Preparing Initial Population of Genetic Algorithm for Region Growing Parameter Optimization

Sándor Szénási*, Zoltán Vámossy**, Miklós Kozlovszky***
* Óbuda University, Doctoral School of Applied Informatics, Budapest, Hungary
** Óbuda University, John von Neumann Faculty of Informatics, Budapest, Hungary
*** MTA SZTAKI, Laboratory of Parallel and Distributed Computing, Budapest, Hungary
szenasi.sandor@nik.uni-obuda.hu; vamossy.zoltan@nik.uni-obuda.hu; m.kozlovszky@sztaki.hu

***Abstract*— The processing of microscopic tissue images is nowadays done more and more using special immunodiagnostic-evaluation software products. Often to evaluate the samples, the first step is determining the number and location of cell nuclei. To do this, one of the most promising methods is the region growing, but this algorithm is very sensitive to the appropriate setting of different parameters. Due to the large number of parameters and due to the big set of possible values setting those parameters manually is a quite hard task, so we developed a genetic algorithm to optimize these values. The first step of the development is the statistical analysis of the parameters, and the determination of the important features, to extract valuable information for a to-be-implemented genetic algorithm that will perform the optimization.**

Keywords: biomedical image processing, nuclei detection, region growing, optimization, genetic algorithm

I. Determining the parameter-intervals required for the optimization

The main steps of the region-growing approach [1] (pre-processing, region growing, post-processing, separating the merged cells) require quite a lot (in our implementation [2], altogether 34) parameters. The sets of possible values vary a lot for the different parameters, but it is obvious that the number of possible variations is so high, that it is hopeless to find the optimal parameters using a manual approach. Due to this, we developed a genetic algorithm [3], that will tries to find the optimal (or very close to optimal) parameter set.

For the genetic algorithm we need an initial population, and these individual elements have to be filled up with a starter set of parameters [4], so we need an approximation for every parameter, or at least we have to be able to set an interval to examine. Some of the parameters are downright technical, these values can be guessed at best or later fine-tuned using empirical methods (these parameters are for example the window sizes used by the different filters). However, other parameters can be easily measured, so we can establish exact intervals for them using the Gold Standard slides.

For the examination, we prepared an application that works according to the following algorithm:

1. Open the next Gold Standard slide.
2. Load the slide annotations. The first of these identifies the region annotated by the doctors (shape with blue contour in Figure 1.).
3. Use this mask for the whole image. The processing of the areas outside this mask is not necessary.

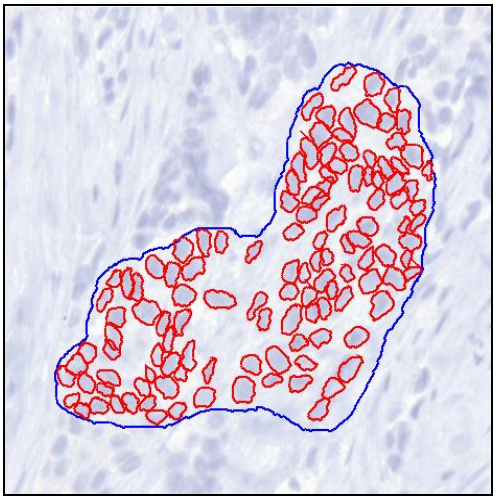

Figure 1. Manually annotated tissue image (shape with blue contour – annotated region; shapes with red contour – cell nuclei)

4. The remaining annotations (shapes with red contour in Figure 1.) mark the cell nuclei marked by the doctors. These areas must be individually examined according to the criteria listed above.
5. Save the gathered statistical data into a database [5].
6. Using the database, the necessary boundaries for the optimization of parameters can be easily extracted.

According to this, using the cell nuclei marked by the doctors, we can identify the minimum and maximum values for the above listed parameters; so we have to search the optimal parameter values in the resulting intervals. Ideally, we would like to have a set of parameters that we can use to get a 100% positive hit rate, but this (using the current region-growing method) is probably not possible. In addition to this, setting the parameters close to the boundaries greatly increases the ratio of false detections, so we will only examine the values within a smaller interval.

For this reason, near the minimum and maximum values we also examine the distribution of the parameters, and we (using another little program of ours) search for the shortest intervals that contain 99.5%, 99%, 95%, 90% and 80% of all the elements. During the optimization, we will work only with these narrowed intervals, which will of course prevent us from finding the perfect solution, but in return, it will greatly speed up the process of searching the optimal parameters.

II. Analysis of the parameters of the region growing algorithm

A. *Examination of the size*

During the region-growing algorithm, one of the most important features is the size of the cell nuclei. It plays an important role several times during the execution of the algorithm:

- During the actual growing of regions, the size of the cell nuclei candidate will grow with one pixel after every new iteration. We can set several limitations for this growth; one of these is that the nucleus itself cannot be bigger than a certain limit. So if the size reaches this limit during the region growing, then the growing loop has to be stopped.
- The size of the cell nuclei is also an important parameter during the post-processing verification that is performed after the end of the region growing. After the region-growing algorithm has determined the possible candidates for cell nuclei, these candidates have to be verified against different necessary conditions. One of these conditions is the minimal size of a cell nucleus (every candidate having less pixels than this limit is rejected) or the maximal size of a cell nucleus (every candidate having more pixels than this limit is also rejected, but due to the aforementioned limitation inside the region growing loop, this case will never happen).

To examine the sizes of the cell nuclei, we used the available Gold Standard slides, and then we selected all annotated nuclei from them. Using the nuclei as a mask, we copied that region to an empty image, and then we counted the number of non-background pixels, thus we got the sizes of the individual cell nuclei.

#	Interval from	Interval to	Number of nuclei	Ratio
1	22.00	124.80	499	6.33%
2	124.80	227.60	2408	30.52%
3	227.60	330.40	2599	32.94%
4	330.40	433.20	1454	18.43%
5	433.20	536.00	512	6.49%
6	536.00	638.80	214	2.71%
7	638.80	741.60	87	1.10%
8	741.60	844.40	59	0.75%
9	844.40	947.20	33	0.42%
10	947.20	1050.00	14	0.18%
11	1050.00	1152.80	5	0.06%
12	1152.80	1255.60	2	0.03%
13	1255.60	1358.40	1	0.01%
14	1358.40	1461.20	0	0.00%
15	1461.20	1564.00	1	0.01%
16	1564.00	1666.80	0	0.00%
17	1666.80	1769.60	0	0.00%
18	1769.60	1872.40	0	0.00%
19	1872.40	1975.20	0	0.00%
20	1975.20	2079.00	1	0.01%

Table 1. – Number of cells by size

We saved this number for every cell nuclei, and based on the 7554 cell nuclei in the database, we determined the following boundaries:

- Minimum size: 22 pixels
- Maximum size: 2079 pixels

Distribution of the cell nuclei according to their sizes is in Table 1.

It is clearly visible from Table 1., that the area of the biggest cell nucleus is indeed 2079 pixels, but this is far away from the average value. After displaying the data in a graph (Figure 2.), it is visible that the distribution is far from the normal distribution:

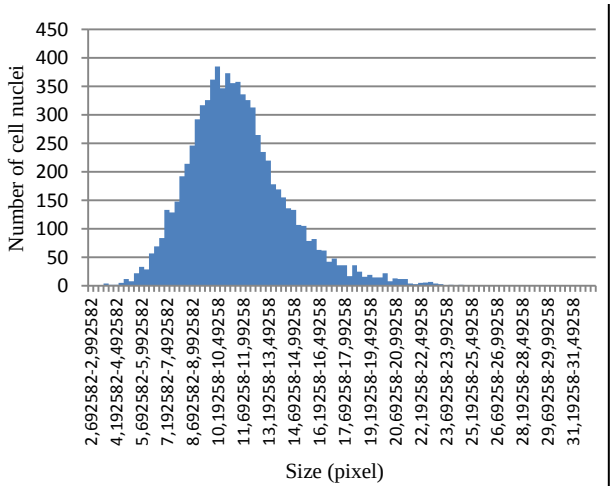

Figure 2. - Number of cells by size

Regarding the sizes, 99.5% of the cell nuclei are between 34 and 882 pixels, so the optimization will use these values as interval boundaries (Table 2.).

Coverage	Interval from	Interval to	Interval length	Length/ (Max-Min)
100.0%	22.00	2079.00	2057.00	100.00%
99.5%	34.00	882.00	848.00	41.23%
99.0%	34.00	801.00	767.00	37.29%
95.0%	65.00	556.00	491.00	23.87%
90.0%	90.00	475.00	385.00	18.72%
80.0%	108.00	397.00	289.00	14.05%

Table 2. – Number of cells by size

B. *Examination of the radius*

The area measured in pixels can only give us information about the size of the recognized object, but the radius will give us (on some level) some hints about its shape as well. For this reason, it is practical to examine the radius of the cell nuclei as well; this can be useful for the region growing in the following cases:

- If during the region growing we reach a given maximal radius, then we can immediately stop the algorithm. Due to the calculation method of the radius it is possible that the radius will decrease during future calculations (in case of the centre of the current region moves), but in the practical approach this will not mean a problem, since it is far more important to close down the unnecessary region growing steps.

- After the region growing has stopped, the radius of the region will also be an important parameter, since this can also be used as a verification condition during the acceptance test of the cell. We introduced a minimal radius value (cell nuclei smaller than this radius will be rejected) and a maximal radius value as well (thought due to the condition that is built into the region growing itself, this radius value will never be reached).

To measure the radius values of the cell nuclei, we took the available Gold Standard slides, and we collected all the annotated cell nuclei. Then the radius was determined with the following steps (where $P[i]$ – ith point of the region; N – number of points of the region):

1. We calculated the centre of mass for the cell:
   $Center_X = \sum P[i]_X / N$
   $Center_Y = \sum P[i]_Y / N$
2. We calculated all the distances between this centre and all pixels of the cell:
   $Distance_i = Distance(Center, P[i])$
3. Out of these distances, we chose the biggest one, and this is the radius of the object:
   $Radius = MAX(Distance_i)$

We saved this number for every cell nuclei, and based on the 7554 cell nuclei in the database, we determined the following boundaries:

- Minimum radius: X pixels
- Maximum radius: X pixels

Distribution of the cell nuclei according to their radius is in Table 3. and Figure 3.

#	Interval from	Interval to	Number of nuclei	Ratio
1	2.69	4.19	8	0.10%
2	4.19	5.69	49	0.62%
3	5.69	7.19	272	3.45%
4	7.19	8.69	816	10.34%
5	8.69	10.19	1543	19.56%
6	10.19	11.69	1819	23.06%
7	11.69	13.19	1475	18.70%
8	13.19	14.69	858	10.88%
9	14.69	16.19	506	6.41%
10	16.19	17.69	251	3.18%
11	17.69	19.19	130	1.65%
12	19.19	20.69	79	1.00%
13	20.69	22.19	44	0.56%
14	22.19	23.69	25	0.32%
15	23.69	25.19	6	0.08%
16	25.19	26.69	2	0.03%
17	26.69	28.19	3	0.04%
18	28.19	29.69	2	0.03%
19	29.69	31.19	0	0.00%
20	31.19	31.93	1	0.01%

Table 3. – Number of cells by radius

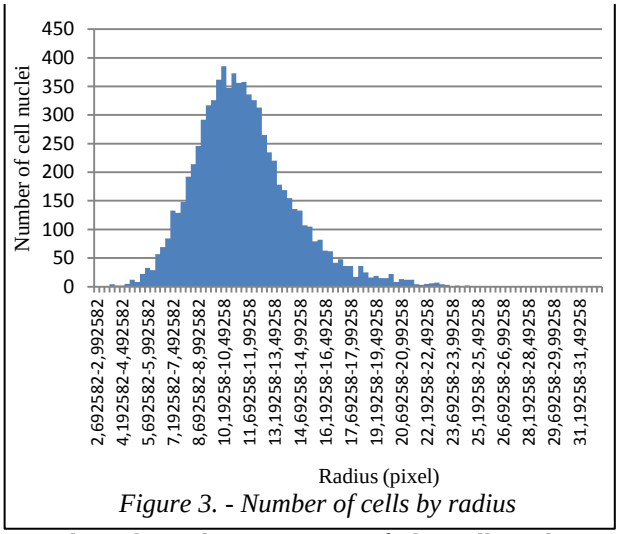

Figure 3. - Number of cells by radius

Regarding the radiuses, 99.5% of the cell nuclei are between 4.80 and 22.83 pixels, so the optimization will use these values as interval boundaries (Table 4.).

Coverage	Interval from	Interval to	Interval length	Length/ (Max-Min)
100.0%	2.69	31.93	29.24	100.00%
99.5%	4.80	22.83	18.03	61.67%
99.0%	5.33	21.57	16.24	55.55%
95.0%	6.29	17.78	11.49	39.30%
90.0%	6.80	16.06	9.26	31.66%
80.0%	7.49	14.49	7.00	23.94%

Table 4. – Number of cells by radius

C. Examination of the circularity

The circularity of the individual cell nuclei is also important information for the search algorithm. Due to their nature, the cell nuclei are usually circle-like objects, which is an important practical feature, because it is important that despite the various size and colour conditions, the mostly round region should be built during the region growing process. To ensure this, the following rules are applied:

- During the region growing, when adding a new point to the region, the utilized fitness function considers the location of the new pixel as well. The closer the potential pixel is to the current centre of the cell, the bigger its fitness value will be. This way, the algorithm will automatically build up as circle-like objects as possible.
- After the region growing has stopped, the circularity is still an important feature, because it is also used during the verification. As a parameter, we can define a lower limit: the objects that are less circle-like will be rejected, and the objects that are more circle-like will be accepted as cell nuclei.

To measure the circularity values of the cell nuclei, we took the available Gold Standard slides, and we collected all the annotated cell nuclei. Then the circularity was determined with the following steps (where $P[i]$ – ith point of the region; N – number of points of the region):

1. We calculated the centre of mass for the cell:
   $Center_X = \sum P[i]_X / N$
   $Center_Y = \sum P[i]_Y / N$

2. We calculated all the distances between this centre and all pixels of the cell:

   $Distance_i = Distance(Center, P[i])$

3. It can be generally said that for a given area, the smaller this sum distance is, the more circle-like the object will be. According to this, we summed up the calculated distances, and then we scaled the sum with the total area of the cell nucleus.

   $Circularity = \sum Distance_i / N$

The resulting value cannot be considered as a perfect measurement number for circularity, there are a lot of even more methods (like circularity measures in [6]), but since this calculation has to be executed at every iteration of the region growing, it is a fairly time critical operation. During the parameter optimization, the execution time is not as critically important factor as it is during the actual detection, but it is practical to use the same calculation method that will be used during the region growing (Table 5., Figure 4.).

- Minimum circularity: 17.32
- Maximum circularity: 103.2

#	Interval from	Interval to	Number of nuclei	Ratio
1	17.32	21.62	4	0.05%
2	21.62	25.92	13	0.16%
3	25.92	30.22	26	0.33%
4	30.22	34.52	51	0.65%
5	34.52	38.82	100	1.27%
6	38.82	43.12	178	2.26%
7	43.12	47.42	296	3.75%
8	47.42	51.72	415	5.26%
9	51.72	56.02	546	6.92%
10	56.02	60.32	666	8.44%
11	60.32	64.62	792	10.04%
12	64.62	68.92	933	11.83%
13	68.92	73.22	917	11.62%
14	73.22	77.52	870	11.03%
15	77.52	81.82	804	10.19%
16	81.82	86.12	658	8.34%
17	86.12	90.42	396	5.02%
18	90.42	94.72	178	2.26%
19	94.72	99.02	39	0.49%
20	99.02	103.02	7	0.09%

Table 5. – Number of cells by circularity

Naturally, in this case as well, the distribution is not completely normal, but it is visible on the diagram that it is close to normal. Since there are quite a few cell nuclei that differ a lot from the average, it is also practical (to speed up the optimization) to narrow down the inspected interval:

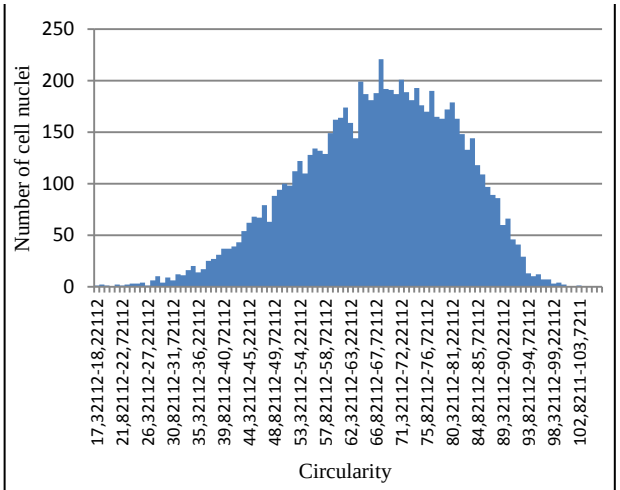

Figure 4. - Number of cells by circularity

Regarding the circularity, 99.5% of the cell nuclei are between 27.66 and 97.10 values, so the optimization will use these values as interval boundaries (Table 6.).

Coverage	Interval from	Interval to	Interval length	Length/ (Max-Min)
100.0%	17.32	103.02	85.70	100.00%
99.5%	27.66	97.10	69.44	81.04%
99.0%	31.72	96.26	64.54	75.31%
95.0%	42.18	93.14	50.96	59.47%
90.0%	46.74	90.99	44.25	51.64%
80.0%	50.96	86.44	35.49	41.41%

Table 6. – Number of cells by circularity

During the optimization, we only have to find the lower limit of the circularity, because the conditions of the cell nuclei detection algorithm are only applied to that one.

D. Examination of the average intensity

Usually we assume that the cell nuclei will appear on the screen as a group of pixels with higher intensity. Because of this, it is practical to introduce a limit that helps us in deciding whether a group of pixels suits this criterion or not. During the post-processing phase after the region growing, there is a step that examines the average intensity of the candidate nucleus' pixels, and then this average is compared with a minimum limit to decide whether the given candidate can be accepted or not (there is no need for a maximum limit). Setting the minimum limit to a too low value would increase the rate of false positive decisions, and a too high value would cause the rejection of candidates that are indeed real cell nuclei.

To measure this parameter, we once again have to examine all annotated cell nuclei in the Gold Standard slides. During the pixel-level examination, we have to check what intensity value an inner pixel has, and then the average of these values will give us the average intensity of the nucleus.

- Minimum average intensity: 26.01
- Maximum average intensity: 214.87

Distribution of the cell nuclei according to their average intensities is in Table 7. and Figure 5.

#	Interval from	Interval to	Number of nuclei	Ratio
1	26.01	35.41	15	0.19%
2	35.41	44.81	127	1.61%
3	44.81	54.21	195	2.47%
4	54.21	63.61	375	4.75%
5	63.61	73.01	703	8.91%
6	73.01	82.41	900	11.41%
7	82.41	91.81	850	10.77%
8	91.81	101.21	636	8.06%
9	101.21	110.61	556	7.05%
10	110.61	120.01	468	5.93%
11	120.01	129.41	408	5.17%
12	129.41	138.81	418	5.30%
13	138.81	148.21	476	6.03%
14	148.21	157.61	593	7.52%
15	157.61	167.01	450	5.70%
16	167.01	176.41	304	3.85%
17	176.41	185.81	181	2.29%
18	185.81	195.21	126	1.60%
19	195.21	204.61	81	1.03%
20	204.61	214.87	27	0.34%

Table 7. - Number of cells by average intensity

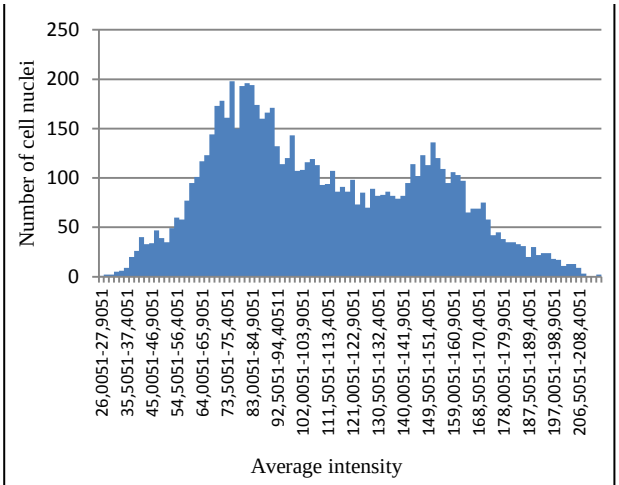

Figure 5. - Number of cells by avg intensity

Regarding the average intensities, 99.5% of the cell nuclei are between 36.59 and 205.01 values, so the optimization will use these values as interval boundaries (Table 8.).

Coverage	Interval from	Interval to	Interval length	Length/ (Max-Min)
100.0%	26.01	214.87	188.87	100.00%
99.5%	36.59	205.01	168.42	89.17%
99.0%	37.28	199.97	162.70	86.14%
95.0%	41.71	180.78	139.08	73.64%
90.0%	54.84	175.36	120.52	63.81%
80.0%	61.44	161.11	99.67	52.77%

Table 8. - Number of cells by average intensity

This value of course depends on the nucleus itself as well, but it is also greatly affected by the sample's illumination, the staining level, etc. However, since these are unknown factors, we had to find a general measurement number that can be used efficiently for every sample.

E. Examination of the seed point's intensity

Before the region growing, we have to locate the pixels on the processed image where the next growing iteration can start. To do this, first we have to select the pixel that currently has the highest intensity (since we are talking about 8 bpp images, there will be probably more than one such pixels). The selected pixel also has to comply with some other rules (it cannot be a part of an already detected cell nucleus and it has to be a local maximum). Since there is only a small possibility that a true cell nucleus will be found after a region growing iteration that started from a pixel that has a lower intensity; it is practical to introduce a minimal intensity limit, this greatly reduces both the ratio of false negatives and the execution time as well.

To determine the proper threshold, we examined how low this seed point intensity threshold should have been set to detect all the cell nuclei in the images that were annotated by the doctors. To do this, we need the pixels of the cell nuclei, and we also need the images themselves, where the algorithm performs the seed point search (these input images are the original images after some modifications by various filters [7]). After this, we checked the colors for all pixels in the nucleus, and we selected the pixel with the biggest intensity. This means that should we choose that color as the lower limit of the seed point intensity, the region growing will surely start on the current nucleus at least from one location (Table 9., Figure 6.):.

- Minimum seed point intensity: 0
- Maximum seed point intensity: 255

#	Interval from	Interval to	Number of nuclei	Ratio
1	0.00	12.80	1	0.01%
2	12.80	25.60	0	0.00%
3	25.60	38.40	0	0.00%
4	38.40	51.20	0	0.00%
5	51.20	64.00	1	0.01%
6	64.00	76.80	0	0.00%
7	76.80	89.60	7	0.09%
8	89.60	102.40	11	0.15%
9	102.40	115.20	31	0.41%
10	115.20	128.00	59	0.78%
11	128.00	140.80	165	2.18%
12	140.80	153.60	485	6.42%
13	153.60	166.40	1114	14.75%
14	166.40	179.20	1356	17.95%
15	179.20	192.00	1232	16.31%
16	192.00	204.80	1171	15.50%
17	204.80	217.60	1015	13.44%
18	217.60	230.40	584	7.73%
19	230.40	243.20	234	3.10%
20	243.20	255.00	88	1.16%

Table 9. - Number of cells by seed intensity

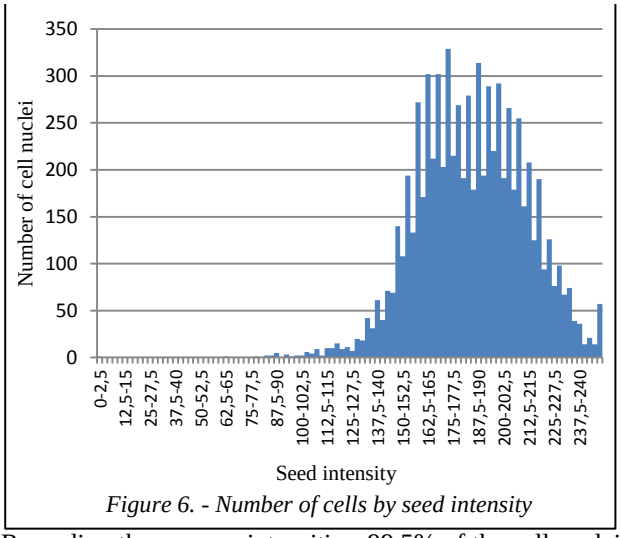

Figure 6. - Number of cells by seed intensity

Regarding the average intensities, 99.5% of the cell nuclei are between 0 and 251 values, so the optimization will use these values as interval boundaries (Table 10.).

Coverage	Interval from	Interval to	Interval length	Length/ (Max-Min)
100.0%	0.00	255.00	255.00	100.00%
99.5%	0.00	251.00	251.00	98.43%
99.0%	0.00	244.00	244.00	95.69%
95.0%	117.00	255.00	138.00	54.12%
90.0%	139.00	233.00	94.00	36.86%
80.0%	148.00	219.00	71.00	27.84%

Table 10. - Number of cells by seed intensity

This shows that even when we consider the 99.5% results, we have to define a fairly high threshold; a lower threshold would only result in unnecessary calculations, or in a worse case, in more false positive hits.

F. Examination of the contour contrast

We are not only evaluating the cell nuclei based on the pixel intensities and their location or size, but it is also an important requirement that the cells must be sharply detached from their surroundings. This is naturally evaluated using a contrast examination, during which we check the relationship between the cells' inner contour (that we consider as the part of the cell nuclei) and the outer contour (that we consider as the part of the environment). We can define a threshold for this ratio as well, and we can set up a condition that ensures that we only accept those cell nuclei candidates where this contrast reaches a certain limit.

The inner and outer contour can be found using various convolution operations [8]:

1. As a first step, we create a mask from the points of the cell nucleus. Then we dilate this mask, and we subtract the pixels of the original cell nucleus from the resulting image, and as a result we get the outer contour.
2. Then we sum up the intensity values of this contour and we divide this sum with the number of contour points. As a result we get the outer contour average intensity.
3. The inner contour can be similarly created using erosion, a dilatation, and another erosion, and if we subtract the pixels of the original cell nucleus from the resulting image.
4. After this, we sum up the intensity values of this contour and we divide this sum with the number of contour points. As a result we get the inner contour average intensity.
5. The ratio of the two values will give us the value we wanted: the contrast value for the contour curves.

We saved this number for every cell nuclei, and based on the 7554 cell nuclei in the database, we determined the following boundaries:

- Minimum contour contrast: 2093
- Maximum contour contrast: 20563

Distribution of the cell nuclei according to their contour contrast is in Table 11. and Figure 7.

#	Interval from	Interval to	Number of nuclei	Ratio
1	2093.00	3016.50	4	0.05%
2	3016.50	3940.00	110	1.39%
3	3940.00	4863.50	318	4.03%
4	4863.50	5787.00	562	7.12%
5	5787.00	6710.50	816	10.34%
6	6710.50	7634.00	1018	12.90%
7	7634.00	8557.50	876	11.10%
8	8557.50	9481.00	704	8.92%
9	9481.00	10404.50	609	7.72%
10	10404.50	11328.00	614	7.78%
11	11328.00	12251.50	557	7.06%
12	12251.50	13175.00	503	6.38%
13	13175.00	14098.50	377	4.78%
14	14098.50	15022.00	266	3.37%
15	15022.00	15945.50	186	2.36%
16	15945.50	16869.00	142	1.80%
17	16869.00	17792.50	111	1.41%
18	17792.50	18716.00	64	0.81%
19	18716.00	19639.50	40	0.51%
20	19639.50	20563.00	12	0.15%

Table 11. - Number of cells by contour contrast

Coverage	Interval from	Interval to	Interval length	Length/ (Max-Min)
100.0%	2093.00	20563.00	18470.00	100.00%
99.5%	3193.00	19130.00	15937.00	86.29%
99.0%	3193.00	18456.00	15263.00	82.64%
95.0%	3707.00	16157.00	12450.00	67.41%
90.0%	4133.00	14730.00	10597.00	57.37%
80.0%	4801.00	13217.00	8416.00	45.57%

Table 12. - Number of cells by contour contrast

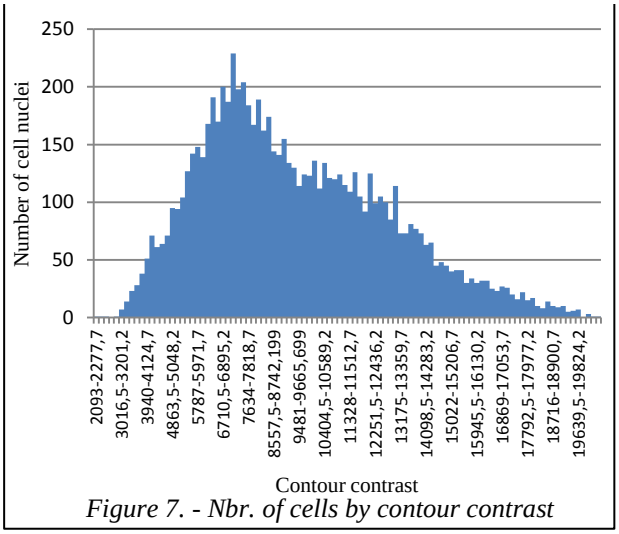

Figure 7. - Nbr. of cells by contour contrast

Regarding the contour contrasts, 99.5% of the cell nuclei are between 3193 and 19130 values, so the optimization will use these values as interval boundaries (Table 12.).

Naturally here we cannot talk about perfect accuracy either, but in our case the fast execution time is a lot more important, even if the price is some smaller inaccuracy. This is because this function is also executed after every iteration of the region growing process.

III. Conclusions

Using the gathered statistical data, we can start developing an optimization algorithm. If we are satisfied by covering 99.5% of the total set of values, then it is practical to use the following intervals to set the initial values in the starter population:

- Size: 34 – 882 pixels
- Radius: 4,8 – 22.832 pixels
- Circularity: 27.659 – 97.103
- Average intensity: 36.593 – 205.009
- Seed point intensity: 0 – 251
- Contour contrast: 3193 – 19130

IV. Plans for further development

Since the optimization itself has fairly high computational needs (the region growing has to be executed on the Gold Standard slides with every parameter set, then we have to execute the evaluation algorithm too [9]), we plan to develop a distributed system. The planned system uses a server (an application to control the genetic algorithm) and arbitrary number of clients (computers that execute the region growing and the evaluation using some predefined set of parameters) in trying to find the most optimal set of parameter values.

Acknowledgment

This work makes use of results produced by the Hungarian National Technology Programme, A1, Life sciences, the "Development of integrated virtual microscopy technologies and reagents for diagnosing, therapeutical prediction and preventive screening of colon cancer "Hungarian National Technology Programme, A1, Life sciences, (3dhist08) project and the ÓE-RH 1104/2-2011 project. Authors would like to thank for their financial support hereby.

The project was realized through the support of the European Union, with co-financing of the European Social Fund TÁMOP-4.2.1.B-11/2/KMR-2011-0001.

References

[1] S. Szénási, Z. Vámossy, M. Kozlovszky, "GPGPU-based data parallel region growing algorithm for cell nuclei detection", 12th IEEE International Symposium on Computational Intelligence and Informatics (CINTI), Budapest, Nov. 2011, pp. 493 - 499, ISBN 978-1-4577-0044-6

[2] Pannon Egyetem, "Algoritmus- és forráskódleírás a 3DHistech Kft. számára készített sejtmag-szegmentáló eljáráshoz", 2009

[3] D. E. Goldberg, "Genetic Algorithm in Search, Optimization and Machine Learning", Addison-Welsey Publishing Company, 1989, ISBN 0-201-15767-5

[4] M. Mitchell, "An Introduction to Genetic Algorithms", MIT Press, 1999, ISBN 0-262-13316-4

[5] Sz. Sergyán, L. Csink, "Consistency Check of Image Databases", Proc. of 2nd Romanian-Hungarian Joint Symposium on Applied Computational Intelligence, Timisoara, Romania, May 12-14, 2005, pp. 201-206., ISBN 963-7154-39-6

[6] R. S. Montero, E. Bribiesca, "State of the Art of Compactness and Circularity Measures", International Mathetical Forum, 4, 2009, no. 27, 1305 – 1335

[7] A. Nagy, Z. Vámossy, "Super-resolution for Traditional and Omnidirectional Image Sequences", Acta Polytechnica Hungarica, vol. 6/1, pp. 117–130, 2009, ISSN 1785 8860

[8] D. A. Forsyth, J. Ponce, "Computer Vision: A Modern Approach", Prentice Hall, 2002, ISBN 0-130-85198-1

[9] S. Szénási, Z. Vámossy, M. Kozlovszky, "Evaluation and comparison of cell nuclei detection algorithms", 16th International Conference on Intelligent Engineering Systems (INES), Lisbon, July. 2012, pp. 469 - 475, ISBN 978-1-4673-2694-0