Region and Contour Based Cell Cluster Segmentation Algorithm for In-Situ Microscopy

A. Sheehy[1], G. Martinez[1], J.-G. Frerichs[2], T. Scheper[2]

[1]Image Processing and Computer Vision Research Laboratory (IPCV-LAB), Universidad de Costa Rica

[2]Institute of Technical Chemistry (TCI), Leibniz Universität Hannover

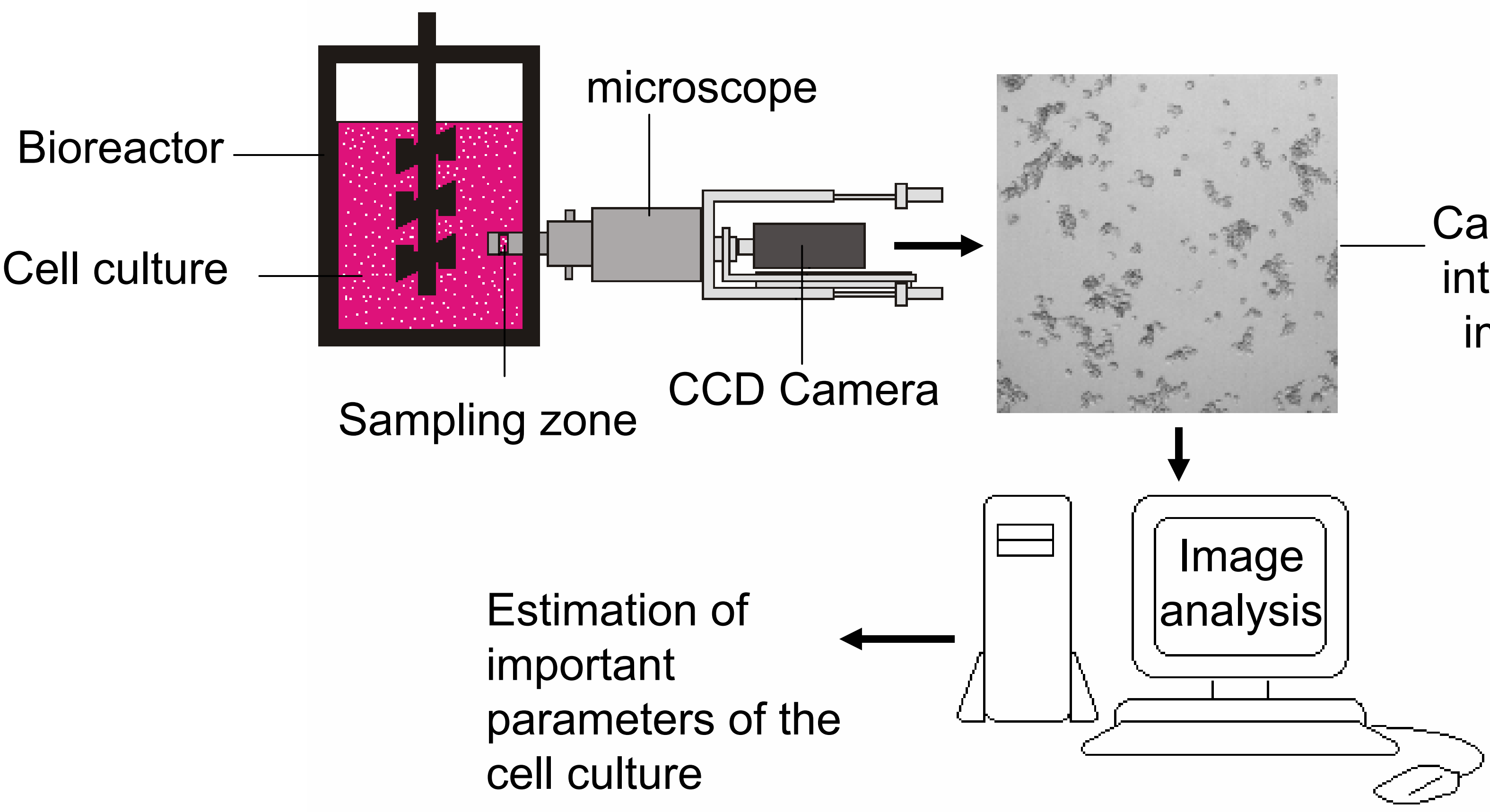
IEEE CCE, Mexico City, Mexico, November 2008

Topics

- Introduction
- Existing algorithms
- Proposed algorithm
- Experimental results
- Summary and conclusions

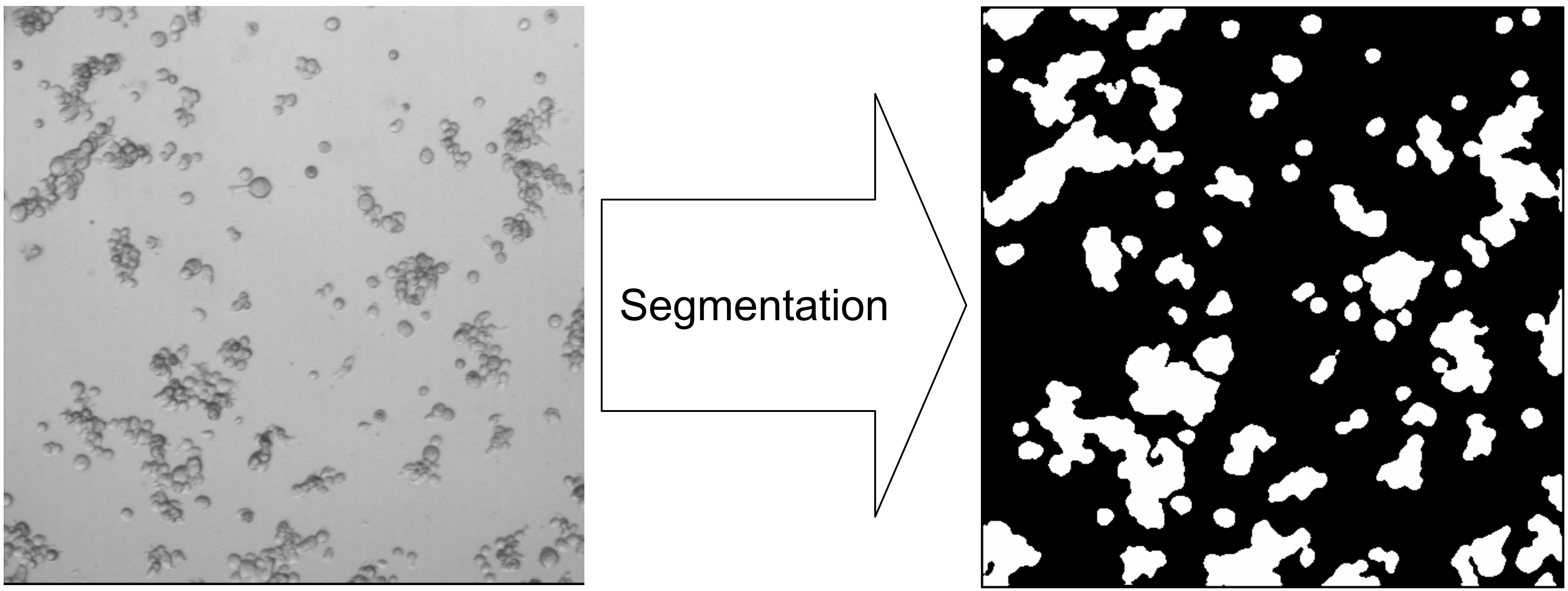
Introduction

In-situ microscopy



Introduction

Cell image segmentation



Intensity image $I(x,y)$, $x:1..L$, $y:1..M$, captured by the in-situ microscope

Binary image $I_B(x,y)$, $x:1..L$, $y:1..M$, with pixels belonging to the cell clusters in white and pixels belonging to the background in black

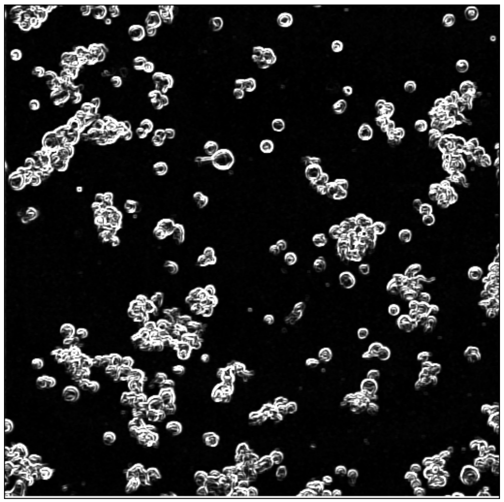
Existing Algorithms

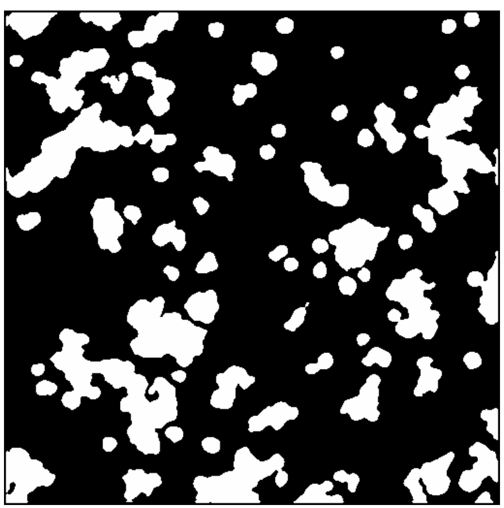
Cell cluster segmentation algorithms can be roughly divided into 4 main groups:

- Thresholding based segmentation algorithms
  - Espinoza’s algorithm: Latest contribution for segmenting BHK cells images captured by an in-situ microscope
- Contour based segmentation algorithms
- Region based segmentation algorithms
- Combinations of the previous

Espinoza's Algorithm

1) Global thresholding of the local variance





a. Estimate the local variance $\sigma^2(x,y)$ at each image position (x,y) using a 3x3 window:

$$\sigma^2(x,y) \cong \frac{1}{50} \cdot \sum_{m=-3}^{3} \sum_{n=-3}^{3} \left[I(x+m, y+n) - m(x,y) \right]^2$$

with:

$$m(x,y) \cong \frac{1}{49} \cdot \sum_{m=-3}^{3} \sum_{n=-3}^{3} I(x+m, y+n)$$

b. Globally threshold the local variance:

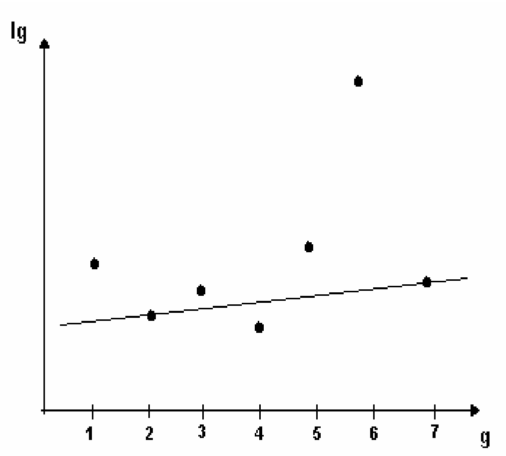
$$I_B(x,y) = \begin{cases} 0, \, if \;\; \sigma^2(x,y) \leq th_g \\ 1, \, if \;\; \sigma^2(x,y) > th_g \end{cases}$$

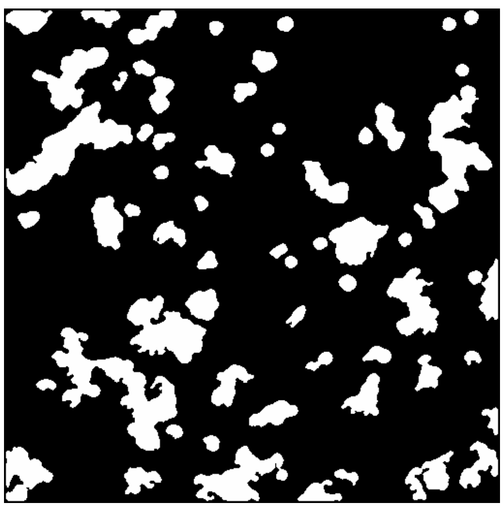
th_g is estimated applying a maximum likelihood algorithm (Kittler & Illingworth)

c. Apply a 5x5 median filter to I_B and eliminate small regions (<0.5%)

Espinoza’s Algorithm

2) Local thresholding of the intensity





a. Find for each segmented region *r* all the border pixels $I_g^{(r)}$, g:0...$G^{(r)}$.

b. Select from all the border pixels $I_g^{(r)}$ only those border pixels $I_h^{(r)}$, j:0...$H^{(r)}$ that have an intensity value similar to the background intensity value by applying a RANSAC algorithm

c. Locally threshold the intensity in each region *r* :

$$I_B(x,y)=\begin{cases}1, if\, I(x,y)\le th^{(r)}\\0, if\, I(x,y)> th^{(r)}\end{cases}\quad,\ \forall\ (x,y)\in r$$

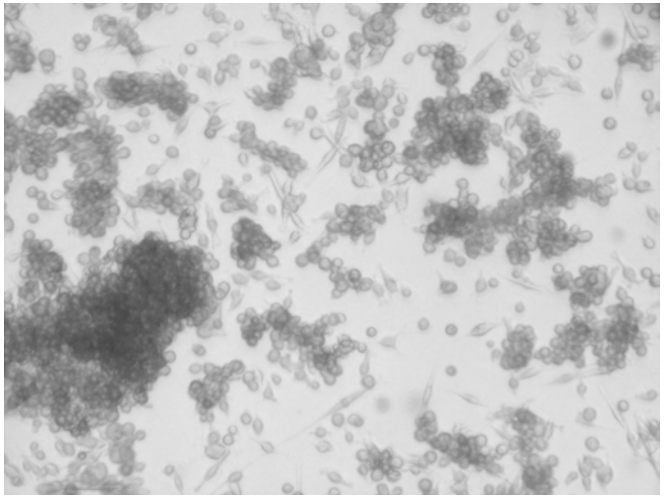
with: $$th_l^{(r)}=\frac{1}{H^{(r)}}\sum_{h=1}^{H^{(r)}} I_h^{(r)}$$

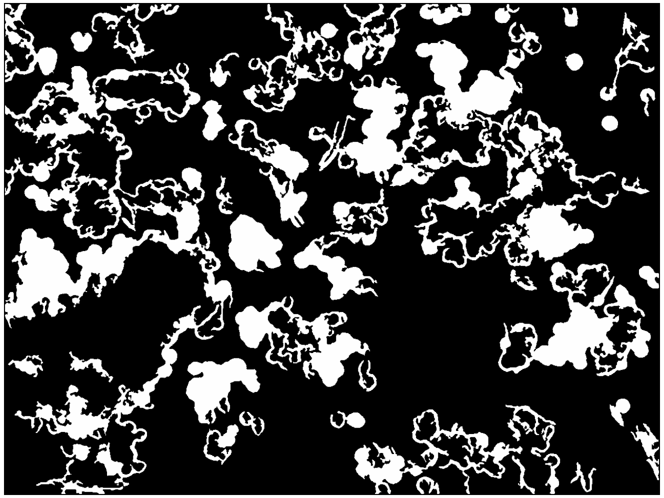
d. Apply a 5x5 median filter to I_B and eliminate small regions (<0.5%)

Espinoza's Algorithm

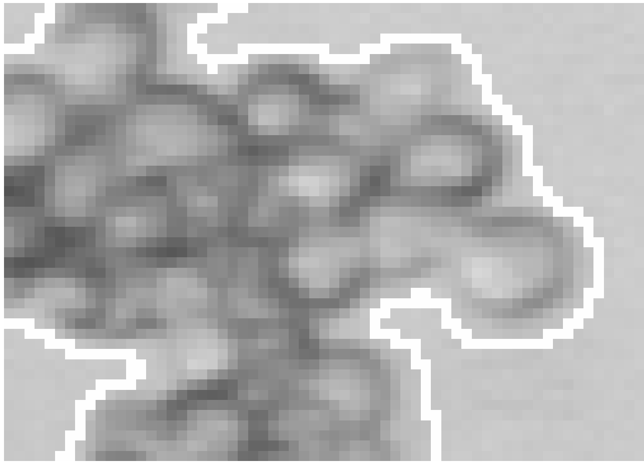
Problems

Low reliability in low contrast images:

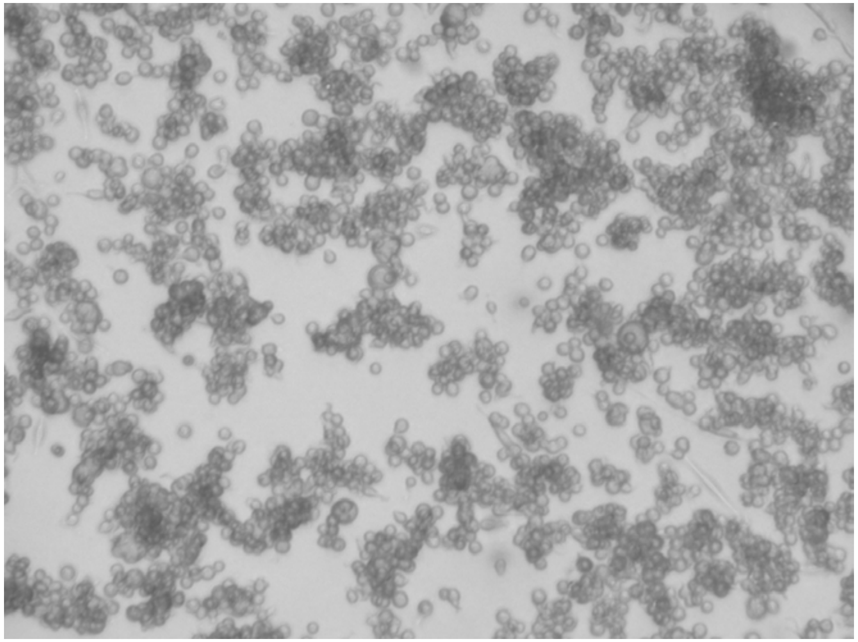




Low precision:



Proposed Algorithm



Because:

1. The background is homogeneous
2. The cell clusters are non homogeneous with darker intensity values and well defined contours

we propose:

To segment the background using a region and contour based approach and then to invert the resulting binary image to get the cell cluster regions

Proposed Algorithm

1) Seed selection

a. Select as background seed candidates those pixles that meet the following rules:

$$I(x,y) > m_2 - \sigma_2$$

$$\sigma^2(x,y) < 0.20 \cdot \sigma_2^2$$

where, m_2 and σ_2 are the mean and standard deviation of the pixels' intensity values in the background

Proposed Algorithm

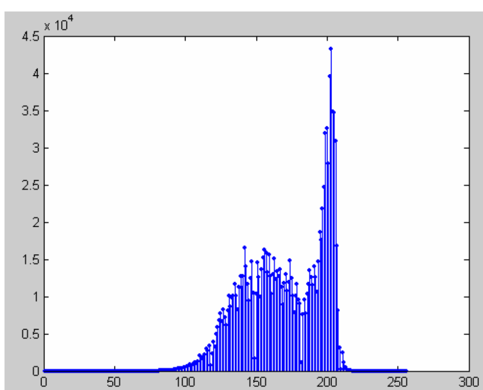
m_2 and σ_2 , as well as the mean m_1 and standard deviation σ_1 of the pixels' intensity values in the cell clusters are estimated modeling the probability density function of the intensity values *p(I)* as a sum of two weighted Gaussian density functions:

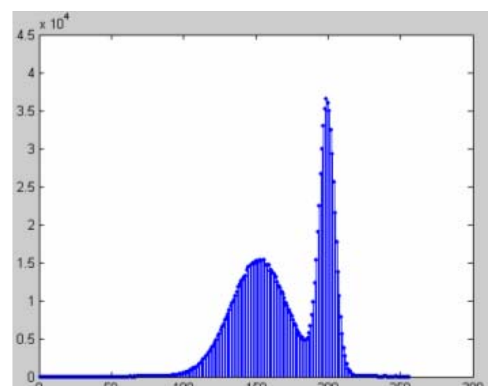
$$c_1 = \sum_{I=0}^{k} \frac{h(I)}{N} \qquad m_1 = \frac{1}{c_1(k)} \cdot \sum_{I=0}^{k} I \cdot \frac{h(I)}{N} \qquad \sigma_1^2 = \frac{1}{c_1(k)} \cdot \sum_{I=0}^{k} (I - m_1)^2 \cdot \frac{h(I)}{N}$$

$$c_2 = \sum_{I=k+1}^{255} \frac{h(I)}{N} \qquad m_2 = \frac{1}{c_2(k)} \cdot \sum_{I=k+1}^{255} I \cdot \frac{h(I)}{N} \qquad \sigma_2^2 = \frac{1}{c_2(k)} \cdot \sum_{I=k+1}^{255} (I - m_2)^2 \cdot \frac{h(I)}{N}$$

where *N* is the total number of pixels, *h(I)* is the number of pixels with intensity *I* and *k* is selected maximizing the following likelihood function:

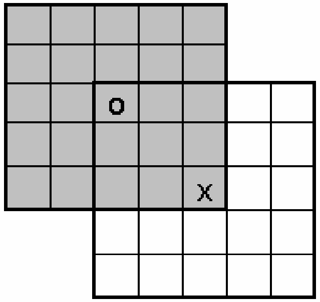
$$L_1(k) = N \cdot \sum_{j=1}^{2} c_j(k) \cdot \log(c_j(k)) - \frac{N}{2} \cdot \log(2\pi) - \frac{N}{2} \cdot \sum_{j=1}^{2} c_j(k) \cdot \log(\sigma_j^2(k)) - \frac{N}{2}$$





Proposed Algorithm

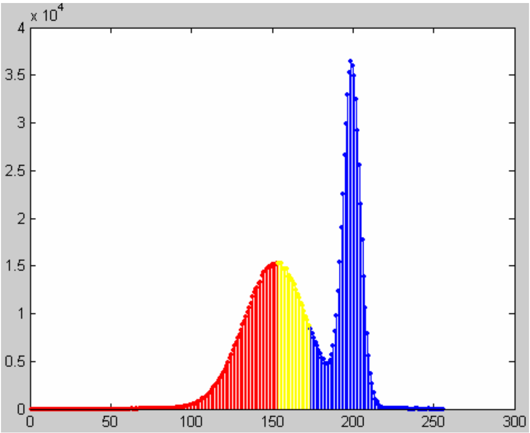
b. Reject those seed candidates that meet at least one of the following rules:



-One of its neighbors is a contour pixel

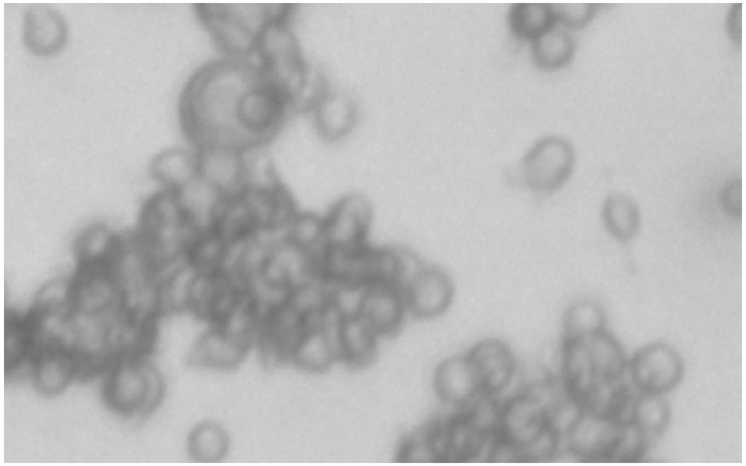
-One of its neighbors has $I(x,y) < m_1$ and also has a neighboring contour pixel

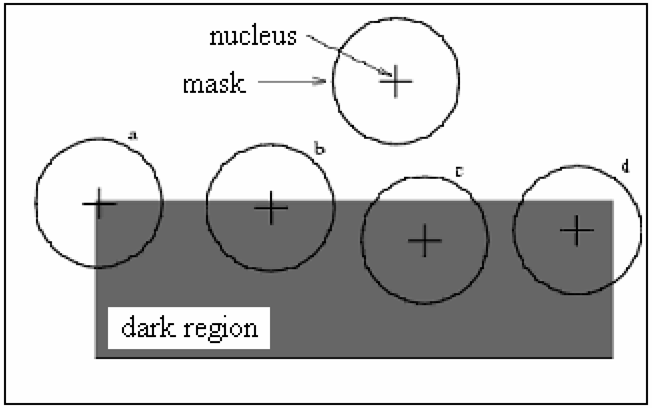
-One of its neighbors has $m_1 < I(x,y) < m_1 + \sigma_1$, $\sigma^2(x,y) > \sigma_2^2$ and a neighboring contour pixel

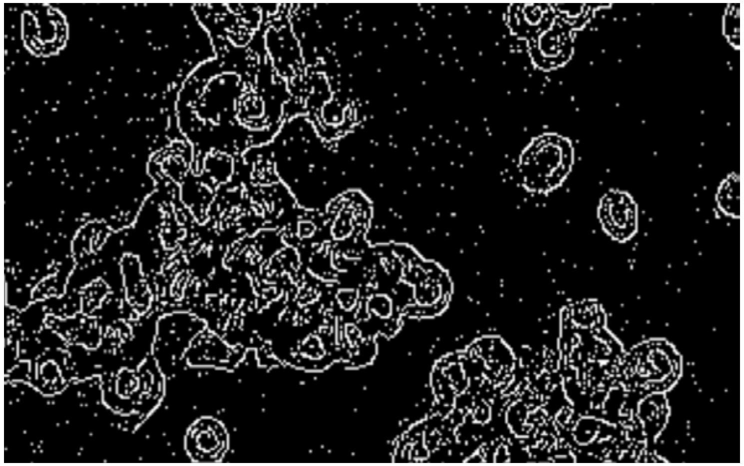


Proposed Algorithm

The contours in the image are obtained with the SUSAN algorithm:





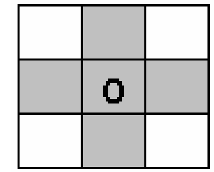


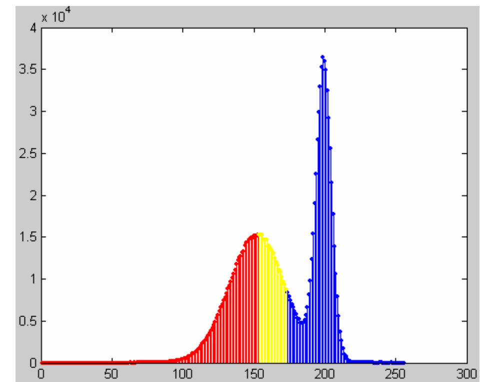
The intensity values inside a circular mask are compared with the intensity value in the center (nucleus). If there is a low amount of pixels inside the mask with an intensity value similar to the intensity value of the nucleus, then the nucleus is a contour pixel.

Proposed Algorithm

2) Region growing

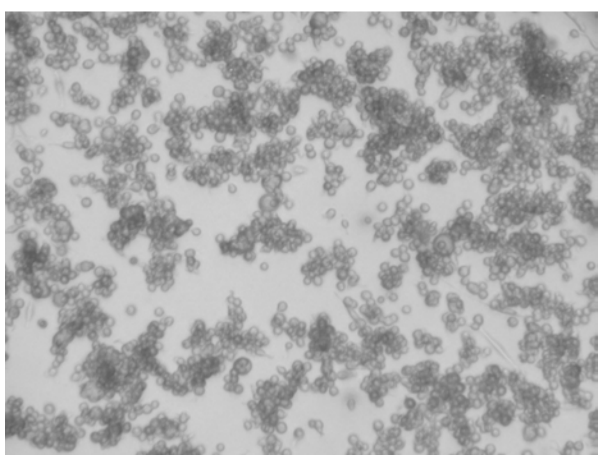
a. Add a pixel to a seed unless it meets at least one of the follwing rules:

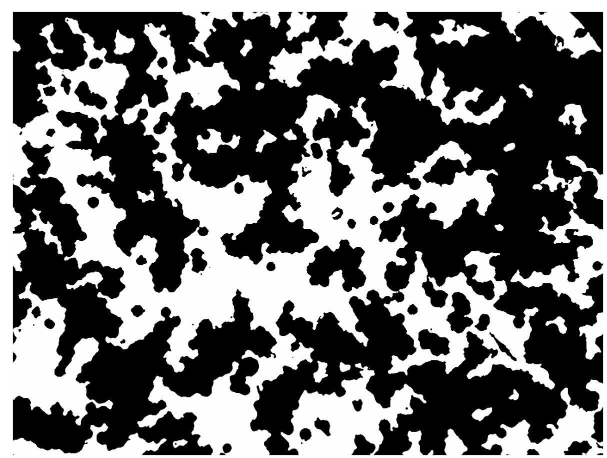




-It is a contour pixel.
-It has $I(x,y) < m_1$ and a neighboring contour pixel.
-It has $m_1 < I(x,y) < m_1 + \sigma_1$, $\sigma^2(x,y) > \sigma_2^2$ and a neighboring contour pixel

b. Repeat step **a.** until no more pixels can be added to the background region

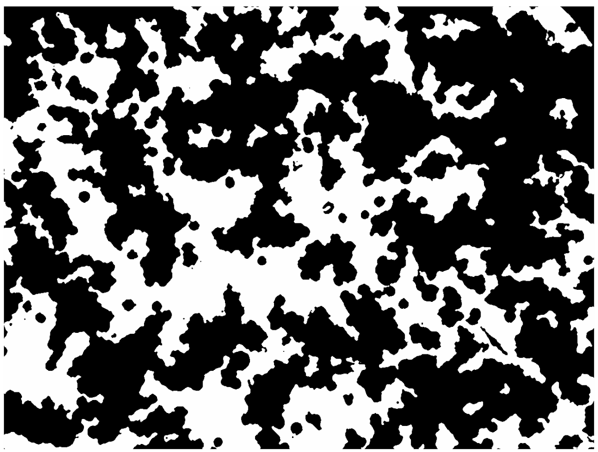

Intensity image


Binary image with segmented background

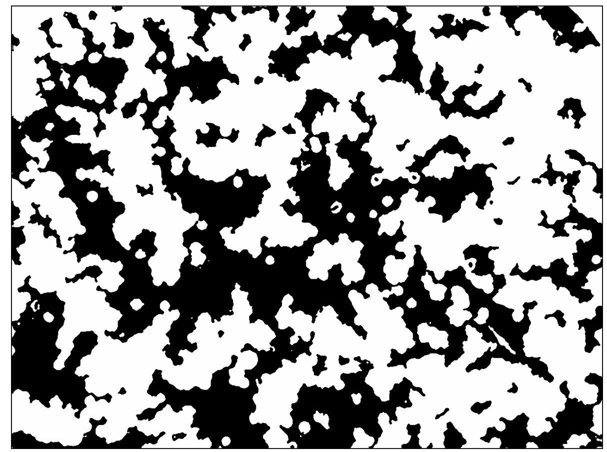
Proposed Algorithm

3) Invert the binary image

a. Invert the binary image with the segmented background.



Binary image with
segmented background



Inverted binary image

b. Apply a 5x5 median filter to the inverted binary image.

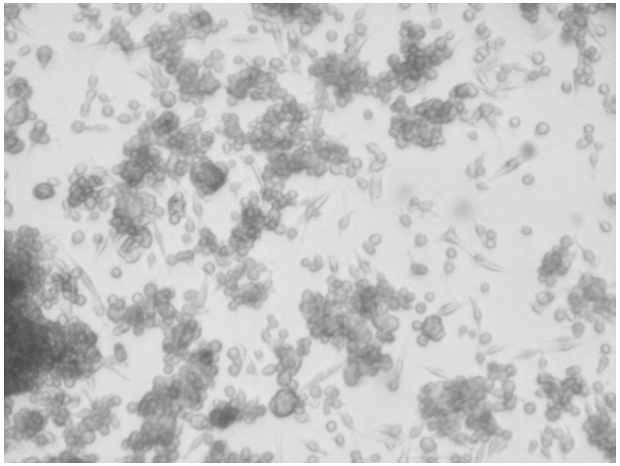
Experimetal Results

Reliability

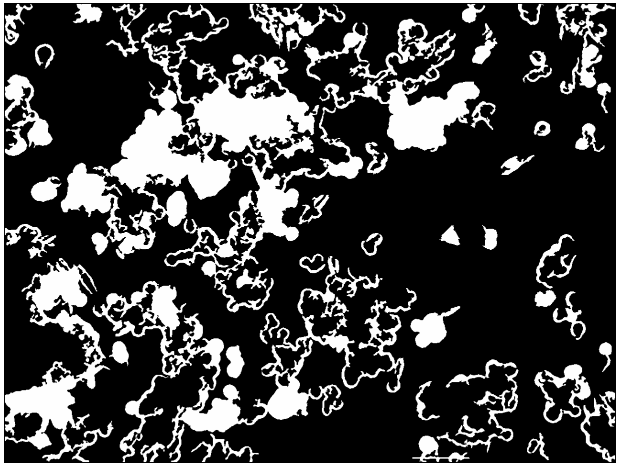
Experiments in 30 different BHK cell images:

*BHK: Baby Hamster Kidney cells

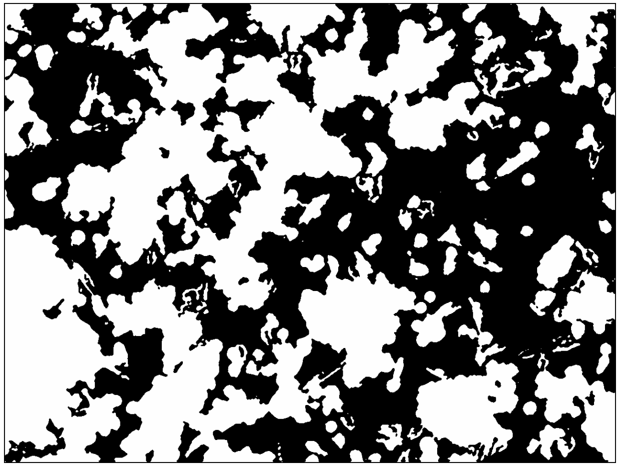
Algorithm	# images with most cell clusters correctly segmented	Percentage of images with most cell clusters correctly segmented
Espinoza's algorithm	20	66%
Proposed algorithm	30	100%



Low contrast intensity image



Espinoza's algorithm segmentation result



Proposed algorithm segmentation result

Experimetal Results

Precision

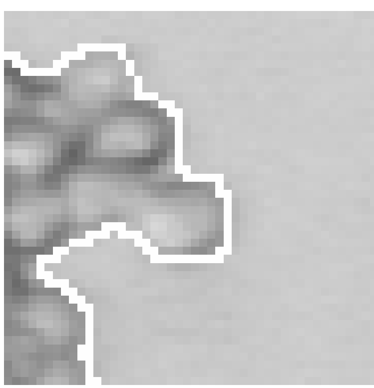
Distance from each boundary pixel *i*, *i:1…n*, of the automatically segmented regions to the closest boundary pixel of the manually segmented regions:

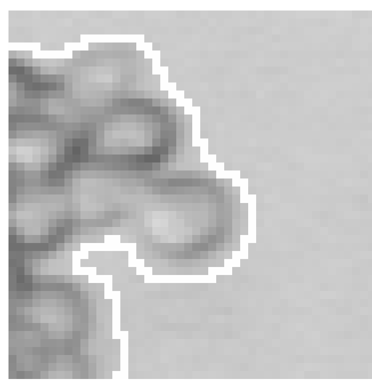
$$d_i^2 = (\Delta x_i)^2 + (\Delta y_i)^2$$

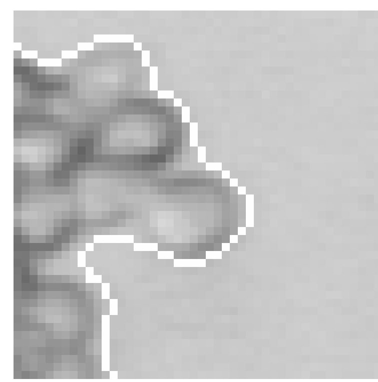
Mean square position error of the boundaries:

$$MSE = \frac{\sum_{i=1}^{n} d_i^2}{n}$$

Average MSE for Espinoza's algorithm	Average MSE for the proposed algorithm	Improvement
4,02 pixel^2	1,8 pixel^2	55,10%


Manual segmentation


Espinoza's algorithm segmentation result


Proposed algorithm segmentation result

Conclusions

The proposed segmentation algorithm is

- 33% more reliable
- 55% more accurate

than the Espinoza's algorithm in low contrast images

Thank you very much for your attention