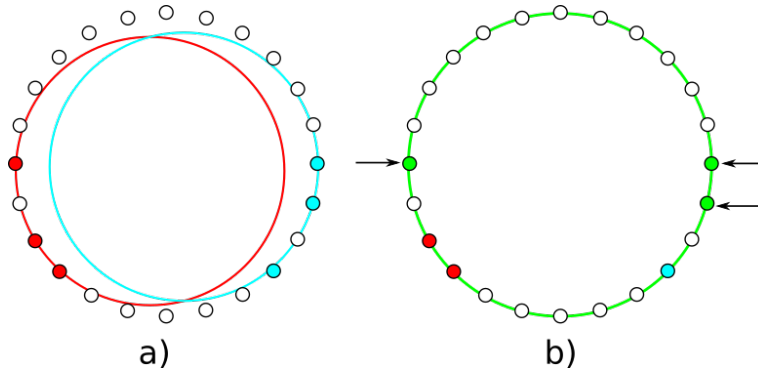
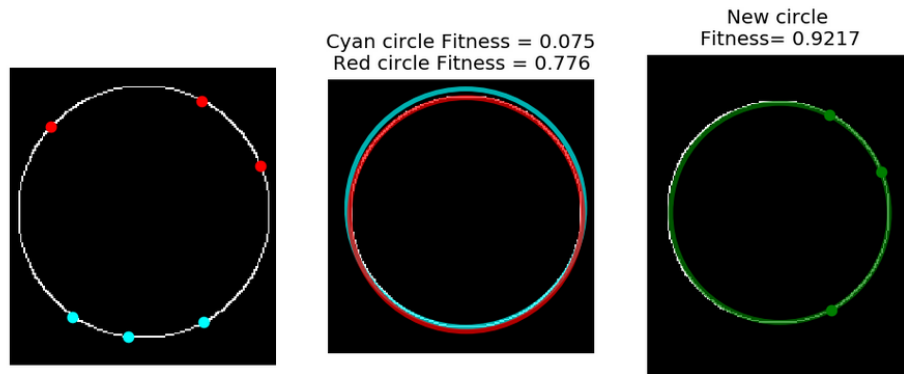


SUPPLEMENTARY MATERIAL

This document presents a justification for the use of GA through the explanation of how the crossover operator works for circle detection. Then the experiments related to the statistical test of the parameter tuning process are presented. As described in [1] illustration a) shows two different circles (red and cyan) each encoded by 3 positions (3 red dots and 3 cyan dots), as can be seen the generated circles do not fit into the point cloud (edge map), however, in illustration b) by taking one of the three dots of red circle, and two dots of the three dots of the cyan circle, together they form a new green circle that fits in all points.



In order to show how this process works, three individuals in a run of the algorithm over a synthetic circle is show in the following figure:

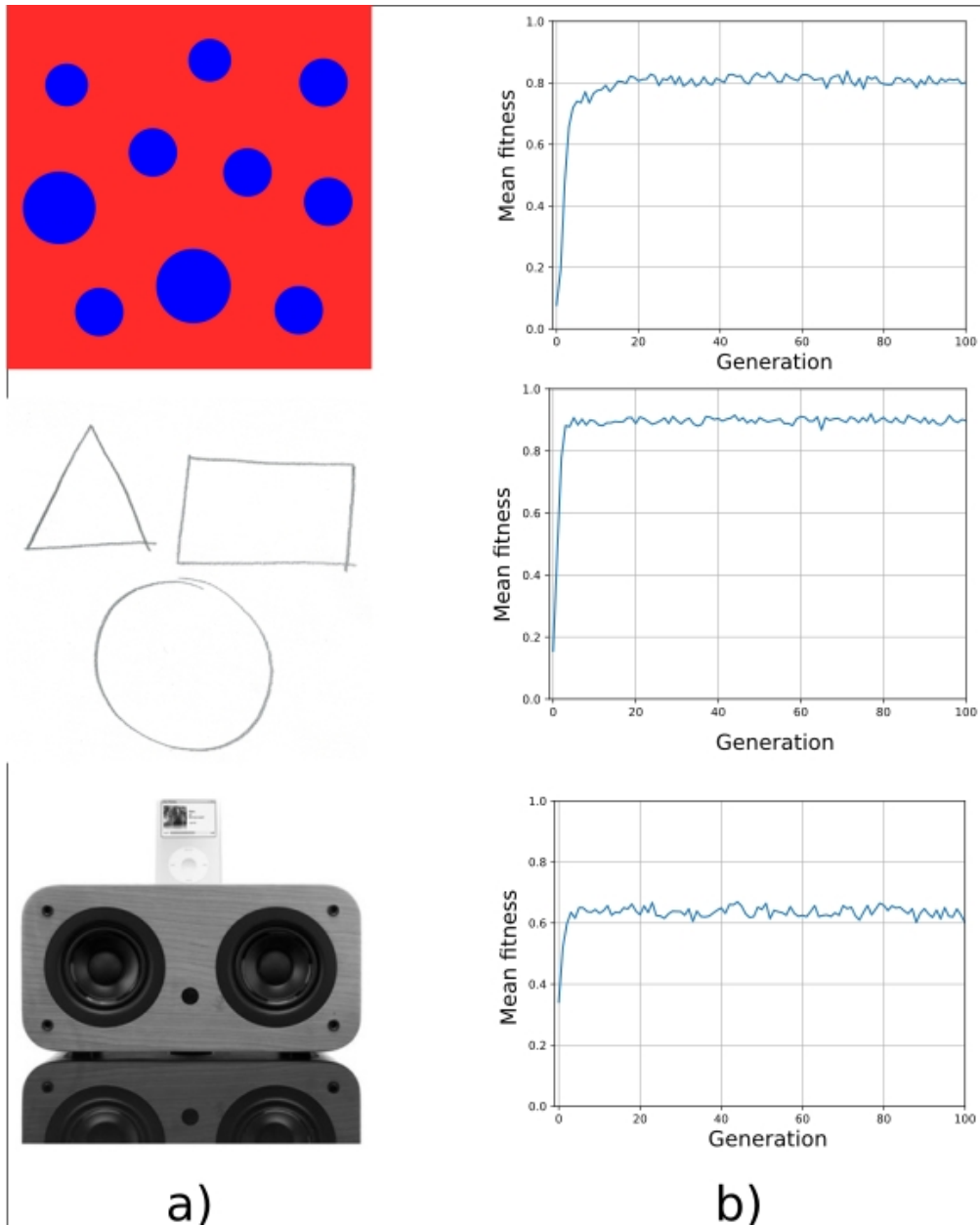


As can be seen, the fitness of the red and cyan circles are 0.776 and 0.075, respectively; by applying a crossover operator a new *better* circle is formed, with a 0.9217 fitness value.

Parameter tuning process

Number of generations

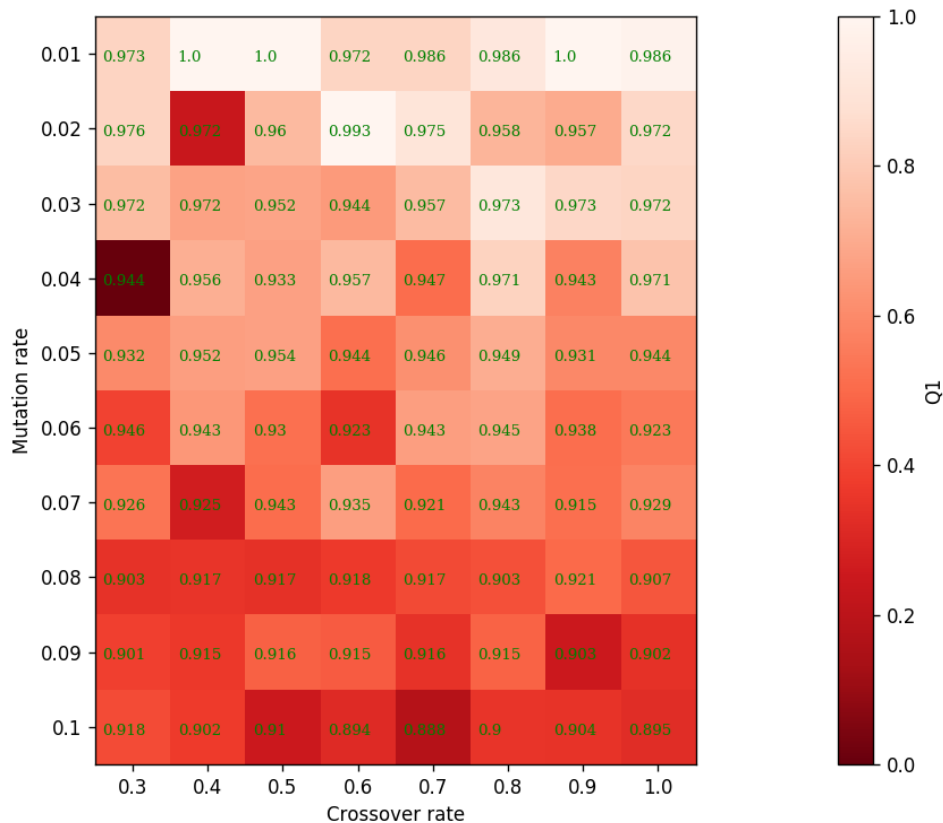
We present a plot of the mean fitness of the population in the evolution process; the input images are: synthetic, hand-drawn and real images. The algorithm is run 30 times.



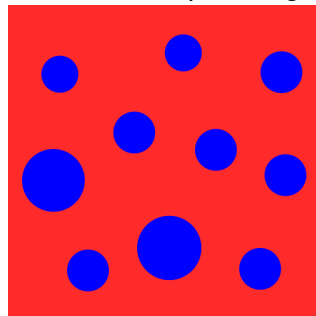
Population size, mutation and crossover rates used for this process were 70, 0.1 and 0.55, respectively, as they were used in the work of Ayala-Ramirez (reference [37] from the paper). In the figure, on column a) we present the input image used, and in column b) the mean fitness value as a function of the generation can be observed. As can be seen, we can conclude that fitness converges at generation 25, however, we suggest using 50 generations to have enough margin to be sure that the convergence is achieved in every case.

Statistical tests for crossover and mutation rates

We present an exhaustive grid search in order to establish the best combination of mutation and crossover rates. To this aim, we ran for 50 generations and a population size of 70 individuals, the experiment is repeated 100 times over 80 different combinations of mutation and crossover rate, the mean fitness value is computed. The mutation rate was explored in range $[0.01, 0.1]$, steps of 0.01, and the crossover rate in range $[0.3, 1]$ in steps of 0.1, resulting in 80 combinations. In order to show all combinations the following grid is presented:



where each cell presents the median value and the color corresponds to the first quartile sample. As can be seen, the combinations: **(0.01, 0.4)**, (0.01, 0.5) and (0.01, 0.9) produce the highest values, so we select the first one. The input image used was the following:

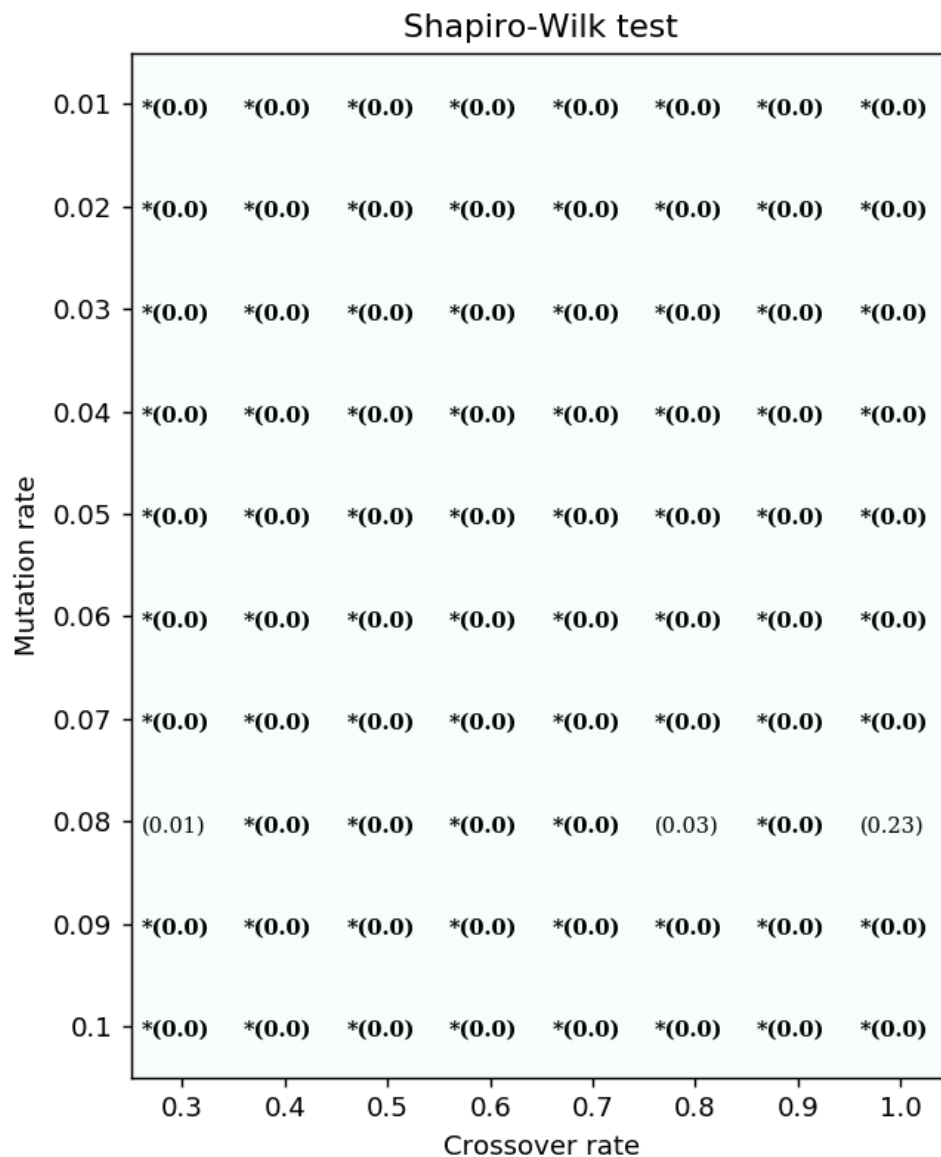


By following reference [2], we use the non-parametric Wilcoxon test, but first we determine whether or not the results follow a normal distribution, in this regard a normality test is performed: Shapiro-Wilk. Remember that the null hypothesis establishes that the sample

under test comes from a normal distribution, while the alternative hypothesis tells us that the sample does not.

Normality tests

Over each grid the resulting p values are presented for each test, the bold ones along with an * represents the ones that reject the null hypothesis with $\alpha = 0.01$, this value is used in all tests.



As can be seen, in Shapiro-Wilk test, only combinations: (0.08,0.3), (0.08,0.8) and (0.08,1) accept the null hypothesis.

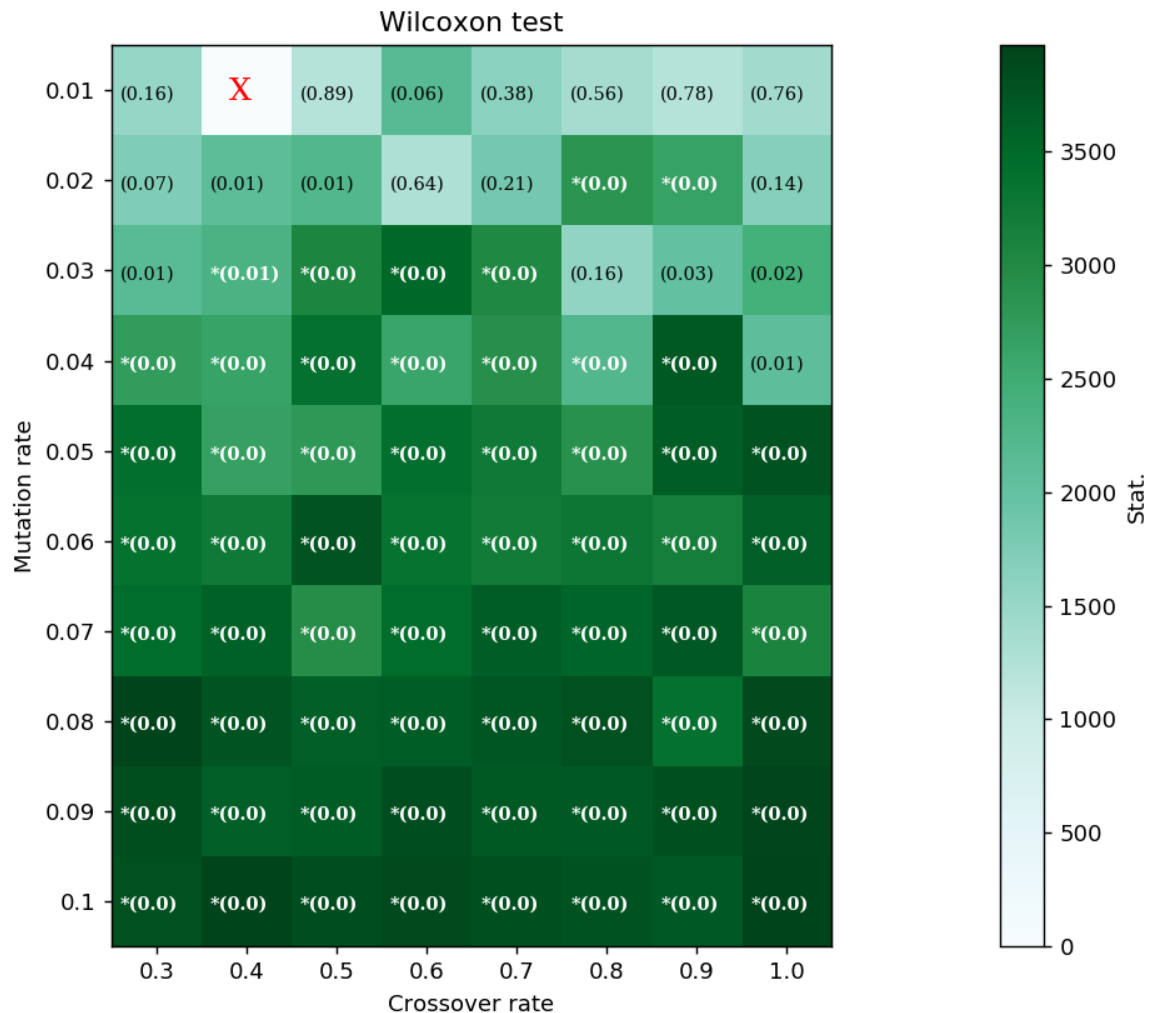
Pairwise Wilcoxon test

Given that the results did not follow a normal distribution we performed a pairwise Wilcoxon test, in order to establish whether there is significant difference between one of the combinations and the rest of them. This non-parametric test establish that:

$H_0: X - Y = 0$

$H_a: X - Y \neq 0$, $H_a: X - Y > 0$, $H_a: X - Y < 0$,

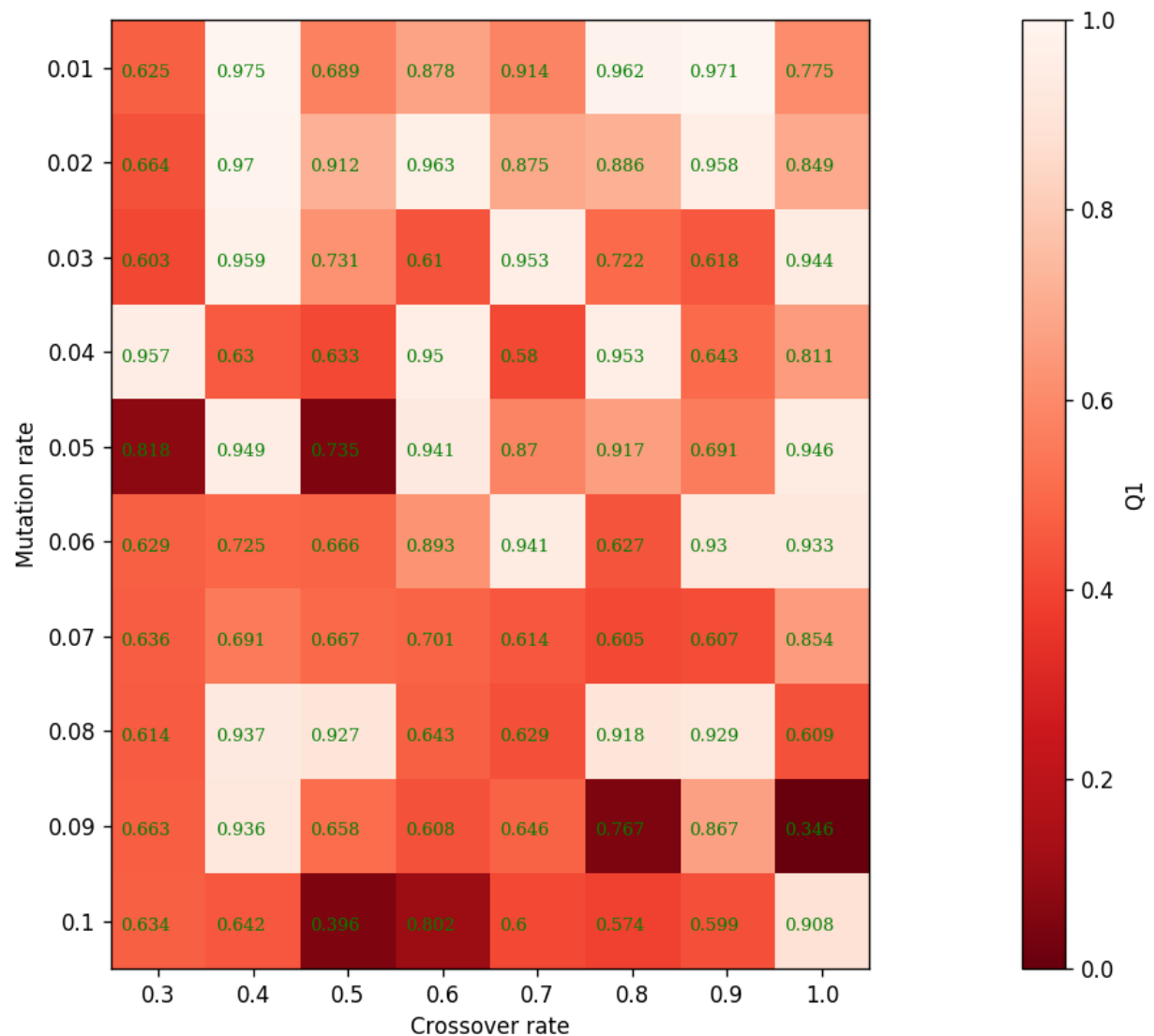
The sample to compare is the one mentioned above (0.01,0.4), the following grid shows the obtained results:



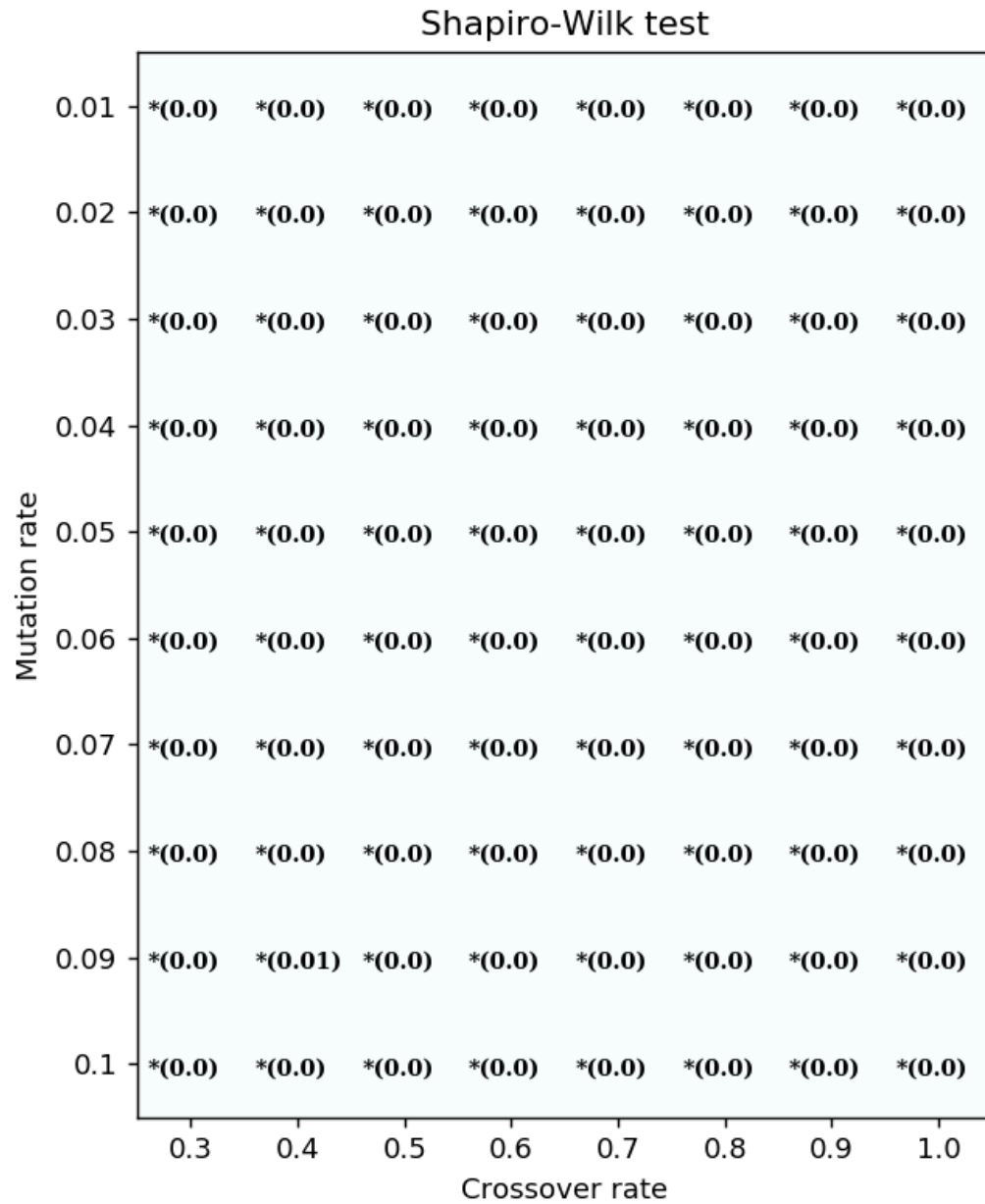
We can conclude that: there is no significant difference over the black values, while white values reject the null hypothesis. In order to extend these results we have performed the same experiment, using a different input image:



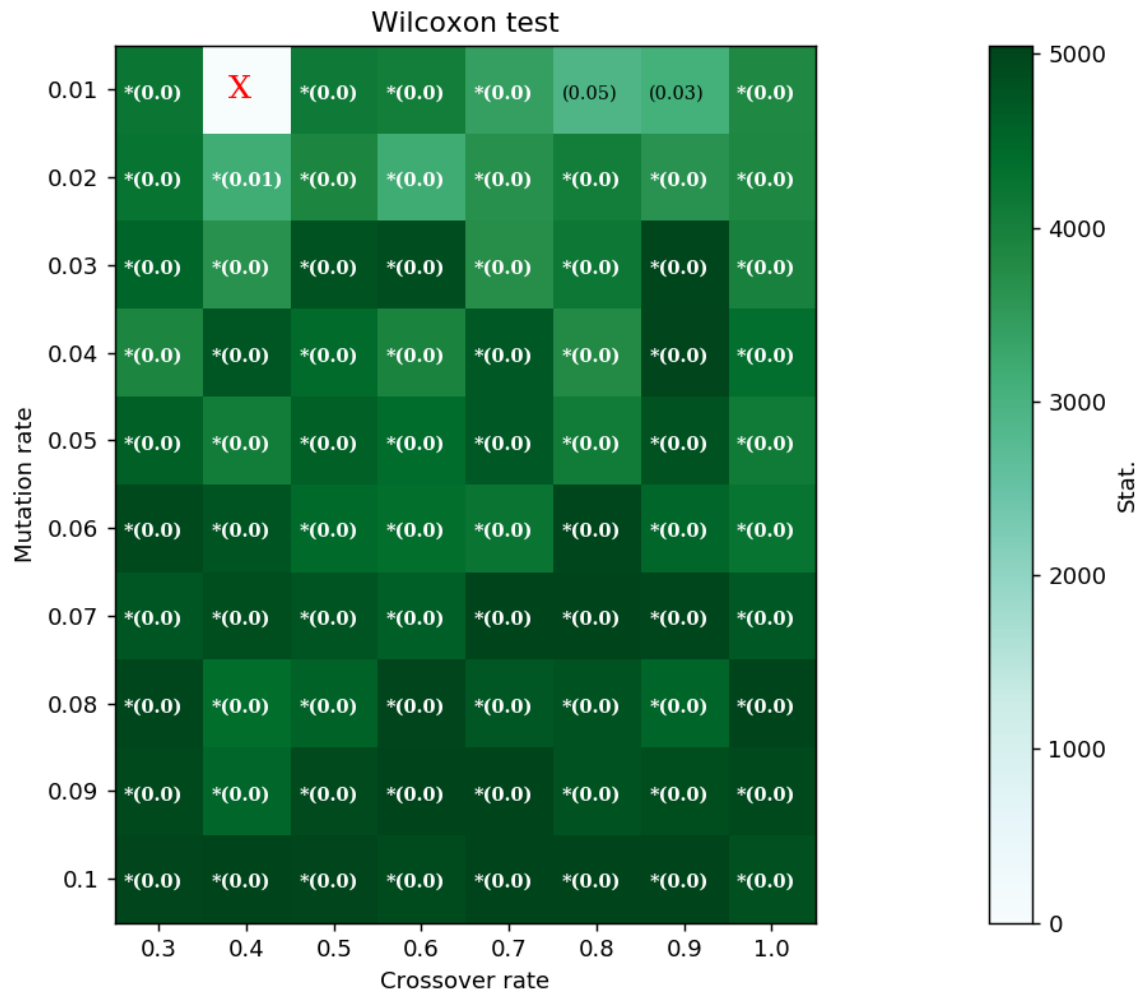
The resulting grid is the following:



We can also see that the combination of (0.01,0.4) is the one with the greatest median fitness value. The normality test results are presented in the following grids:



As can be seen, Shapiro-Wilk test rejects the null hypothesis in all cases. Last, we repeat the pairwise Wilcoxon test, the obtained results are the following:



Here we can observe that only 2 combinations (the black values) accept the null hypothesis, while for the other values there is a significant difference from the result produced by the combination (0.01,0.4).

Concluding that the values for each rate are the following:

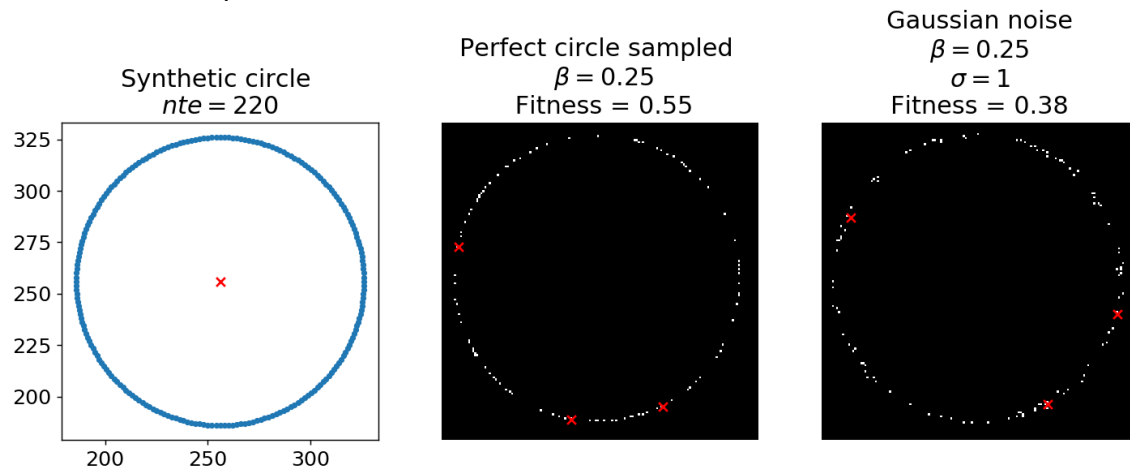
1. **Mutation rate = 0.01**
2. **Crossover rate = 0.4**

All statistical tests were performed in R with the following functions:

- wilcox.test and
- shapiro.test

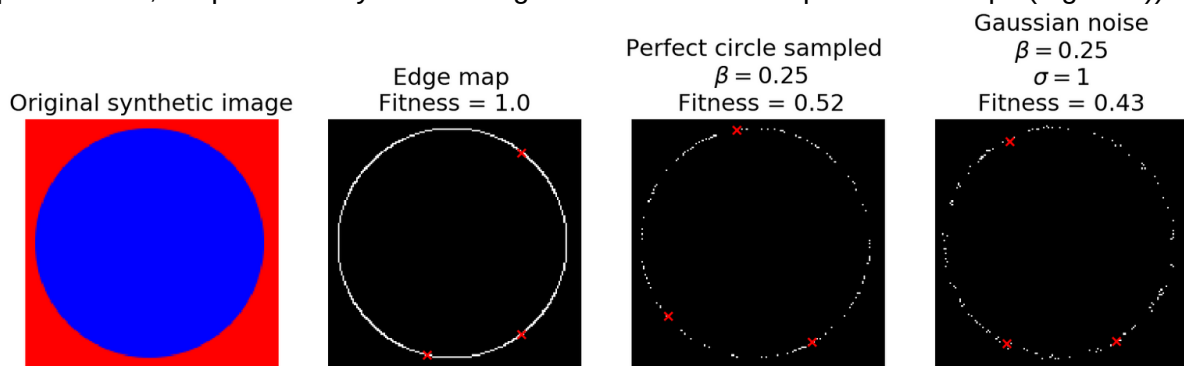
Fitness threshold

The fitness threshold (used as a stopping criterion for detection process) is fixed to 0.5, we can see how this parameter works:



a)	b)	c)
----	----	----

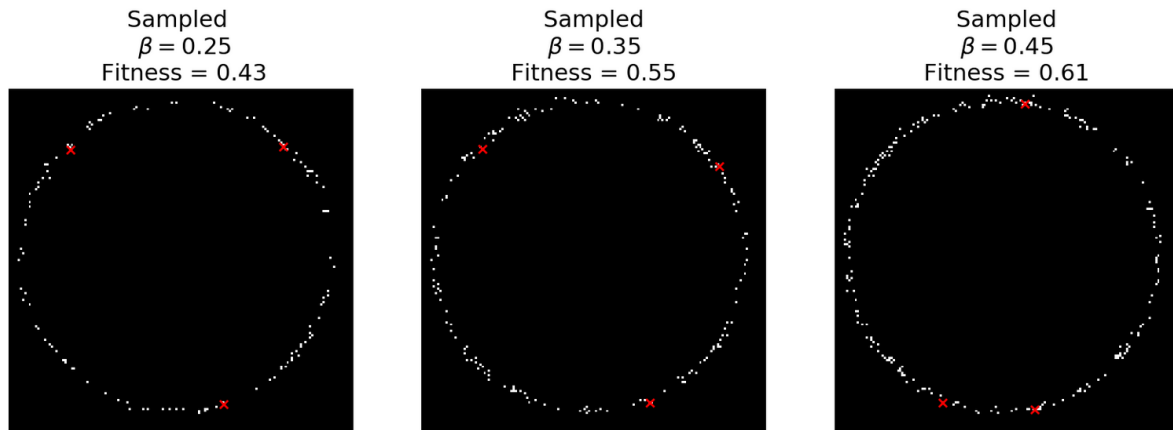
Figure a) shows a generated circle with the TP method by using the following parameters: $r = 70$, $\alpha = 0.5$ and $n_{te} = 220$ (α represents the factor that controls the number of points to generate). Figure b) presents also a TP generated circle with $r = 70$, but with $\alpha = 1$, $n_{te} = 220$, and by randomly missing 75% of its points ($\beta = 0.25$), its corresponding fitness equals to 0.55; this fitness is computed from the individual (3 points) displayed in Figure b) as 3 red marks ("x"). Then, in Figure c) the same circle is presented but with Gaussian noise-added to mimic a deformation; this process consists of adding a random Gaussian 2D array with factor noise $\sigma = 1$ and a center mean value of zero, to the circle presented in Figure b); as can be seen the fitness drops to 0.38. With a fitness threshold of 0.5 we can recover circles with 75% of missing points from its perimeter. Notice that the β factor is not mentioned in the paper, this value was only used to illustrate what happens when the circles present missing points. Now, we present a synthetic image created with inkscape with $r = 70$ px (Figure d)):



d)	e)	f)	g)
----	----	----	----

Figure e) presents the edge map obtained from Figure d) and the fitness of the individual displayed (red "x") is equal to 1. Figures f) and g) present the same conditions as in figures b) and c), respectively, but using the points from the edge map in Figure e). As can be seen,

the fitness values are equal to 0.52 and 0.43, respectively. Concluding that a fitness threshold 0.5 can recover non deformed circles with 75% of missing points from its perimeter. Additionally, we present the following figures:



In figures h), i) and j), also the points from edge map in figure e) were used, and the same configuration noise as in Figure g) was used. We decrease the amount of missing points: from 75%, to 65% and to 55%, respectively ($\beta = 0.25, 0.35$ and 0.45). We can observe that circles with deformation (Gaussian noise-added) and between 55% and 65% of its perimeter points missing (as can be seen in figures i) and j)) can be recovered by using a fitness threshold of 0.5. However, circles that are defined by only 25% of the total points in the circumference will not be recovered.

References

- [1] G. Roth and M. D. Levine, "Geometric primitive extraction using a genetic algorithm," in *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 9, pp. 901-905, Sept. 1994, doi: 10.1109/34.310686.
- [2] García, Salvador, et al. "A study on the use of non-parametric tests for analyzing the evolutionary algorithms' behaviour: a case study on the CEC'2005 special session on real parameter optimization." *Journal of Heuristics* 15.6 (2009): 617. doi: 10.1007/s10732-008-9080-4(2008)