all these methods had some disadvantages. PGSA, being inspired from plant phototropism has many advantages that can be used to solve the circle detection problem. The steps of the algorithm has been outlined in Fig. 2.

Our approach is to use the above efficient naturally occurring technique to solve the leukocyte detection problem the overview of which is presented hereunder (Figs. 3, 4).

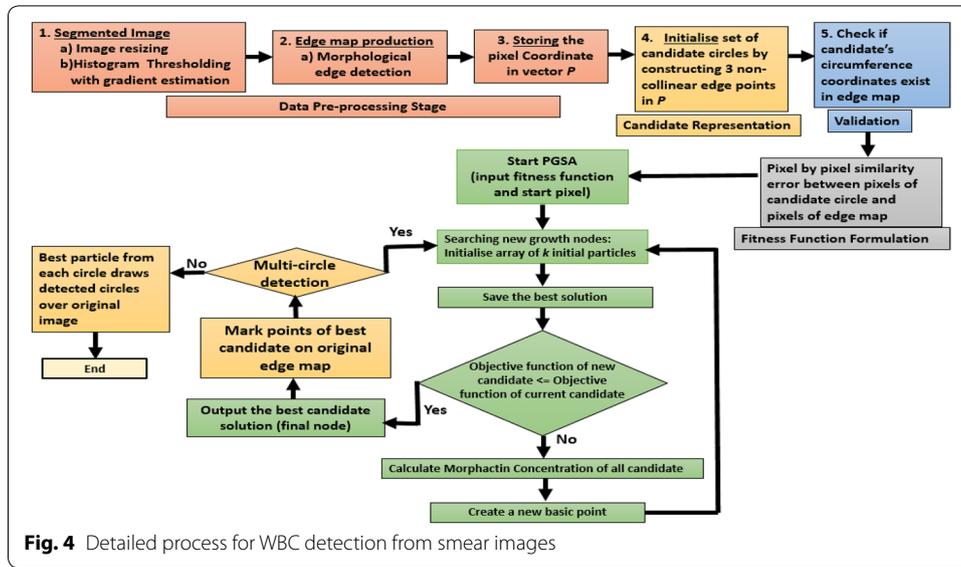
Data pre-processing

To employ the proposed scheme with respect to leukocyte detection, the smear images are pre-processed to obtain two new images. (1) The segmented image and (2) The edge



pixel map of the segmented image. For the segmentation pre-processing part, the WBCs are isolated from other structures including red blood cells and the background pixels. Information of colour, brightness, and gradients are used with a corresponding threshold to generate classes to classify each pixel. A histogram thresholding has been incorporated to segment the WBCs.

Now that the segmentation is done, the corresponding edge map is produced. The edge map maintains the total object structure while being just a simple representation of the original image. There many different methods to detect the edges, but for our work the morphological edge detection procedure has been used (Fu and Han 2012; Chandrasiri and Samarasinghe 2014) where erosion followed by inversion of the original image is carried out to ultimately compare it pixel-by-pixel with the original image. This results in the detection of pixels which are present in both the images. This gives the calculated edge map.



Thereafter, the (x_i, y_i) coordinates for every pixel p_i defining the image edge is stored in the image edge pixel vector $P = \{p_1, p_2, \dots, p_{N_p}\}$, with N_p being the total number of pixels defining the edge of the analysed image.

Particle representation for candidate solutions

For the implementation of the algorithm the set of candidate solutions needs to be initialized. For this purpose, the circle candidates need to be constructed using the three non-collinear points on the edge of the circle, previously stored in the vector P. The indices which represent three edge pixel points are grouped assuming that the circle’s contour map connects them together. Let the indices be $e_i, i = \{1, 2, 3\}$, then the circle C passing through these points can be the potential candidate solution for the circle detection problem. $C = \{p_{e_1}, p_{e_2}, p_{e_3}\}$. The centre and radius of the circle C are given by the well-known second degree equation as seen in Eq. (6).

$$(x - x_0)^2 + (y - y_0)^2 = r^2 \tag{6}$$

Here x_0 and y_0 are computed using

$$x_0 = \frac{\det(A)}{4((x_j - x_i)(y_k - y_i) - (x_k - x_i)(y_j - y_i))}, \tag{7}$$

$$y_0 = \frac{\det(B)}{4((x_j - x_i)(y_k - y_i) - (x_k - x_i)(y_j - y_i))}$$

where,

$$A = \begin{bmatrix} x_j^2 + y_j^2 - (x_i^2 + y_i^2) & 2(y_j - y_i) \\ x_k^2 + y_k^2 - (x_i^2 + y_i^2) & 2(y_k - y_i) \end{bmatrix}$$

and

$$B = \begin{bmatrix} 2(x_j - x_i) & x_j^2 + y_j^2 - (x_i^2 + y_i^2) \\ 2(x_k - x_i) & x_k^2 + y_k^2 - (x_i^2 + y_i^2) \end{bmatrix}$$

and

$$r = \sqrt{(x - x_b)^2 + (y - y_b)^2}$$

Here $\det(\cdot)$ stands for the determinant and $b \in \{e_1, e_2, e_3\}$. Therefore, a set of parameters is represented for each circle $[x_0, y_0, r]$ as a transformation T for all edge vector indices e_1, e_2, e_3 which yields:

$$[x_0, y_0, r] = T(e_1, e_2, e_3) \quad (8)$$

By considering each index as a particle in the search space, the continuous search space is explored by using PGSA for a lookup of circle parameters $[x_0, y_0, r]$.

Fitness function for the circle detection problem

Before defining the fitness function of the proposed scheme in context of the problem at hand, it is necessary to validate whether the circumference coordinates of the candidate circle C exists in the edge image. Upon this validation, the fitness function can be calculated. The coordinates spanning the circle circumference are $J = \{j_1, j_2, \dots, j_N\}$, where N is the total number of pixel points used for validating the circle coordinates i.e. whether the candidate circle is in the edge map of the image. Thereafter, the fitness function $f(C)$ is defined as the pixel by pixel similarity error between the above set of pixels J of the circle candidate C (particle) and the pixels of the edge map, giving a greater fitness value for a higher resemblance. This function is given as follows,

$$f(C) = 1 - \frac{\sum_{i=1}^N E(j_i)}{N} - \frac{W_p}{B_p} \quad (9)$$

where, $E(j_i)$ is the expectation function of the presence of the candidate circle pixel at j_i . Thus, $E(j_i)$ has maximum value of 1 if the pixel j_i is an edge pixel point. For all other pixel points, the expectation function has a zero value. Also, W_p is the amount of white pixel falling inside the candidate circle represented by C and B_p is the amount of black pixels falling inside C (Cuevas et al. 2012a, b, c). These two parameters have been taken into the calculation of the objective function as analysis of smear images cannot be done by directly applying the PGSA algorithm to such images. Smear images present different imaging conditions and staining intensities, which result in noisy edge maps. Thus, in order to use PGSA based circle detector within the context of WBC detection, the fitness function requires these parameters.

PGSA implementation

The PGSA has the following steps.

Step 1 The Canny filter finds and stores the edges in the vector P as discussed in the pre-processing step, where P contains the set of all edge pixels of the image. The iteration index is set to 1.

Step 2 k initial particles are generated ($C_{a,iteration=1}, a \in [1.k]$) in the plant growth environment state space.

Step 3 The fitness function $f(C_{a,iteration=i})$ is evaluated to find the best candidate solution like B_{M2} as mentioned in the PGSA discussion in the previous section. This best candidate solution is named as $C^{best} \leftarrow \arg \min\{f(C_{a,iteration=i})\}$.

Step 4 As per the PGSA discussed in the previous section, the constraint satisfaction at each of these particles (nodes) is checked, that is their morphactin concentration is calculated as per the Eq. (1). The particle with the higher morphactin concentration has a higher probability to branch out or move to the next iteration as an evolved candidate solution.

Step 5 The new branch position, which is the new particle's position is stored and the morphactin concentrations of all the particles are calculated again according Eq. (4) and Eq. (5) except for C^{best} as it has already produced a branch i.e. it is the best solution and hence is the current local optimum solution.

Step 6 For every new particle, a maximum number of q particles are generated as per $j = 1, 2, \dots, q$ discussed previously and based on the newly calculated morphactin concentration in *Step 5*, the new best candidate is found on the current generated branch in the previous step. This accounts for a neighbourhood like search for optimal candidate solutions. This process of generating new candidate solutions continues till a better minimized objective function is achieved and stops till there is no improvement in the fitness value of the generated candidate solutions.

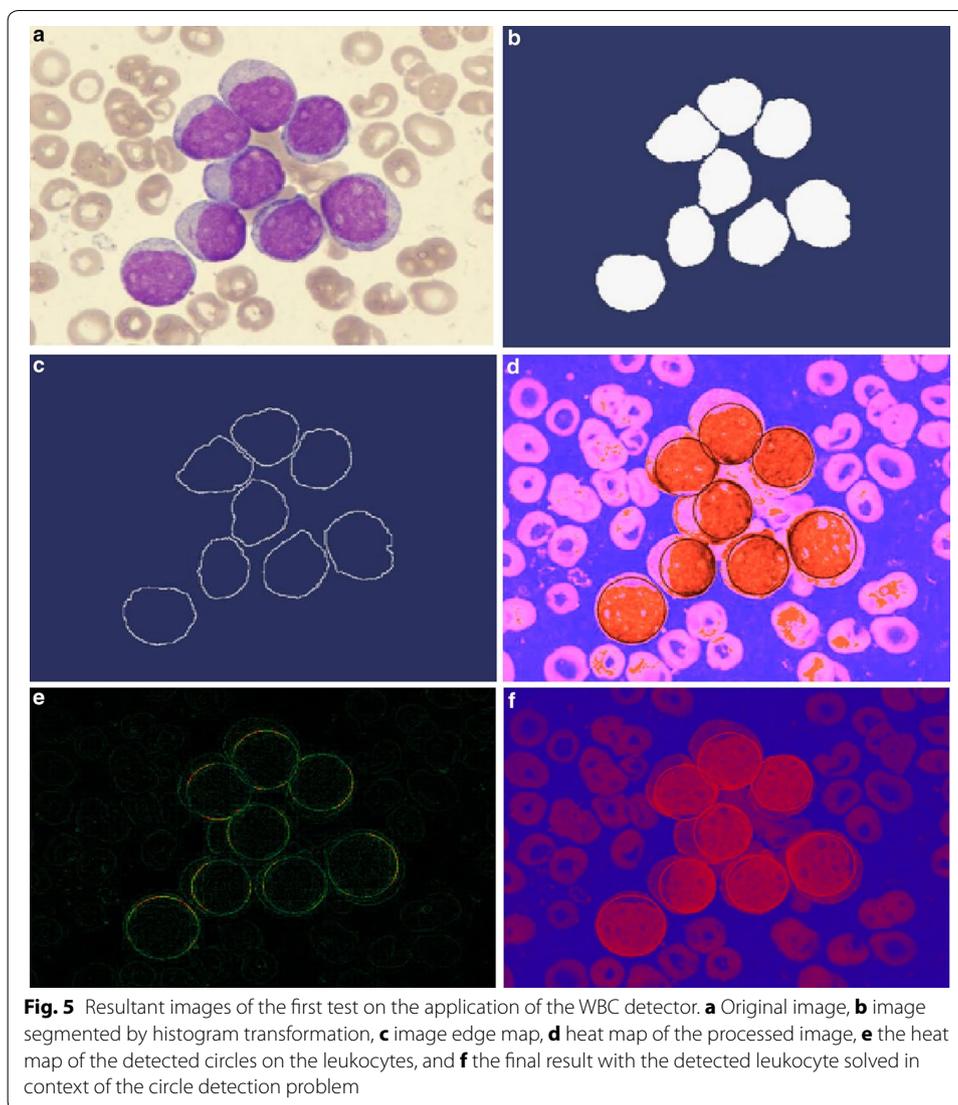
Step 7 The set of all nodes that have branched out are the possible candidate solution with the final node C^{best} being the global best solution and others the local optimal solutions.

Step 8 From the original edge map, the algorithm marks the points corresponding to C^{best} . In case of multi-circle detection it jumps to *Step 2*.

Step 9 Finally, the best particle $C_{N_c}^{best}$ from each circle is used to draw (over the original image) the detected circles, where N_c is the number of circles detected.

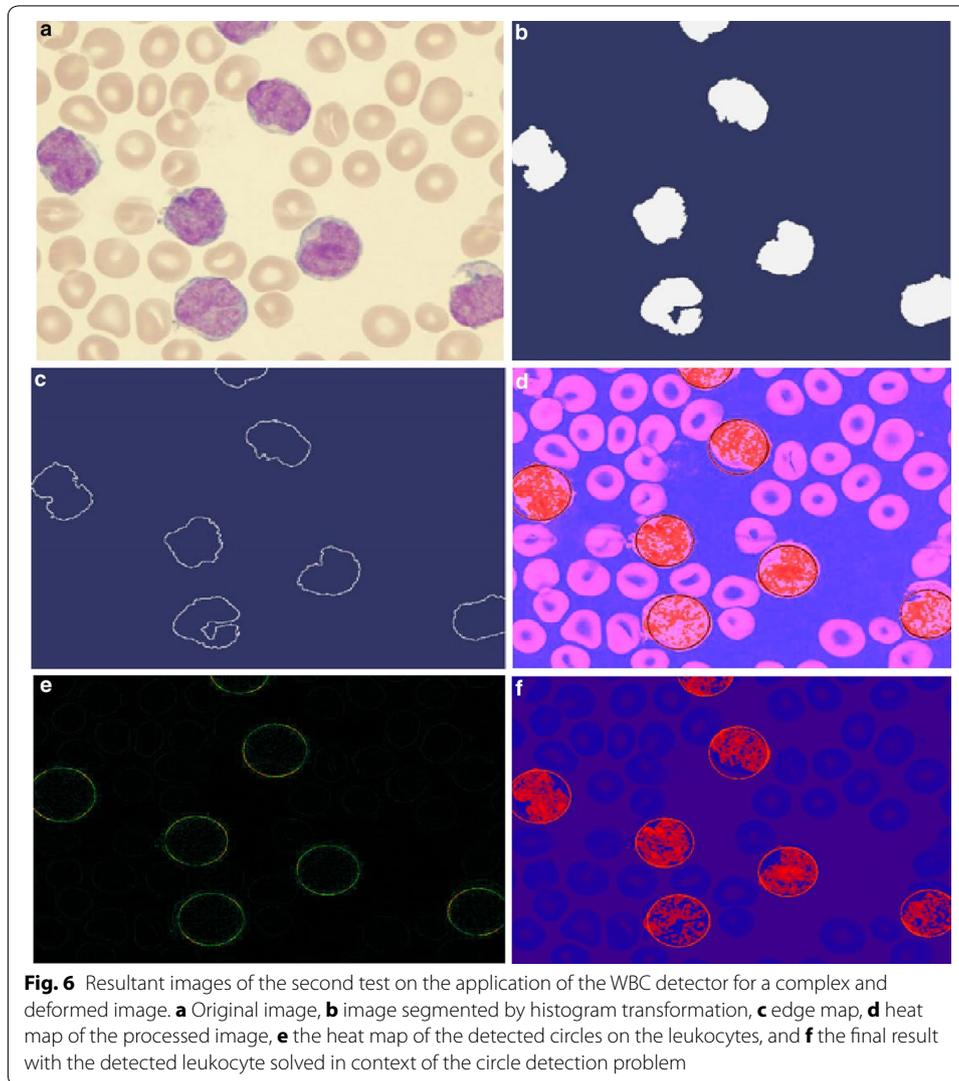
Experimental results

In order to validate the proposed scheme, experimental tests were carried out and thereafter the algorithm performance was evaluated. The proposed method for the detection of WBC was tested over 80 microscopic images of blood smear with a resolution of 360×363 pixels. The precision and sensitivity of the algorithm has been tested under such challenging conditions. The images that were used for the experiments had several deformities, occlusions and overlaps with other images that pose a significant challenge for the circle detection. Figure 5a shows an input blood smear image sample tested with the proposed method for circle detection. Figure 5b shows the segmented WBCs obtained by histogram transforming. Figure 5c is the edge map of the segmented image and Fig. 5d is the heat map of the processed image followed by the heat map of the detected leukocytes as seen in Fig. 5e and finally Fig. 5f gives the detected WBCs.



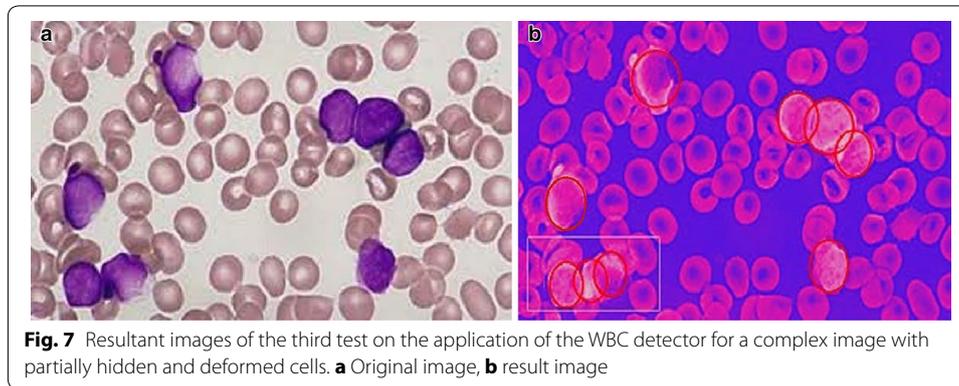
The proposed algorithm is tested on further complicated smear images having highly complex features due to an almost deformed cell. The result is presented in Fig. 6. It represents an example with an image of deformed cells. Clearly, detecting true circles in such complex images becomes tougher. Despite such imperfections, the proposed approach can effectively detect the cells as it is shown in Fig. 6f.

Further, the proposed scheme was tested for detecting partial or hidden leukocytes in blood smear images. Detecting partial and hidden leukocytes are utmost challenging for any detection method and thus, experiments were carried out to check if such leukocytes can be correctly detected by the proposed scheme. The result is shown in Fig. 7. Although, the problem is quite complex in nature, PGSA is seen to be quite successful in detecting such leukocytes. However, the leukocytes that were detected were partially still visible with 68 % hidden surface and not for completely hidden leukocytes. Such detections will be part of the future scope of this paper.



Thus, the proposed method can successfully detect damaged, complex and partially hidden leukocytes correctly.

The dataset of smear-blood test images for evaluating the proposed scheme is downloaded from the website Cellavision.com. The dataset includes 80 images from the Cellavision public dataset which were in JPG format of size 360×363 pixels, with a resolution of 10 pixels per $1 \mu\text{m}$. These images were medically graded and had 463 white blood cells (256 bright leukocytes and 207 dark leukocytes as per the blood smear conditions), all detected by a haematologist—a human medical expert. These numbers were taken as graded standards for all experimentations. For testing the proposed scheme over these images, the true positive rate (known as the number of correctly detected leukocytes over the number of leukocytes detected by the expert) and the false positive rate (known as the number of non-leukocytes that have been wrongly identified as leukocytes over the number of leukocytes which have been actually detected by the medical expert) have been evaluated. The results of the experiments show that the proposed method, achieves 98.28 % leukocyte True Positive Rate with 1.72 % False Positive Rate, and is therefore,



arguably better than the other existing methods. To establish this statement, the proposed scheme has been further discussed and compared with other existing methods in context of the leukocyte detection problem as a circle detection problem hereunder.

Discussion

As this circle detection problem can be considered as a binary classification problem, where the detections can be classified as either positive or negative, the classification can be represented as a confusion matrix which has the following four classes. (1) True Positives for circles correctly detected as positives; (2) False Positives are wrongly detected circles as positives; (3) True Negatives are non-circles correctly detected as such; (4) False Negatives are positive circles incorrectly labelled as non-circles (Parikh et al. 2008, Davis and Goadrich 2006). According to the problem of circle detection, the circle can be either correctly detected (true positive), wrongly detected (false positive) or not detected at all (false negative). Thus there is no true negative in a circle detection task. A true negative condition would have arisen if there isn't a circle in the image and the algorithm also hasn't detected it. As it is a detection problem, non-detections of non-circles are not taken into consideration. Further, based on the confusion matrix, the parameters of true positive, false positive and false negative have been taken to evaluate the performance indices like true positive rate, false positive rate, false discovery rate and positive predictive value for further analysis of the proposed scheme. However, other performance indices like specificity and negative predictive value have not been used as they require true negatives by definition for the calculations (Parikh et al. 2008). As mentioned above, the dataset of 80 images from Cellavision Reference Library was used to identify both bright and dark leukocytes using our proposed scheme. The total number of detected leukocytes, wrongly detected leukocytes and undetected leukocytes have been shown in the Table 1 below. The detection rate and other performance indices in the confusion matrix have been compared among four algorithms. The images from the Cellavision dataset are processed by hough transform (traditional method), modified genetic algorithm with ant colony optimization (both GA and swarm intelligence method are analysed here), fuzzy cellular neural network (which is a member of neural networks) and the proposed method of PGSA; as discussed previously in the background section. All the four algorithms are applied by maintaining their own settings. Thus the images were processed by making no further adjustments to the existing methods, thereby facilitating

Table 1 Comparative performance of HT, FCCN, GA+ACO and PGSA for leukocyte detection with respect to true positive rate, false positive rate, false discovery rate and positive predictive value

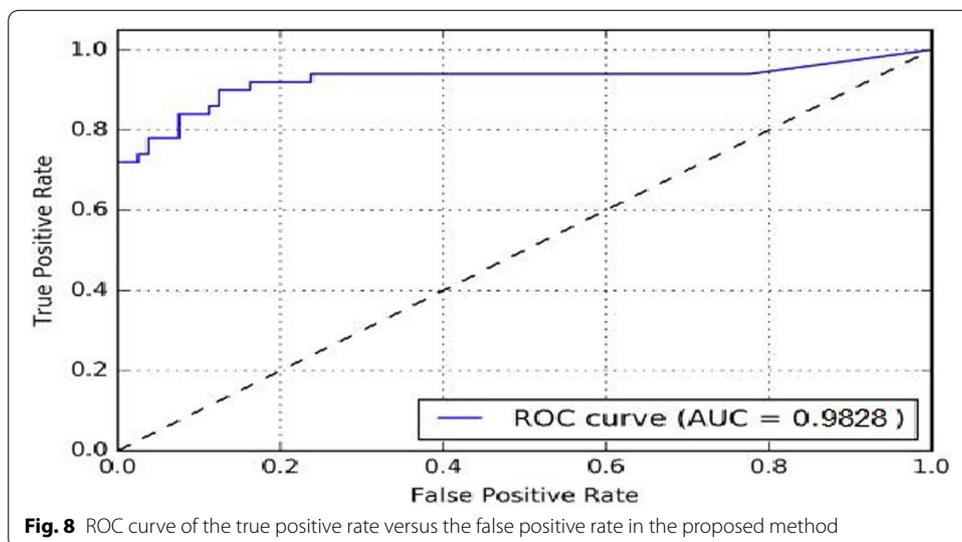
Leuko-cyte type	Method	Leuko-cytes detected (true positives)	Missing leuko-cytes (false negatives)	Missing leuko-cytes (false negatives)	True positive rate (%)	False positive rate (%)	False discovery rate (%)	Positive predictive value (%)
Bright leukocyte (256)	HT	135	121	67	46.85	30.18	59.90	66.83
	FCCN	206	50	55	78.83	24.77	19.16	78.93
	GA + ACO	217	39	42	83.78	18.92	15.06	83.78
	PGSA	242	14	10	98.01	1.99	5.56	96.03
Dark leukocyte (207)	HT	100	107	54	48.04	26.47	69.48	64.94
	FCCN	168	39	49	81.37	24.02	17.97	77.42
	GA + ACO	183	24	38	88.72	18.63	10.86	82.81
	PGSA	202	5	6	98.55	1.45	2.40	97.12
Overall (463)	HT	235	228	121	47.42	28.40	64.04	66.01
	FCCN	374	89	104	80.05	24.41	18.62	78.24
	GA + ACO	400	63	80	86.15	18.78	13.13	83.33
	PGSA	444	19	16	98.28	1.72	4.13	96.52

Italic values signifies the % of true and false positive

the performance comparison of all the methods over different blood smear images with varying background and target object conditions. These resultant images are then checked with medically graded standards by a medical expert.

It has been already seen that as far as the detection comparison is concerned the PGSA technique fairs extremely well with a detection rate of 98.28 % and is better when compared to other leukocyte detection algorithms. The detection rate is depicted from the ROC curve of the True Positive Rate or the Detection Rate versus the False Positive Rate as shown in Fig. 8.

Thereafter all the results are analysed with the help of statistically calculated precision and sensitivity graphs as shown in Figs. 9 and 10. The algorithm is tested out and

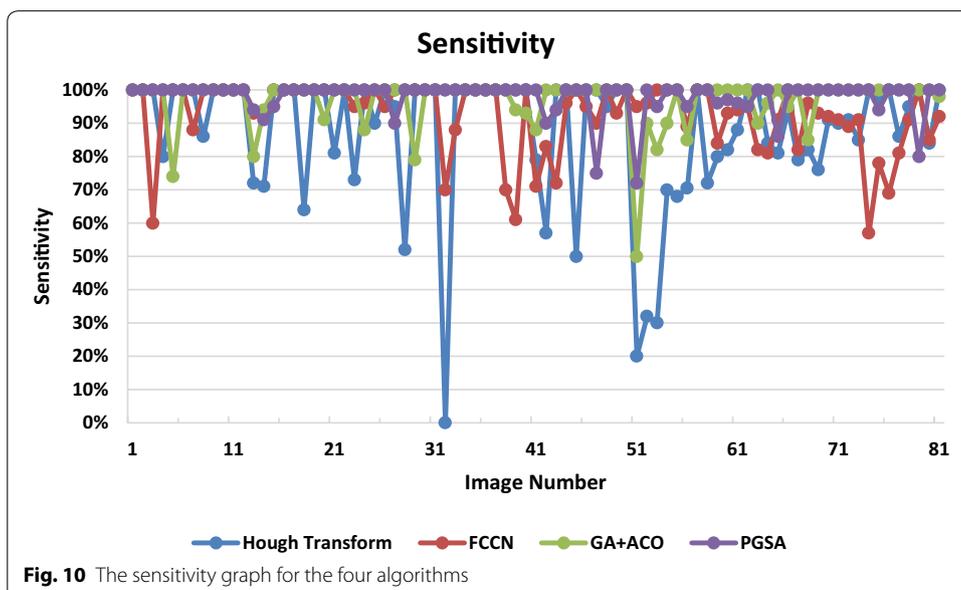
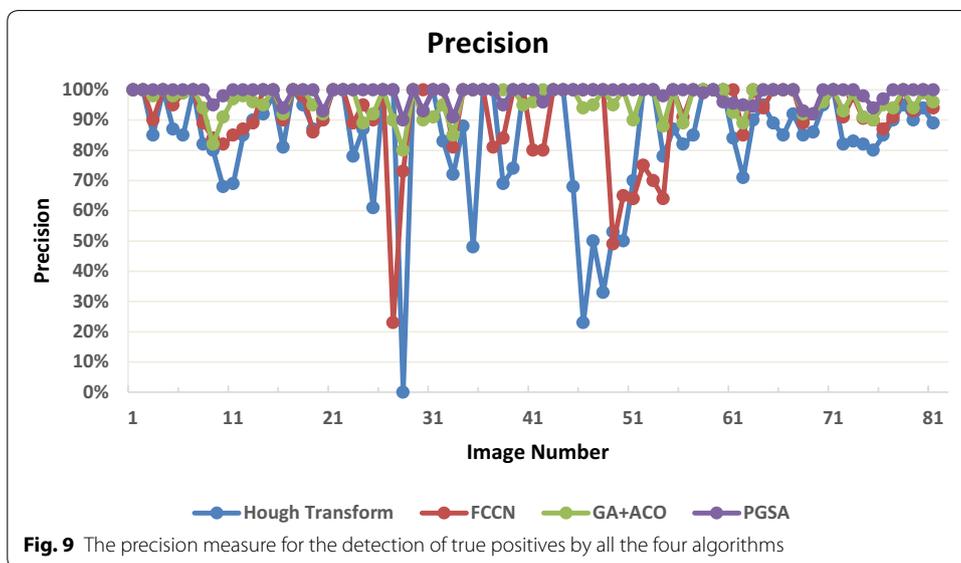


evaluated for its efficiency based on two parameters named as precision and sensitivity as defined below.

$$Precision = \frac{\text{correctly detected circles}}{\text{correctly detected circles} + \text{wrongly detected circles}} \tag{10}$$

$$Sensitivity = \frac{\text{correctly detected circles}}{\text{correctly detected circles} + \text{undetected circles}} \tag{11}$$

Under ideal conditions, the precision and sensitivity both are 100 %, thus the circle detection technique under such scenario would be able to find all the circles in the image without any false positives or un-detected circles. But under normal conditions



precision and sensitivity is seen to decrease because when detecting such circles from the images the circle detection method has to find both the centre of the circle as well as the radius of the circle.

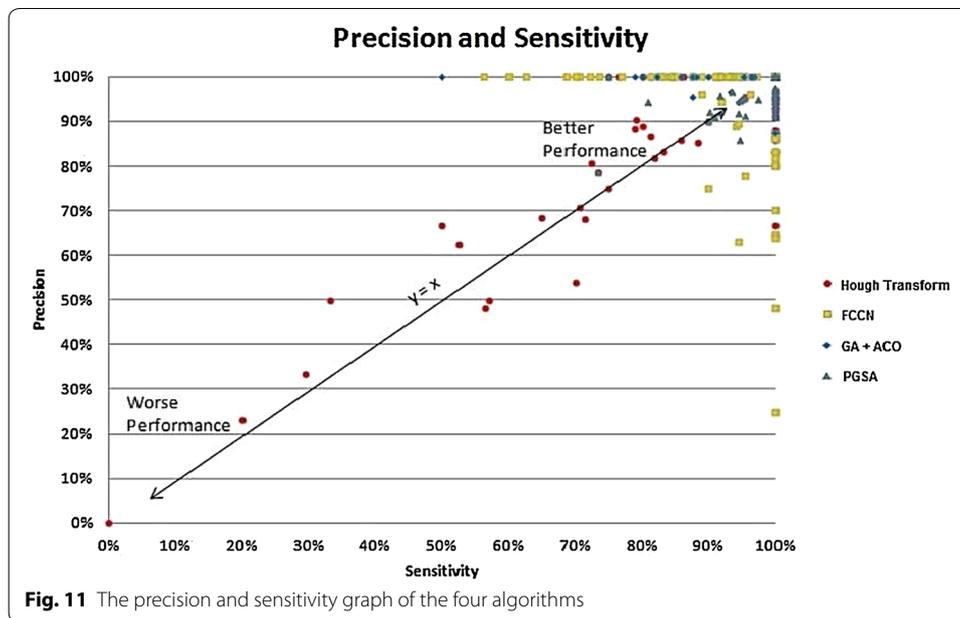
Precision is a measurement of the rate of correct detection over all detected circles. A perfect situation arrives when there is a straight line all along 100 % of y-axis in Fig. 6 for all the images. It is apparent that Hough Transform based method has the worst results, which is not beyond expectation. A completely extreme case is a 0 % precision, implying no detection at all. Regular Hough Transform suffers a number of false detections with many images. The Modified Genetic Algorithm with Ant Colony Optimization and PGSA are seen to achieve correct detection above 90 % with small standard deviation, which are superior to the Fuzzy Cellular Neural Network (FCCN) and evidently the conventional Hough Transform. The possible problem with the FCCN is that it requires a lot of computational time and memory to train its network and the detection is achieved over a large number of generations. Also, as shown in Cuevas et al. (2012a, b, c), when the number of iterations increases, the possibility to cover other structures increases too. Thus, if the image has a complex background like in smear images, the method gets confused because of which finding the correct contour configuration from the gradient magnitude becomes highly difficult.

Sensitivity evaluates the detection rate over all existing circular shapes. On one hand, a low sensitive setting would result in good precision numbers because detected circles must be true and distinctive, while sensitivity rate would be low because non-perfect shapes are missing. On the other hand, high sensitivity means reducing missing incidence but increasing the false detection rate at the same time. The PGSA model initializes nodes on all edge segments, which gives an almost perfect 100 % overall sensitivity, due to every edge segment being checked for constraint satisfaction during each iteration as seen in Fig. 10. Additionally, with the inherent randomness and ideal load balancing capabilities to give optimal configurations, this stochastic method converges very fast as well (Table 2).

In Fig. 11, the results of both the precision and sensitivity have been combined at the same time from the previous two figures. Under the ideal situation, the result would be that all the values would have been centred at the top right corner giving 100 % detection rate. In Fig. 11, the line of $y = x$ indicates an equal probability detection rate based on hit and trial where the probability of such detections become 0.5 or 50 %. Thus, after both precision and sensitivity were combined, the data samples were seen to be scattered for the algorithm with the least satisfactory performance while the data points were seen to incline towards the upper right corner or 100 % detection rate with an improved

Table 2 Statistics of sensitivity for the four analysed algorithms

Approach	Maximum	Minimum	Mean	SD
Hough transform	1	0	0.87	0.198
FCCN	1	0.56	0.93	0.104
GA + ACO	1	0.5	0.97	0.075
PGSA	1	0.73	0.98	0.049



algorithm like FCCN, Genetic Algorithm with Ant Colony Optimization and PGSA, in that order. This shows the advantage of PGSA model over the other methods.

In addition to the above performance measures, the proposed algorithm was also tested for its resistance towards noise. Most of the blood smear images are noisy because of defects or noise in the system that acquires such images, the haematology equipment or other factors. Thus, the circle detection task also depends on how noise-resistant the applied algorithm is. In order to test this performance characteristic of the proposed algorithm the images from the used dataset were corrupted using the (1) Gaussian noise, (2) Salt and pepper noise, which are the most pedestrian noise found in smear images of blood samples. All the 80 images containing 463 leukocytes (both bright and dark) were corrupted with these two noise types in varying levels and were analysed by applying the four algorithms. The Gaussian noise level was used at $\sigma = 10$ and $\sigma = 15$ and the salt and pepper noise level was used at 10 and 15 %. Examples of such corrupted images used for the noise-resistance experiment are shown in Fig. 12. The performance of these algorithms are shown in Tables 3 and 4 respectively.

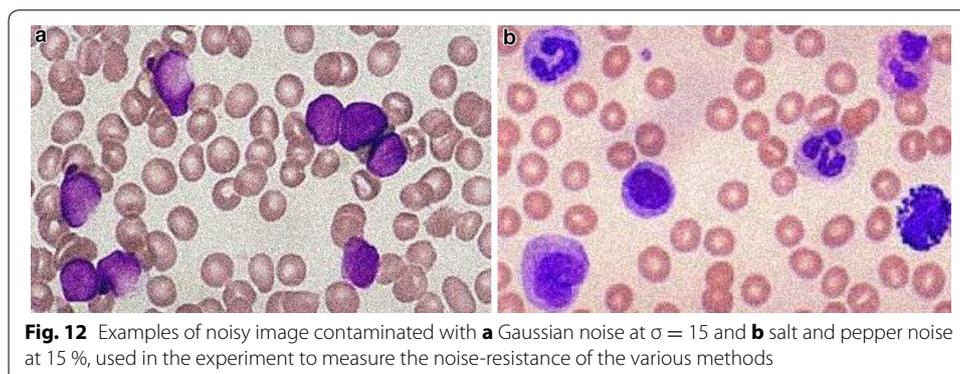


Table 3 Comparative performance of leukocyte detection over 80 images contaminated by various levels of gaussian noise

Noise level	Method	Leukocyte detected (true positives)	Missing Leukocyte (false negatives)	Wrongly detected leukocyte (false positives)	True positive rate (%)	False positive rate (%)	False discovery rate (%)	Positive predictive value (%)
Gaussian noise 463 leukocytes $\sigma = 10$	HT	206	257	77	40.37	18.07	90.81	72.79
	FCCN	343	120	71	72.53	16.67	28.99	82.85
	GA + ACO	335	128	65	70.66	15.26	32.00	83.75
	PGSA	431	32	21	93.19	4.93	7.08	95.35
Gaussian noise 463 leukocytes $\sigma = 15$	HT	177	285	106	33.57	24.88	98.77	62.54
	FCCN	315	148	89	65.96	20.89	36.63	77.97
	GA + ACO	298	165	102	61.97	23.94	41.25	74.50
	PGSA	414	49	32	89.20	7.51	10.99	92.83

Table 4 Comparative performance of leukocyte detection over 80 images contaminated by various levels of salt and pepper noise

Noise level	Method	Leukocyte detected (true positives)	Missing leukocyte (false negatives)	Wrongly detected leukocyte (false positives)	True positive rate (%)	False positive rate (%)	False discovery rate (%)	Positive predictive value (%)
Salt and pepper Noise level 10 % 463 leukocytes	HT	182	281	114	34.74	26.76	94.93	61.49
	FCCN	304	159	106	63.38	24.88	38.78	74.15
	GA + ACO	284	179	118	58.68	27.70	44.53	70.65
	PGSA	424	39	30	91.55	7.04	8.59	93.39
Salt and pepper Noise level 15 % 463 leukocytes	HT	135	328	120	23.71	28.17	128.63	52.94
	FCCN	274	189	78	56.34	18.31	53.69	77.84
	GA + ACO	218	245	123	43.19	28.87	71.85	63.93
	PGSA	408	55	35	87.79	8.21	12.42	92.10

Thus, even under noisy conditions the PGSA is the most robust method to detect the leukocytes with the best detection rate, best positive predictive value, least false positive rate and the least false discovery rate.

Conclusion

In this paper, a new bionic random search algorithm has been proposed that makes use of the objective function's value as an input to the learning model while simulating a plant's phototropism for the automatic detection of WBCs that are embedded into complicated, obscure and cluttered smear images by considering the WBC detection problem as a circle detection problem. The PGSA has been applied to solve this circle detection problem which gives the location of the WBCs in the images using three non-collinear edge points on the segmented edge map of the image as candidate circles. The resemblance of the encoded candidate circles to the actual WBC is evaluated by the objective function which uses the edge map and segmentation results for calculating the

resemblances. Based on the calculated value of the objective function, the set of encoded candidate circles (branch nodes) are evolved by using the PGSA so that they can fit into the actual blood cells that are contained in the edge map. The experimental results and the performance of the PGSA has been compared with other existing WBC detection algorithms which demonstrate the high performance of the proposed method in terms of detection accuracy, precision, and sensitivity and also under noisy conditions. Although, there has been quite some research done to solve the circle detection problem when processing images, it has not been applied in the context of medical image processing. Moreover, PGSA has never been applied to solve such a problem. This evolutionary algorithm is highly efficient and is new to the field of computing intelligence. Thus, it offers a lot of scope for applications, implementations and further extension of this algorithm.

Authors' contributions

Both DB and AP designed the research. DB gathered the data, processed and implemented it while also drafting the manuscript, which was proofread by AP, JHK and MK. All authors read and approved the final manuscript.

Author details

¹ Department of Computer Science and Engineering, Kyungpook National University, Daegu, South Korea. ² The Department of Media Software, Sungkyul University, Anyang, South Korea.

Acknowledgements

This study was supported by the Brain Korea 21 Plus project (SW Human Resource Development Program for Supporting Smart Life) funded by Ministry of Education, School of Computer Science and Engineering, Kyungpook National University, Korea (21A20131600005). This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government (MSIP) [No. 10041145, Self-Organized Software platform (SoSp) for Welfare Devices]. This research was also supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (NRF-2016R1A2A1A05005459).

Competing interests

The authors declare that they have no competing interests.

Ethics approval and consent to participate

No human or animal participant were used in the study, including human tissues. A publicly available database of human blood smear images in digital format has been used.

Received: 19 May 2016 Accepted: 29 September 2016

Published online: 13 October 2016

References

- Atherton TJ, Kerbyson DJ (1993) Using phase to represent radius in the coherent circle Hough transform. In: Proceedings of the IEEE colloquium on the hough transform. IEEE, London
- Ayala-Ramirez V, Garcia-Capulin CH, Perez-Garcia A, Sanchez-Yanez RE (2006) Circle detection on images using genetic algorithms. *Pattern Recognit Lett* 27(6):652–657
- Becker JM, Grousson S, Coltuc D (2002) From Hough transform to integral geometry. In: Proceedings of IEEE international geoscience and remote sensing symposium (IGARSS'02), vol 3, pp 1444–1446
- Bhattacharjee D, Paul A (2016) A hybrid search optimization technique based on evolutionary learning in plants. In: Proceedings of 2016 7th international conference of swarm intelligence, pp 271–279
- Campi C, Perasso A, Beltrametti MC, Massone AM, Sambuceti G, Piana M (2013) Proceedings of 8th IEEE international symposium on image and signal processing and analysis (ISPA 2013), pp 280–283
- Chandrasiri S, Samarasinghe P (2014) Morphology based automatic disease analysis through evaluation of red blood cells. In: Proceedings of 2014 fifth international conference on intelligent systems, modelling and simulation, pp 318–323
- Chattopadhyay K, Acharya A, Banerjee A, Basu J, Konar A (2008) Fast and efficient circle detection schemes for digital image. In: Proceedings of first international conference on emerging trends in engineering and technology, pp 128–133
- Crepinsek M, Liu S, Mernik M (2013) Exploration and exploitation in evolutionary algorithms: a survey. *ACM Comput Surv* 45(3):1–33
- Cuevas E, Ortega-Sanchez N, Zaldivar D, Pérez-Cisneros M (2012a) Circle detection by harmony search optimization. *J Intell Robot Syst* 66:359–376
- Cuevas E, Oliva D, Zaldivar D, Perez-Cisneros M, Sossa H (2012b) Circle detection using electro-magnetism optimization. *Inf Sci* 182(1):40–55

- Cuevas E, Sención-Echauri F, Zaldivar D, Pérez-Cisneros M (2012c) Multi-circle detection on images using artificial bee colony (ABC) optimization. *Soft Comput* 16(2):281–296
- Das S, Dasgupta S, Biswas A, Abraham A (2008) Automatic circle detection on images with annealed differential evolution. In: Proceedings of eighth international conference on hybrid intelligent systems, Sept 2008, pp 684–689
- Dasgupta S, Biswas A, Das S, Abraham A (2008) Automatic circle detection on images with an adaptive bacterial foraging algorithm. In: Proceedings of the 10th annual conference on genetic and evolutionary computation, pp 1695–1696
- Davis J, Goadrich M (2006) The relationship between precision-recall and ROC curves. In: Proceedings of the 23rd international conference on Machine learning, pp 233–240
- Fischler MA, Bolles RC (1981) Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun ACM* 24(6):381–395
- Fu Z, Han Y (2012) A circle detection algorithm based on mathematical morphology and chain code. In: Proceedings of 2012 international conference on computing, measurement, control and sensor network, pp 253–256
- George YM, Zayed HH, Roushdy I, Elbagoury BM (2014) Remote computer-aided breast cancer detection and diagnosis system based on cytological images. *IEEE Syst J* 8(3):949–964
- Isa N, Sabri NM, Jazahanim KS, Taylor NK (2010) Application of the clonal selection algorithm in artificial immune systems for shape recognition. In: Proceedings of 2010 international conference on information retrieval and knowledge management, pp 223–228
- Karkavitsas G, Rangoussi M (2007) Object localization in medical images using genetic algorithms, vol 2. World Academy of Science, Engineering and Technology, Istanbul, pp 499–502
- Kumar RM, Thanushkodi K (2013) Network reconfiguration and restoration in distribution systems through opposition based differential evolution algorithm and PGSA. In: Proceedings of international conference on current trends in engineering and technology, pp 284–290
- Li T, Wang Z-t (2008) Application of plant growth simulation algorithm on solving facility location problem. *Syst Eng Theory Pract* 28(12):107–115
- Lu S, Yu S (2013) A middleware-based model for redundant reader elimination using plant growth simulation algorithm. In: Proceedings of 2013 ninth international conference on computational intelligence and security, pp 36–40
- Luo WQ, Yu JT (2008) Bionic algorithm for solving nonlinear integer programming. *Comput Eng Appl* 44(7):57–59
- Muammar H, Nixon M (1989) Approaches to extending the Hough transform. In: Proceedings of the international conference on acoustics, speech, and signal processing (ICASSP'89), vol 3, pp 1556–1559
- Pan QK, Suganthan PN, Fatih TM, Liang JJ (2010) A self-adaptive global best harmony search algorithm for continuous optimization problems. *Appl Math Comput* 216:830–848
- Pan QK, Suganthan PN, Liang JJ, Fatih Tasgetiren M (2011) A local-best harmony search algorithm with dynamic sub-harmony memories for lot-streaming flow shop scheduling problem. *Expert Syst Appl* 38:3252–3259
- Parikh R, Mathai A, Parikh S, Sekhar GC, Thomas R (2008) Understanding and using sensitivity, specificity and predictive values. *Indian J Ophthalmol* 56(1):45–50
- Seo SW, Kim M (2015) Efficient architecture for circle detection using Hough transform. In: Proceedings of 6th IEEE international conference on information and communication technology convergence, pp 570–572
- Shaked D, Yaron O, Kiryati N (1996) Deriving stopping rules for the probabilistic Hough transform by sequential analysis. *Comput Vis Image Underst* 63(3):512–526
- Tong L, Wang C, Wang W, Su W (2005) A global optimization bionics algorithm for solving integer programming—plant growth simulation algorithm. *Syst Eng Theory Pract* 25:76–85
- Wang C, Cheng HZ (2007a) A plant growth simulation algorithm and its application in power transmission network planning. *Autom Electr Power Syst* 31(7):24–28
- Wang C, Cheng HZ (2007b) Reconfiguration of distribution network based on plant growth simulation algorithm. *Proc CSEE* 27(19):50–55
- Wang M, Chu R (2009) A novel white blood cell detection method based on boundary support vectors. In: Proceedings of IEEE international conference on systems, man and cybernetics (SMC'09), pp 2595–2598, San Antonio, TX, USA
- Wu J, Zeng P, Zhou Y, Olivier C (2006) A novel color image segmentation method and its application to white blood cell image analysis. In: Proceedings of the 8th international conference on signal processing (ICSP'06), pp 235–239
- Xu L, Oja E, Kulhanen P (1990) A new curve detection method: randomized Hough transform (RHT). *Pattern Recognit Lett* 11(5):331–338
- Xu L, Tao M, Ming H (2012) A hybrid algorithm based on genetic algorithm and plant growth simulation algorithm. In: Proceedings of 2012 international conference on measurement, information and control, pp 445–448
- Yadav VK, Batham S, Acharya AK, Paul R (2014) Approach to accurate circle detection: circular hough transform and local maxima concept. In: Proceedings of 2014 international conference on electronics and communication systems (ICECS-2014), pp 489–576
- Zhou X, Ito Y, Nakano K (2014) An efficient implementation of the one-dimensional Hough transform algorithm for circle detection on the FPGA. In: Proceedings of 2014 second international symposium on computing and networking, pp 44–452