How to Cite:
Veena, R. S., Shravan, S., Biradar, V. G., & Prashanth, V. (2022). Circle fitting of boundaries of microarray spots. *International Journal of Health Sciences*, *6*(S1), 4736–4745.
https://doi.org/10.53730/ijhs.v6nS1.6023

Circle fitting of boundaries of microarray spots

Dr. Veena R. S.
Assistant Professor, Department of Computer Science and Engineering, RV Institute of Technology and Management, Bengaluru, India
Email: veenars1975@gmail.com

Dr. Shalini Shravan
Assistant Professor, Department of Electronics and Communication Engineering, RV Institute of Technology and Management, Bengaluru, India
Email: shalini.iy@gmail.com

Dr. Vidyadevi G. Biradar
Associate Professor, Department of Information Science and Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, India
Email: vidyadevi.g.biradar@nmit.ac.in

Dr. Prashanth V
Associate Professor, Department of Electrical and Electronics Engineering, Nitte Meenakshi Institute of Technology, Bengaluru, India
Email: prasanth.v@nmit.ac.in

***Abstract*---**A new circle fitting algorithm is presented that generates the best circle to approximate the boundary of a given microarray spot. Here, the best approximation means, the total differential area between the circumference of the circle and the boundary of the spot is kept minimum. The center and the radius of the circle are determined to satisfy this condition. The differential area is used as a quantitative metric to represent the boundary quality of the corresponding microarray spot.

***Keywords*---**Circle fitting boundaries, microarray spots, Nearest Approximate Circle (NAC).

Manuscript submitted: 27 Feb 2022, Manuscript revised: 18 March 2022, Accepted for publication: 09 April 2022

Introduction

Boundary shapes of Microarray Spots (MAS) reveal their spatial distortion. The boundary of an ideal spot appears as shown in Fig. 1(a) whereas that of a distorted spot looks as shown in Fig. 1(b).

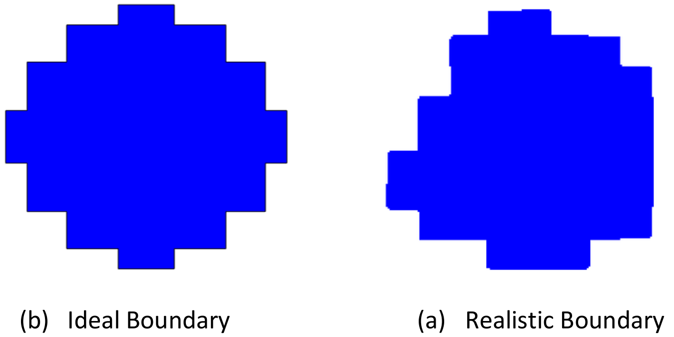
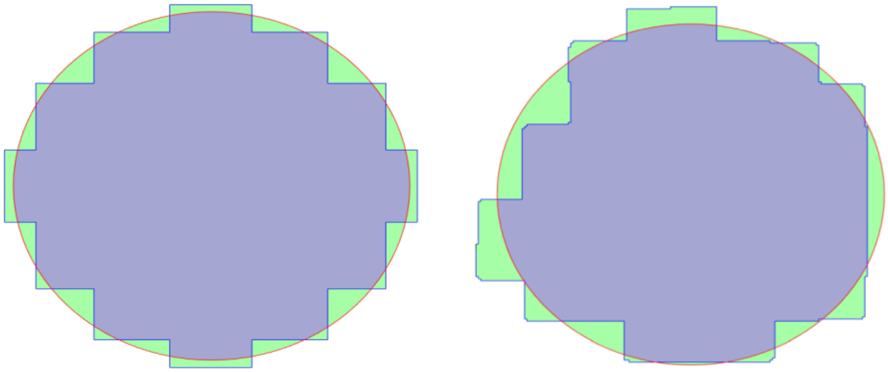


Fig. 1. Ideal and Realistic Boundaries of a Microarray Spot

The boundary of an ideal spot has horizontal as well as vertical symmetry and the boundary formation is uniform throughout as shown in Fig. 1(a). This criterion is not satisfied by distorted spots. For better characterization of the distortion in spot boundaries, we propose a circle fitting scheme and an associated metric to measure the quality of the spot boundary. The proposed optimized fitting circle is designated as the *Nearest Approximate Circle*.

Nearest Approximate Circle

Nearest Approximate Circle (NAC) is that circle whose circumference is the best approximation of the boundary of the given spot. Here, the best approximation means, the total differential area between the circumference of the NAC and the boundary of the spot is kept minimum. The NAC's of an ideal spot and a distorted spot are shown in Fig. 2(a) and 2(b) respectively. The circumferences of the circles are shown in red and the spot boundaries are shown in blue. The differential areas are shown in green.



(a) NAC Ideal Spot (b) NAC Distorted Spot

Fig. 2. Nearest Approximate Circles for Ideal and Distorted Spots

Differential Area

From Fig. 2, it can be seen that the NAC circumference intersects the spot boundary at several points. Between consecutive intersection points, the arc connecting them and the spot boundary covered by this arc, form a closed region. The area of this closed region is the differential sub-area. The total of all the differential sub-areas encompassing all the intersection points is the differential area (DA). Since the value of DA for a spot depends on the scale and the unit of representation, it is normalized to get the *differential area ratio*.

Differential Area Ratio as a Boundary Distortion Metric

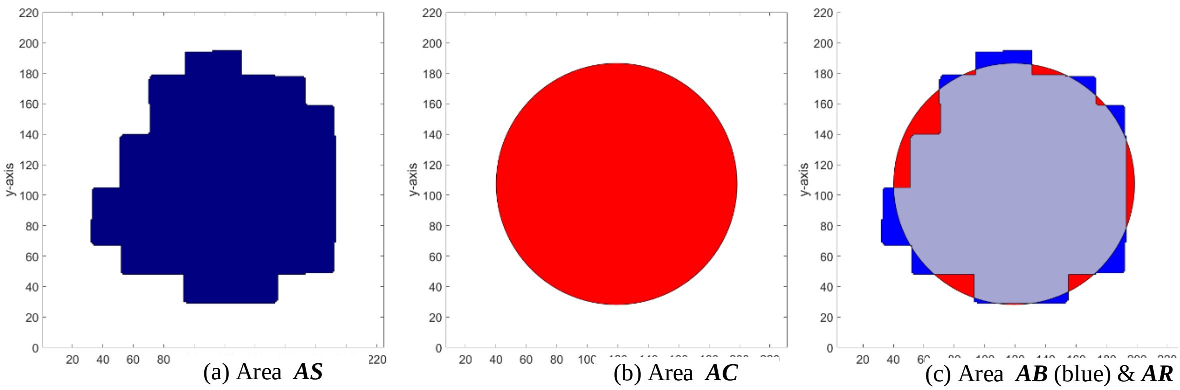
Differential area is normalized with respect to the area of the NAC. The normalized DA value is designated by DAR which is defined as,

$$Differential\ Area\ Ratio = DAR = \frac{DA}{Area\ of\ the\ NAC} \quad (1)$$

Smaller the DAR, better is the quality of the spot boundary. Higher the distortion level of a microarray spot, higher is the value of DAR. Thus DAR is a **Boundary Distortion Metric** based on circle fitting.

Areas as Sets and the Differential Area

Let the segmented microarray spot in its matrix format be denoted by *spotImage* which is a binary image. Let symbol ***AS*** represent the area (blue) of the given *spotImage*, as a collection of the pixels in terms of their (x, y) coordinates as shown in Fig. 3(a). Similarly, let ***AC*** represent the circular area (red) of the NAC, of that spot, as a collection of the (x, y) coordinates of the pixels enclosed by it as shown in Fig. 3(b). Here ***AS*** and ***AC*** are two sets of pixels.



(a) Area ***AS*** (b) Area ***AC*** (c) Area ***AB*** (blue) & ***AR*** (red)

Fig. 3. Areas ***AS***, ***AC*** and the differential area

Now, let us consider the set subtraction {***AS*** – ***AC***} represented by the set ***AB*** as,

$$\boldsymbol{AB} = \{\boldsymbol{AS} - \boldsymbol{AC}\} \quad (2)$$

Area ***AB*** is shown in Fig. 4(c) in blue. Similarly, the set subtraction {***AS*** – ***AC***} is represented by the set ***AR*** as,

$$\boldsymbol{AR} = \{\boldsymbol{AC} - \boldsymbol{AS}\} \quad (3)$$

Area ***AR*** is shown in Fig. 3(c) in red. Now, the total differential area is given by the union of AB and AR as,

$$\boldsymbol{DA} = \boldsymbol{AB} \cup \boldsymbol{AR} \quad (4)$$

The resulting area ***DA*** is marked in red and blue in Fig. 3(c). In (4), ***DA*** is the set of pixels represented by their (x, y) coordinates. Let *area**DA*** be the scalar that represents the number of pixels in DA or $|\boldsymbol{DA}|$. Thus, *area**DA*** is the quantitative representation of ***DA***.

Objective

Our objective is to determine the circle ***AC*** for a given microarray spot, such that the *area**DA*** is minimum. Then that circle is the nearest fit circle for that microarray spot whose area is represented by ***AS***. Such a circle is called the Nearest Approximate Circle (NAC) of that spot. Minimization of *area**DA*** is achieved using a multi-variable non-linear function minimization solver. Minimum areaDA is essentially quantitative metric to represent the quality of the boundary line of an already segmented spot.

Related work

ScanAlyze software developed by Eisen (1999) uses fixed circles for microarray spot segmentation. Here, it is assumed that all the spots of the microarray have the same radius. In practice, constant radius assumption is not true and hence, the next modification is the adaptive circle fitting used in GenePix software developed by Axon Instruments (1999). This method also assumes that the spots have circular shape but their radii are different. This approach is more representative of the actual nature of the microarray spots. Thus the adaptive circle fitting provides higher accuracy compared to the fixed circle method. In [1], the authors have used Circular Hough Transformation (CHT) to represent the segmented spots. CHT provides dissimilar circularly segmented spots according to the actual sizes and locations of the spots. But, in reality, the actual shapes of the segments need not be circular. Therefore, non-circular and more accurate shapes are adopted to represent the spot boundaries. In [2], integro-differential operators have been used for accurate spot segmentation. This method combines both the contour line based as well as the region based approaches. Several non-circular methods [3-9] like watershed algorithm, seed region growing, histogram based segmentation, cluster based segmentation, *etc.*, are available. Since our main objective is circle fitting of the given spot and to get its DAR metric, the above mentioned methods are not discussed in this section.

Determination of the nearest approximate circle

Let us designate the algorithm that determines the NAC of a given spot by 'GET-NAC'. Let the (x-y) coordinates of the center of NAC be denoted by (u, v) and its radius by *r*. Thus the circle is specified by three parameters, (u, v, r) in the Cartesian system. This circle is denoted as *circle*(*u, v, r*). In GET-NAC algorithm, for the given spot, (*u, v, r*) are determined using **fminsearch(...)[ref1]** such that the differential area between the *circle*(*u, v, r*) and given spot is minimum. Then the optimized circle(u, v, r) is the NAC of that spot. **fminsearch(...)**is a nonlinear programming solver which finds the minimum of unconstrained nonlinear multivariable function using derivative-free method. **fminsearch(...)**is formulated as [10]**,**

$$\boldsymbol{x} = \text{fminsearch}(\text{fun}, \boldsymbol{x0}) \qquad (5)$$

In our case, ***x*** is a three element vector [x(1), x(2), x(3)] that represents (u, v, r). Therefore, x(1) = u, x(2) = v and x(3) = r. The solver minimizes the return value (output) of the nonlinear function 'fun' specified in (5). The input parameters to the 'fun' are the decision variable vector *x* and the given microarray spot image, The 'fun' is to be scripted such that its output is the differential area (areaDA) that is to be minimized. Therefore the 'fun' is formulated as,

$$area\boldsymbol{DA} = \text{fun}(\boldsymbol{x}, spotImage) \qquad (6)$$

In (5), x0 specifies the initial value of x. Therefore, x0 = (u0, v0, r0) which are the initial values of (u, v, r). Since NAC is the nearest circle approximating the *spotImage,* the initial *x0* = (u0, v0, r0) values are selected based on the *spotImage* to achieve faster result while using fminsearch(fun,***x0***). Let the spot image be as shown in Fig. 4.

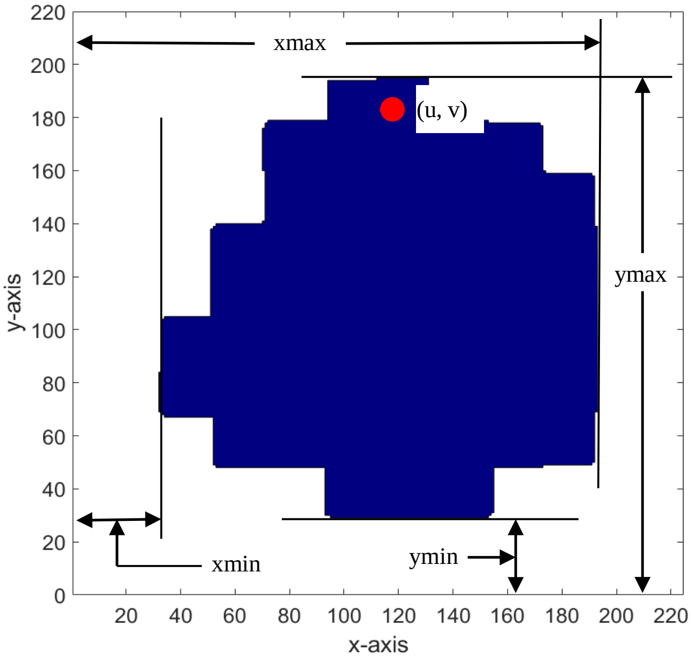


Fig. 4. Extremum values of the spotImage

The extremum values of the spot, [xmin, ymin, xmax, ymax] are marked in Fig. 4
The initial x and y coordinates of the center are selected as,

$$x0(1) = u0 = (xmin+xmax)/2 \quad (7)$$

$$x0(2) = v0 = (ymin+ymax)/2 \quad (8)$$

Now, (u0, v0) is at the center of the spot image. The initial radius r0 is taken as,

$$x0(3) = r0 = (ymax-ymin)/2 \quad (9)$$

(It is found that, x0(3) = r0 = (xmax–xmin)/2 would also give the correct optimum result).
The set ***AS***, which is the collection of the coordinates of the pixels which cover the *spotImage* can be obtained using the ***find***(…) function as,

[**row, col**] = *find*(*spotImage*) (10)

Here, [**row, col**] are the row and column subscripts of each non-zero element of *spotImage* matrix. In terms of (x, y) coordinates, **col** vector gives the collection of x-coordinates and **row** vector gives the collection of y-coordinates in inverted order. To get the straight order for y-coordinates, we can use,

$$[\boldsymbol{XS}, \boldsymbol{YS}] = find\big(flipud(spotImage)\big) \quad (11)$$

In (11), function *flipud* (…) flips the spotImage in up-down direction. Now, vector ***YS*** gives the collection of y-coordinates in the normal order so that the pixel co-ordinates are in the first-quadrant of the (x-y) coordinate system. Similarly, for a circle with given (u, v, r) the filled circle can be drawn [11] and then its ***AC*** can be obtained using the ***find***(…) function.

Function fun(*x*, *spotImage*)

The the pseudo-code for function *areaDA* = fun(***x***, *spotImage*) is formulated as follows.

----

function *areaDA* = fun(*x*, *spotImage*)

----

1. Set u = x(1), v = x(2), r = x(3).
2. Generate the circle with (u, v, r).
3. Get ***AC***, the set of coordinates of the pixels which cover the circle.
4. Place the *spotImage* inside a suitable axis.
5. Get ***AS***, the set of coordinates of the pixels which cover the *spotImage* using (11).
6. Get set ***AB*** = {***AS*** – ***AC***} // Set subtraction
7. Get set ***AR***= {***AC*** – ***AS***} // Set subtraction
8. Get set ***DA*** = ***AB*** ∪ ***AR*** // Set Union
9. Get *areaDA* = **|*DA*|** = the number of elements in set ***DA***.
10. Return *areaDA*

GET-NAC algorithm uses fminsearch(fun,x0)where x0 = [x0(1), x0(2), x0(3)] which are given by (7), (8), (9) and the 'fun', as described in section 2.1 that returns *areaDA*. The GET-NAC algorithm is given below.

Algorithm GET-NAC

Algorithm GET-NAC

Input: correctly segmented *spotImage* which is a binary image.

Output: NAC represented by x = [x(1), x(2), x(3)] = [u, v, r].

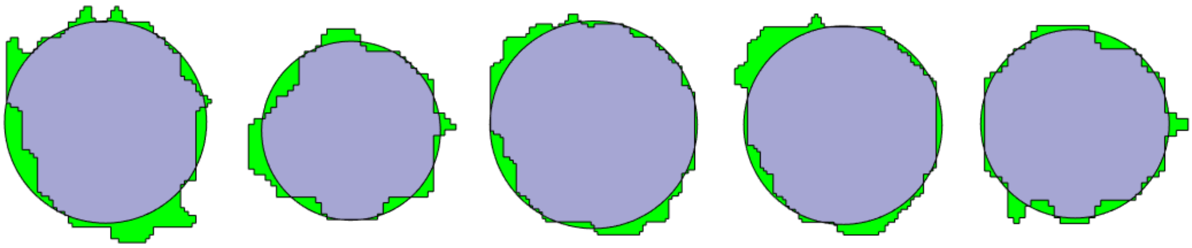
1. Get the vector ***x0*** = [x0(1), x0(2), x0(3)] as given by (7), (8), (9) from the *spotImage*.
2. Keep the function *areaDA* = *fun*(***x***, *spotImage*) ready.
3. Execute the solver function, ***x*** = fminsearch(fun,***x0***)that minimizes *areaDA*.
4. From ***x*** = [x(1), x(2), x(3)], get [u, v, r] as, [u, v, r] = [x(1), x(2), x(3)]
5. Get NAC from the circle specified by parameters (u, v, r).
6. Over.

Differential Area Ratio

Once the NAC of a *spotImage* is obtained, ***DA*** is obtained using (2), (3) and (4). Then, *areaDA* is obtained by counting the elements of DA as |DA| and the DAR is obtained as specified by (1). In (1), the area of NAC is given by $\pi * r^2$.

Experimental Results
Example 1

Example 1: Here, five different MA image spots, which are in binary matrix form, are resized to [100, 100] and then approximated by their NAC's using GET_NAC, as shown in Fig. 5.



(a) DAR = 0.1856 (b) DAR = 0.1532 (c) DAR = 0.1090 (d) DAR = 0.1365 (e) DAR = 0.1072

Fig. 5. Five different MA spots and their DAR's

The calculated values of areaDA, area of NAC and DAR for these spots are shown in shown in Table 1. From Table 1, it can be seen that spot (e) has the lowest DAR.

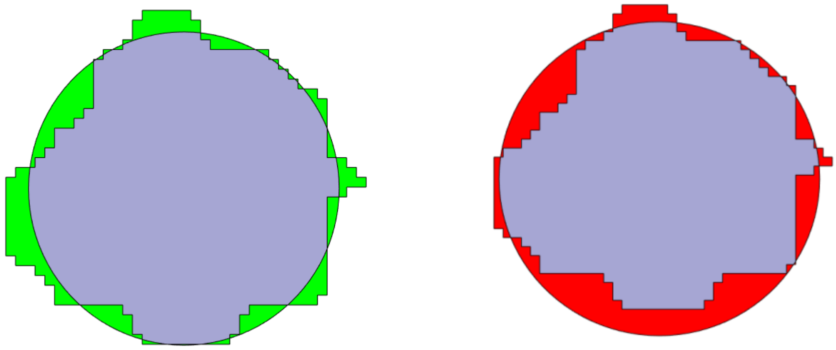
Table 1
areaDA, area NEC and DAR are for five dissimilar spots

	Spot (a)	Spot (b)	Spot (c)	Spot (d)	Spot (e)
areaDA	397.702	122.305	349.884	411.833	88.356
Area NAC	2142.8	798.5	3209.3	3016.0	823.9
DAR	0.1856	0.1532	0.1090	0.1365	0.1072

Comparison with Circular Hough Transformation Fitting

Circular Hough Transformation (CHT) method [12] can be used to approximate the segmented spot with a circular fit using the function imfindcircles(…)[13]. Here we compare the circle fittings by our method, GET-NAC with imfindcircles(…)method.

Example 2: Here, spot 5(b) is used for comparing GET-NAC and CHT methods of circle fitting. The resulting circles and differential areas are shown in Fig. 6.



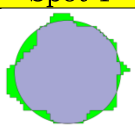
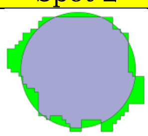
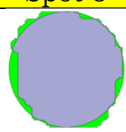
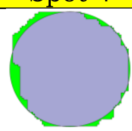
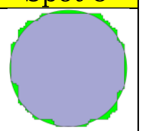
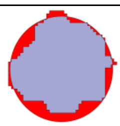
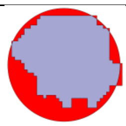
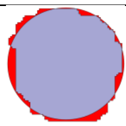
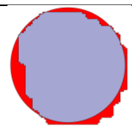
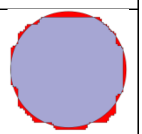
(a) GET_NAC Method
DAR = 0.1532

(b) CHT Method
DAR = 0.2024

Fig. 6. Comparison view of circles obtained using GET_NAC and CHT methods

From Fig. 6, it can be seen that the radius of CHT circle is slightly bigger than that of GET-NAC circle. Correspondingly, the DAR(CHT) is greater than DAR(GET-NAC). Therefore we can say that GET-NAC is superior to CHT method for the circular representation of MA spots. Table 2 gives the different parameters (in proper units) of circle fitting with respect to GET-NAC and CHT methods for the five dissimilar spots.

Table 2
Different parameters (in proper units) of circle fitting using GET-NAC and CHT methods

		Spot 1	Spot 2	Spot 3	Spot 4	Spot 5
Fitting Circle	GET_NAC					
	CHT					
areaDA	GET_NAC	122.3	137.4	347.0	309.2	216.8
	CHT		301.1	357.5	522.5	266.4
Circle	GET_NAC	15.90	14.70	32.60	32.70	32.50

Radius	CHT	17.50	17.80	32.30	33.20	32.52
DAR	GET_NAC	0.1532	0.2033	0.1042	0.0920	0.0654
	CHT	0.2024	0.3034	0.1091	0.1509	0.0802

From Table 2, we see that GET_NAC method performs better than CHT method in terms of DAR. This means GET-NAC circle fits the spot more tightly than CHT circle. When the spot itself is nearly circular, then also, GET-NAC method is slightly better than CHT method as seen in the last column of Table. 2.

Conclusion

A new method of circle fitting for microarray spots has been proposed. The generated circle is the nearest approximation to the spot boundary and the differential area between the circle and the spot boundary represents the quality of the spot boundary. The normalized differential area denoted by *differential area ratio* is the metric to represent the distortion in spot boundaries. Smaller the *differential area ratio* lesser is the spot boundary distortion. Nonlinear optimization solver is used to get the optimal circle that results in minimum *differential area ratio*. Minimized *differential area ratio* can be used to compare the DNA microarray spots of healthy and unhealthy tissues. Compared to the Circular Hough Transformation method, the proposed method is found to be superior in tightly fitting the microarray spot by the optimized circle.

References

[1] Farouk RM, SayedElahl MA. (2015). “Robust cDNA microarray image segmentation and analysis technique based on Hough circle transform,” Human Frontier Science Program, HFSP journal, part.1, ISSN:1955-2068,Vol.9, Issue.7, 2015. pp. 1-13.

[2] R. M. Farouk and M. A. SayedElahl, Microarray spot segmentation algorithm based on integro-differential operator, Egyptian Informatics Journal, https://doi.org/10.1016/j.eij.2019.04.001.

[3] R. Karim, M. Khaliluzzaman and S. Mahmud, "A review of image analysis techniques for gene spot identification in cDNA Microarray images," The 2nd International Conference on Next Generation Information Technology, 2011, pp. 36-41.

[4] Giannakeas, Nikolaos & Fotiadis, Dimitrios. (2009), “Image processing and machine learning techniques for the segmentation of cDNA microarray images,” Handbook of Research on Advanced Techniques in Diagnostic Imaging and Biomedical Applications. 294-306. 10.4018/978-1-60566-314-2.ch019.

[5] V. Uslan and İ. Ö. Bucak, "Clustering-based spot segmentation of cDNA microarray images," 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, 2010, pp. 1828-1831, doi: 10.1109/IEMBS.2010.5626430.

[6] D. Maroulis and E. Zacharia, "Microarray image segmentation using spot morphological model," 2009 9th International Conference on Information Technology and Applications in Biomedicine, 2009, pp. 1-4, doi: 10.1109/ITAB.2009.5394330.

[7] H. Yijun and W. Guirong, "Segmentation of cDNA Microarray Spots Using K-means Clustering Algorithm and Mathematical Morphology," 2009 WASE International Conference on Information Engineering, 2009, pp. 110-113, doi: 10.1109/ICIE.2009.17.

[8] P. Ye and G. Weng, "Microarray Image Segmentation Using Region Growing Algorithm and Mathematical Morphology," 2009 Fifth International Conference on Information Assurance and Security, 2009, pp. 373-376, doi: 10.1109/IAS.2009.188.

[9] J. Ho and W. Hwang, "Segmenting Microarray Image Spots using an Active Contour Approach," 2007 IEEE International Conference on Image Processing, 2007, pp. VI - 273-VI - 276, doi: 10.1109/ICIP.2007.4379574.

[10] https://www.mathworks.com/help/matlab/ref/fminsearch.html.

[11] https://www.mathworks.com/matlabcentral/answers/87111-how-can-i-draw-a-filled-circle.

[12] H.K Yuen, J. Princen, J. Illingworth, and J. Kittler. "Comparative study of Hough transform methods for circle finding." Image and Vision Computing. Volume 8, Number 1, 1990, pp. 71–77. https://www.mathworks.com/help/images/ref/imfindcircles.html