

Image Thresholding Techniques for Localization of Sub-Resolution Fluorescent Biomarkers

Julien Ghaye,^{1*} Madhura Avinash Kamat,¹ Linda Corbino-Giunta,² Paolo Silacci,² Guy Vergères,² Giovanni De Micheli,¹ Sandro Carrara¹

¹Laboratory of Integrated Systems (LSI), Swiss Federal Institute of Technology, EPFL, Lausanne, Switzerland

²Federal Department of Economic Affairs Education and Research EAER, Agroscope, Bern, Switzerland

Received 11 July 2012; Revised 29 April 2013; Accepted 16 July 2013

Grant sponsor: NanoSys Project, Grant number: ERC-2009-AdG-246810

Grant sponsor: NutriChip Project (Financed by the Swiss Nano-Tera.ch Initiative and Evaluated by the Swiss National Science Foundation)

Correspondence to: Julien Ghaye, EPFL IC ISIM LSI1 INF 336 (Bâtiment INF), Station 14, CH-1015 Lausanne, Switzerland. E-mail: julien.ghaye@epfl.ch

Published online 16 September 2013 in Wiley Online Library (wileyonlinelibrary.com)

DOI: 10.1002/cyto.a.22345

© 2013 International Society for Advancement of Cytometry

• Abstracts

In this article, we explore adaptive global and local segmentation techniques for a lab-on-chip nutrition monitoring system (NutriChip). The experimental setup consists of Caco-2 intestinal cells that can be artificially stimulated to trigger an immune response. The eventual response is optically monitored using immunofluorescence techniques targeting toll-like receptor 2 (TLR2). Two problems of interest need to be addressed by means of image processing. First, a new cell sample must be properly classified as stimulated or not. Second, the location of the stained TLR2 must be recovered in case the sample has been stimulated. The algorithmic approach to solving these problems is based on the ability of a segmentation technique to properly segment fluorescent spots. The sample classification is based on the amount and intensity of the segmented pixels, while the various segmenting blobs provide an approximate localization of TLR2. A novel local thresholding algorithm and three well-known spot segmentation techniques are compared in this study. Quantitative assessment of these techniques based on real and synthesized data demonstrates the improved segmentation capabilities of the proposed algorithm. © 2013 International Society for Advancement of Cytometry

• Key terms

image processing; adaptive thresholding; fluorescence microscopy; Caco-2 cell line; toll-like receptor 2; nutrition

A significant part of our knowledge about biological processes, cell structures, functions, and mechanisms is acquired through direct optical observations. In the area of bioimaging, one of the common and principal tools used to make observations is fluorescence microscopy. This microscopy technique, combined with state-of-the-art signal processing methods (1–7), forms a powerful tool for cell analysis. Fluorescence bioimaging is extensively used because of two main characteristics. First, specific biological details can be targeted and highlighted through the use of molecule labeling by using specific fluorescent probes or dyes (8). Second, light microscopy has the advantage of being noninvasive. Thus, it allows us to observe live samples in vitro and study intracellular structures in situ. However, fluorescence microscopy has an inherent limitation. The spatial resolution of the imaging system is physically limited by the diffraction of light (1,2).

The NutriChip project is an example of a biological application that uses fluorescence microscopy (9–11). This project proposes a lab-on-chip (LoC) platform to investigate the effects of dairy food ingestion by feeding an artificial and miniaturized gastrointestinal track (μ GIT). Fluorescence microscopy is used to observe various sub-resolution biomarkers within the immune cell layer of the μ GIT. Finally, conclusions are drawn based on measurements made using dedicated image processing techniques.

For this study, an emulation of the μ GIT has been created to develop the image processing sub-system of NutriChip. Cell samples have been cultured and separated

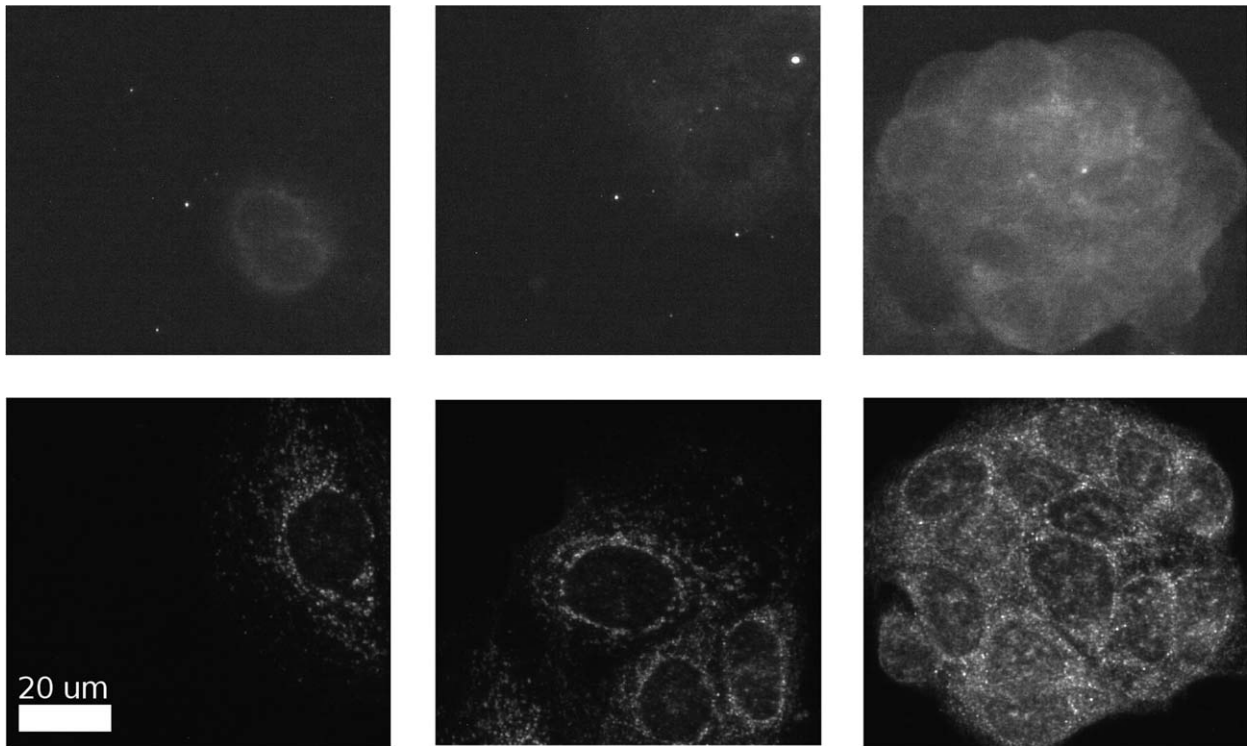


Figure 1. Example of stained Caco-2 cells expressing TLR2. Images on the top row belong to the negative control group (NCG) set. Their intensity has been quadrupled for display purposes. Images on the bottom row belong to the stimulated group (SG) set.

into two groups, a negative control group (NCG) and a stimulated group (SG). Unlike NCG cell samples, SG cell samples are stimulated to trigger a response from the innate immune system. This response is in turn monitored through fluorescence microscopy. Practically, the Caco-2 cell line is used and the stimulation is done using bacterial lipopolysaccharide (LPS) to set off the immune response, inducing the expression of toll-like receptor 2 (TLR2) (12). Fluorescent spots appear on SG images after dyeing the TLR2 with fluorophores, which are the resulting patterns created by many isolated sub-diffraction-sized fluorescent groups. Examples of SG and NCG images are presented in Figure 1.

In this study, we are addressing the two following problems. First, we need to identify if an observed sample has been stimulated or not. This biologically indicates whether the sample has had its immune system activated (i.e., the Caco-2 cells have been exposed to LPS). Second, if the sample is recognized as stimulated, recovering the spatial location of the TLR2 within the images is of general interest and can be used for instance in quantitative colocalization analysis (13). We are, thus, investigating the use of image thresholding and spot extraction techniques for answering these problems. Such techniques provide a mask of segmented pixels for any given image. From this mask, we are addressing the two problems that are at hand. The resulting amount of segmented pixels and their intensity are classification features used for sample identification, while segmentation blobs provide an approximate TLR2 localization.

Recent studies comparing segmentation algorithms (4), and more specifically spot detection methods (5,6), give us a good overview of the state-of-the-art and common practices in quantitative fluorescence microscopy. Global thresholding is a common and basic technique in which a threshold value must be computed for each image. Histogram-based algorithms such as Otsu's method (14), Ridler's method or isodata (15), maximum entropy (16), and the T-point algorithm (17) are often used to extract this threshold value (18–21). When applied to fluorescent spot extraction, these algorithms often work on images preprocessed by signal enhancement methods (5,6). These signal enhancement techniques are often based on wavelets (22) and mathematical morphology (MM) (23–26). In particular, the MM-based top-hat (TH) filter is widely used for removing low-frequency contributions such as background fluorescence, out-of-focus fluorescence, and cytoplasm auto-fluorescence (18,19,27,28). Aside from global thresholding, local thresholding techniques are also used for extraction of spots (29) by computing a specific threshold for each pixel based on its surroundings. Niblack's method (30) and its enhanced version by Sauvola and Pietikainen (31) fall in this category. Another popular method in cell segmentation is the watershed algorithm mainly used for separating overlapped objects of interest (28,32–34).

In this study, we have selected a subset of these methods to segment TLR2 in our images. The Otsu and T-point algorithms have been selected for their ability to separate bimodal and unimodal gray-level image histograms, respectively. Local

thresholding approaches are represented by Sauvola's technique. We also study the use of the TH filter as a preprocessing step to leverage the segmentation results in our application. On top of these methods, we propose a novel local thresholding technique using the TH filter as a way to identify local signal-to-noise ratios (LSNRs) and extract objects below a user-defined size. In total, we test and quantitatively compare nine different segmentation schemes, resulting from the combination of a TH preprocessing step and a global or local thresholding algorithm. On one hand, we compare their influence on the classification performances of the naïve Bayes classifier trained for separating NCG from SG images using experimental data. On the other hand, we compare the quality of the produced segmentation masks from synthetic images, using the various segmentation blobs as approximations of TLR2 locations. Note that in the course of the study, we also tested Ridler's algorithm as an alternative to Otsu's method. As the obtained results between these two methods were not significantly different, we have decided to leave out the results obtained using Ridler's algorithm. Thus, Otsu's method is used as the representative algorithm to improve and avoid redundant results and analysis in the article.

MATERIALS AND METHODS

Sample Preparation

Caco-2 cells were seeded in Lab-Tek chamber slides (Thermo Scientific Nunc, Waltham, MA) at a cell density of 1.2×10^{-5} cells/cm² in DMEM/F12 medium supplemented with 10% fetal bovine serum. After 21 days, differentiated cells were stimulated over a 24-h treatment period with lipopolysaccharide (LPS) from *Escherichia coli* bacteria (L4391, Sigma-Aldrich, St-Louis, MO) at a final concentration of 1 µg/ml. After treatment, the cells were rinsed with phosphate-buffered saline (PBS), pH 7.4, and fixed with 4% paraformaldehyde in PBS for 10 min at room temperature (RT). After rinsing, the cells were permeabilized with 0.1% Triton X-100 in PBS for 3 min, then rinsed twice with PBS, and treated with 10% goat serum albumin in PBS for 20 min at RT. The cells were then labeled with an anti-TLR2 antibody (dilution 1:200, H175, Santa Cruz Biotechnologies, Dallas, TX) for 1 h at RT in 10% goat serum in PBS. A negative control where the LPS stimulation was omitted has also been performed. The cells were then washed thrice with 0.1% Tween in PBS. After additional incubation with a FITC-conjugated goat anti-rabbit IgG (dilution 1:100, F2911, Santa Cruz Biotechnologies) for 30 min at RT, the cells were mounted with Vectamount mounting medium for fluorescence (Vector Laboratories, Burlingame, CA). Immunostaining was visualized and confirmed with Olympus Bx41 microscope equipped with a Color view III camera (Soft Imaging System, Münster, Germany).

Fluorescence Imaging System

An Eclipse Ti-S inverted microscope combined with a Plan Fluor objective (40X, NA=0.6; Nikon, Tokyo, Japan) was used to image the samples. A 465–495 nm band pass excitation filter, a 505 nm dichroic mirror, and 515–555 nm emission filter were used in the epi-fluorescence setup. The images were

taken using a noncooled black and white charged coupled device image sensor (ICX274AL, Sony, Tokyo, Japan) with a dynamic range of 12 bits. The imager photosites, or pixels, are 4.4 µm by 4.4 µm in size and form an array of 1,628 by 1,236 active pixels. Taking into consideration the 0.7× magnification of the relay lens in front of the imager, a distance of 4.4 µm in the image plane corresponds to ~157 nm in the object plane.

The NIS-Elements software (Nikon, Tokyo, Japan) was used to gather images and the image processing was performed using Matlab (R2009b, MathWorks, Natick, MA).

Segmentation Schemes

In image processing, thresholding is a basic tool used to segment objects from the background in a raster image. It provides as an output a map of the same size as the input image. This map indicates which pixels are considered part of foreground or part of the background. In our application, the pixels classified as foreground, or segmented pixels, are also referred to as fluorescent pixels as they are assumed to carry the fluorescent signal of the image.

In this study, we are considering nine different segmentation schemes; eight taken from the state-of-the-art and the proposed method. The group of eight state-of-the-art schemes consists of a set of four algorithms, working either on raw images or on TH prefiltered images. Algorithms embodying this set can either be:

- a global thresholding technique for which the threshold value is computed from the image histogram using the T-point algorithm (17), Otsu's algorithm (14), or Otsu's algorithm applied recursively twice (Otsu's algorithm is being run on the histogram of the pixel intensities extracted by a first run of the algorithm).
- a local thresholding algorithm, namely Sauvola's thresholding technique (31).

The proposed novel local thresholding algorithm is to be compared against these schemes. It processes raw images, sweeping the threshold value from low to high to extract relevant blobs of fluorescent pixels. Such blobs are characterized by a suitable pixel count and a high enough LSNR.

Top-Hat Transformation

It is generally admitted that segmentation algorithms designed for bright fluorescent spot extraction have to deal with the problem of uneven background (18,19,27,28). When objects of interest have similar shapes and sizes within an image, such as fluorescent spots, the TH filter (23–26) can be applied to enhance the signal.

The TH filter is based on the theory of MM. Two of the fundamental grayscale transformations of MM are the dilation and the erosion, whose definitions are reproduced hereunder

$$g = I \oplus B, \quad (1)$$

$$g = I \ominus B, \quad (2)$$

where \oplus denotes the dilation operator, \ominus the erosion operator, I is a gray-scale image, and B a binary structuring element.

The dilation operator replaces the value of each pixel in an image by the maximum value of its neighboring pixels. The structuring element B is responsible for defining where the neighbor pixels are located with respect to the processed pixel. Similarly, the erosion operator replaces the processed pixel value by the minimum value of the neighboring pixels. The resulting image after applying the TH filter is obtained by removing from the original image its lower envelope (26)

$$g = I - I \circ B, \quad (3)$$

where the lower envelope $I \circ B$ is computed using the MM opening operator \circ . The opening operator is defined by successively performing the erosion and then the dilation operations using a common structuring element B :

$$I \circ B = (I \ominus B) \oplus B. \quad (4)$$

The structuring element B used by the TH filter is usually a square window or a disk. The size of the structuring element, either defined by the width of the window or the disk radius, is related to the largest feature we are interested in. Practically, the TH filter acts as a background removal tool. Features in the original image that are typically smaller than the structuring element are kept. In the case of fluorescent spot extraction, we need to make sure that this size is greater than the typical spot size. As our application uses fluorophores bound to TLR2 observed through wide-field microscopy, a typical spot size is defined by the diffraction pattern observed on the imager generated by a sub-diffraction-sized point source.

T-Point Algorithm

The T-point algorithm (17) proposes an approach to setting thresholds for images with unimodal histograms. The proper operation of this automated method is based on two assumptions about the histogram. The background noise is the main pixel population that contributes to the single, major histogram peak in the low intensities. The remaining pixels of interest form the tail of the histogram. Such histograms can be decomposed into three parts: a steep rising slope, a steep descending slope, and a slow descending slope or tail.

The goal of the T-point binarization method is to model the descending slopes using two lines, one for the steep part and one for the tail. Consider a unimodal histogram featuring L bins indexed by $[0, \dots, L-1]$ where the peak has been identified at the bin index m . The steep and slow descending slopes can be located on the bin indexes $[m, \dots, k]$ and $[k, \dots, L-1]$, respectively, where $k \in [m+1, L-2]$. Two lines are fitted on the histogram using a least mean square approach, L_{steep} and L_{slow} for the steep and the slow descending slope, respectively. The T-point algorithm is searching for the threshold T matching the bin index k such that the total error of the fittings is minimized

$$T = \arg \min (e(k)), \quad (5)$$

where the total error $e(k)$ is the sum of the fitting errors of the lines L_{steep} and L_{slow}

$$e(k) = \varepsilon_{\text{steep}}(k) + \varepsilon_{\text{slow}}(k) \quad (6)$$

with

$$\varepsilon_{\text{steep}}(k) = \sum_{i=m}^k (h_i - \hat{h}_i)^2 \quad (7)$$

$$\varepsilon_{\text{slow}}(k) = \sum_{i=k}^{L-1} (h_i - \hat{h}_i)^2 \quad (8)$$

where h_i is histogram value at index i and \hat{h}_i is the estimated value of the histogram by either of the fitted lines.

Otsu's Algorithm

This method is a histogram-based, nonparametric method to automatically select a threshold level for a grayscale image. It aims at selecting a threshold by maximizing a criterion measure that evaluates the goodness of that threshold.

The original mathematical formulation and discussion can be found in Ref. (14); the following presents the equations and concepts behind the algorithm used for software implementation. The only input of the method is the normalized gray-level histogram, or probability distribution, generated from the image to be segmented. It has L bins that are dichotomized in two classes: C_0 gathering the bins indexed by $[0, k-1]$ and C_1 gathering the bins indexed by $[k, L-1]$. The gray level corresponding to the bin k indicates the selected threshold that must have its goodness evaluated.

The input histogram is regarded as a probability distribution of the pixels within the image. P_i is the probability that a pixel from the image falls into the bin number i , where $i \in [0, L-1]$. Let us also give the definition of the probabilities of class occurrence and the class mean level

$$\omega_0 = \Pr(C_0) = \sum_{i=0}^{k-1} p_i \quad (9)$$

$$\omega_1 = \Pr(C_1) = \sum_{i=k}^{L-1} p_i \quad (10)$$

$$\mu_0 = \sum_{i=0}^{k-1} \frac{i p_i}{\omega_0} \quad (11)$$

$$\mu_1 = \sum_{i=k}^{L-1} \frac{i p_i}{\omega_1} \quad (12)$$

We also define the total mean level of the original image by

$$\mu_T = \sum_{i=0}^{L-1} i p_i \quad (13)$$

Following a discriminant analysis (14), the following criterion measure is used to evaluate the goodness or separability of both classes at bin k

$$\eta = \frac{\sigma_B^2}{\sigma_T^2} \quad (14)$$

where σ_B^2 and σ_T^2 are the between-class variance and the total variance of levels, respectively, which are computed using

$$\sigma_B = \omega_0 \omega_1 (\mu_1 - \mu_0)^2 \quad (15)$$

$$\sigma_T = \sum_{i=0}^{L-1} (i - \mu_T)^2 p_i \quad (16)$$

Finding the optimal threshold k^* that maximizes η , or equivalently maximizes σ_B^2 , is reduced to a simple optimization problem

$$\max_{S^*} (\sigma_B^2(k)) \quad (17)$$

where S^* is the range of k over which the maximum is sought

$$S^* = \{k; \omega_0 \omega_1 > 0, \text{ or } 0 < \omega_0 < 1\} \quad (18)$$

Sauvola's Thresholding Technique

Sauvola's algorithm (31) acts as an edge-detection method based on a sliding window. It is targeting objects with a size similar to that of the window. Practically, the algorithm provides a threshold value for each pixel based on the mean and standard deviation of the neighboring pixel intensities. Consider an image I as an array of W by H pixels and a given pixel $p(x, y)$ within the image (i.e., $x \in [0, W-1]$ and $y \in [0, H-1]$). The threshold $t(x, y)$ is a function of the mean $m(x, y)$ and standard deviation $s(x, y)$ of the pixel intensities within a square window of size w by w around $p(x, y)$ and is computed using

$$t(x, y) = m(x, y) + (m(x, y) - 1)k \left(\frac{s(x, y)}{R} - 1 \right), \quad (19)$$

where k is a positive parameter, usually within the range $[0.2, 0.5]$, and R is the dynamic range of $s(x, y)$ (e.g., $R=0.5$ on normalized pixel intensities and $R=128$ on 8-bit images).

In Eq. (19), the standard deviation has a direct effect on the computed threshold value. Consider a high standard deviation among the pixels in the window, approaching R . The computed threshold value will then tend to the mean intensity. On the contrary, the threshold of the central pixel will tend to $m(x, y) + (1 - m(x, y))k$ for a window with pixel intensities having a low variance. In this case, the threshold is raised above the local mean of the pixel intensities, and the central pixel will more likely be classified as background.

Note that Sauvola's original binarization equation was designed for images featuring dark objects on a light background, while Eq. (19) is modified to account for light objects on a dark background, as fluorescence microscopy images require.

Novel Local Thresholding Technique

We propose a novel thresholding method for the extraction and localization of fluorescent spots that deal with the

limitations observed while working with other thresholding methods from the literature. Global thresholding methods using Otsu's algorithm or the T-point algorithm have a limited efficiency in extracting bright spots. First, fluorescent images principally feature unimodal histograms; this implies a trade off when selecting a threshold value as the signal and the background histogram contributions are merged. Second, the fluorescent signal is corrupted by auto-fluorescence of the cytoplasm and out-of-focus contributions of some fluorescent probes. This results in a random and uneven background in the fluorescent images.

As a consequence, we have developed the following thresholding method by taking advantage of the background removal effect of the TH filter, drawing inspiration from the watershed segmentation approach (32).

The purpose of this new method is to extract a certain amount of pixels around local maxima (i.e., identify blobs of fluorescent pixels). Ideally, the number of extracted blobs would equal the amount of TLR2 as each blob would only enclose the location of a receptor. Practically, these values cannot be achieved because of diffraction limits present in microscopy. Moreover, the possible proximity between two or more TLR2 indicates that more than one fluorophore can contribute to a single spot on an image (1,2).

The core concept of this method lies in sweeping threshold values starting from the lowest value in the image I and searches for blobs (four-connectivity connected components) of a suitable size containing local maxima of the image. The fluorescent pixels composing a given blob are characterized by an intensity value greater than a particular threshold, specific to this blob and the local maxima it is enclosing. The user specifies a maximum size for the blobs. This indicates if a blob needs to be broken down. Details of the algorithm flow can be found in Figure 2 and are explained hereunder.

1. Input and initialization: This method takes as input the original image I and sets a starting threshold value T equal to the lowest pixel intensity. The original image is thresholded with T and a fluorescent-background pixel map is generated, labeling the pixels either as fluorescent (white) or background (black). A maximal size S_{\max} in pixels for the blob size is also input.
2. Pixel categorization: The pixel categorization is the core of the algorithm. It is an iterative process. Each iteration is detecting the blobs, or connected components, within the fluorescent-background pixel map. Each detected blob is analyzed and categorized depending on its size. If the amount of pixels forming the blob is lower than S_{\max} , it is considered as having a suitable size and assumed to contain a local maximum. It is then removed from the pixel map and will be part of the output mask. Otherwise, it is kept in the pixel map for the next iteration. Before iterating, the threshold value T is increased. The iteration process stops once the amount of fluorescent pixels in the pixel map reaches 0.
3. Output: The output mask is a binary image. It contains only blobs of fluorescent pixels with a pixel count inferior

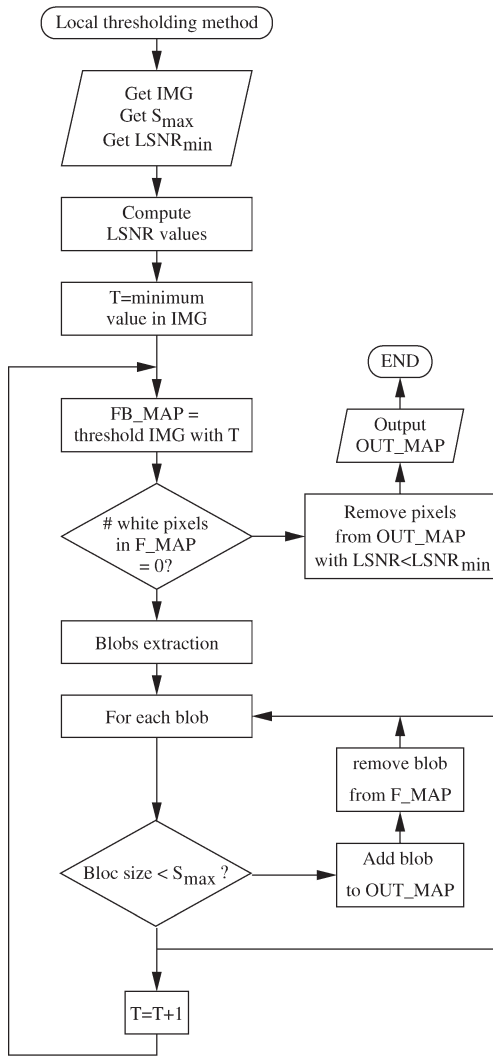


Figure 2. Flowchart of the proposed local thresholding technique.

to S_{max} each containing at least one local maximum. The output map locates the bright fluorescent spots within the image and provides an approximation of their location.

Our method leverages the capabilities of the MM operators by using them to compute an estimation of the LSNR

$$LSNR = \frac{I_{sig}}{I_{bg}}, \quad (20)$$

where I_{bg} is an estimation of the background approximated by the lower envelope of an original image I , and I_{sig} is an estimation of the useful signal of I using the TH filter.

$$I_{bg} = I \circ B \quad (21)$$

$$I_{sig} = I - I \circ B = I - I_{bg} \quad (22)$$

This estimation of the LSNR is used to filter the extracted output mask. Fluorescent pixels corresponding to a LSNR inferior to a given value $LSNR_{min}$ are trimmed out.

Thresholding Schemes Comparison Approach for Classification

To determine what segmentation schemes are the best for providing selected classification features used for distinguishing NCG from SG samples, we need two image datasets. An SG dataset featuring images of Caco-2 cells with induced TLR2 and a NCG dataset featuring images of unstimulated cells.

Provided with a sample image, a thresholding technique produces a segmentation mask and an observation $t_l = (x_l, c_l)$ is made, where c_l is the class of the sample (i.e., SG or NCG) and x_l is a measurement made on the image via the mask. We are considering two classification features for making the measurements:

- Amount of fluorescent pixels per cell: This value is the average number of fluorescent pixels per cell. In this work, the amount of cells in each image is assumed to be known. It can be measured in a preprocessing step, for example, through DNA staining. The amount of fluorescent pixels per cell is expected to be higher for SG images than for NCG images. Ideally, a segmentation scheme processing a NCG image should not segment any pixels and discard them all as background.
- Mean pixel intensity: This value is the normalized average pixel intensity of all the fluorescent pixels.

For a given feature, we obtain a set of observations $t = \{t_l, l=1, \dots, N\}$, which is the union $t = t_{SG} \cup t_{NCG}$ between the set of observations made on the SG dataset $t_{SG} = \{t_i : t_i = (x_i, SG), i=1, \dots, N_{SG}\}$ and the set of observations made on the NCG dataset $t_{NCG} = \{t_j : t_j = (x_j, NCG), j=1, \dots, N_{NCG}\}$.

The first point of comparison between the segmentation schemes is an evaluation of the effect size between the NCG and SG group using the d statistic (Hedge's g) (35). Practically, the effect size is evaluated using

$$d = \frac{\overline{x_{SG}} - \overline{x_{NCG}}}{S_{pooled}} \quad (23)$$

$$s_{pooled}^2 = \frac{(N_{SG} - 1)s_{SG}^2 + (N_{NCG} - 1)s_{NCG}^2}{N_{SG} + N_{NCG} - 2}, \quad (24)$$

where $\overline{x_{SG}}$ and $\overline{x_{NCG}}$ are the sample means made on the observation sets t_{SG} and t_{NCG} , respectively, and s_{SG}^2 and s_{NCG}^2 are their associated variance. The standard error on this d statistic is computed using

$$se_d = \sqrt{\frac{N_{SG} + N_{NCG}}{N_{SG} N_{NCG}} + \frac{d^2}{2(N_{SG} + N_{NCG} - 2)}} \quad (25)$$

Such an effect size can be computed for each segmentation scheme based on the observations sets. As a high effect size indicates a good separation of the two classes (SG and NCG), we can identify for each feature the best performing segmentation schemes for classifying the Caco-2 cell samples by putting the effect sizes side by side.

We are also considering receiver operating characteristic (ROC) analysis as a comparison tool. Consider the naive Bayes classifier trained with the observation sets \mathbf{t}_{SG} and \mathbf{t}_{NCG} . A common performance metric of the resulting trained classifier is the area under the ROC curve (AUC). It is used in this study to compare the performances of the various Bayes classifiers trained using the observation sets generated from the segmentation results of the algorithms.

The method used to estimate the various AUC is that described in Ref. 36). It provides a nonparametric estimator for the AUC and an estimate for its variance based on leave-pair-out bootstrap scheme, which makes it ideal for datasets with few samples. Practically, B bootstrap replications $\mathbf{T} = \{\mathbf{t}_b^*, b=1, \dots, B\}$ are obtained by resampling the set of observations \mathbf{t} using a balanced bootstrap mechanism (37). Each replication \mathbf{t}_b^* is then used to train a naive Bayes classifier, effectively producing B scoring functions $h_{\mathbf{t}_b^*}(x)$. The procedure selects a pair of observations (t_i, t_j) —one from each class SG and NCG—and averages its contribution to the Wilcoxon–Mann–Whitney statistics over all the possible bootstrap replications. Formally, we have

$$\widehat{\text{AUC}} = \frac{1}{N_{\text{SG}} N_{\text{NCG}}} \sum_{i=1}^{N_{\text{SG}}} \sum_{j=1}^{N_{\text{NCG}}} \sum_{b=1}^B \frac{I_i^b I_j^b H(h_{\mathbf{t}_b^*}(x_i) - h_{\mathbf{t}_b^*}(x_j))}{\sum_{b'=1}^B I_i^{b'} I_j^{b'}} \quad (26)$$

where H is the Heaviside step function and $I_i^b I_j^b$ equals 1 unless either of the observations forming the pair (t_i, t_j) is contained in \mathbf{t}_b^* , in which case it equals 0 and voids the contribution. On the other hand, a standard error on $\widehat{\text{AUC}}$ can be computed using

$$se_{\widehat{\text{AUC}}} = \frac{1}{N_{\text{SG}}^2} \sum_{i=1}^{N_{\text{SG}}} \hat{U}_{1i}^2 + \frac{1}{N_{\text{NCG}}^2} \sum_{j=1}^{N_{\text{NCG}}} \hat{U}_{2j}^2, \quad (27)$$

where \hat{U}_{1i}^2 and \hat{U}_{2j}^2 can be obtained from the bootstrap replications \mathbf{t}_b^* , similar to $\widehat{\text{AUC}}$. Details of their calculation are left out of this article for clarity and we refer the reader to Ref. 36) for more details.

Similar to the effect size, an estimation of the AUC can be computed for each segmentation scheme. The closer the area is to 1, the better the trained naive Bayes classifier is. In turn, for each classification feature (amount of fluorescent pixels per cell or mean pixel intensity), we can evaluate the performances of the segmentation schemes by directly comparing the effect sizes and classifier performances.

Thresholding Schemes Comparison Approach for Localization

After a sample image has been classified as SG, we are interested in using the blobs of fluorescent pixels as an approximation of the stained TLR2 locations. Comparing algorithms on this ground based on our SG datasets is not feasible as we do not precisely know the location of every fluorescent probes bound to the TLR2. To overcome this, we are using an image simulator (38). It generates synthetic fluorescently stained cell populations and simulates the imaging pro-

cess. This simulator has been configured so that it matches the experimental setup and imaging conditions.

Consider a synthetic SG dataset $\mathbf{I}_{\text{syn}} = \{I_n, n=1, \dots, N\}$ built from N synthetic images. Such dataset contains images each featuring a random number of stained cells within a given range for which the precise location (i.e., pixel location within the image) of each fluorescent probe is known. Overall the dataset features N_{cells} , each stained by N_p probes over N images. A segmentation scheme processing \mathbf{I}_{syn} generates a set of segmentation masks $M_s = \{M_n, n=1, \dots, N\}$ that can be partitioned in N_{blobs} segmentation blobs using four-connectivity component labeling. These segmentation results are analyzed and compared using the following metrics:

- Recovered probes: Consider the number of probes N_{rec} that are recovered by M_s over the synthetic dataset, knowing that a probe is deemed recovered if its pixel location within the image belongs to a blob of fluorescent pixels. The recovered probes value corresponds to the relative amount of recovered probes over the synthetic dataset N_{rec}/N_p .
- Blobs per cell: This value is a direct measure of the average amount of segmentation blobs per cell over the synthetic dataset $N_{\text{blobs}}/N_{\text{cells}}$.
- Blobs without probes: This value indicates the fraction of blobs that do not recover any probes over the synthetic database.
- Blob size: This is the average size in pixels of the segmentation blobs over the synthetic dataset.
- Probes per blob: Knowing the location of the probes, we can determine the average amount of probes that are recovered by a single segmentation blob.
- χ^2 : This value is the chi-squared (χ^2) histogram distance metric (39) used to quantify whether an image segmentation of a sample featuring fluorescent probes is relevant and contains useful information.

Consider the N_p probes and their location within each image of the set \mathbf{I}_{syn} . One can define a set of masks $\mathbf{M}_p = \{L_n, n=1, \dots, N\}$, where L_n is a mask identifying the location of the probes within I_n using

$$L_N(x, y) = \begin{cases} 1, & \text{if } p(x, y) \text{ encloses at least one probe} \\ 0, & \text{otherwise} \end{cases}, \quad (28)$$

where $p(x, y)$ is a pixel of I_n . In other words, \mathbf{M}_p is an ideal set of segmentation masks featuring the minimum number of segmentation pixels such that the amount of recovered probes equals 100% with 0 blobs without probes. From \mathbf{M}_p , one can compute the following set of images

$$\mathbf{D}_p = \{I_n * L_n, n=1, \dots, N\}, \quad (29)$$

where $*$ is a pixel-wise multiplication between the image I_n and the location mask L_n . Finally, we obtain the intensity distribution of the pixels where the probes are located by taking the histogram P of the non-null pixel values of \mathbf{D}_p . That is, the histogram P represents the distribution of pixels selected by the ideal set of segmentation masks \mathbf{M}_p .

Similarly, provided with a set of segmentation masks M_s generated by an algorithm, one can extract the intensity histogram S of the segmented pixels using

$$D_S = \{I_n * M_n, n=1, \dots, N\} \quad (30)$$

For our purpose, the χ^2 histogram distance value is defined as the bin-to-bin distance between the ideal histogram P and the histogram S . It is computed using

$$\chi^2(S, P) = \sum_i \frac{(S_i - P_i)^2}{S_i + P_i}, \quad (31)$$

where P_i and S_i are the frequencies reported in the i th bin of P and S , respectively.

This χ^2 histogram distance gives us a single figure of merit to evaluate the quality of a set of segmentation masks. The less a set of segmentation masks M_s is recovering basic information carried by the fluorescent probes, the greater the value of χ^2 .

Practically, we are interested in recovering as many probes as possible. Thus, we favor any algorithm that generates the greatest recovered probe values, other things being equal. However, we consider segmentations with small, spatially contained blobs better in localizing fluorescent spots compared to segmentations featuring big, spread blobs. So, having a small average blob size, characterized by few probes per blobs is favored compared to having a great number of recovered probes. Also, the amount of blobs without probes and χ^2 distances are monitored to ensure that the segmentation masks are not extracting irrelevant pixels.

RESULTS

Algorithms Comparison on Real Samples for Classification

Following the segmentation schemes comparison approach for classification presented in the Material and Methods section, we have created a SG dataset of 22 images and a NCG dataset of 15 images. Each image features a variable amount of imaged Caco-2 cells ranging from 1 to 20, and they all have been taken in the same experimental and imaging conditions. The first row of Figure 3 depicts representative sample images from both datasets.

The various algorithms used in the considered segmentation schemes are parameterized as follows:

- Top-hat transformation: The structuring element used for this operator is a disk of 21 pixels in diameter.
- T-point algorithm: The histograms used to set thresholds are computed using a bin width of 32, as smaller bins produce big discontinuities in the histograms frequencies, which in turn prevent proper lines fitting on the descending slopes. As our images have a 12-bit resolution, this divides the histograms into 128 bins.
- Otsu's algorithm: The histogram bin width used is 16, resulting in histograms divided into 256 bins. The purpose of this selection is that it has no significant impact on the

results compared to a bin size of 1 while enabling the method to be less computationally intensive.

- Sauvola's local thresholding technique: The algorithm parameters are set as follows: dynamic range $R = 0.5$ (our images have normalized pixel values), window width $w = 41$ (radius = 20), and $k = 0.34$ (as suggested in Ref. 31)).
- Proposed local thresholding technique: The maximum size for any fluorescent blob is $S_{max} = 15$ and only fluorescent pixels with a LSNR greater than 1.7 are kept.

Each row in Figure 3 displays the resulting segmentation masks of one of the nine considered segmentation schemes, except for the first row, which depicts the original, unprocessed images. Note that the Otsu-based schemes are only represented by the second row for space consideration. When dealing with SG images, the masks produced by Otsu (not recursively)-based schemes resemble the ones produced by T-point-based schemes. As for the one using Otsu recursively, the masks are similar to the one presented on the second row, only the blobs of fluorescent pixels appear slightly "fatter." The most important results on the second row are the masks of the NCG images. They are typical of all the Otsu-based schemes and give an idea of how these schemes are handling image with low fluorescence and few spots.

Table 1 presents the average amount of fluorescent pixel recovered by each segmentation scheme over both datasets. The resulting effect sizes and performance of trained naive Bayes classifiers (AUC) are reported in Table 1 with a 95% confidence interval.

Similarly, Table 2 deals with the mean pixel intensity and related effect sizes and AUC. In this case, we also consider an extra method called mean, which considers all the pixels of an image as segmented. That is, the resulting mean pixel intensity (first row in Table 2) using this method is the average image intensity divided by the number of cells.

Finally, Figures 4 and 5 report graphically the various computed effect sizes along with the various estimated AUC from Tables 1 and 2 for an easy comparison.

Algorithms Comparison on Synthetic Images for Localization

We apply the comparison approach for localization presented in the Material and Methods section on a synthetic SG dataset built from 100 images. For each considered scheme, we perform a parameter space exploration and analyze how the six comparison metrics are influenced. A segmented sample of a synthetic cell is depicted in Figure 6, featuring randomly distributed TLR2 stained with fluorescent probes, uneven background fluorescence, and cytoplasm autofluorescence. Similar to the real datasets, each image from the synthetic dataset features a varying amount of cells, between 1 and 10 in this case.

We are interested in comparing the segmentation quality on SG images of the best-performing segmentation schemes for classification. Thus, the three schemes we are considering here are (a) global thresholding on TH filtered images, (b) Sauvola's thresholding technique, and (c) the proposed local

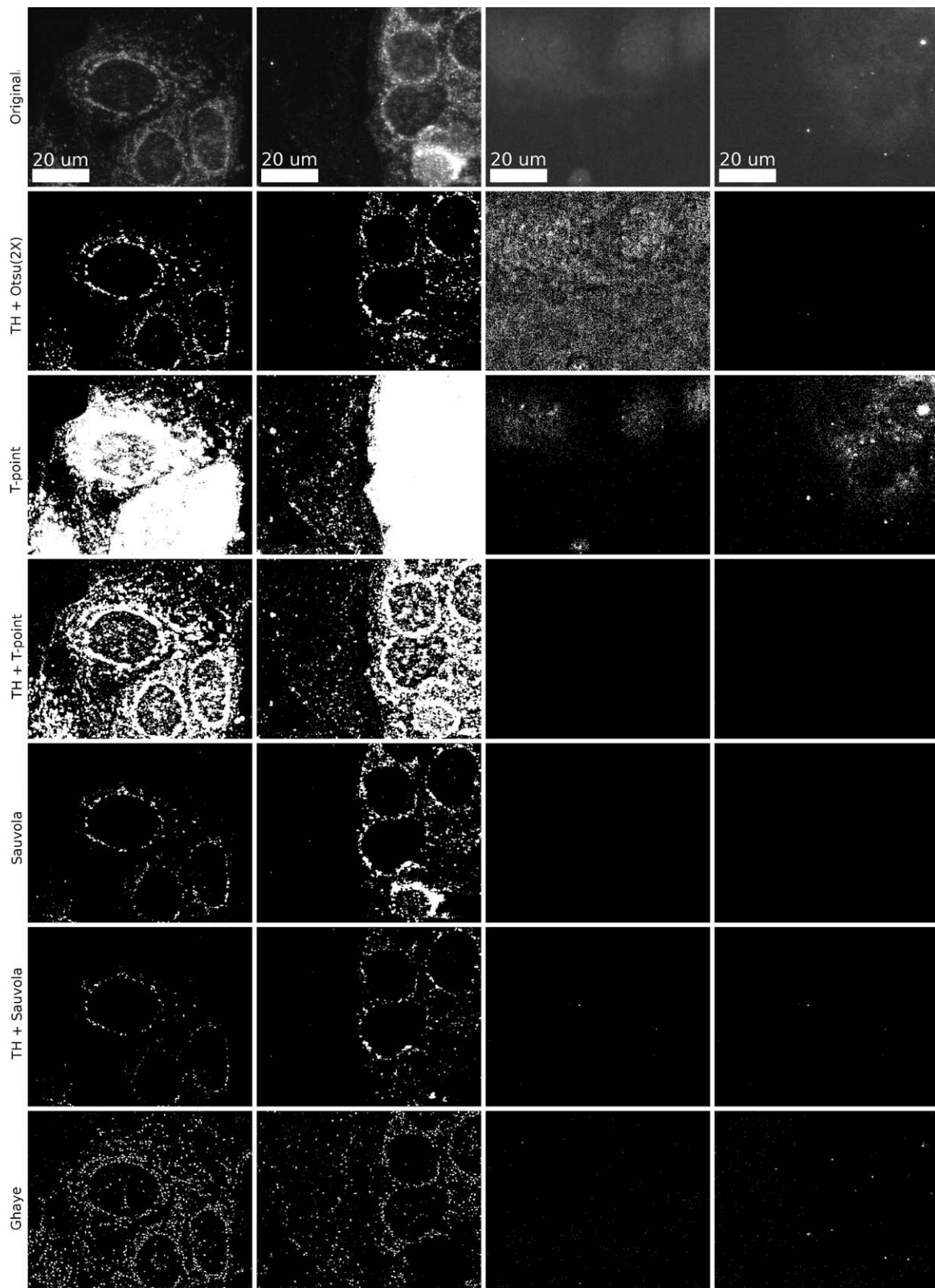


Figure 3. The first row depicts unprocessed sample images of stained Caco-2 cells. Two representative images from the SG dataset and the NCG dataset (their intensity has been quadrupled in the figure for display purposes) are shown. Subsequent rows depict segmentation masks computed by different segmentation schemes.

Table 1. Amount of fluorescent pixels per cell on each dataset, related effect size, and AUC of trained naive Bayes classifiers for different segmentation schemes.

METHOD	SG PIXEL PER CELL (SD)	NCG PIXEL PER CELL (SD)	EFFECT SIZE (1.96*SE)	AUC (1.96*SE)
Mean	–	–	–	–
T-point	28,332 (19,752)	2,626 (4,354)	1.69 (0.26)	0.925 (0.026)
Otsu	15,914 (9,291)	33,004 (26,254)	−0.97 (0.23)	0.753 (0.028)
Otsu(2X)	3,178 (1,826)	10,758 (8,939)	−1.33 (0.25)	0.839 (0.030)
Sauvola	819 (689)	2 (5)	1.56 (0.25)	0.987 (0.030)
TH + T-point	17,471 (8,556)	973 (732)	2.54 (0.31)	0.993 (0.030)
TH + Otsu	7,618 (3,188)	43,676 (44,449)	−1.31 (0.24)	0.925 (0.026)
TH + Otsu(2X)	1,244 (566)	14,260 (17,579)	−1.20 (0.24)	0.998 (0.030)
TH + Sauvola	327 (260)	2 (4)	1.65 (0.27)	0.985 (0.028)
Ghaye	2,593 (1,504)	75 (118)	2.20 (0.29)	0.981 (0.028)

Table 2. Mean pixel intensity on each dataset, effect size, and AUC of trained naive Bayes classifiers for different segmentation schemes.

METHOD	SG PIXEL INTENSITY (SD)	NCG PIXEL INTENSITY (SD)	EFFECT SIZE (1.96*SE)	AUC (1.96*SE)
Mean	0.114 (0.048)	0.059 (0.012)	1.48 (0.25)	0.938 (0.037)
T-point	0.216 (0.058)	0.096 (0.027)	2.55 (0.31)	0.970 (0.033)
Otsu	0.271 (0.087)	0.073 (0.018)	2.96 (0.35)	1.000 (0.036)
Otsu(2X)	0.422 (0.128)	0.105 (0.085)	2.87 (0.34)	0.964 (0.034)
Sauvola	0.554 (0.054)	0.806 (0.167)	−2.27 (0.29)	0.863 (0.034)
TH + T-point	0.242 (0.074)	0.091 (0.021)	2.62 (0.32)	0.989 (0.033)
TH + Otsu	0.316 (0.089)	0.092 (0.104)	2.40 (0.30)	0.955 (0.032)
TH + Otsu(2X)	0.490 (0.117)	0.236 (0.339)	1.12 (0.24)	0.972 (0.038)
TH + Sauvola	0.444 (0.036)	0.769 (0.159)	−3.18 (0.24)	0.843 (0.032)
Ghaye	0.316 (0.094)	0.111 (0.029)	2.79 (0.33)	1.000 (0.036)

thresholding technique. The results are presented in Figures 7–9, respectively. Figure 7 details the behavior of the global thresholding method applied on TH filtered images by plotting the various comparison metrics with respect to the thresholding value. The average normalized threshold values computed using the T-point algorithm, Otsu’s algorithm, and Otsu’s algorithm applied recursively are 0.02, 0.14, and 0.33, respectively. Similarly, Figure 8 presents the response of the metrics using Sauvola’s thresholding techniques by varying the k parameter and the window radius. Finally, Figure 9

presents the response of the metrics using the proposed technique by varying the S_{max} and $LSNR_{min}$ parameters.

DISCUSSION

Classification Performances on Real Images

In this section, we discuss the usage of the amount of pixel per cell and mean pixel intensity measurements as classification features to distinguish between SG and NCG. The results of each segmentation scheme are analyzed and a

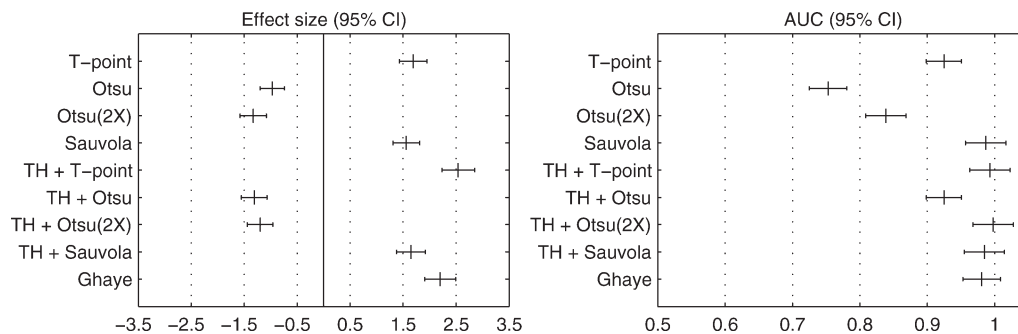


Figure 4. Effect size and AUC computed from the amount of fluorescent pixels per cell measurements made by each segmentation scheme on both SG and NCG datasets. The effect size is a measure of the distance between both datasets and the AUC estimates are measures of the performance of naive Bayes classifiers trained on the measurements.

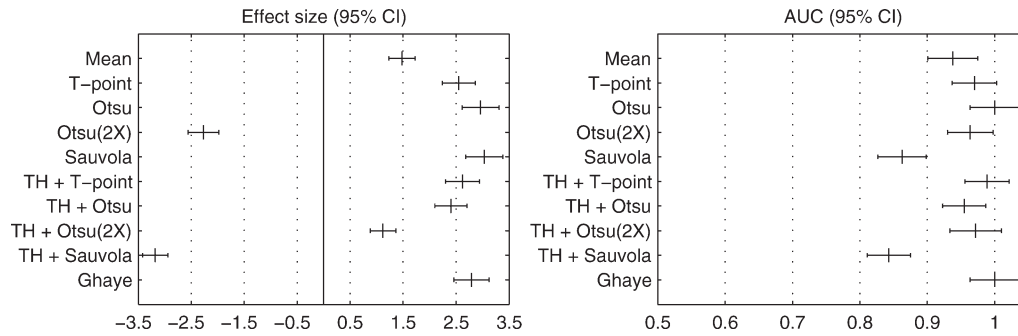


Figure 5. Effect size and AUC computed from the mean pixel intensity measurements made by each segmentation scheme on both SG and NCG datasets. The effect size is a measure of the distance between both datasets and the AUC estimates are measures of the performance of naive Bayes classifiers trained on the measurements.

conclusion on the best-performing approaches for our application is drawn. The following discussion is supported by the various masks presented in Figure 3 and by the numerical results presented in Tables 1 and 2.

- Mean: Using only the average pixel intensity per cell, without segmentation, as a classification feature is not reliable. Table 2 shows that the average pixel intensity is 0.114, which is the smallest value compared with the other

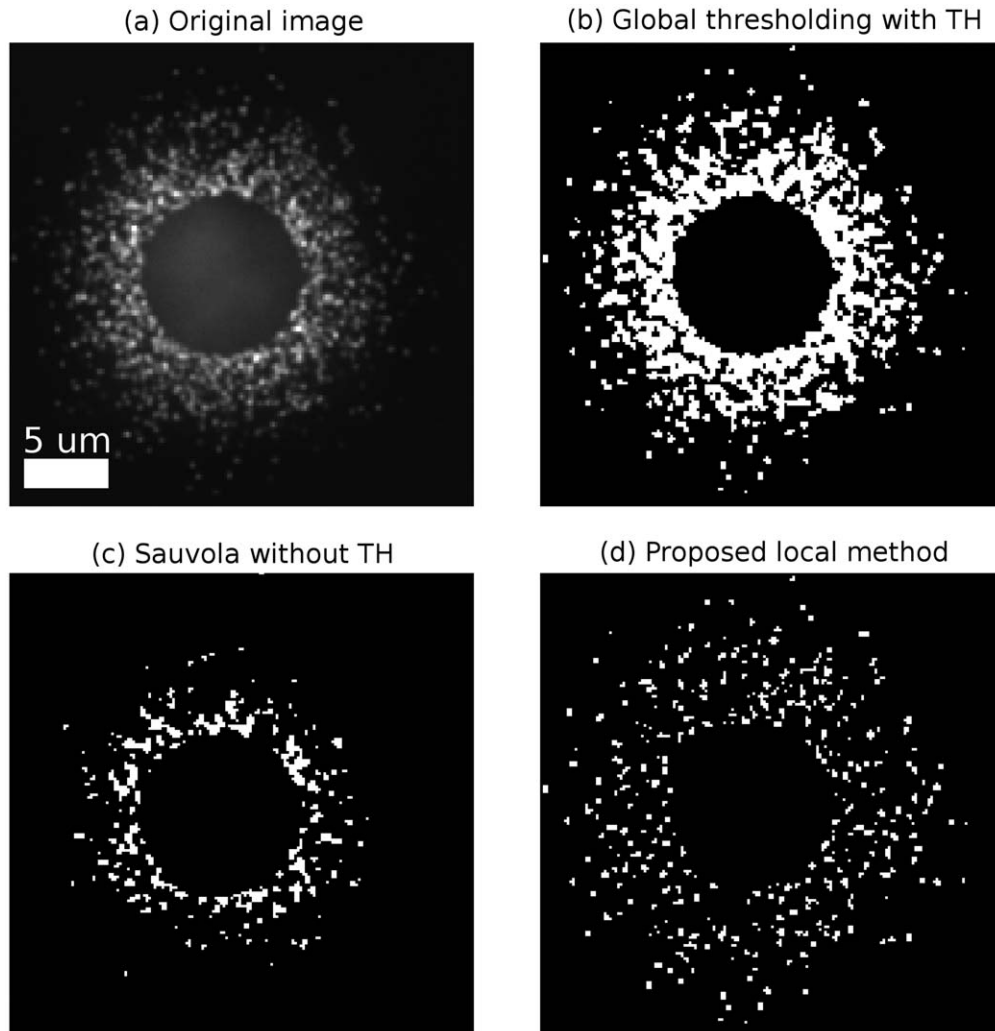


Figure 6. Close-up on segmentation maps of an imaged synthetic cell (a) with global thresholding ($th = 0.09$) on the top-hat prefiltered image (b), with Sauvola's thresholding technique ($k = 0.34$, $radius = 20$) (c), and with the proposed local thresholding technique ($S_{max} = 15$, $LSNR_{min} = 1.7$) (d).

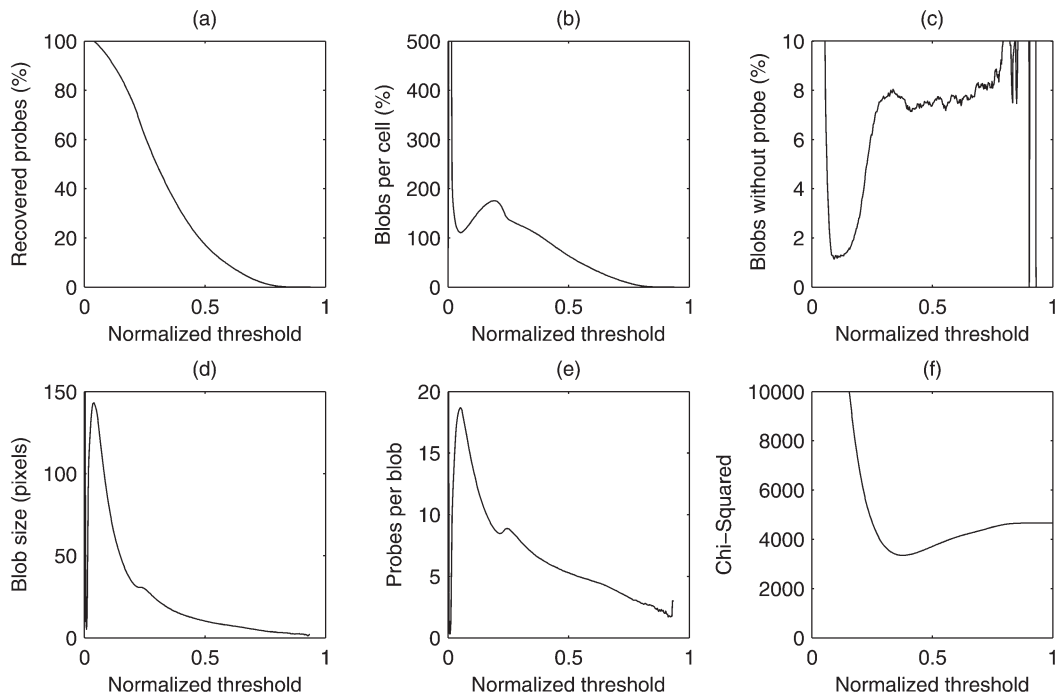


Figure 7. Parameter exploration for the global thresholding technique with top-hat prefiltering. This technique has only one parameter, the threshold value, represented on the x-axis in a normalized manner.

methods. Fluorescent images from our SG dataset feature few bright pixels and a lot of background, low-intensity pixels. This effectively decreases the mean pixel intensity when all the pixels are segmented, and reduces the effect size between the two datasets.

- T-point algorithm: On the SG dataset without TH prefiltering, using the threshold value computed by the T-point algorithm results in segmenting most of the cell cytoplasm. Similarly, on the NCG dataset images, clouds of segmented pixels appear where the cells are located. This can be

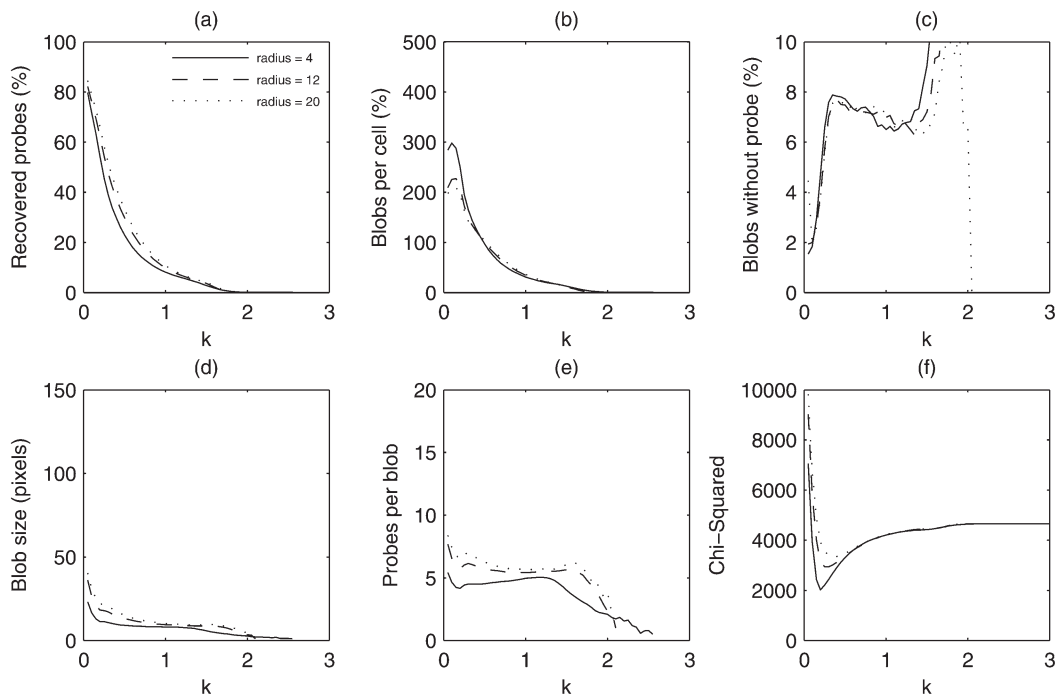


Figure 8. Parameters exploration for Sauvola's thresholding technique without top-hat prefiltering. The x-axis represents the k parameter ranging from 0 to 3. The solid (radius = 4), dashed (radius = 12), and dotted (radius = 20) curves show the influence of the window radius.

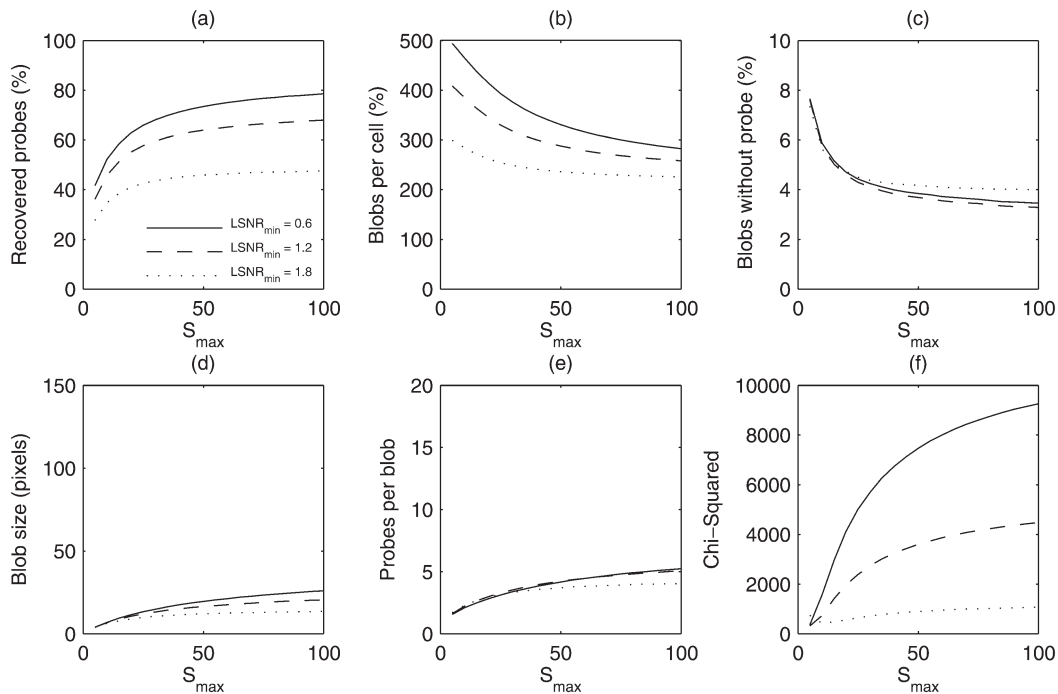


Figure 9. Parameters exploration for the proposed local thresholding technique. The x-axis represents the S_{\max} parameter ranging from 0 to 100. The solid ($LSNR_{\min} = 0.6$), dashed ($LSNR_{\min} = 1.2$), and dotted ($LSNR_{\min} = 1.8$) curves show the influence of the filtering based on the local SNR.

observed in the third row of Figure 3. Looking at the amount of fluorescent pixels in Table 1, we can observe a huge variability in the results, which limits the effect size (1.69) and the classifier performance ($AUC = 0.925$) with respect to the best methods. Table 2, referring to the mean pixel intensity classification feature, offers the same conclusion. Only this time the relatively low mean pixel intensity (0.216) of the SG dataset is the cause, as the segmented areas include a lot of background pixel. Prefiltering the images with TH before computing the threshold value is beneficial. As the fourth row of Figure 3 shows, this segmentation scheme prevents the background pixels from the cell cytoplasm from being segmented in SG images. Furthermore, few pixels are segmented in NCG images compared with other schemes, which is expected behavior. In Table 1, we can notice the reduced variability in the amount of fluorescent pixels per cell classification feature introduced by the TH filter, by comparing both the T-point and TH followed by T-point segmentation schemes. This results in a higher classification performance as indicated by the increased effect size ($1.69 < 2.54$) and AUC ($0.925 < 0.993$). From Table 2, we can see that using the TH filter helps segmenting the fluorescent spots. The mean pixel intensity increases for the SG dataset, whereas it remains almost unchanged for the NCG dataset. This results in slightly improved classification performances. Practically, the T-point algorithm is particularly well suited as most fluorescent images feature a unimodal histogram, characteristic of a lot of background pixels and few pixels of interest. The TH filter further enhances the unimodality by removing

slow variations of background. Considering the amount of fluorescent pixels per cell or the mean pixel intensity classification features independently, the segmentation scheme using the T-point algorithm on TH filtered images is among the best.

- Otsu's algorithm: Unlike the T-point algorithm, Otsu's method is designed to separate the image histogram into two classes using a threshold having the highest separability. This makes it ideal for bimodal histograms. Fluorescent images rarely feature this behavior unless the amount of fluorescently stained cells is high enough to balance the background contribution to the histogram. As predicted, Otsu's method is ill-suited for our application, even when applied recursively twice. While the extracted information on the SG dataset appears to be good (i.e., low amount of fluorescent pixel per cell and high mean pixel intensity), every segmentation scheme using Otsu behaves poorly when facing the NCG dataset. Table 1 presents negative effect sizes for these schemes as the average amount of fluorescent pixels per cell is higher for the NCG dataset than for the SG dataset. Furthermore, the variability of the amount of fluorescent pixels per cell for the NCG dataset is considerable. The masks, example of NCG images provided in the second row of Figure 3, show the two typical outcomes from our NCG dataset. Either a lot of noise is segmented, or only bright macro-objects (e.g., particles) are segmented. The former comes from Otsu selecting a threshold somewhere in the middle of the single mode of the histogram that represents unstained cell samples and the background.

The latter comes from the proper separation between the unstained cells/background mode in the low intensities and the mode added by the high-intensity unwanted objects. To sum up, Otsu-based segmentation schemes cannot be recommended for our application. Despite the relative good classification performances obtained when using the mean pixel intensity classification feature on our datasets, these schemes provided unwanted segmentation results for NCG images where a low (ideally null) amount of fluorescent pixel per cell is expected.

- **Sauvola's thresholding technique:** Unlike the previously discussed methods, Sauvola's approach provides a specific threshold value for each pixel in an image based on its immediate surroundings. This practically removes the influence of slow varying background intensity by extracting regions with a local high contrast. Compared with other methods, Sauvola's thresholding technique is extracting fewer pixels of higher intensities with a limited standard deviation when processing the SG dataset. Notably, the extracted amount of fluorescent pixels per cell on the NCG dataset is statistically null, showing that this technique can be set to exclude most of the background noise. As a result, the effect size and AUC figures presented in Table 1 confirm that the amount of fluorescent pixels per cell can be used to obtain a reliable classification. Conversely, a classification based on the mean pixel intensity is not reliable owing to the very low amount of extracted pixels in the NCG dataset. When the TH filter preprocessing step is used, the results and observations practically remain unchanged, which is no surprise considering that the TH filter filters out low frequencies that are already ignored in Sauvola's thresholding technique by design. This is confirmed by the effect sizes and AUC (Table 1), which show that TH prefiltering does not enhance the classification results. The observed drop in the amount of fluorescent pixels per cell can be reduced by increasing the size of the structuring element used in TH filtering compared with the size of the sliding window used in Sauvola's thresholding technique.
- **Proposed local thresholding technique:** This method was specifically designed to extract fluorescent spots by searching for blobs of fluorescent pixels of limited size having intensity higher than that of their surroundings. With Sauvola's thresholding technique, this method is one of the two methods that extract the fewer pixels when no fluorescence is present in an image. Looking at the effect sizes and AUC obtained in Tables 1 and 2, this method is comparable to the global thresholding scheme using T-point and TH prefiltering.

After having analyzed the real images, we can already sum up a few important points for classifying fluorescent sample images. First, the simple average image intensity and Otsu's segmentation method are not reliable. The best segmentation schemes for this task are the global thresholding of TH filtered images using the T-point algorithm, Sauvola's approach, and the proposed local thresholding method.

Fluorescent Probes Localization on Synthetic Images

In this section, we analyze the various curves plotted in Figures 7–9 to determine what method is best suited for properly segmenting fluorescent spots using synthetic images.

- **Global thresholding on TH filtered images:** Starting at a threshold value of 0, the whole filtered image is segmented. In this case, we have a single blob having the same size as the image enclosing all the fluorescent probes. As the threshold value increases, the average amount of blobs per cell decreases drastically (Fig. 7b) and their average size drops just above 5 pixels (Fig. 7d). In this case, many blobs just represent noise from the background. Thus, the relative amount of blobs failing to recover fluorescent probes is high (Fig. 7c). Proper fluorescent dot segmentation happens for a normalized threshold value high enough so that background noise is not segmented. In our test case, this happens for a normalized threshold value of 0.09, where the amount of blobs not recovering fluorescent probes is minimum (Fig. 7c). Further increase of the threshold will eventually trigger some blobs to be broken down into many blobs (local maximum in Fig. 7b) alongside with the amount of blobs not enclosing any probes (local maximum in Fig. 7c). This indicates that global thresholding applied on TH filtered images has an optimal threshold value which maximizes the segmentation and spot extraction efficiency. Below this optimal threshold, the segmentation masks include background noise. This situation occurs when using the T-point algorithm to determine the threshold value (0.02) of the TH filtered synthetic dataset. Above the optimal threshold, some information is lost. With an average normalized threshold value of 0.14 and 0.33 for Otsu's method and recursive Otsu's method, respectively, the latter appears less optimal.
- **Sauvola's thresholding technique:** A very low value of the parameter k comes down to thresholding a given pixel with the average pixel intensity within its surrounding window. This results in many blobs segmenting noise. The amount of blobs is very high (Fig. 8b), just like the amount of blobs not enclosing any fluorescent probes (Fig. 8c). By design, Sauvola's method removes segmentation noise for a high enough value of k . In our test case, this happens for a value of k greater than 0.2, independently from the radius. From this value of k and higher, which corresponds to peaks found in the average blob size (Fig. 8d), all the metrics from Figure 8 are decreasing except for the blobs without probes and χ^2 histogram distance. This means blobs are getting fewer and smaller, effectively locating fluorescent dots but leaving behind some useful information contained in the images. Looking at the influence of the window radius, we observed that the various curves seem to converge as the radius increases. Increase of the window radius seems to favor a bigger average size of the blobs, which inherently favors the amount of fluorescent probes enclosed per blob.
- **Proposed local thresholding technique:** We are analyzing the effect of the maximum allowed size for a blob S_{\max} (x-

axis on Fig. 9) and of the $LSNR_{\min}$ parameter. As we can see on Figure 9, a value of S_{\max} smaller than the average size of a fluorescent spot cannot be considered. If S_{\max} is smaller than 5 pixels, the results are meaningless. However, as we sweep S_{\max} up until 100 pixels, we can observe that more fluorescent spots are segmented, while the amount of blobs not segmenting probes is decreasing and the χ^2 histogram distance is increasing. This indicates that our blobs are becoming bigger and collecting more and more probes per blob. The $LSNR_{\min}$ parameter practically reduces the average size of the blobs as it increases. This has the exact opposite effect as the S_{\max} parameter on the metrics in Figure 9. Note that in our test case, a $LSNR_{\min}$ value smaller than 0.6 allowed background noise to be segmented.

Knowing that the segmentation goal is to locate the fluorescent probes, we are interested in segmenting an image so that we have as many small blobs as possible, each enclosing a minimum amount of fluorescent probes. The proposed local thresholding method performs best in that aspect. Looking at Figure 9 and within the parameter range described just above for the proposed method, we are able to provide 225–300 blobs per cell between 4 and 13 pixels in size enclosing from two to four fluorescent probes. In other words, we are recovering 27–47% of the probes while keeping the χ^2 histogram distance between 400 and 1,000. Sauvola's method is able to provide similar blob sizes but the blobs are fewer per cell, between 75 and 85, extracting up to 15% of the probes only for a χ^2 distance ranging from 3,600 to 4,600. A smaller χ^2 distance could be achieved at the expense of the blob size and the amount of probes per blob. In contrast, the global thresholding of TH filtered images is not able to provide a χ^2 distance smaller than 3,350, which results in blobs of 16 pixels segmenting 35% of the probes only. A lower threshold value increases drastically the blob sizes and a smaller threshold value further reduces the accuracy. As a result, the proposed method is preferred as it is able to recover relevant fluorescent pixels in a greater number of smaller blobs compared with other methods, while keeping the amount of failed blob segmentations contained.

CONCLUSION

In this work, we have applied commonly used global threshold computation algorithms (T-point and Otsu) and segmentation techniques (Sauvola) combined with the TH MM filter for localizing sub-resolution fluorescent biomarkers and classifying fluorescence microscopy images. We then introduced a novel local thresholding technique and used the previously cited methods as points of comparison.

The proposed local thresholding method was proven to be the best for classifying the fluorescent images from our real sample image datasets, followed by Sauvola's method and the T-point algorithm applied to TH filtered images. Considering the amount of segmented pixels and their intensity as classification features, these three methods were separating the SG and NCG dataset better than any of the other segmentation

approaches. These results provide leads for our μ GIT system called NutriChip, as well as LoC systems in general, because a combination of these three methods used in parallel can provide a robust image processing system for detecting and monitoring fluorescent signals.

In a second part, we also quantitatively analyzed the capacity of these three methods for extracting useful fluorescent signal and hopefully localize the various stained TLR2. This analysis was done on computer-generated images as real images lack the metadata of biomarker locations in order to evaluate the algorithms efficiency. The global thresholding applied on images filtered by the TH filter was able to recover the most information by recovering up to 94% of the stained TLR in the best case. However, it performs poorly at localizing them because the segmenting blobs are big compared with typical fluorescent spots. The proposed local thresholding method, which forces segmenting blobs below a given size, recovers fewer biomarkers but provides better localization results.

LITERATURE CITED

- Huang B, Bates M, Zhuang X. Super resolution fluorescence microscopy. *Annu Rev Biochem* 2009;78:993–1016.
- Huang B, Babcock H, Zhuang X. Breaking the diffraction barrier: Super-resolution imaging of cells. *Cell* 2010;143:1047–1058.
- Mascetti G, Vergani L, Diaspro A, Carrara S, Radicchi G, Nicolini C. Effect of fixatives on calf thymocytes chromatin as analyzed by 3D high-resolution fluorescence microscopy. *Cytometry* 1996;23:110–119.
- Dima AA, Elliott JT, Filliben JJ, Halter M, Peskin A, Bernal J, Kocielek M, Brady MC, Tang HC, Plant AL. Comparison of segmentation algorithms for fluorescence microscopy images of cells. *Cytometry Part A* 2011;79A:545–559.
- Smal I, Loog M, Niessen W, Meijering E. Quantitative comparison of spot detection methods in fluorescence microscopy. *IEEE Trans Med Imaging* 2010;29:282–301.
- Smal I, Loog M, Niessen W, Meijering E. Quantitative comparison of spot detection methods in live-cell fluorescence microscopy imaging. In: *Proceedings of the IEEE International Symposium on Biomedical Imaging: From Nano to Macro* 2009. pp 1178–1181.
- Bolte S, Cordelières FP. A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc* 2006;224:213–232.
- Mascetti G, Carrara S, Vergani L. Relationship between chromatin compactness and dye uptake for in situ chromatin stained with DAPI. *Cytometry* 2001;44:113–119.
- Ramadan Q, Jafarpoorchehab H, Huang C, Silacci P, Carrara S, Koklū G, Ghaye J, Ramsden J, Ruffert C, Vergères G, et al. NutriChip: Nutrition analysis meets microfluidics. *Lab Chip* 2013;13:196–203.
- Vergères G, Bogicevic B, Buri C, Carrara S, Chollet M, Corbino-Giunta L, Egger L, Gille D, Kopf-Bolanck K, Laederach K, et al. The nutrichip project—Translating technology into nutritional knowledge. *Br J Nutr* 2012;142:245–250.
- Ramadan Q, Jafarpoorchehab H, Bolanz K, Schwander F, Silacci P, Carrara S, Ramsden J, Vergères G, Gijs M. Nutrichip: An integrated microfluidic system for in vitro investigation of the immunomodulatory function of dairy products. *MicroTAS* 2011. pp 650–652.
- Chen J, Rao JN, Zou T, Liu L, Marasa BS, Xiao L, Zeng X, Turner DJ, Wang J-Y. Polyamines are required for expression of toll-like receptor 2 modulating intestinal epithelial barrier integrity. *Am J Physiol* 2007;293:G568–G576.
- Triantafilou M, Manukyan M, Mackie A, Morath A, Hartung T, Heine H, Triantafilou M. Lipoteichoic acid and toll-like receptors 2 internalization and targeting to the Golgi are lipid raft-dependent. *J Biol Chem* 2004;279:40882–40889.
- Otsu N. A threshold selection method from gray-level histograms. *IEEE Trans Syst Man Cybern* 1979;9:62–66.
- Ridler T, Calvard S. Picture thresholding using an iterative selection method. *IEEE Trans Syst Man Cybern* 1978;8:630–632.
- Kapur JN, Sahoo PK, Wong AKC. A new method for gray-level picture thresholding using the entropy of the histogram. *Comp Vis Graph Image Process* 1985;29:273–285.
- Coudray N, Buessler J-L, Urban J-P. Robust threshold estimation for images with unimodal histograms. *Pattern Recogn Lett* 2010;31:1010–1019.
- Ram S, Rodríguez JJ, Bosco G. Segmentation and detection of fluorescent 3D spots. *Cytometry Part A* 2012;81A:198–212.
- Netten H, van Liet LJ, Vrolijk H, Sloos WCR, Tanke HJ, Young IT. Fluorescent dot counting in interphase cell nuclei. *Bioimaging* 1996;4:93–106.
- Pham TD, Crane DI, Tran TH, Nguyen TH. Extraction of fluorescent cell puncta by adaptive fuzzy segmentation. *Bioinformatics* 2004;20:2189–2196.

21. Zhang G, Jia X, Pham TD, Crane DI. Multistage spatial property based segmentation for quantification of fluorescence distribution in cells. *Comput Models Life Sci* 2010; 1210:3–12.
22. Olivo-Marin J-C. Extraction of spots in biological images using multiscale products. *Pattern Recogn* 2002;35:1989–1996.
23. Serra J. *Image Analysis and Mathematical Morphology*. London: Academic Press; 1982.
24. Haralick RM, Sternberg SR, Zhuang X. Image analysis using mathematical morphology. *IEEE Trans Pattern Anal Mach Intell* 1987;PAMI-9:532–550.
25. Meyer F, Beucher S. Morphological segmentation. *J Visual Comm Image* 1990;1:21–46.
26. Verbeek PW, Vrooman HA, Vliet LJ. Low-level image processing by max-min filters. *Signal Process* 1988;15:249–258.
27. Kimori Y, Baba N, Morone N. Extended morphological processing: A practical method for automatic spot detection of biological markers from microscopic images. *BMC Bioinformatics* 2010;11:373.
28. Reljin B, Paskas M, Reljin I, Konstany K. Breast cancer evaluation by fluorescent dot detection using combined mathematical morphology and multifractal techniques. *Diagn Pathol* 2011;6:S21.
29. Nielsen B, Albregtsen F, Danielsen HE. Automatic segmentation of cell nuclei in Feulgen-stained histological sections of prostate cancer and quantitative evaluation of segmentation results. *Cytometry Part A* 2012;81A:588–601.
30. Niblack W. *An Introduction to Image Processing*. Prentice-Hall: Englewood Cliffs, NJ; 1986. pp 115–116.
31. Sauvola J, Pietikainen K. Adaptive document image binarization. *Pattern Recogn* 2000;33:225–236.
32. Vincent L, Soille P. Watersheds in digital spaces: An efficient algorithm based in immersion simulations. *IEEE Trans Pattern Anal Mach Intell* 1991;23:583–598.
33. Gniadek T, Warren G. Watershed3D: A new method for segmenting and counting punctate structures from confocal image data. *Traffic* 2007;8:339–346.
34. Sezgin M, Sankur B. Survey over image thresholding techniques and quantitative performance evaluation. *J Electron Imaging* 2004;13:146–165.
35. Nakagawa S, Cuthill IC. Effect size, confidence interval and statistical significance: A practical guide for biologists. *Biol Rev* 2007;82:591–605.
36. Yousef WA, Wagner RF, Loew MH. Estimating the uncertainty in the estimated mean area under the ROC curve of a classifier. *Pattern Recogn Lett* 2005;26:2600–2610.
37. Davison AC, Hinkley DV, Schechtman E. Efficient bootstrap simulation. *Biometrika* 1986;73:555–556.
38. Ghaye J, De Micheli G, Carrara S. Simulated biological cells for receptor counting in fluorescence imaging. *BioNanoScience* 2012;2:94–103.
39. Pele O, Werman M. The quadratic-chi histogram distance family. In: Daniilidis K, Maragos P, Paragios N, editors. *Computer Vision, ECCV. 2nd part*. Springer: Berlin Heidelberg; 2010. LNCS 6312, pp 149–762.