IDENTIFICATION OF ENDOTHELIAL CELL MORPHOLOGY IN CORNEA USING EVOLUTION STRATEGIES

Lisa Obritzberger [(a)], Susanne Schaller [(a)], Viktoria Dorfer [(a)], Claudia Loimayr [(b)], Simone Hennerbichler [(b)], Stephan Winkler [(a)]

[(a)] University of Applied Sciences Upper Austria
Bioinformatics Research Group
Softwarepark 13, 4232 Hagenberg, Austria

[(b)] Red Cross Blood Transfusion Service for Upper Austria
Austrian Cluster for Tissue Regeneration
Krankenhausstraße 9, 4020 Linz, Austria

[(a)] lisa.obritzberger@fh-hagenberg.at, susanne.schaller@fh-hagenberg.at, viktoria.dorfer@fh-hagenberg.at,
[(b)] claudia.loimayr@o.roteskreuz.at, simone.hennerbichler@o.roteskreuz.at, [(a)] stephan.winkler@fh-hagenberg.at

ABSTRACT

In this paper we describe an algorithm for identifying shapes of endothelial cells in images of human cornea. In order to analyze the quality of cornea samples that shall be transplanted, endothelial cells have to be counted and the morphology of cells has to be identified.

The algorithm for shape detection presented in this paper is based on evolution strategies (ES): After several pre-processing steps an initial shape is fitted into the detected cells and repeatedly evaluated and mutated in order to create new candidates from which the best ones are promoted to the next generation. In the experimental section of this paper we analyze the performance of this approach for identifying cell morphologies in images of human cornea samples using different fitness functions.

Keywords: structure identification, image analysis evolution strategies, human cornea analysis

1. INTRODUCTION

Cornea transplantation may be the only recovery chance for diseases such as Fuchs' Dystrophy, Keratoconus or similar disorders (Kampik and Grehn 2002). Here the cornea of the donor is taken and fixed at the right position of the recipient eye with few sutures. The post mortem enucleation and cornea processing according to European Community Directives 2004/23/EC, 2006/17/EC, and 2006/86/EC can be performed within 48 hours without loss of quality. To minimize risks of cornea transplantations, the tissue donation has to pass strict quality checks. Besides tests to exclude the possibility of HIV, hepatitis, and syphilis, corneas are also evaluated concerning their quality.

The quality of corneas can be defined by two criteria: the endothelial cell count and the morphology of consisting cells. As cornea endothelial cells are not able to undergo cell division and consequently cannot be replicated, the number of endothelial cells decreases over the years. Therefore, higher donor age leads to a reduced endothelial cell count (Loimayr et al. 2012).

The minimum of endothelial cell count should be 2000 cells per mm² to be adequate for transplantation use (Dichtl et al. 2010). Quality determination through cell counting is in many tissue banks still a manual process. Usually, cornea endothelial cells have a hexagonal cell form as apparent in Figure 1.

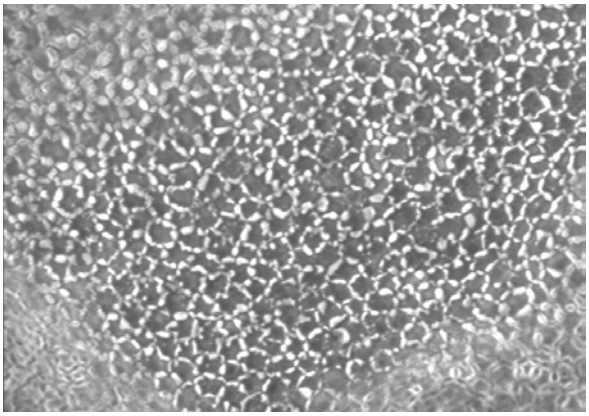

Figure 1: Endothelial cells of which most have hexagonal shapes.

The typical form of endothelial cells is lost during apoptosis. In this process, cells lose their original form and size and cannot be counted as hexagonal. The higher the number of cells with hexagonal shape, the better is the quality of a cornea and the higher is the possibility to be used for transplantation.

Thus it is not only important to identify the endothelial cell count, but also to analyze the shape of cells on the cornea that are to be used for transplantation. The goal of the research work presented in this paper is to develop an algorithm that is able to automatically identify the number of cells and the classification according to their shape.

2. STATE OF THE ART

Currently, there are only a view tools for analyzing images of corneas. There is for example the Endothel Analysis System of the German company Rhine-Tec, which rather concentrates on counting the cells and makes it not clear how the identification of the cell morphology is calculated. Another software specialized on endothelial cell count is the NAVIS Cell Count-Advanced Vision Information System of the German company Nidek Technologies, which offers also functionalities to identify the cells shape. Researchers have compared these two software tools against manual endothelial cell counts and their studies have shown that the fully automated analysis of cornea cells remains problematic (Hirneiss et al. 2007).

There are also several cell counting tools that are not geared towards endothelial cell count and hence not optimally appropriate for this purpose.

Due to the lack of appropriate cell counting tools, in many tissue banks the cell count and cell morphology identification is still a manual progress, for which an alternative shall be shown in this paper.

3. DATA PRE-PROCESSING

In this section we describe the pre-processing steps that have been performed before applying evolution strategy on processed images. The following pre-processing steps are applied: identification of sharp areas in images, dilation, transformation to binary images, and cell detection using a flood fill algorithm.

3.1. Identification of Sharp Areas

Microscopy images of cornea samples are out of focus at the edges, which is caused by curvature of the cornea. In these areas a proper analysis of endothelial cells is not possible and would falsify the results. Therefore, the focused area of cornea images has to be detected. This is achieved by using the change of intensity in the images: Areas that are out of focus have smoother transitions than focused areas hence change of intensity is higher in focused areas of the image.

First the intensity of each pixel of the image is calculated by using the RGB-data of the pixel p. The thereby used definition can be seen in Equation 1, where $p.R$ represents the red component, $p.G$ the green component, and $p.B$ the blue component of pixel p (Nagabhushana, 2005).

$$intensity = 0.3 \times p.R + 0.59 \times p.G + 0.11 \times p.B \quad (1)$$

In the next step a radius is defined in which each pixel is compared to its neighbors with regard to the intensity value. If the difference between the maximum and the minimum intensity value in the neighborhood exceeds a pre-defined threshold, the centered pixel is marked as focused.

To identify not only sharp pixels, but also whole areas as focused, a second radius is defined, in which a minimum number of pixels that have been marked as sharp has to be reached. If the number of sharp pixels in the neighborhood is at least as high as the pre-defined minimum, then the pixel belongs to a focused area. An example for this is shown in Figure 2.

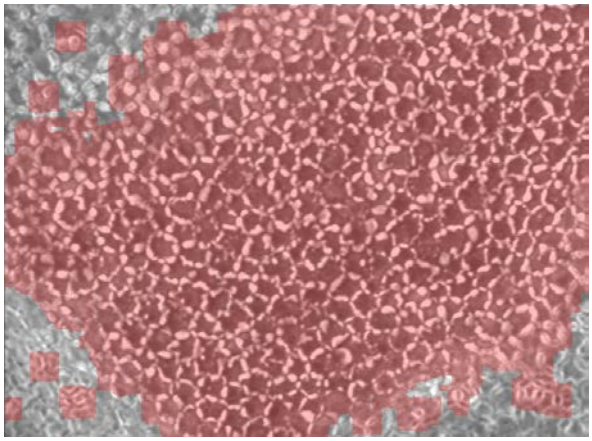

Figure 2: Focused area identified in a cornea image.

3.2. Dilation for Cell Margin

In order to identify pixels belonging to a cell area proper cell margin has to be performed. Therefore, the holes between the cell membrane of endothelial cells have to be closed. To achieve the closing of cell membrane dilation is processed on the cornea cell image. Dilation uses a mask that causes objects to increase and thereby fills small holes between objects (Jähne 2012) as depicted in Figure 3.

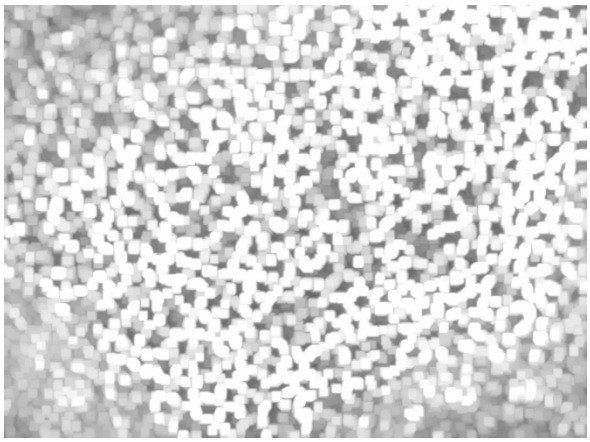

Figure 3: Cornea image after performing dilation.

3.3. Transformation to Binary Images

To distinguish cell areas from cell membrane, we consider the different intensity levels in the images. A transformation to a binary image is achieved using a pre-defined threshold; intensity levels below the

threshold are transformed to black representing cell pixels, others to white ones.

3.4. Flood Fill Algorithm for Cell Separation

To identify pixels that belong to the same cell a recursive flood fill algorithm is applied. A point within the cell area is selected as seed from where all 4-connected pixels, if they are also part of the cell, are colored. The process is repeated recursively for each of the connected pixels until there is no neighboring pixel left that can be colored (Pachghare 2008). This leads to images as the one shown in in Figure 4:

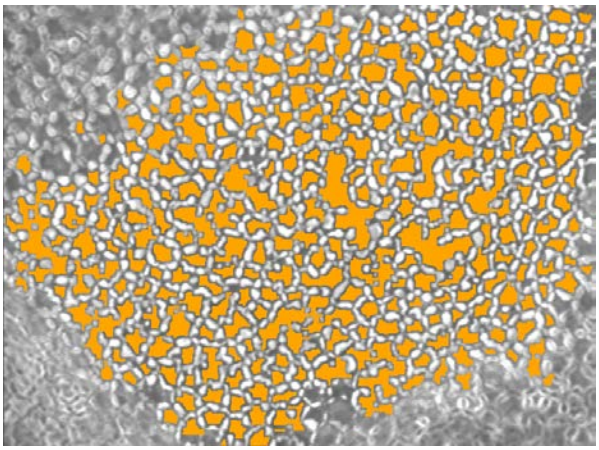

Figure 4: Detected cells after performing a flood fill algorithm.

4. IDENTIFICATION OF ENDOTHELIAL CELL MORPHOLOGY IN IMAGES OF CORNEA SAMPLES BASED ON EVOLUTION STRATEGIES

In this section we describe an approach for identifying cell forms in images using evolution strategies: The fitness of several created shapes is repeatedly evaluated and mutated in order to generate new solution candidates.

The initial shapes are created with a randomly chosen number of corners in the range [4, 7]. This form is fitted into the cell and evaluated. Then, mutation is executed by changing the corners of the initial shape, followed by re-evolution; solution candidates with the best fitness are chosen as new parents for the next generation.

4.1. Evolution Strategies

Evolutionary algorithms are inspired by the Darwinian paradigm of evolution, and use the principle of variation and selection combined with generational changes (Beyer 2001). The two main types of evolutionary algorithms are genetic algorithms and evolution strategies.

Evolution strategies were developed since the 1960s, primarily by a German research community around Rechenberg and Schwefel at the Technical University of Berlin, and have been extensively studied in Europe (see for example (Rechenberg 1973) and (Schwefel 1994)).

As one of the most representative of evolutionary computation ES executes the optimization process by applying operators in a loop. Each step of the algorithm's execution is known as a generation. Over generations main operations are applied on the solution candidates repeatedly until a given termination criterion is met. One of the termination criteria can for example be the reaching of a pre-defined number of generations.

An ES works with a population of individuals which are also called solution candidates and are characterized by their parameter vector. The vector is used to calculate the individual's fitness value.

In each generation the old population is replaced by successful child individuals. Individuals are represented as real-valued vectors in which each position corresponds to a feature of the individual.

In contrast to genetic algorithms, where mutation is mainly used for avoiding stagnation, mutation is the main reproduction operator in evolution strategies. In each generation each component of the vector is mutated individually, where small mutations are more likely than big ones.

As the algorithm converges with increasing quality of the solution candidates, the mutation width δ (convergence rate) should be adapted. For the purpose of self-adaptation, the 1/5 success rule proposed by Rechenberg (for Details see Rechenberg 1973) is processed. Here the percentage of successful mutations *s* is recorded over *n* generations, to calculate the convergence rate as seen in Equation 2, whereas c_d = 0.82 and $c_i = 1 - 0.82$.

$$\delta(t) = \begin{cases} c_d \times \delta(t-n) & if\ s < \frac{1}{5} \\ c_i \times \delta(t-n) & if\ s > \frac{1}{5} \\ \delta(t-n) & if\ s = \frac{1}{5} \end{cases} \quad (2)$$

Another possibility to create new individuals can be recombination where a child is created by recombining two solution candidates (parents). This can for example be processed by calculating their geometric average. Strategies for choosing candidates for recombination and recombination techniques can be very different. Typically, parent selection in ES is performed uniformly and random, with no regard to fitness.

The algorithm described in this paper does not use recombination.

Examples for mutation and recombination in the context of ESs are shown in Figure 5.

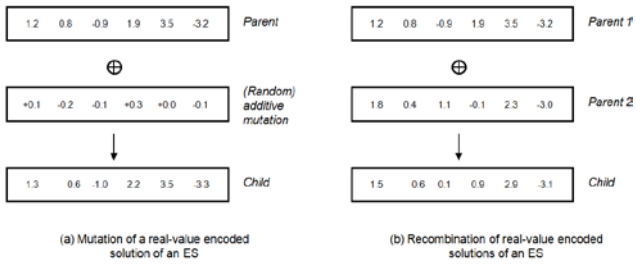

Figure 5: Exemplary solution candidates and the effect of genetic operations in ES (Winkler 2009).

The selection of individuals for the next generation is a deterministic method that uses the survival of the fittest principle: only the best individuals remain.

In each generation of an ES algorithm, λ children are produced by μ parent individuals. By selection, the best children are chosen and become the parents of the next generation. The μ best individuals are based on the relative ordering of fitness values.

Basically, there are two selection strategies for ESs: the (μ,λ)-strategy and the (μ+λ)-strategy. We have focused on the (μ+λ)-strategy, which is also called plus-selection and allows not only the λ produced children, but also the μ parents to be included in the pool of potential new parents. The best individuals of parents and children are chosen as the next generation's parents.

The main procedure steps of the execution of ES is summarized and graphically shown in Figure 6.

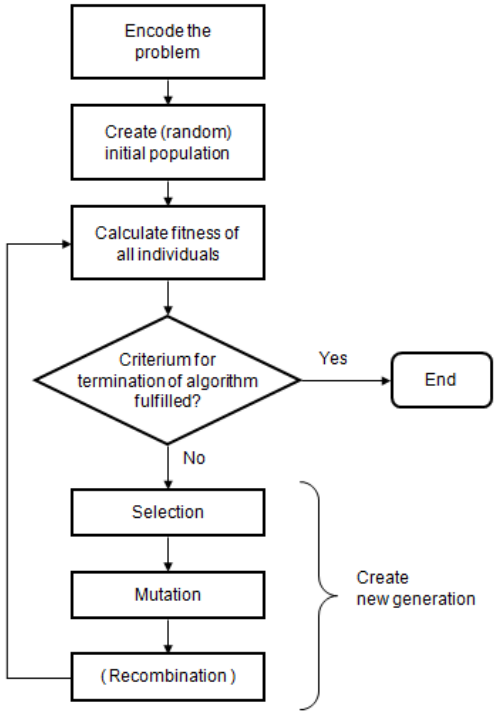

Figure 6: Workflow of the standard evolution strategy (ES) algorithm (Winkler 2009).

4.2. Cell Morphology Identification using Evolution Strategies

Solution Candidates

A solution candidate for the ES performed to detect the correct cell morphologies is represented as a list of points which describe the corners of a shape. Figure 7 shows an example:

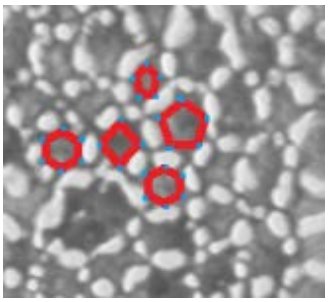

Figure 7: Idea of fitting shapes into cells; shapes are defined by their corners.

Initial Candidates

Initial solution candidates have a uniformly distributed possibility of being quadrangular, pentagonal, hexagonal or heptagonal. As the best fit can be created by using pixels that are part of the cell border, the cell border pixel with the minimum distance from the original corner of the initial solution candidate is calculated for each corner and assigned as the new corner. The so created solution candidates are the first individuals of the ES' population.

Mutation

After calculating the fitness of each parent they are mutated in order to create children as new candidates from whom the best ones are promoted to the next generation. Mutation is here accomplished in two ways:

- Mutation of points: Each corner of a parent solution candidate has a pre-defined possibility of being mutated. This mutation causes a small movement of the edges and thereby changes the shape and area of the original solution candidate.
- Change of the number of points: There can either be a new corner added to the shape by mutating an existing corner and adding it as an additional one, or a randomly chosen corner can be deleted.

Fitness Function

The fitness function has to return a value that reflects how well the shape fits into the original cell area. To get the fitness f of a solution candidate s, the area of the created shape is calculated and compared with the area of the cell. Overlapping areas lead to

improvement of the fitness value, whereas area that does not fit leads to fall off in fitness quality. Thereby the overlapping areas are weighted with a value $\propto_1$ and points outside the shape are weighted with a value $\propto_2$ as defined in Equation 3:

$$f(s) = \left(\sum_{p \in ca(s)} I(p)\right) \times \propto_1 - \left(\sum_{p \notin ca(s)} I(p)\right) \times \propto_2 \quad (3)$$

Where *ca(s)* is the area covered by the shape *s*. The aim of this approach is to maximize the values returned by the fitness function. The degree of reward / punishment depends on the grayscale intensity *I* of the pixels. In the experiments summarized in this paper $\propto_1$ was always greater than $\propto_2$.

As it is easier for shapes with a higher number of corners to fit the original cell, fitness has to be weighted by the number of corners.

5. EXPERIMENTAL RESULTS

Several variants of this solution evaluation approach have been evaluated and are described in the following section.

5.1. Using Fitness Function with Non-Weighted Corners

Initial ES runs have been performed without weighting the fitness by the number of corners of a shape. As it is easier for shapes with higher number of corners to fit a cell, a non-corner weighted fitness function favored shapes with higher corner number which is shown in Figure 8.

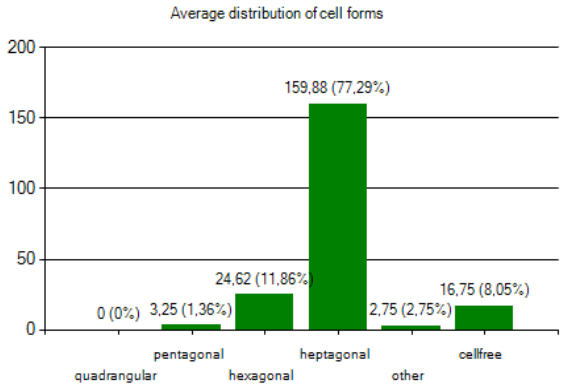

Figure 8: Distribution of identified cells using the non-corner weighted fitness function.

To further evaluate the quality of the results a manual analysis has been executed where the shape of the cells (originally defined by a human expert) was compared to the result of the algorithm. We have used 8 images of corneas provided by the Red Cross Blood Transfusion Service for Upper Austria. The evaluation of the results of five independent identification runs per image (ES with 100 generations and 1 parent with 10 offspring) are summarized in Table 1. The overall classification of calculated forms compared to the original shape classification can be seen in the confusion matrix shown in Table 2.

Image Sample	**Classified correctly**	
	Average	**Standard deviation**
1	32.50%	0.07
2	12.86%	0.03
3	14.67%	0.03
4	13.33%	0.05
5	9.09%	0.00
6	4.62%	0.04
7	13.85%	0.08
8	1.33%	0.03
Mean Quality	**12.78%**	**0.10**

Table 1: Average results achieved with five runs, each with 100 ES iterations per cell, using a fitness function with non-weighted corners.

expected / identified	4	5	6	7	
4	0.00%	0.00%	0.00%	0.00%	
5	0.00%	0.20%	0.20%	0.00%	
6	0.41%	2.24%	2.86%	0.61%	
7	2.65%	25.10%	57.14%	8.57%	
					11.63%

Table 2: Confusion matrix with average classification ratios for cells in 8 cornea images. Classifications computed in 5 independent test runs per cell are compared to original cell shapes classifications defined by a human expert. The results were achieved by using a fitness function with non-weighted corners.

5.2. Using Fitness Function with Non-Linear Functions

To avoid drifting to shapes with higher numbers of corners, tests with a weighted fitness function have been performed: The calculated fitness values have been weighted by dividing them by the number of corners, by the logarithm of number of corners and by the square root of number of corners.

The results obtained using these functions are similar, they all perform better than the non-corner weighted fitness function; still, shapes with higher number of corners are favored which can be seen in Figure 9. Though results improved they were still not satisfying which can be seen in Table 3 and Table 4.

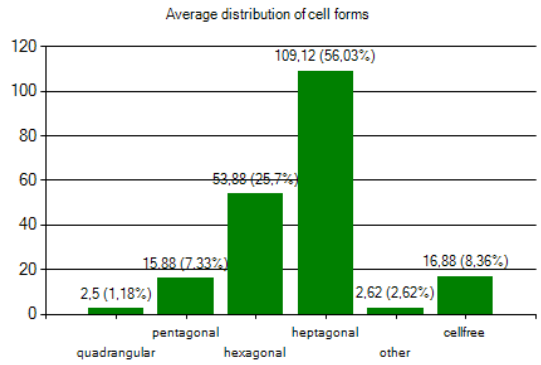

Figure 9: Distribution of identified cells using the fitness function weighted by logarithm of corner number.

Image Sample	Classified correctly	
	Average	Standard deviation
1	40.00%	0.21
2	24.29%	0.10
3	34.67%	0.16
4	26.67%	0.13
5	21.82%	0.08
6	16.92%	0.03
7	23.08%	0.05
8	20.00%	0.11
Mean Quality	**25.93%**	**0.13**

Table 3: Average results achieved with five runs, each with 100 ES iterations per cell, with a fitness function weighted by logarithm of corner number.

expected / identified	4	5	6	7	
4	0.00%	0.20%	0.82%	0.20%	
5	0.20%	1.63%	6.73%	0.82%	
6	1.22%	10.61%	18.57%	3.06%	
7	1.63%	15.10%	34.08%	5.10%	
					25.31%

Table 4: Confusion matrix with average classification ratios for cells in 8 cornea images. Classifications computed in 5 independent test runs per cell are compared to original cell shapes classifications defined by a human expert. The results were achieved by using a fitness function weighted by logarithm of corner number.

5.3. Using Fitness Function with Linear Weighting

As the previous methods did not show the expected success, the fitness was weighted by a pre-defined value depending on corner numbers which are depicted in Table 5. This approach has led to better results with mainly hexagonal cell forms, which can be seen in Figure 10.

Numbers of corners	Weight
4	1.6
5	1.4
6	1.2
7	1.0

Table 5: Weights for shapes depending on corners.

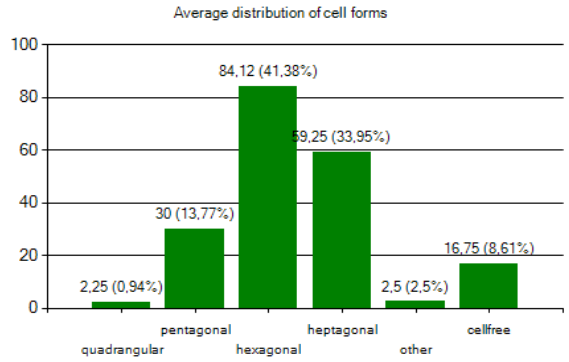

Figure 10: Distribution of identified cells using fitness function weighted by pre-defined values.

To evaluate the performance of this approach in more detail, the change of fitness over generations was observed and showed that the shapes' fitness improves over time. This can be seen in Figure 11:

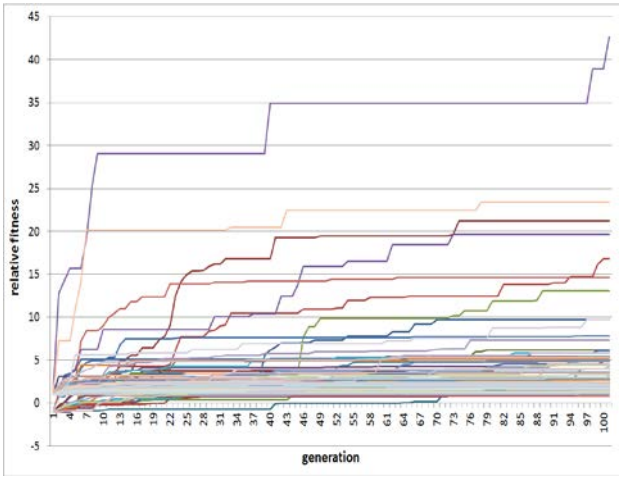

Figure 11: Improvement of fitness of shapes over generations for randomly selected cells.

The shape with the best fitness is selected to represent the morphology of the cell and plotted into the original image which is shown in Figure 12:

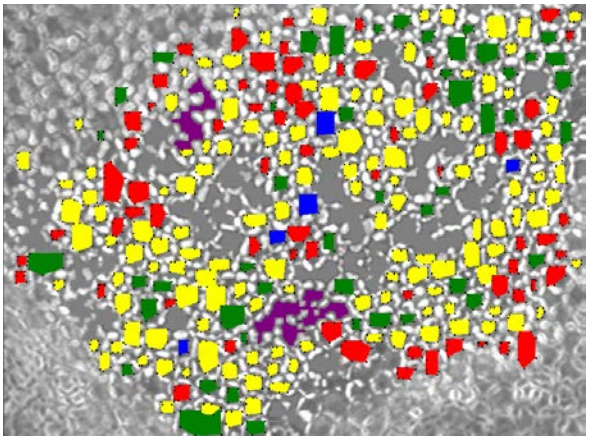

Figure 12: Identified shapes with the best fitness fitted into endothelial cells.

The evaluation of the results of five independent identification runs per image (ES with 100 generations) are summarized in Table 6. The overall classification of calculated forms compared to the original (human) shape classification can be seen in the confusion matrix shown in Table 7.

With this approach 70.03% of the samples were classified correctly. By accepting an error of + / - one corner the number of correctly classified samples is even 92.80%.

Image Sample	Classified correctly	
	Average	Standard deviation
1	62.10%	0.13
2	62.90%	0.04
3	72.17%	0.06
4	67.94%	0.08
5	74.07%	0.12
6	74.05%	0.11
7	70.98%	0.11
8	69.92%	0.07
Mean Quality	**69.27%**	**0.09**

Table 6: Average results achieved with five runs, each with 100 ES iterations per cell, with a fitness function with linear weighted corners.

expected / identified	4	5	6	7	
4	0.31%	0.32%	1.47%	0.30%	
5	0.60%	8.77%	7.67%	1.32%	
6	1.03%	6.86%	51.62%	3.64%	
7	0.14%	2.93%	3.68%	9.33%	
					70.03%

Table 7: Confusion matrix with average classification ratios for cells in 8 cornea images. Classifications computed in 5 independent test runs per cell are compared to original cell shapes classifications defined by a human expert. The results were achieved by using a fitness function with linear weighted corners.

6. CONCLUSION

In this paper we have described an evolutionary algorithm that is able to automatically identify the morphology of endothelial cells of human cornea in microscopy images. In future work this approach shall be included in a fully automated image analysis framework that is designed for biomedical research and production.

ACKNOWLEDGMENTS

The work described in this paper was done within the FIT-IT project number 835918 “NanoDetect: A Bioinformatics Image Processing Framework for Automated Analysis of Cellular Macro and Nano Structures” sponsored by the Austrian Research Promotion Agency (FFG).

REFERENCES

Beyer, H., 2001. *The Theory of Evolution Strategies*. Heidelberg: Springer Verlag.

Dichtl M., Gabriel C., Hennerbichler S., Seitz B., Priglinger SG., *EU konformes Hornhautbanking – eine Bestandsaufnahme Hornhautbank Linz*, Spektrum der Augenheilkunde, 2010;24(3): 166-173

Hirneiss C., Schumann RG., Grüterich M., Welge-Luessen UC., Kampik A., Neubauer AS., *Endothelial cell density in donor corneas: a comparison of automatic software programs with manual counting.* PubMed, 02/2007; 26(1):80-3

Jähne, B., 2012. *Digitale Bildverarbeitung und Bildgewinnung*. Heidelberg: Springer Verlag.

Kampik, A., Grehn, F., 2002. *Augenärztliche Therapie.* Stuttgart: Georg Thieme Verlag.

Loimayr, C., Schuster, G., Gabriel, C., Hennerbichler-Lugscheider, S., 2012. *Comparison of Cornea Donor Age and post-mortem Time to Endothelial Cell Count.* Poster presented at the European Association of Tissue Banks, Vienna.

Nagabhushana, S., 2005. *Computer Vision and Image Processing.* New Delhi: New Age International(P) Ltd., Pulishers.

Pachghare, V. K., 2008. *Comprehensive Computer Graphics*. Laxmi Publications

Rechenberg, I., 1973. *Evolutionsstrategie*. Friedrich Frommann Verlag.

Schwefel, H.-P., 1994. *Numerische Optimierung von Computer-Modellen mittels der Evolutionsstrategie.* Basel: Birkhäuser Verlag.

Winkler, S., 2009. *Evolutionary System Identification: Modern Concepts and Practical Applications*. Linz: Trauner Verlag.

AUTHORS BIOGRAPHIES

LISA OBRITZBERGER received her Bachelor of Science in Medical and Bioinformatics in 2012 from the University of Applied Sciences Upper Austria (UAS). She is a research associate at the Research & Development GmbH, in Hagenberg. Her research interests include medical image analysis, tissue engineering, machine learning and data mining. 2014 she received her Masters of Science in Biomedical Informatics from UAS.

SUSANNE SCHALLER received her Masters of Science in Biomedical Informatics in 2009 and 2010 from the University of Skoevde and Upper Austrian University of Applied Sciences. She is a research and teaching associate at the Research Center Hagenberg of UAS. Her research interests include system biology, medical image analysis, tissue engineering, data mining and machine learning.

VIKTORIA DORFER is a senior researcher in the field of bioinformatics at the Research Center Hagenberg, School of Informatics, Communications and Media. After finishing the diploma degree of bioinformatics in 2007 she was a team member of various projects and is currently pursuing her PhD studies in bioinformatics.

CLAUDIA LOIMAYR works at the Red Cross Blood Transfusion Service of Upper Austria in Linz. Her research interests include cornea transplantations and tissue engineering.

SIMONE HENNERBICHLER is Head of Production at the Red Cross Blood Transfusion Service of Upper Austria in Linz since 2006. She holds a master's and a PhD degree in Pharmacy from Karl-Franzens-University Graz and is responsible for development and production of blood and tissue products. Previously she was involved in stem cell and tissue regeneration research and teaching.

STEPHAN M. WINKLER received his PhD in engineering sciences in 2008 from Johannes Kepler University (JKU) Linz, Austria. His research interests include genetic programming, nonlinear model identification and machine learning. Since 2009, Dr. Winkler is professor at the Department for Medical and Bioinformatics at the University of Applied Sciences (UAS) Upper Austria at Hagenberg Campus; since 2010, Dr. Winkler is head of the Bioinformatics Research Group at UAS, Hagenberg.