Globally optimal segmentation of cell nuclei in fluorescence microscopy images using shape and intensity information

L. Kostrykin[a,*], C. Schnörr[b], K. Rohr[a]

[a]*Biomedical Computer Vision Group, BIOQUANT, IPMB,*
Heidelberg University and DKFZ, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany
[b]*Image and Pattern Analysis Group, Heidelberg University, 69120 Heidelberg, Germany*

Abstract

Accurate and efficient segmentation of cell nuclei in fluorescence microscopy images plays a key role in many biological studies. Besides coping with image noise and other imaging artifacts, the separation of touching and partially overlapping cell nuclei is a major challenge. To address this, we introduce a globally optimal model-based approach for cell nuclei segmentation which jointly exploits shape and intensity information. Our approach is based on implicitly parameterized shape models, and we propose single-object and multi-object schemes. In the single-object case, the used shape parameterization leads to convex energies which can be directly minimized without requiring approximation. The multi-object scheme is based on multiple collaborating shapes and has the advantage that prior detection of individual cell nuclei is not needed. This scheme performs joint segmentation and cluster splitting. We describe an energy minimization scheme which converges close to global optima and exploits convex optimization such that our approach does not depend on the initialization nor suffers from local energy minima. The proposed approach is robust and computationally efficient. In contrast, previous shape-based approaches for cell segmentation either are computationally expensive, not globally optimal, or do not jointly

---

*Corresponding author
  Email addresses: `leonid.kostrykin@bioquant.uni-heidelberg.de` (L. Kostrykin),
`schnoerr@math.uni-heidelberg.de` (C. Schnörr), `k.rohr@dkfz-heidelberg.de` (K. Rohr)

exploit shape and intensity information. We successfully applied our approach to fluorescence microscopy images of five different cell types and performed a quantitative comparison with previous methods.

Keywords: Fluorescence microscopy, cell segmentation, cell-cluster splitting, model fitting, global energy minimization, convex optimization

1. Introduction

In fluorescence microscopy imaging, chemical compounds are used to label cellular structures. Upon light excitation, these compounds re-emit light of a characteristic color that is captured during image acquisition. Cell nuclei typi-
⁵ cally appear in the resulting images as roughly elliptical regions of brighter image intensity. The segmentation of individual cell nuclei is important for many biological studies, including the analysis of cell morphology and cell movement. In multi-channel image data, cell nuclei segmentation is required to determine regions of interest for analyzing fluorescence signals in the other image channels.
¹⁰ Automatic segmentation of cell nuclei is challenging for many reasons such as imaging artifacts like image noise, image blur, and intensity inhomogeneities, as well as separation of touching and partially overlapping objects (see Figure 1). Intensity inhomogeneities occur across multiple objects (inter-object inhomogeneities) and within individual cell nuclei (intra-object inhomogeneities).

¹⁵ In previous work, **classical segmentation** approaches have often been used for cell segmentation. These approaches include intensity thresholding, region growing, morphological analysis, and combinations of these methods (e.g., Singh et al., 2011; Plissiti et al., 2011; Singh et al., 2017), which, however, are generally sensitive to texture, image noise, and intensity inhomogeneities. Clustering-
²⁰ based methods, variants of the watershed transform, and random walker algorithms (e.g., Wählby et al., 2002; He et al., 2014) require accurate initialization to prevent false object merging and splitting.

Many cell segmentation methods are based on a **variational framework**, where object contours are represented as **level sets** of functions (e.g., de Solorzano
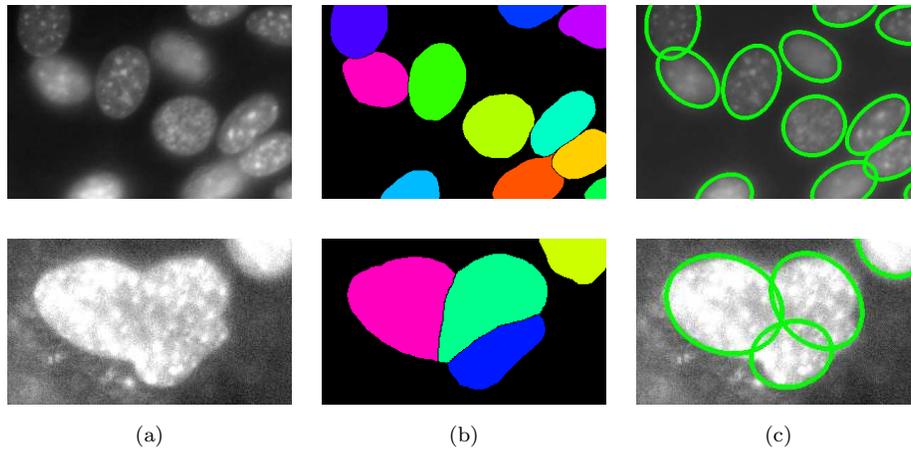
Figure 1: Separation of touching (first row) and overlapping (second row) cell nuclei. (a) Original image section (NIH3T3 cells). (b) Corresponding ground truth data. (c) Segmentation result of the proposed approach (green contour).

25   et al., 2001; Dufour et al., 2005; Li et al., 2008). The initial choice of the level set functions is crucial for these methods, since non-convex energies are used and only *local* energy minima are generally determined. Bergeest and Rohr (2012) described a *globally optimal* approach based on *convex* level set energies. Such approaches are robust, since they guarantee finding the best admissible solution

30   for any initialization. Thus, these approaches are less data-dependent.

To better cope with partial object overlap and image distortions like strong image noise, cell segmentation methods were developed which exploit *shape information*. Many of these **shape-based methods** use the variational level set model of Chan and Vese (2001) in combination with different kinds of shape

35   regularization. Ali and Madabhushi (2012) and Kong et al. (2016) employed statistical shape priors based on similarity transformations for segmentation of overlapping cells in histological images. Nosrati and Hamarneh (2015) introduced a star-shape prior for the segmentation of the cytoplasm of cervical cells. Xing and Yang (2015) exploited shape information to guide the evolution of the

40   level set functions. Gharipour and Liew (2016) used level sets for cell nuclei

3

segmentation and identified individual cell nuclei by analyzing the morphology, but shape and intensity information were used in consecutive steps and not jointly exploited. Zhang et al. (2017) used sparse shape priors in conjunction with object overlap penalty terms. *None of the above shape-based methods yield*
₄₅ *globally optimal solutions.*

Combinatorial approaches have gained increased interest for cell segmentation (e.g., Al-Kofahi et al., 2010; Lou et al., 2012; Maška et al., 2013; Poulain et al., 2015; Soubies et al., 2015), since the solution can be computed close to global optimality in many cases using graph cuts (Boykov and Funka-Lea,
₅₀ 2006). However, shape information is difficult to handle in a purely combinatorial framework. Therefore, **hybrid approaches** were introduced, which use classical approaches to generate object-based segmentation candidates, and combinatorial optimization to select the overall best coherent subset of candidates for the final segmentation. Poulain et al. (2015) used a sequence of graph
₅₅ cuts to determine the best maximum subset of segmentation candidates, but properties of global convergence were not studied. In Akram et al. (2016, 2017) and Türetken et al. (2017), candidates are selected using a tracking approach based on combinatorial search for the most plausible trajectories in temporal image sequences, either using shortest-path formulations (Akram et al., 2016,
₆₀ 2017) or integer linear programming (Türetken et al., 2017). However, tracking-based methods are not applicable for the segmentation of individual images.

Cell segmentation approaches based on candidate selection were also used in conjunction with explicitly parameterized shape models and **marked point processes** (MPP). Dong and Acton (2007) employed elliptical models and used
₆₅ an MPP with a pairwise interaction model to cope with object overlap. Soubies et al. (2015) used an MPP and a contrast-invariant data term based on the image gradient, but did not directly exploit the image intensities. Descombes (2017) embedded an MPP into a simulated annealing scheme. In principle, this scheme converges to a global solution, but it is computationally expensive
₇₀ and requires careful calibration of the cooling parameters. Markowsky et al. (2017) employed an MPP to determine a low-cardinality set of circles in order

4

to split touching and overlapping cell nuclei in binarized images. Iteratively updated parameter distributions were used to guide the MPP towards a global solution, but image intensities were not exploited. Panagiotakis and Argyros (2018) proposed a similar approach based on elliptical models and expectation maximization instead of MPP. However, no guarantees were made on global optimality and only binarized images are used to separate individual objects. In our previous work (Kostrykin et al., 2018), we described a globally optimal approach which does not require binarization of an image, but global optimality was only achieved locally within prior detected image regions.

Learning-based methods for cell segmentation determine model parameters by training using *annotated* data prior to applying the methods. Ronneberger et al. (2015) proposed a deep convolutional neural network which is widely used for biomedical images (e.g., Akram et al., 2017; Böhm et al., 2018; Fan and Rittscher, 2018). Böhm et al. (2018) described an extension for overlapping objects. Fan and Rittscher (2018) introduced an approach which fuses the segmentation results from two deep neural networks by a global probabilistic model. Shape and intensity information are employed, but the computational complexity of the used annealing scheme was not mentioned. Generally, training deep networks requires large amount of annotated data, and although Xie et al. (2018) demonstrated that synthetic data can be used, training is computationally expensive. Goodfellow et al. (2015) showed that many learning-based methods (including neural networks) are prone to *adversarial* perturbations. Fawzi et al. (2017) found that many deep neural network architectures are prone to image noise and *universal* perturbations, and conclude that developing provably robust deep neural networks is an open issue.

Contribution. In this work, we introduce a new *globally optimal* approach for cell nuclei segmentation, which *jointly* exploits shape and intensity information. The approach is based on implicitly parameterized elliptical models and global energy minimization. Our proposed shape parameterization leads to a convex energy for single objects, that is optimally minimized using robust numerical methods. The optimization does not depend on the initialization and

5

does not suffer from local minima. To avoid prior detection of image regions corresponding to the individual cell nuclei, we generalize the single-object model to the multi-object case. Our multi-object model consists of multiple collaborating ellipses, which represent a whole image. This leads to a non-convex energy, yet we have found that model fitting using the multi-object model corresponds to the *min-weight set-cover* problem (e.g., Vazirani, 2001). Assuming that individual objects are roughly elliptical, simply connected, and correspond to one or more local intensity peaks, the result is determined close to the *global solution* using an efficient combination of combinatorial and second-order convex optimization schemes.

The proposed energies are contrast-invariant and thus our approach is robust to *inter-object* intensity inhomogeneities. The joint exploitation of shape and intensity information enables our approach to intrinsically cope with *intra-object* intensity inhomogeneities and partial object overlap (see Figure 1). Thus, for splitting of clustered cell nuclei, our approach neither requires an object interaction model (e.g., Poulain et al., 2015; Soubies et al., 2015; Descombes, 2017; Markowsky et al., 2017) nor prior image binarization (e.g., Panagiotakis and Argyros, 2018). In contrast to our previous single-object approach (Kostrykin et al., 2018), the proposed single-object scheme leads to a convex energy, which can be directly (exactly) minimized and does not require an approximation. In addition, we propose a multi-object scheme. The structure of our approach inherently permits effective parallelization. In contrast to learning-based approaches, our approach does not require data-driven training nor annotated data. A main advantage of our model-based approach is that the explicit model assumptions allow designing well-defined algorithms that facilitate both reproducibility and predictability. To the best of our knowledge, we propose the first globally optimal model-based approach which jointly exploits shape and intensity information and is computationally tractable in practical applications.

We have evaluated our approach using fluorescence microscopy datasets of five different cell types, including publicly available benchmark datasets, and performed a quantitative comparison with previous methods. It turned out

that the proposed approach generally improves the performance.

135   This paper is organized as follows. Section 2 presents our intensity-based segmentation approach using convex energies and an implicitly parameterized single-object model. Section 3 describes the generalized multi-object model and the corresponding scheme for global energy minimization. Experimental results and a comparison with previous methods are provided in Section 4. We discuss
140   the results of our work in Section 5.

2. Single-object model and convex energy formulation

In this section, we describe our globally optimal approach for cell nuclei segmentation, which jointly uses shape and intensity information based on a single-object model. We define a *shape model* as the zero-level set $\mathcal{C}(s) =$
145     $\left\{x \in \mathbb{R}^2 \middle| s(x; \theta) = 0\right\}$ of a model function s, which maps an image point to a real value. More specifically, we parameterize s as a second-order polynomial,

$$s(x; \theta) = x^\top A\, x + b^\top x + c, \tag{1}$$

where the symmetric 2×2 matrix A, the vector $b \in \mathbb{R}^2$, and $c \in \mathbb{R}$ are represented by the shape parameter vector θ. The shape of the zero-level set $\mathcal{C}(s)$ is then confined to an ellipse, a parabola, hyperbola, line, or a stripe,
150   unless $\mathcal{C}(s)$ corresponds to the whole image plane, the empty set, or a single dot. The model function s induces two disjoint image regions, the zero-sublevel set $\mathcal{I}^-(s) = \left\{x \in \mathbb{R}^2 \middle| s(x; \theta) < 0\right\}$ of s and its zero-superlevel set $\mathcal{I}^+(s) = \left\{x \in \mathbb{R}^2 \middle| s(x; \theta) > 0\right\}$. If the parameters θ are chosen such that the shape model $\mathcal{C}(s)$ is elliptical, then the regions $\mathcal{I}^+(s)$ and $\mathcal{I}^-(s)$ correspond to
155   the interior and exterior of the ellipse, respectively.

Given an image $g : R \to \mathbb{R}_{\geq 0}$ of an object and its background in an image region $R \subset \mathbb{R}^2$, which are *roughly* separable using an intensity offset τ, the τ-superlevel set $\mathcal{I}^+(g(x) - \tau)$ indicates the imaged object, whereas the τ-sublevel set $\mathcal{I}^-(g(x) - \tau)$ corresponds to the image background. To segment the image

7

160 region R, we thus consider the *intensity* model

$$y_x = (g(x) - \tau) \cdot \gamma \tag{2}$$

and seek those shape parameters θ^*, for which $\mathcal{I}^+(s)$ covers $\mathcal{I}^+(y)$ while $\mathcal{I}^-(s)$ covers $\mathcal{I}^-(y)$. The fixed factor $\gamma > 0$ in Eq. (2) governs the contrast between the imaged object and the background. More formally, we minimize

$$J_{R,L}(\theta) = \sum_{x \in R} L(y_x; s(x;\theta)), \quad L(y;s) = \begin{cases} 1 & \text{if } y \cdot s < 0, \\ 0 & \text{else,} \end{cases} \tag{3}$$

which penalizes each image point x with $\operatorname{sgn} y_x \neq \operatorname{sgn} s(x;\theta)$. Our approach
165 exploits both shape and image intensity information. An extension to other image features (e.g., texture) is possible by including additional terms in Eq. (2).

Since the energy $J_{R,L}$ is non-smooth, we instead determine the optimal shape parameters θ^* as the global minimizer of the energy function

$$J_R(\theta) = J_{R,\phi}(\theta) = \sum_{x \in R} \phi(y_x; s(x;\theta)), \quad \phi(y;s) = \ln(1 + \exp(-y \cdot s)), \tag{4}$$

where the loss function ϕ is smooth and convex in s. In addition, ϕ is a minimal
170 convex upper-bound of the 0/1-loss L in Eq. (3), if L is weighted by the constant factor $\phi(y;0) = \ln 2$. Thus, the minimization of the energy J_R also minimizes the energy $J_{R,L}$, since the minimizers of a function are invariant to positive constant factors ($\ln 2$). The energy formulation (4) is analogous to logistic regression using polynomial basis function expansion (e.g., Murphy, 2012), but
175 in contrast to logistic regression, y_x in our approach is not limited to binary values.

For analyzing the convexity of the energy J_R in Eq. (4), the following composition rule is central: The composition $\phi(y_x; s(x;\theta))$ of the convex loss function $\phi(y_x; \cdot)$ and the model parameterization $s(x;\theta)$ is convex in the model
180 parameters θ for all y_x, if the parameterization s is affine in θ (e.g., Boyd and Vandenberghe, 2004, Section 3.2.2). This condition is true for our parameterization (1). Since the sum of convex functions is convex, the energy J_R in Eq. (4) is also convex.

An alternative shape model parameterization was used in our previous work

185 (Kostrykin et al., 2018): $s\left(x;\theta\right) = \left(x - b\right)^{\top} A\left(x - b\right) + c$. In contrast to Eq. (1) the vector b in this parameterization represents the center of the shape model. However, the parameterization is not affine in θ, and thus the corresponding energy is *non-convex*. In our previous work, we hence used a sequential approximation scheme for global energy minimization. In contrast, for the new param-

190 eterization (1), the energy J_R is *convex* and can be directly globally minimized without requiring an approximation. In our proposed approach, we determine the globally optimal parameters θ^* by robust numerical methods and an arbitrary initialization, as detailed in Section 3.3. From a theoretical point of view, for certain image data, the convex energy J_R might not possess a minimizer,

195 for example, if the image region contains no foreground ($y_x < 0$ for all $x \in R$). In this case, however, the energy value $J_R\left(\theta^*\right)$ of the numerically determined minimizer θ^* will be arbitrarily close to $\inf J_R$. Thus, from a practical point of view, the convex energy is globally minimized for any image data.

Another advantage of the new parameterization (1) is that it is homogeneous.

200 Consequently, the optimal zero-level set $\mathcal{C}\left(s\left(\cdot;\theta^*\right)\right)$ is invariant to the factor γ in the intensity model (2), since the set of feasible parameters θ is unbounded, and thus closed under scalar multiplication. Using the parameterization (1) we can assume $\gamma = 1$ without loss of generality. Model fitting using the energy J_R is hence invariant to the image contrast.

205 ## 3. Multi-object model and globally optimal energy minimization

The single-object model described above represents a single elliptical object. Hence, the model can only be fitted to an image region, which contains at maximum one single object. However, since cell microscopy images generally contain multiple objects, we generalize the single-object model (1) to the multi-

210 object case. Below, we describe the multi-object model, the method for global energy minimization, and implementation details.

9

3.1. Multi-object model

Recall that the single-object model function s, described in the Section 2 above, induces two disjoint image regions, which are defined by the zero-superlevel set $\mathcal{I}^+(s)$ and the zero-sublevel set $\mathcal{I}^-(s)$. If the zero-level set $\mathcal{C}(s)$ has an elliptic shape, then the two regions correspond to the interior and exterior of the ellipse. For the multi-object case, we extend the model function s such that it represents *multiple* elliptical objects.

In the multi-object case, we seek to cover the image points $x \in R$ of the image foreground ($y_x > 0$) by the *union* of the foreground of multiple models $s_1, \ldots, s_m$ of the form (1). At the same time, the image background (image points x with $y_x < 0$) is covered by the *intersection* of the background of these models. The union of the foreground of the models $s_1, \ldots, s_m$ can be expressed as $\cup_{k \in [m]} \mathcal{I}^+(s_k) = \mathcal{I}^+(s)$, which is the zero-superlevel set of

$$s(x;\theta) = \max_{k \in [m]} s_k(x;\theta), \tag{5}$$

since, for fixed x and θ, $s(x;\theta) > 0$ occurs if and only if there is a $k \in [m]$ with $s_k(x;\theta) > 0$. The intersection of the background of the models is given by $\cap_{k \in [m]} \mathcal{I}^-(s_k) = \mathcal{I}^-(s)$, since $s(x;\theta) < 0$ occurs if and only if $s_k(x;\theta) < 0$ for all $k \in [m]$. Using the formulation (5), the models $s_1, \ldots s_m$ thus *collaboratively* represent the image foreground and background. Below, to improve the readability, we will skip the explicit dependence of s_k and s on θ.

An example illustrating the multi-object model is provided in Figure 2. Naturally, at any given image point x, the pointwise maximum in Eq. (5) does not depend on models s_k with $s_k(x) < s(x)$. Hence, models s_k with $s_k(x) = s(x)$ are of major interest. In our approach, such a model is denoted to be *active* at x. The set of all image points, where this model is active, forms an *activity region*. Closer characterization of these regions proves to be advantageous, as detailed in Section 3.2 below. Figure 2(e) shows the activity regions, which correspond to the multi-object model depicted in Figure 2(d).

The multi-object model (5) is homogeneous in the parameters θ, since the pointwise maximum and s_k are homogeneous functions. Thus, the multi-object
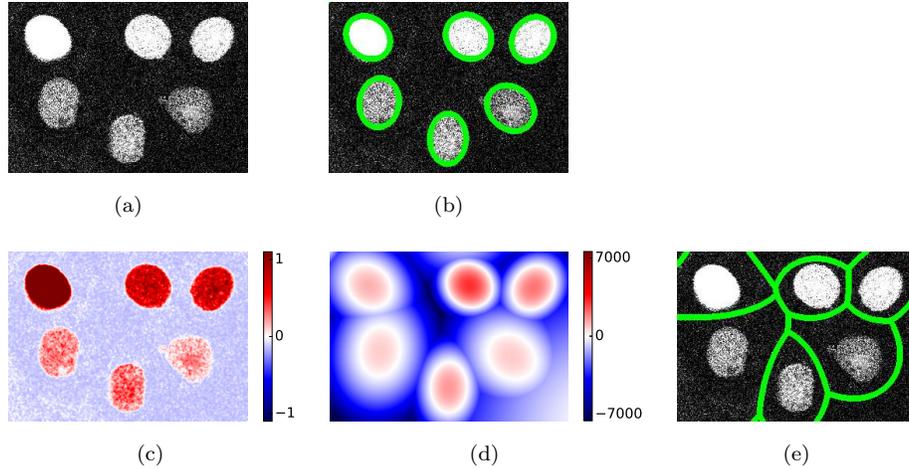
10

Figure 2: Example illustrating the multi-object model with optimally chosen model parameters θ^*. (a) Original image section (GOWT1 cells). (b) Zero-level set $\mathcal{C}\left(s\left(\cdot;\theta^*\right)\right)$ (green) of the multi-object model s. (c) Intensity model y_x as a function of x. (d) Multi-object model $s\left(x;\theta^*\right)$ as a function of x. (e) Corresponding model activity regions (green contour).

model preserves the contrast invariance property of the single-object model and we may hence assume $\gamma = 1$ in Eq. (2) without loss of generality for the multi-object model (5). Segmentation of multiple objects using the multi-object model (5) is performed by minimizing the energy function

$$J\left(\theta\right) = \sum_{x \in R} \phi\left(y_x; \max_{k \in [m]} s_k\left(x\right)\right), \qquad (6)$$

245   where the image intensities are incorporated via Eq. (2) using an intensity offset τ. The offset is computed by analyzing the local image intensities, as detailed in Section 3.3 below. Image binarization is not required.

Minimization of the energy function (6) determines the parameters θ of the multi-object model (5) which define the *globally optimal collaborating ellipses*
250   (GOCELL) representation of the image. In the energy function (6), the pointwise maximum of the family of linear functions $s_1, \ldots, s_m$ is convex in θ, but not affine. Thus, in Eq. (6), the concatenation of the convex loss function $\phi\left(y_x; \cdot\right)$, as given by Eq. (4), and the multi-object model (5) is non-convex (e.g., Boyd
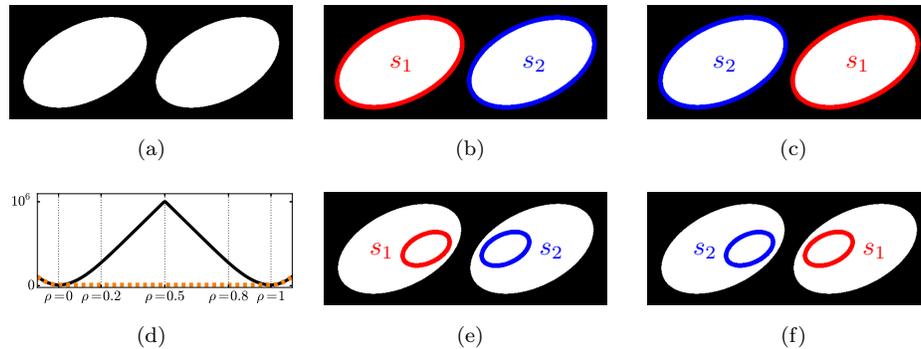
11

Figure 3: Example illustrating the non-convexity of the energy (6) using $m = 2$ object models. (a) Synthetic image. (b) Parameters θ_0 corresponding to the zero-level sets $\mathcal{C}\left(s_1\left(\cdot;\theta_0\right)\right)$ (red) and $\mathcal{C}\left(s_2\left(\cdot;\theta_0\right)\right)$ (blue) of the two models. (c) Parameters θ_1, which are symmetric to θ_0 ($\mathcal{C}\left(s_1\left(\cdot;\theta_1\right)\right)$ red and $\mathcal{C}\left(s_2\left(\cdot;\theta_1\right)\right)$ blue). (d) Energy function (6) along the ρ-parameterized line $\theta_\rho = \theta_0 \cdot (1 - \rho) + \theta_1 \cdot \rho$ between θ_0 and θ_1 (black, solid) and its corresponding convex envelope (orange, dashed). (e) Zero-level sets for the parameters $\theta_{0.2}$ and (f) $\theta_{0.8}$.

and Vandenberghe, 2004). An example of this non-convexity is provided in

²⁵⁵ Figure 3, which illustrates the energy function along a straight line in the parameter space. The minimization of non-convex energies is generally difficult. In our case, convex envelope-based reformulations (e.g., Mollenhoff et al., 2016) are not applicable, because the convex envelope of the energy (6) possesses an infinite number of global minimizers, which are far from being optimal with re-

²⁶⁰ spect to the energy function (e.g., $\rho = 0.5$ in Figure 3(d)). Also, the sequential approximation scheme in Kostrykin et al. (2018) is not applicable, since the multi-object model (5) is non-polynomial. However, global minimization of the energy (6) is tractable if the model activity regions are assumed to be unions of adjacent subregions, which we call *region fragments*, as detailed below.

²⁶⁵ *3.2. Global energy minimization*

In this section, we derive a global minimization scheme for the energy function (6). Since the loss function $\phi\left(y_x;\cdot\right)$ is monotonously decreasing for $y_x > 0$, the maximization of s_k with respect to $k \in [m]$ is equivalent to the minimization of $\phi\left(y_x;s_k\right)$ for non-negative y_x. For negative y_x, the minimization is equivalent

12

270 to the maximization of $\phi(y_x; s_k)$, since then $\phi(y_x; \cdot)$ increases monotonously. Thus, the energy (6) of the multi-object model can be written as

$$
J(\theta) = \sum_{x \in R} \Big([y_x \geq 0] \cdot \min_{k \in [m]} \{\phi(y_x; s_k(x))\} \Big)
$$
$$
+ \Big([y_x < 0] \cdot \max_{k \in [m]} \{\phi(y_x; s_k(x))\} \Big),
$$
(7)

where $[\text{statement}] = \{1 \text{ if statement} = \text{true}; 0 \text{ else}\}$ are the Iverson brackets.

The loss function ϕ only attains non-negative values. Thus, the sum over $\phi(y_x; s_k(x))$ for $k \in [m]$ is an upper bound of the maximal $\phi(y_x; s_k(x))$. Using

275 this upper bound to reformulate Eq. (7) leads to

$$
\bar{J}(\theta) = \sum_{x \in R} \Big([y_x \geq 0] \cdot \min_{k \in [m]} \{\phi(y_x; s_k(x))\} \Big) + \Big([y_x < 0] \cdot \sum_{k \in [m]} \phi(y_x; s_k(x)) \Big)
$$
(8)

with the property $0 \leq J(\theta) \leq \bar{J}(\theta)$ for all θ. Hence, the minimization of the upper bound $\bar{J}$ also minimizes the energy J.

The pointwise minimum of $\phi(y_x; s_k(x))$ over $k \in [m]$, as given in Eq. (8), can be written as a pointwise minimization with respect to a binary indicator

280 vector $z(x)$. We hence introduce a binary vector $z(x) \in \{0, 1\}^m$ for each image point x, where $z_k(x) = 1$ means that the k-th model is active at x. Then, the energy (8) can be expressed as a pointwise minimization with respect to z, that is

$$
\bar{J}(\theta) = \min_{z \in \mathcal{Z}} \bar{J}_z(\theta)
$$
(9)

with

$$
\bar{J}_z(\theta) = \sum_{x \in R} \sum_{k \in [m]} [y_x \geq 0] \cdot z_k(x) \cdot \phi(y_x; s_k(x)) + [y_x < 0] \cdot \phi(y_x; s_k(x)).
$$
(10)

285 The constraint $z \in \mathcal{Z}$, where $\mathcal{Z} = \{z : \mathbb{R}^2 \to \{0, 1\}^m | \mathbb{1}_m^\top z(x) \geq 1 \, \forall x \in R\}$ and $\mathbb{1}_m^\top$ is an $1 \times m$ vector of values one, enforces that there must be an active model for each $x \in R$. Reordering the two sums in Eq. (10) leads to

$$
\bar{J}_z(\theta) = \sum_{k \in [m]} \sum_{x \in R} [x \in Z_k \vee y_x < 0] \cdot \phi(y_x; s_k(x)),
$$
(11)

13

where $Z_k = \{x \in R | z_k(x) = 1\}$ is the activity region of the k-th shape model. The minimization of $\bar{J}_z(\theta)$ with respect to z and θ also minimizes the upper

290 bound $\bar{J}$ of the energy J and thus determines the optimal fitting of the multi-object model (5) to the image data. The energy $\bar{J}_z(\theta)$ is convex in θ for fixed z, but it is non-smooth in z. However, since the order of minimization with respect to z and $\theta_1, \ldots, \theta_m$ is interchangeable, the minimization of Eq. (11) boils down to

$$\inf_\theta \min_{z \in \mathcal{Z}} \bar{J}_z(\theta)) = \min_{z \in \mathcal{Z}} \left\{ \sum_{k \in [m]} \inf_{\theta_k} \sum_{x \in R} [x \in Z_k \vee y_x < 0] \cdot \phi(y_x; s_k(x)) \right\}. \quad (12)$$

295 Notably, if $Z_k = \emptyset$, then $\inf_{\theta_k} \sum_{x \in R} [x \in Z_k \vee y_x < 0] \cdot \phi(y_x; s_k(x)) = 0$. We use an indicator vector $u \in \{0, 1\}^m$ and define $u_k = 0$ as an equivalent representation of $Z_k = \emptyset$. With this substitution ($u_k = [Z_k = \emptyset]$), we then obtain

$$\inf_\theta \min_{z \in \mathcal{Z}} \bar{J}_z(\theta) = \min_{z \in \mathcal{Z}} \langle u(z), f(z) \rangle, \quad (13)$$

as a reformulation of Eq. (12), where $f \in \mathbb{R}_{\geq 0}^m$ is a vector with the components

$$f_k = \inf_{\theta_k} J_{R_k}(\theta_k), \quad R_k = Z_k \cup \{x \in R | y_x < 0\} \quad (14)$$

and the convex energy J_{R_k} is given by Eq. (4).

300   Below, we describe the solution of Eq. (13) by characterizing the model activity regions $Z_1, \ldots, Z_m$ as a subset of finitely many fixed region prototypes. Formally, let $Z_1^*, \ldots, Z_m^*$ be the model activity regions, which correspond to the optimal $z \in \mathcal{Z}$ with respect to Eq. (13). Then, each Z_k^*, which is non-empty ($u_k = 1$), is unique among the regions $Z_1^*, \ldots, Z_m^*$, because otherwise the

305 optimality assumption is contradicted due to $f_1, \ldots, f_m \geq 0$. Hence, $Z_1^*, \ldots, Z_m^*$ form a subset of an overcomplete set

$$\mathcal{S} = \{Z_1, \ldots, Z_n\} \quad (15)$$

of $n \geq m$ region prototypes, as detailed in the next paragraph. Using a slightly different connotation of the vector u, each region prototype $Z_k \in \mathcal{S}$ is either not included in the solution ($u_k = 0$) or it is included once ($u_k = 1$). Thus, the

14

310 minimization in Eq. (13) can be solved by minimizing with respect to u directly instead of z,

$$\inf_{\theta} \min_{z \in \mathcal{Z}} \bar{J}_z \left(\theta \right) = \min_{u \in \mathcal{U}} \left\langle u, f \right\rangle, \tag{16a}$$

where u and f are now vectors in $\mathbb{R}^n$ and

$$\mathcal{U} = \left\{ u \in \{0,1\}^n \middle| \sum_{k \in [n]} \left[x \in Z_k \right] \cdot u_k \geq 1 \, \forall x \in R \right\}. \tag{16b}$$

The vector f in Eq. (16a) is independent of z and u, since it now represents the energies (14) of the region prototypes $Z_1, \ldots, Z_n$, which are fixed. The

315 constraint $\sum_{k \in [m]} \left[x \in Z_k \right] \cdot u_k \geq 1 \, \forall x \in R$ in Eq. (16b) enforces that the union of the included prototypes equals R, that is, covers the whole image.

The vector f is invariant to the values of $z \left(x \right)$ at image points $x \in R$ with $y_x < 0$, since in Eq. (14), the energy J_{R_k} is minimized for the *union* of the region prototype Z_k and all image points x with $y_x < 0$. For the computation of the

320 set (15), it is hence sufficient to consider only those region prototypes $Z_1, \ldots, Z_n$, which differ with regard to the image foreground (image points x with $y_x > 0$). Thus, each optimal region Z_k^* covers all image points of a *single* cell nucleus and an *arbitrary* part of the image background. This observation motivates two mild assumptions, which prove to be convenient for the characterization of $\mathcal{S}$ as

325 a set of a computationally tractable cardinality:

1. *Model activity regions are simply connected.* This assumption is reasonable since cell nuclei are physically connected objects without holes, although they might *appear* as objects with holes due to staining.

2. *Each model activity region contains one or more local intensity peaks.* This

330 is generally the case, since cell nuclei are brighter than the image background in fluorescence microscopy images. Correspondingly, *the boundary of an activity region is located at an intensity valley.* However, not all intensity valleys represent boundaries of activity regions.

Based on the first assumption, we characterize each region prototype Z_k in $\mathcal{S}$ as

335 a simply connected union of region *fragments* Ω. Each fragment $\omega \in \Omega$ is formed

15

around a local intensity peak $r\left(\omega\right)$ and the fragments are separated by intensity valleys, due to the second assumption. We propose Algorithm 1 to generate Ω and the prototype set $\mathcal{S}$. The algorithm constructs the fragment adjacency graph $(\Omega, \mathcal{E})$, where $\mathcal{E} \subset \Omega \times \Omega$ and the adjacency of two fragments $\omega, \omega' \in \Omega$

340 with sufficiently close local intensity peaks $r\left(\omega\right), r\left(\omega'\right)$ is represented by the edge $\{\omega, \omega'\} \in \mathcal{E}$. The region prototypes $\mathcal{S}$ are then obtained as locally confined subgraphs. The cardinality of the generated set $\mathcal{S}$ is smaller than $|\Omega| \cdot 2^{\Delta^h}$, where Δ is the maximum degree of the fragment adjacency graph $(\Omega, \mathcal{E})$ and h is the maximum search depth. Thus, for any fixed h and degree Δ, the cardinality of

345 $\mathcal{S}$ grows linearly with the number $|\Omega|$ of fragments. The number of fragments is controlled either by choosing the smoothing strength σ of the Gaussian filter or by limiting the set Π in Algorithm 1 to a fixed number of most significant local intensity maxima. Although we do not control Δ directly, the average degree of the fragment adjacency graph is strictly smaller than 6, since the graph is

350 planar. The runtime complexity of Algorithm 1 is at worst quadratic in the number of generated prototypes, since duplicate prototypes must be sorted out to form the prototype set $\mathcal{S}$.

Incorporation of region fragments enables formulating the minimization in Eq. (16a) as an integer linear program (ILP):

$$\min_{u \in \{0,1\}^n} \quad \langle u, f \rangle \quad \text{subject to} \quad \sum_{Z_k \in \mathcal{S}:\omega \subseteq Z_k} u_k \geq 1 \ \ \forall \omega \in \Omega. \qquad (17)$$

355 Since the region prototypes $Z_1, \ldots, Z_n$ are independent of u, the computation of each component of the vector f amounts to solving an individual convex program, as detailed in Section 3.3 below. The combinatorial minimization by the ILP (17) yields the vector u. This determines, which of the n prototypes are to be used to form the model activity regions ($u_k = 1$), subject to the constraint

360 that an active model exists at each point of the image. Thus, our derivation shows that model fitting using the multi-object model (5) corresponds to solving the NP-hard *min-weight set-cover* problem (17) (e.g., Vazirani, 2001).

Due to the overcompleteness of the prototype set $\mathcal{S}$, the number m of models being fitted by the ILP (17) only has the upper-bound $n \geq m$, which is the

16

Algorithm 1: Generating the fragments Ω and the prototype set $\mathcal{S}$.

---

input: Image $g : R \to \mathbb{R}_{\geq 0}$, smoothing strength $\sigma \geq 0$, relative intensity

threshold $0 \leq \varepsilon \leq 1$, minimum seed distance $\delta \geq 1$, maximum

search depth $h \geq 0$, maximum fragments distance $d \geq 0$.

Apply Gaussian filter with standard deviation σ to the image g;

Let $B_\delta(p)$ be $\{q \in R | \|p - q\| \leq \delta\}$;

$\Pi \leftarrow$

$\quad \left\{ p \in R \big| g(p) = \max_{q \in B_\delta(p)} g(q) \wedge (1 - \varepsilon) \cdot g(p) \geq \min_{q \in B_\delta(p)} g(q) \right\}$;

$\Omega \leftarrow \{\omega \in \mathbb{P}(R) | \omega$ is a region of Π-seeded watershed transform of $g\}$;

Let $r(\omega)$ be the centroid of $\Pi \cap \omega$;

$\mathcal{E} \leftarrow \left\{ \{\omega, \omega'\} \in \Omega^2 \big| \omega$ is adjacent to $\omega' \wedge \|r(\omega) - r(\omega')\| \leq d \right\}$;

$\mathcal{S} \leftarrow \emptyset$;

for $\omega_0 \in \Omega$ **do**

$\quad$ **for** *each simply connected subgraph* $(\Omega', \mathcal{E}')$ *of* $(\Omega, \mathcal{E})$ *induced by*

$\quad\quad \Omega' \subseteq \Omega$ *with* $\omega_0 \in \Omega'$ *and* $\max_{\omega \in \Omega'} \mathrm{dist}_{\mathcal{E}'}(\omega_0, \omega) \leq h$ **do**

$\quad\quad\quad \mathcal{S} \leftarrow \mathcal{S} \cup \{Z\}$ where $Z = \cup_{\omega \in \Omega'} \omega$;

return $\Omega, \mathcal{S}$;

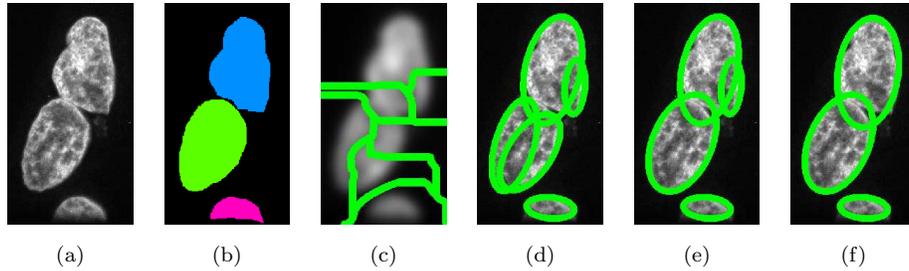(a)   (b)   (c)   (d)   (e)   (f)

Figure 4: Influence of the tolerance for shape irregularities α in the ILP (18) on the segmentation result. (a) Original image section (HeLa cells). (b) Ground truth. (c) Model activity region fragments Ω (green contour) generated from intensity peaks of the smoothed image (Algorithm 1). (d) Segmentation result (green contour) using $\alpha = 0$, (e) $\alpha = 0.6\,\tilde{f}$, and (f) $\alpha = 1\,\tilde{f}$, where $\tilde{f}$ is the median of the components of the vector f.

cardinality of the prototype set $\mathcal{S}$. Since this number is usually larger than the number of cell nuclei in an image, this likely leads to oversegmentation of non-ideally elliptical cell nuclei, as the example in Figure 4 shows. The two large, slightly irregularly shaped cell nuclei in Figure 4(a) are falsely split (Figure 4(d)). The reason is that, for our multi-object model, these cell nuclei are rather cases of two overlapping cell nuclei, than individual objects. To cope with that, we incorporate a tolerance for slight shape irregularities by demanding additional sparsity for the solution u of the ILP (17). This is done by introducing a penalty $\alpha > 0$ for each selected prototype ($u_k = 1$) into the objective function of the ILP (17) which leads to the ILP

$$
\begin{aligned}
&\min_{u \in \{0,1\}^n} \quad \langle u, f \rangle + \alpha \cdot \langle u, \mathbb{1}_n \rangle \\
&\text{subject to} \quad \sum_{Z_k \in \mathcal{S}: \omega \subseteq Z_k} u_k \geq 1 \ \ \forall \omega \in \Omega.
\end{aligned}
\tag{18}
$$

The ILP (18) is identical to (17) for $\alpha = 0$. It is beneficial to choose a value of α in the same range as the components of the vector f. Throughout this work, we specified α as a multiple of the median $\tilde{f}$ of the components of the vector f. Figures 4(e) and 4(f) show that the segmentation result improves by increasing α up to $\alpha = \tilde{f}$. Choosing a too large value for α might cause false merges of closely located cell nuclei.

18

The ILP (18) incorporates both shape and intensity information through the components of the vector f defined in Eq. (14). Solving this ILP boils down to merging adjacent region fragments if this improves the value of the objective function, as described in Section 3.3 below. This enables coping not only with touching and partially overlapping cell nuclei, but also with intensity inhomogeneities of the cell nucleus due to staining.

3.3. Optimization methods and implementation details

To compute the components of the vector f according to Eq. (14), the energy (4) needs to be minimized for each region prototype $Z_k \in \mathcal{S}$. This energy depends on the intensity offset τ of the intensity model (2). In fluorescence microscopy images, we need to cope with cell nuclei of varying intensities and non-homogeneous image backgrounds. Thus, instead of using a global intensity offset τ for the whole image, we computed τ adaptively for each region Z_k. The value of τ was selected either using Otsu thresholding (Otsu, 1979) or by determining the first mode in the intensity distribution of the image region, obtained by kernel density estimation (Parzen, 1962) using Gaussian kernels.

In order to accelerate the computation of the vector f, we exploited that the vector components are independent of each other by computing the components in parallel. For each component, we solved the minimization of the energy (4) using an iterative second-order solver for non-linear convex programming (Andersen et al., 2015). The numerical scheme determines the zeros of the first-order energy derivative ∇J_R with respect to θ and exploits the positive definiteness of the Hessian matrix $\nabla^2 J_R$ for rapid convergence. The first- and second-order derivatives of the energy (4) are

$$\nabla J_R \left(\theta\right) = - \sum_{x \in R} y_x \cdot \kappa_{x,\theta} \cdot \nabla s \left(x\right), \quad \kappa_{x,\theta} = \frac{1}{1 + \exp \left(y_x \cdot s \left(x;\theta\right)\right)},$$
$$\nabla^2 J_R \left(\theta\right) = \sum_{x \in R} y_x^2 \cdot \left(\kappa_{x,\theta} - \kappa_{x,\theta}^2\right) \cdot \nabla s \left(x\right) \nabla^\top s \left(x\right), \tag{19}$$

where ∇s is the gradient of the linear model (1) with respect to its parameters θ. Choosing the basis $E_1 = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right]$, $E_2 = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, $E_3 = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$ to represent the matrix

19

$A = \sum_{i=1}^{3} a_i \cdot E_i$ and $\theta^{\top} = [a_1 \ a_2 \ a_3 \ b_1 \ b_2 \ c]$ for the vectorial representation of the parameters θ, we obtain the gradient of the linear model as $\nabla^{\top} s(x) = [x_1^2 \ x_2^2 \ 2x_1 x_2 \ x_1 \ x_2 \ 1]$. We used the zero-vector for initialization. However, the initialization can be arbitrary, since the Hessian matrix $\nabla^2 J_R$ in Eq. (19) is positive-semidefinite for all θ.

The min-weight set-cover problem (18) is NP-hard. In general, it thus cannot be expected that an exact solution for u is obtained in polynomial time. However, a simple greedy heuristic is known to determine an approximate solution within an approximation guarantee of factor $H(|\Omega|)$ or better (Johnson, 1974), where $H(t) = \sum_{i=1}^{t} 1/i$ is the t-th harmonic number. We used the Algorithm 2 to solve the ILP (18), where the greedy heuristic is combined with a local search. The local search merges adjacent regions if this decreases the energy $P = \langle u, f + \mathbb{1}_n \cdot \alpha \rangle$ of the solution. Hence, the approximation ratio of Algorithm 2 is *never worse* than $H(|\Omega|)$. This conservative lower bound can be tightened *a posteriori* by solving the linear programming (LP) relaxation. Let P_{LP}^* be the exact solution (energy) of the LP relaxation of the ILP (18). Since P_{LP}^* is a lower bound of the unknown optimal solution P^* of the ILP (18), the achieved approximation ratio of the solution P is at least P_{LP}^*/P. We found that Algorithm 2 succeeded in determining a de-facto exact solution ($P_{\mathrm{LP}}^*/P \geq 99\%$) in at least 91.5% of our experiments (Section 4) and the obtained worst lower bound of the ratio was 88.7%. The runtime complexity of Algorithm 2 is at worst quadratic in the cardinality of $\mathcal{S}$.

The final segmentation result is given by the subset of the shape models $\mathcal{C}(s_1), \ldots, \mathcal{C}(s_n)$ which are identified by $u_k = 1$, as determined by Algorithm 2. Ellipses with a significant overlap (larger than 40% or 50%, depending on the image data) were considered as single objects.

4. Experimental results

We have applied our multi-object model-based approach GOCELL (globally optimal collaborating ellipses) to 2-D fluorescence microscopy image data. Our

20

Algorithm 2: Determining the global solution of the ILP (18) within an approximation guarantee of $H\left(|\Omega|\right)$ or better.

---

input: Vector f, sparsity α, region fragments Ω, region prototypes $\mathcal{S}$.

Initialize $u \leftarrow 0 \cdot \mathbb{1}_n$; $\mathcal{V} \leftarrow \Omega$; $\mathcal{Z} \leftarrow \mathcal{S}$;

while $\mathcal{V} \neq \emptyset$ **do**   // greedy

$\quad\vert\quad c_k \leftarrow \frac{f_k+\alpha}{|Z_k \cap \mathcal{V}|}$ for all $k = 1, \ldots, n$ where $\mathcal{S} = \{Z_1, \ldots, Z_n\}$;

$\quad\vert\quad$ Choose k' such that $c_{k'} = \min_k c_k$;

$\quad\vert\quad u_{k'}, \mathcal{V} \leftarrow 1, \mathcal{V} \setminus \{Z_{k'}\}$;

while $\mathcal{Z} \neq \emptyset$ **do**   // local search

$\quad\vert\quad$ Choose k' such that $Z_{k'} \in \mathcal{Z} \wedge f_{k'} = \min_{k:Z_k \in \mathcal{Z}} f_k$;

$\quad\vert\quad$ **if** $u_{k'} = 0 \wedge \exists v \in \{0,1\}^{|\mathcal{S}|} : v \leq u \wedge \cup_{k:v_k=1} Z_k = Z_{k'}$ **then**

$\quad\vert\quad\quad\vert\quad$ **if** $f_{k'} + \alpha < \langle v, f + \alpha \cdot \mathbb{1}_n \rangle$ **then**

$\quad\vert\quad\quad\vert\quad\quad\vert\quad$ Update $u \leftarrow u - v$; and $u_{k'} \leftarrow 1$;

$\quad\vert\quad \mathcal{Z} \leftarrow \mathcal{Z} \setminus \{Z_{k'}\}$;

return u;

21

experiments comprise image datasets of five different cell types. We studied the segmentation accuracy as well as the computation time, and performed a comparison with previous methods. To quantify the segmentation accuracy, we used region-based and contour-based measures. The *region-based* measures described below are based on the the set of all ground truth objects $\mathcal{R}$ within an image and the set of all segmented objects $\mathcal{S}$ in that image:

Dice similarity coefficient (Dice). The Dice value is computed as

$$\text{Dice}\left(\mathcal{R}, \mathcal{S}\right) = \frac{2 \cdot |(\cup \mathcal{R}) \cap (\cup \mathcal{S})|}{|\cup \mathcal{R}| + |\cup \mathcal{S}|}. \tag{20}$$

The value ranges from 0 to 1, where 0 means that there is no overlap between the foreground of the ground truth and the foreground of the segmentation result, whereas 1 means perfect agreement.

Rand index (Rand). The Rand index measures the similarity between the segmentation result and the ground truth based on both the image foreground and background (Coelho et al., 2009). The value ranges from 0 (no overlap) to 1 (perfect agreement).

Object-based Jaccard index (SEG). The SEG measure was used in the IEEE ISBI Cell Tracking Challenge (Maška et al., 2014). For each ground truth object $R \in \mathcal{R}$, the measure is defined as

$$\text{SEG}\left(R, \mathcal{S}\right) = \begin{cases} \frac{|R \cap S|}{|R \cup S|} & \text{if } \exists S \in \mathcal{S} : |R \cap S| > 0.5 \cdot |R|, \\ 0 & \text{else} \end{cases} \tag{21}$$

and attains values between 0 and 1. The SEG value is only non-zero, if there is a sufficient overlap between the ground truth object and the segmented object. SEG is thus sensitive to occurrences of falsely split, merged, and undetected objects. The upper bound 1 is only attained if a ground truth object is perfectly segmented.

Notably, the Dice and Rand measures are invariant to falsely merged and falsely split objects, which is not the case for SEG. On the other hand, SEG is invariant

22

460 to false-positive detections, which is not the case for Dice and Rand. Thus, using SEG in combination with Dice or Rand well reflects the segmentation performance from a region-based point of view.

In addition, we used two *contour-based* performance measures defined below, which are based on $D_{\partial R}(x)$ denoting the minimal Euclidean distance of an image

465 point x to the contour ∂R of the ground truth object $R \in \mathcal{R}$:

Hausdorff distance (HSD). The Hausdorff distance,

$$\text{HSD}(R, S) = \max_{x \in \partial S} D_{\partial R}(x), \tag{22}$$

is the maximum distance of the object contour ∂R to the contour ∂S of the corresponding segmented object $S \in \mathcal{S}$ (Bamford, 2003). The HSD is 0 if the segmented contour perfectly matches the ground truth contour.

470 **Normalized sum of distances (NSD).** The NSD measure,

$$\text{NSD}(R, S) = \sum_{x \in S \triangle R} D_{\partial R}(x) \Big/ \sum_{x \in S \cup R} D_{\partial R}(x), \tag{23}$$

is the ratio of the number of image points, which are either only in R or only in S, where each image point is weighted by its distance to the ground truth object contour (Coelho et al., 2009). NSD attains 0 if the ground truth contour perfectly matches the contour of the corresponding

475 segmented object. The NSD is 1 if there is no overlap.

4.1. Computational complexity

First, we studied the computational complexity of our approach using an example microscopy image of DAPI-stained HeLa cell nuclei (Figure 5(a)). The size of the image is 741×1000 pixels. As described in Section 3, the compu-

480 tational complexity of our approach crucially depends on the cardinality of the prototype set $\mathcal{S}$. Therefore, we varied the smoothing strength σ of the Gaussian filter (within the range $[2, 50]$) when applying Algorithm 1 (using $\delta = 20$, $\varepsilon = 1\%$, $h = 2$, $d = 100$). For smaller values of σ (little smoothing), more activity region fragments occur. Assuming that the maximum number of fragments

23

485 adjacent to any single fragment in Ω (i.e., the maximum degree of the fragment adjacency graph $(\Omega, \mathcal{E})$) is constant, the number of prototypes scales linearly with the number of the fragments. However, when the smoothing strength is lowered, the fragments become more irregularly shaped, and thus the maximum degree of the fragment adjacency graph grows. Hence, the cardinality of the

490 prototype set increases somewhat faster with increasing number of fragments (Figure 5(b)). This is tolerable, since the overall runtime grows almost linearly with the number of prototypes (Figure 5(c)), where $\alpha = 2\tilde{f}$ was used for Algorithm 2 and τ was determined by Otsu thresholding. Although the computational complexity of Algorithms 1 and 2 is at worst quadratic in the number

495 of the prototypes, they terminate rapidly, since both algorithms consist of only few instructions per iteration. Thus, the overall runtime is dominated by the computation of the energy values of the region prototypes and grows linearly with the number of the prototypes (controlled by σ). For each value of σ, eight prototypes were processed in parallel using a regular consumer CPU (Intel(R)

500 Core(TM) i7 860 2.80GHz). The overall runtime performance, as a function of the number of region fragments, is shown in Figure 5(d). Our approach terminated after 1 minute for 16 fragments (using $\sigma = 16$) and after 9.7 minutes for 41 fragments (using $\sigma = 6$). In both cases, a Dice value of 94.6 % was achieved. For less than 16 fragments (corresponding to $\sigma > 16$), the fragments become

505 too coarse and the segmentation accuracy reduces (cf. Dice value for less than 1 minute in Figure 5(e)). For more than 16 fragments (corresponding to $\sigma < 16$), the runtime increases but the segmentation accuracy remains high (cf. Dice value for more than 1 minute in Figure 5(e)). Thus, for this example image, $\sigma = 16$ is an optimal trade-off between segmentation accuracy and runtime.

510 For comparison, we also applied another globally optimal approach for cell nuclei segmentation (Descombes, 2017). This approach is also based on a parameterized shape model, but uses a marked point process, that is embedded into a simulated annealing scheme. We applied this approach using circular (MPP) or elliptical (MPPELL) shape models. MPP converged after 38.2 min-

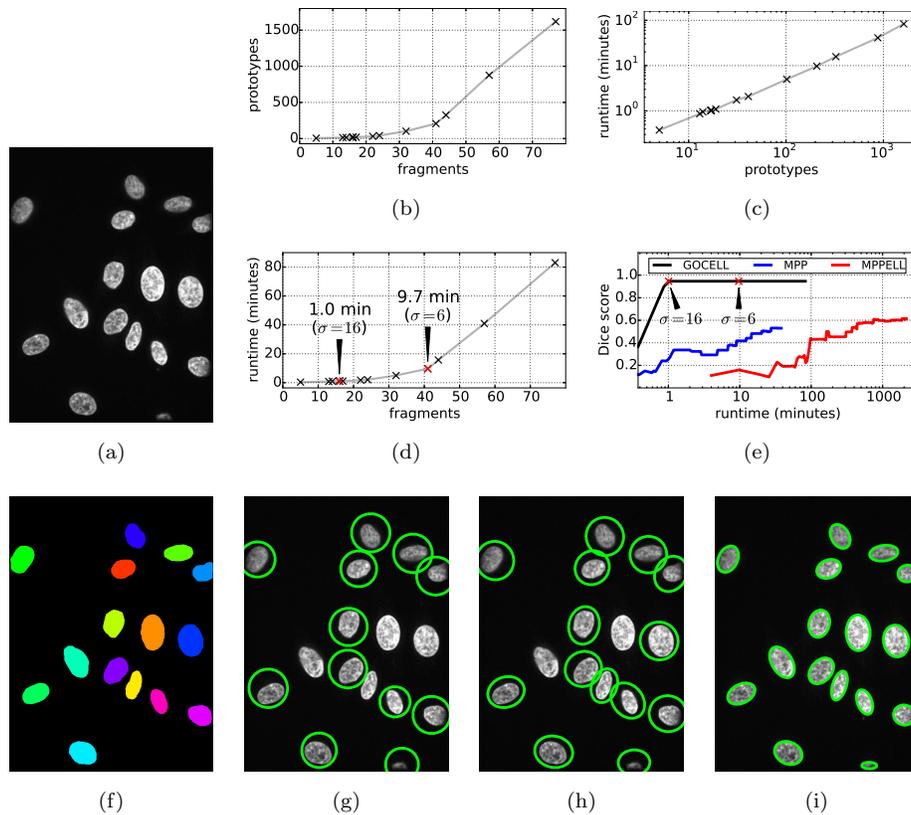515 utes, achieving a Dice value of 52.9 % (Figure 5(g)). MPPELL converged after

24

Figure 5: Runtime performance for an example image. (a) Original image (HeLa cells). (b) Cardinality of the prototype set as a function of the number of activity region fragments. (c) Runtime as a function of the cardinality of the prototype set. (d) Runtime as a function of the number of activity region fragments. (e) Dice score as a function of the runtime of our approach (GOCELL) and an approach based on marked point processes using circular (MPP) or elliptical models (MPPELL). (f) Ground truth segmentation. (g) Segmentation result (green contour) for MPP, (h) MPPELL, (i) and GOCELL.

25

35.7 hours with an improved Dice value (61.6 %, Figure 5(h)). In comparison, our approach (GOCELL) yielded a better Dice value of 94.6 % (Figure 5(i)) after only 1 minute. We note that the computation time of our approach can be straightforwardly reduced by parallelization (e.g., using more than eight CPU threads as in our case). This is possible since computing the energy values of the region prototypes (which dominates the runtime) can be performed independently from each other. Typically, we obtained a few hundred prototype regions in our experiments (553 or less in 95 % of the images). The ratio between the number of region prototypes and the final number of segmented cell nuclei in an image was between 4 and 8 in most cases, the median was 5.4.

4.2. Macrophage, HeLa cell, and fibroblast datasets

Next, we studied the segmentation performance of our GOCELL approach using DAPI-stained images of three different cell types. The first dataset (Ljosa et al., 2012) consists of 20 images of murine bone-marrow derived macrophage cells. The images have a size of 1388×1040 pixels and show 30–50 cell nuclei per image. Half of the images are strongly out of focus, but blurring artifacts also exist in those images which are in focus. The second dataset comprises 25 images of HeLa cells, where each image has a size of 1200×1620 pixels and contains 10–15 cell nuclei. The third dataset (Kalinin et al., 2018) contains 175 3-D stacks of 35–50 images of human fibroblast cells, where each image has a size of 1024×1024 pixels. For most of the 3-D stacks (112 out of 175), the cells were serum-starved and forced to arrest in the same cell cycle phase. For our evaluation, we used the image slice with the largest foreground portion from each of the 175 stacks. In total, 965 cell nuclei are included this dataset.

We applied our approach to all three datasets and performed a comparison with the standard approaches described below.

Global intensity thresholding (Otsu, 1979). An image is binarized using a global intensity threshold based on the histogram.

Blob detection-based level sets (BLB-LS). The level set model of Chan

26

545   and Vese (2001) is used. For initialization, image blobs are detected by
         multi-scale Laplacian of Gaussian filtering (Lindeberg, 1998).

Blob detection-based random walker (BLB-RW). First, a multi-scale Lapla-
         cian of Gaussian-based blob detector (Lindeberg, 1998) is used. For each
         detected blob, a circular foreground marker is initialized (half the radius
550   of the detected blob). The background marker is determined as the wa-
         tershed of the negative image intensities between the foreground markers.
         Foreground and background markers are then expanded using the random
         walker algorithm (Grady, 2006).

For all three methods, we applied pre-processing by Gaussian filtering and post-
555 processing by morphological closing. We optimized the parameters of the three
methods as well as their respective pre/post-processing steps individually for
each dataset. This was accomplished by an automatic grid search scheme, which
maximizes the average Dice and SEG values using two randomly chosen images
from each dataset. In contrast, for our GOCELL approach, we did not adapt the
560 parameters individually for each dataset but used the same set of parameters
for all three datasets (as described in Section 4.1, using $\sigma = 11$).

The macrophage dataset is difficult due to partially strong image blur (see
Figure 6). The quantified segmentation results are provided in Table 1. It turns
out that our approach yields the best result for all three performance measures
565 (SEG, Dice, NSD). Compared to the second-best method (BLB-LS), SEG is
improved by 3.5 % and NSD by 27 %, while for Dice we have a small degradation
of 0.7 %. The highest improvement for SEG is obtained compared to Otsu
(6 %). For this dataset, we also computed results for our approach (GOCELL)
when adapting the parameters (as for the other three methods) by reducing
570 the tolerance for shape irregularities to $\alpha = 0.6\,\tilde{f}$ and using kernel density
estimation instead of Otsu thresholding to determine the intensity threshold τ
(GOCELL*). This improved the SEG value by 2.5 %, the Dice value by 1.4 %,
and the NSD value by 23 %.

For the HeLa dataset, our approach (GOCELL) performed better than all

27

Approach	SEG	Dice	NSD	Parameter sets
Macrophages				
Otsu	66.7 %	81.3 %	0.268	Adapted
BLB-LS	69.2 %	81.4 %	0.227	Adapted
BLB-RW	67.8 %	79.0 %	0.176	Adapted
GOCELL	72.7 %	80.7 %	0.166	Same for all
GOCELL*	**75.2 %**	**82.1 %**	**0.127**	Adapted
HeLa cells				
Otsu	85.4 %	93.7 %	0.077	Adapted
BLB-LS	85.4 %	93.2 %	0.063	Adapted
BLB-RW	68.3 %	81.3 %	0.146	Adapted
GOCELL	87.9 %	**94.3 %**	0.037	Same for all
GOCELL*	**89.0 %**	**94.3 %**	**0.030**	Adapted
Fibroblasts				
Manual	92.3 %	89.5 %	0.008	
Otsu	78.3 %	86.4 %	0.135	Adapted
BLB-LS	71.5 %	83.4 %	0.178	Adapted
BLB-RW	29.3 %	63.8 %	0.281	Adapted
GOCELL	**93.1 %**	**90.9 %**	**0.012**	Same for all

Table 1: Segmentation performance of our approach using the same parameter configuration for all three datasets (GOCELL) or using dataset-specific parameter configurations (GO-CELL*) compared to manual segmentation and standard approaches, which were optimized for each dataset. The best results for each dataset are highlighted in bold.
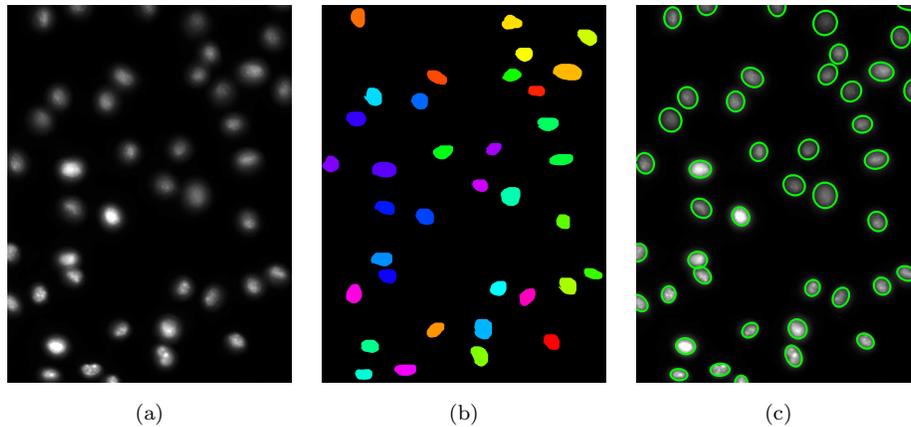
Figure 6: Example results for the macrophage dataset. (a) Original image. (b) Ground truth segmentation. (c) Segmentation result (green contour) using our GOCELL approach.

other methods. Using dataset-specific adaptations by increasing the tolerance for shape irregularities ($\alpha = 3$) further improved SEG by 1.1 % and NSD by 19 % (GOCELL*). The other parameters remained the same as in Section 4.1.

For the fibroblast dataset, our approach (GOCELL) performed better than the other approaches for all three performance measures. An example of a segmentation result is shown in Figure 7. It can be seen that our approach effectively separates touching cell nuclei since shape information is exploited. In contrast, Otsu thresholding, which performed second-best on this dataset, falsely merges closely located cell nuclei (Figure 7(b)). BLB-LS and BLB-RW performed worse since their initialization is prone to the densely located and non-elliptical cell nuclei. For a comparison with the performance of manual segmentation, a human expert manually segmented 34 images, yielding a SEG value of 92.3 % and Dice value of 89.5 %. Thus for our approach it turns out that the SEG and Dice values are higher compared to manual segmentation.

Overall, our approach performed best for all three datasets. The best results were obtained using dataset-specific adaptations for our approach (GOCELL*). However, more importantly, using fixed parameters for all three datasets for our approach (GOCELL) yielded better or comparable results than the other
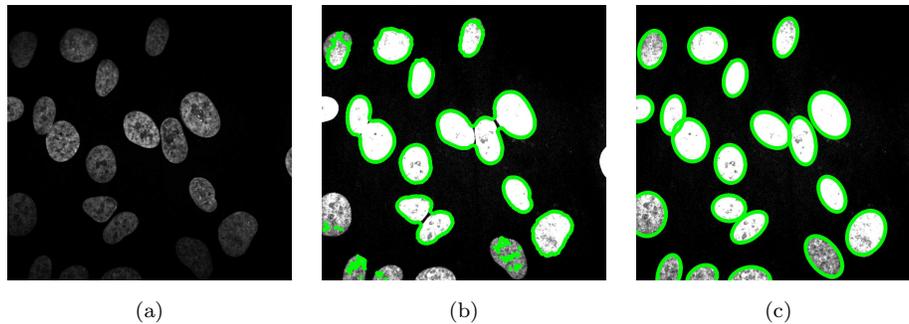
29

Figure 7: Example results for the fibroblast dataset. (a) Original image. (b) Segmentation result (green contour) using Otsu thresholding. (c) Segmentation result (green contour) using our GOCELL approach.

approaches, despite of the heterogeneity of the datasets (cf. Figures 7 and 6).

4.3. NIH3T3 dataset

595   We also applied our approach to the publicly available NIH3T3 dataset (Coelho et al., 2009), which consists of 49 Hoechst-stained images with a size of 1344×1024 pixels. In total, the dataset includes 2209 cell nuclei. The dataset is difficult because of visible artifacts and strong intensity differences between cells. Figure 8 shows an example image with corresponding ground truth data

600   and segmentation results. For our approach, we used $\sigma = 6$ and $\alpha = 0.3\,\tilde{f}$ for all images of the dataset, the other parameters remained the same as in Section 4.1. To better cope with the strong background intensity inhomogeneities in this dataset, we employed local background subtraction based on the minimal intensities of the Gaussian-filtered image (standard deviation 1) within circu-

605   lar neighborhoods (50 pixels radius) of each pixel (see Figure 9). In addition, segmented objects with a radius smaller than 22 pixels were discarded to eliminate the visible debris objects. We quantitatively compared the segmentation performance of our GOCELL approach with our previous convex ellipses-based approach (CVXELL, Kostrykin et al., 2018), which is based on a single-object

610   model, as well as the following state-of-the-art methods, which were reported to achieve the best results on this dataset:

30

Convex variational level sets (CVX-LS). Bergeest and Rohr (2012) use a two-step approach based on convex optimization of variational energies without regarding shape information.

615 **Template matching (TM).** Chen et al. (2012) perform supervised learning to build a filter bank of templates, which is used for the detection of cell nuclei. Non-rigid registration is used for local alignment of the templates within each detected image region.

Graph cuts with blob-like shape prior (GC-BLB). Lou et al. (2012) de620 termine the locations and sizes of cell nuclei based on second-order image statistics. Segmentation is performed by solving a Potts model, where cuts perpendicular to the object edges of the detected cell nuclei are penalized.

Region-based progressive localization (RPL). Song et al. (2013) use a progressive contrast enhancement scheme and pre-trained classifiers for 625 the detection of salient image regions, followed by cluster splitting based on binary classification of local image features.

Bayesian risk-based variational level sets (BR-LS). Gharipour and Liew (2016) use a level set functional based on the binary Bayesian classification risk to segment the image foreground. Individual cell nuclei are identified 630 by morphological analysis and separated using a shortest-path formalism based on the image intensities.

Adaptive thresholding and decremental ellipses (ATDELL). Panagiotakis and Argyros (2018) binarize an image by locally adaptive thresholding and discard falsely detected cell nuclei using thresholds for the appear635 ance. Clusters of cell nuclei are split by approximating each cluster with a number of ellipses, which is decremented as long as the approximation error is sufficiently small.

The results of the different approaches are given in Table 2. The performance values for TM, GC-BLB, RPL, and BR-LS were reported in publications by the
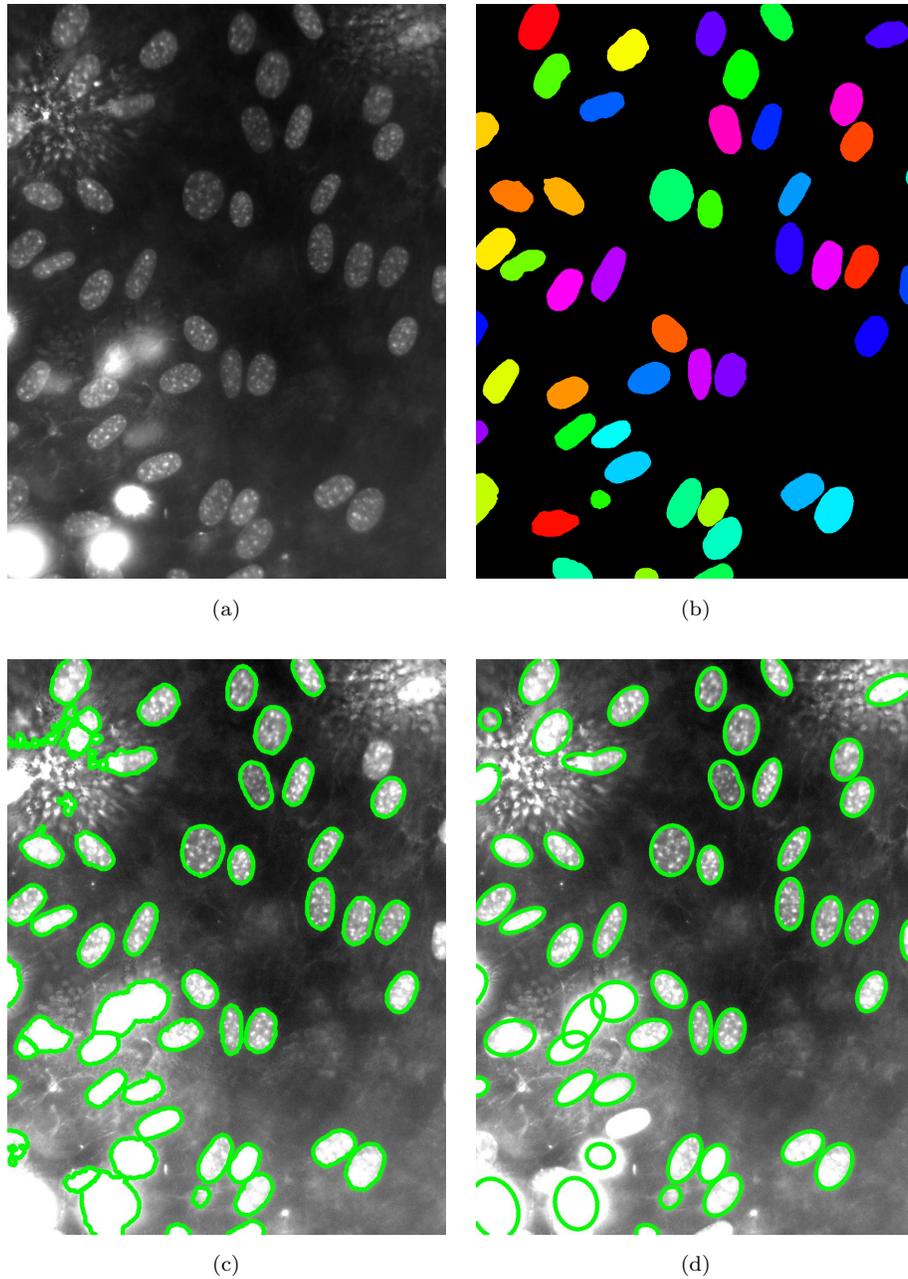
31

Figure 8: Example results for the NIH3T3 dataset. (a) Original image (contrast-enhanced). (b) Ground truth segmentation. (c) Segmentation result (green contour) using the ATDELL approach. (d) Segmentation result (green contour) using our GOCELL approach.
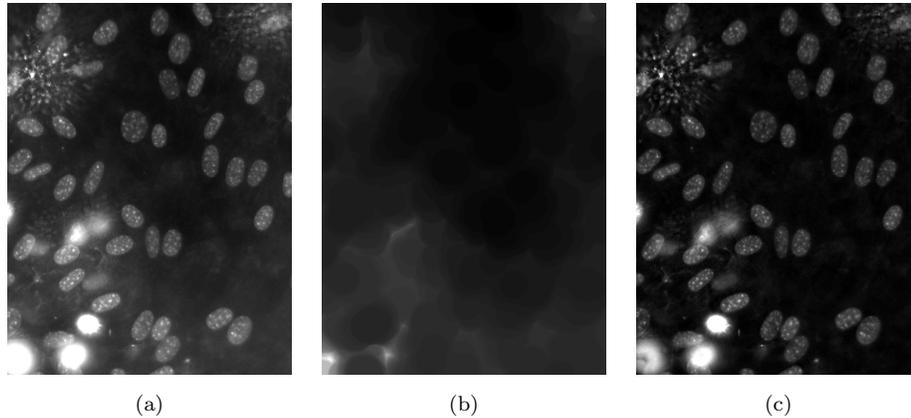
Figure 9: Example for preprocessing of the NIH3T3 dataset. (a) Original image. (b) Computed background. (c) Results after local background subtraction.

authors. It turns out that our approach (GOCELL) achieved the best results for all region- (SEG, Dice, Rand) and contour-based (HSD, NSD) measures. As additional object-based performance measures we determined the average numbers of falsely merged/split cell nuclei per image (as defined in Coelho et al., 2009). Our GOCELL approach yielded the lowest number of falsely merged cell nuclei and the second-lowest number of falsely split cell nuclei. Although ATDELL yielded only slightly more falsely merged cell nuclei per image (0.8 compared to 0.7), much more falsely split cell nuclei were obtained (1.3 compared to 0.4). Also, the overall performance of ATDELL was worse. The lowest number of falsely split cell nuclei was achieved by CVX-LS and TM, but these approaches performed worse with regard to all other performance measures. In particular, falsely merged cell nuclei occurred more than twice as often as for our approach. Compared to RPL, our approach yields a slight improvement for the region-based measures (Dice improved by 0.9%, Rand by 0.4%, SEG was not reported for RPL), but a significant improvement for the contour-based measures (HSD improved by 41%, NSD by 33%). Considering all performance measures, our approach performed overall best on this dataset.

33

Approach	SEG	Dice	Rand	HSD	NSD	Merged	Split
CVX-LS	65.2 %	85.3 %	90.5 %	14.2	0.12	1.6	**0.0**
TM	–	–	88 %	134.1	0.29	4.0	**0.0**
GC-BLB	–	–	91.5 %	10.3	–	1.6	2.4
RPL	–	90.6 %	93.2 %	14.1	0.09	–	–
BR-LS	–	86.3 %	–	14.3	–	–	–
ATDELL	79.7 %	88.7 %	91.8 %	9.5	0.08	0.8	1.3
CVXELL	74.5 %	87.4 %	90.6 %	14.3	0.14	1.4	0.3
GOCELL	**83.7 %**	**91.5 %**	**93.6 %**	**8.3**	**0.06**	**0.7**	0.4

Table 2: Segmentation performance of our GOCELL approach for the NIH3T3 dataset compared to previous approaches. Not available results are indicated by "–". The best results are highlighted in bold.

4.4. GOWT1 challenge datasets

We also applied our approach to two image sets of mouse embryonic stem cells (GOWT1) from the IEEE ISBI Cell Tracking Challenge training data (Maška et al., 2014). The two image sets are temporal image sequences, where each image has a size of 1024×1024 pixels. The sequences consist of 31 and 20 images, which contain 150 and 128 cell nuclei, respectively. The ground truth consists of four fully annotated images from each dataset and partial annotations for the other images. Since for the partially annotated images ground truth is not available for all objects, using a performance measure which is not invariant to false-positive detections would yield misleading results. In previous work (e.g., Akram et al., 2017), only SEG was used as performance measure for the whole dataset, since it is invariant to false-positive detections and reflects the object-based segmentation performance. In our evaluation, we also used SEG for the whole dataset, but additionally used Dice for the fully-annotated images of the dataset.

The GOWT1 datasets are challenging due to a partially low signal-to-noise ratio, the visible presence of the cell nucleoli (distinct dark regions within indi-
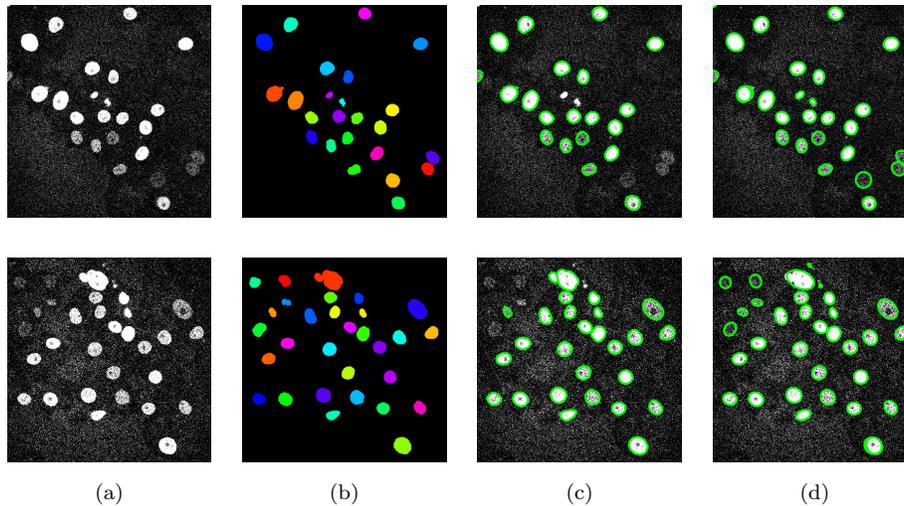
34

| (a) | (b) | (c) | (d) |

Figure 10: Example results for the GOWT1-1 (first row) and GOWT1-2 (second row) datasets. (a) Original images (contrast-enhanced). (b) Ground truth segmentations. (c) Segmentation results (green contour) using the ATDELL approach (Panagiotakis and Argyros, 2018). (d) Segmentation results (green contour) using our GOCELL approach.

vidual cell nuclei), and since for many images only the difficult cell nuclei were annotated in the ground truth. An example image from each dataset, the corresponding ground truth, and the segmentation result of our GOCELL approach are shown in Figure 10. Due to the nucleoli, the cell nuclei often appear rather as bright rings than as ellipses. Therefore, we pre-processed the images using a Laplacian of Gaussian filter to detect small dark regions and decreased the contrast based on the mean intensities inside and outside these regions. For both datasets, we used $\sigma = 10$ and $h = 1$. We used kernel density estimation to determine the intensity offset τ. To reliably separate very noisy, but almost ideally elliptical nuclei (e.g., Figure 10, first row, bottom-right), we reduced the tolerance for shape irregularities to $\alpha = 0.1\,\tilde{f}$ and $\alpha = 0$ for the GOWT1-1 and GOWT1-2 datasets, respectively. All other parameters remained the same as in Section 4.1.

We compared the performance of our approach on the two GOWT1 datasets

35

to ATDELL (see Section 4.3 above) and the following other state-of-the-art methods:

690 **KTH.** Magnusson and Jaldén (2012) use band pass filtering and thresholding for initial segmentation and a watershed transform for subsequent cluster splitting. This method performed overall best for segmentation in the IEEE ISBI Cell Tracking Challenge (Maška et al., 2014).

Blob detection-based approach (BLOB). Akram et al. (2016) use ellipti-

695 cal filter banks to detect cell nuclei and a watershed transform for segmentation. The final segmentation result consists of segmentation candidates obtained from different filter banks. Temporal information from the image sequence is used for candidate selection.

Cell Proposal Network (CPN). Akram et al. (2017) use a region proposal

700 network in conjunction with a convolutional neural network to generate segmentation candidates. Temporal information from the image sequence is used for candidate selection.

Since both BLOB and CPN rely on temporal information to determine the final segmentation result, they are not applicable to individual images.

705 The results for all approaches are given in Table 3. The performance values for KTH, BLOB, and CPN were provided in Akram et al. (2017). For the GOWT1-1 dataset and the SEG measure, GOCELL performed not only better than BLOB (+10.3 %) and ATDELL (+32.4 %), but also significantly better than KTH (+16 %), which achieved the best overall result for segmentation in

710 the ISBI challenge (Maška et al., 2014). GOCELL yielded a slightly worse result compared to CPN (SEG −0.6 %), which, however, exploits temporal information.

For the GOWT1-2 dataset, our approach (GOCELL) achieved a slightly lower SEG value (−0.3 %) than our previous CVXELL approach, but a signif-

715 icantly better Dice value (+5.1 %). More importantly, our approach outperformed ATDELL by 9.2 %, KTH by 1.6 %, and the tracking-based approaches

36

Dataset	Approach					
	Tracking-based		Single-image segmentation			
	BLOB	CPN	KTH	ATDELL	CVXELL	GOCELL
GOWT1-1						
SEG	74.2 %	**85.1 %**	68.5 %	52.1 %	82.1 %	84.5 %
Dice	–	–	–	88.7 %	63.7 %	**94.0 %**
GOWT1-2						
SEG	90.5 %	87.3 %	89.4 %	81.8 %	**91.3 %**	91.0 %
Dice	–	–	–	91.8 %	89.4 %	**94.5 %**

Table 3: Segmentation performance of our GOCELL approach for the GOWT1 datasets compared to previous approaches. The tracking-based approaches (BLOB and CPN) exploit information from several images (temporal coherence) in contrast to the other approaches, which use only the information of a single image. The Dice measure was computed for those images only, for which fully labeled ground truth was available (four images per dataset). Not available results are indicated by "–". The best results are highlighted in bold.

CPN and BLOB by 3.7 % and 0.5 %, respectively. Thus, GOCELL performed overall best on this dataset.

5. Discussion

⁷²⁰ We have introduced a new globally optimal approach for cell nuclei segmentation in fluorescence microscopy images. The approach is based on implicitly parameterized elliptical shape models and incorporates intensity information by a contrast-invariant energy function. An advantage of the single-object model is that the corresponding energy is convex. This means that the energy can be ⁷²⁵ directly globally minimized using an arbitrary initialization. However, since this model represents a single object, prior extraction of image regions is required, which contain at most one cell nucleus. To perform segmentation, which is globally optimal with respect to the *entire* image, we generalized the model such that multiple shape models collaboratively represent *all* objects of an image. ⁷³⁰ The corresponding energy function is non-convex and global minimization is

37

challenging. However, we have derived a global minimization scheme, which is based on activity regions of the individual shape models. Our theoretical considerations have shown that the global solution is invariant with respect to specific non-identical regions. We exploited this observation to reduce the set of possible

735 model activity regions to a computationally tractable size using an overcomplete set of region prototypes. Each region prototype is associated with a non-negative energy, which is the infimum of a convex function. We showed, that the non-convex multi-object model energy is minimized by choosing an energy-minimal subset of region prototypes, which covers the whole image. Since computing

740 this *min-weight set-cover* is NP-hard, a fast approximation algorithm has been used, which is *guaranteed* to determine a solution close to global optimality. In addition, global optimality was checked *a posteriori* and we found, that the global solution was *exactly* determined in at least 91.5% of our experiments comprising 380 images.

745 Previous approaches, which jointly exploit shape and intensity information (e.g., Descombes, 2017), assemble segmentation results by selecting object segmentation masks. In contrast, the model activity regions used in our approach only coarsely subdivide an image compared to the final segmentation result. Hence, the solution space for determining the optimal subset of region proto-

750 types is smaller than if considering directly the segmented objects as in previous approaches. This is advantageous, since combinatorial optimization is computationally challenging. Our combinatorial formulation is also fundamentally different on the conceptual level. Previous approaches (Poulain et al., 2015; Soubies et al., 2015; Descombes, 2017) identified favorable segmentation candidates by

755 *negative* energy values while performing energy *minimization*. Hence, object interaction models (mutual exclusion constraints) were required to prevent non-meaningful solutions, such as the trivial solution (selection of all candidates with negative energy values). Since, however, the energies in our approach are *non-negative*, mutual exclusion constraints are not required, and an object in-

760 teraction model (e.g., maximum allowed object overlap) is not needed. Our non-negative energy minimization scheme intrinsically favors sparse solutions

38

and enables our approach to naturally cope with touching and overlapping cell nuclei. To better cope with non-elliptical cell nuclei, we included a parameter α in our energy minimization scheme, which controls the tendency of recognizing such cell nuclei as single objects.

The computational complexity of our approach depends on the number of the region prototypes. We used *region fragments* to approximate the set of all permissible region prototypes by a set of computationally tractable cardinality. By controlling the coarseness of the fragments, the error introduced by the approximation is balanced against the computation time. Ideally, the fragments are as coarse as possible, but no fragment should cover more than one cell nucleus. Thus, the choice is intuitive and can be adapted in advance. The runtime is dominated by the computation of the energies of the individual region prototypes, which can be highly reduced by parallelization. In addition, adaptation of the parameter α does not require recomputing these energies and is thus fast. Hence, our approach is suitable for high-throughput applications and large datasets.

We applied our approach to fluorescence microscopy images of five different cell types and performed a quantitative comparison with previous methods. We demonstrated the robustness of our approach for datasets of three different cell types (macrophages, HeLa cells, and fibroblasts), achieving equally good or improved results using a fixed set of parameters compared to standard approaches using individually optimized parameters for each of the three datasets. For the NIH3T3 benchmark dataset (Coelho et al., 2009), our approach performed best, achieving a relatively low number of falsely merged/split cell nuclei compared to previous approaches. This highlights the effectiveness of our approach which performs joint segmentation and cluster splitting, as opposed to explicit cluster splitting (e.g., Gharipour and Liew, 2016). Our approach exploits both shape and intensity information jointly, while in Panagiotakis and Argyros (2018) the image intensities were not directly exploited for cluster splitting. In our approach, elliptical models are fitted directly to the image intensities. For the two GOWT1 datasets (Maška et al., 2014), our approach achieved compet-

39

itive or improved results compared to state-of-the-art methods, including two tracking-based approaches which exploit the temporal coherence of the datasets. Moreover, our approach performed overall best among those methods which do not exploit temporal information and are applicable to individual images.

Future research will concentrate on improving the global minimization schemes. In particular, we will study other optimization methods to further increase the robustness by improving the global optimality.

Acknowledgements

References

References

Akram, S., Kannala, J., Eklund, L., Heikkilä, J., 2016. Joint cell segmentation and tracking using cell proposals, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 920–924.

Akram, S., Kannala, J., Eklund, L., Heikkilä, J., 2017. Cell tracking via proposal generation and selection. arXiv preprint arXiv:1705.03386.

Al-Kofahi, Y., Lassoued, W., Lee, W., Roysam, B., 2010. Improved automatic detection and segmentation of cell nuclei in histopathology images. IEEE Trans. Biomed. Eng. 57, 841–852.

Ali, S., Madabhushi, A., 2012. An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery. IEEE Trans. Med. Imag. 31, 1448–1460.

Andersen, M., Dahl, J., Vandenberghe, L., 2015. CVXOPT: A Python package for convex optimization, version 1.1.8. URL: http://cvxopt.org.

40

Bamford, P., 2003. Empirical comparison of cell segmentation algorithms using an annotated dataset, in: Proceedings of the International Conference on Image Processing (ICIP), IEEE. pp. II–1073–1076.

Bergeest, J.P., Rohr, K., 2012. Efficient globally optimal segmentation of cells in fluorescence microscopy images using level sets and convex energy functionals. Med. Image Anal. 16, 1436–1444.

Böhm, A., Ücker, A., Jäger, T., Ronneberger, O., Falk, T., 2018. ISOO DL: Instance segmentation of overlapping biological objects using deep learning, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1225–1229.

Boyd, S., Vandenberghe, L., 2004. Convex Optimization. Cambridge University Press.

Boykov, Y., Funka-Lea, G., 2006. Graph cuts and efficient ND image segmentation. Internat. J. Computer Vis. 70, 109–131.

Chan, T.F., Vese, L.A., 2001. Active contours without edges. IEEE Trans. Image Process. 10, 266–277.

Chen, C., Wang, W., Ozolek, J.A., Lages, N., Altschuler, S.J., Wu, L.F., Rohde, G.K., 2012. A template matching approach for segmenting microscopy images, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 768–771.

Coelho, L., Shariff, A., Murphy, R., 2009. Nuclear segmentation in microscope cell images: A hand-segmented dataset and comparison of algorithms, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 518–521.

Descombes, X., 2017. Multiple objects detection in biological images using a marked point process framework. Methods 115, 2–8.

41

Dong, G., Acton, S., 2007. Detection of rolling leukocytes by marked point processes. J. Electron. Imag. 16(3), 033013.

Dufour, A., Shinin, V., Tajbakhsh, S., Guillén-Aghion, N., Olivo-Marin, J.C., Zimmer, C., 2005. Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces. IEEE Trans. Image Process. 14, 1396–1410.

Fan, M., Rittscher, J., 2018. Global probabilistic models for enhancing segmentation with convolutional networks, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1234–1238.

Fawzi, A., Moosavi Dezfooli, S.M., Frossard, P., 2017. The robustness of deep networks – a geometric perspective. Signal Processing Magazine 34, 50–62.

Gharipour, A., Liew, A.W.C., 2016. Segmentation of cell nuclei in fluorescence microscopy images: An integrated framework using level set segmentation and touching-cell splitting. Pattern Recognit. 58, 1–11.

Goodfellow, I.J., Shlens, J., Szegedy, C., 2015. Explaining and harnessing adversarial examples, in: Proceedings of the International Conference on Learning Representations (ICLR), Computational and Biological Learning Society.

Grady, L., 2006. Random walks for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 28, 1768–1783.

He, Y., Meng, Y., Gong, H., Chen, S., Zhang, B., Ding, W., Luo, Q., Li, A., 2014. An automated three-dimensional detection and segmentation method for touching cells by integrating concave points clustering and random walker algorithm. PloS One 9, e104437.

Johnson, D.S., 1974. Approximation algorithms for combinatorial problems. J. Computer Syst. Sciences 9, 256–278.

Kalinin, A.A., Allyn-Feuer, A., Ade, A., Fon, G.V., Meixner, W., Dilworth, D., Jeffrey, R., Higgins, G.A., Zheng, G., Creekmore, A., et al., 2018. 3D cell nuclear morphology: Microscopy imaging dataset and voxel-based morphometry

42

classification results, in: Proceedings of the Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), IEEE. pp. 2272–2280.

Kong, J., Zhang, P., Liang, Y., Teodoro, G., Brat, D., Wang, F., 2016. Robust cell segmentation for histological images of Glioblastoma, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1041–1045.

Kostrykin, L., Schnörr, C., Rohr, K., 2018. Segmentation of cell nuclei using intensity-based model fitting and sequential convex programming, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 654–657.

Li, C., Kao, C.Y., Gore, J.C., Ding, Z., 2008. Minimization of region-scalable fitting energy for image segmentation. IEEE Trans. Image Process. 17, 1940–1949.

Lindeberg, T., 1998. Feature detection with automatic scale selection. Internat. J. Computer Vis. 30, 79–116.

Ljosa, V., Sokolnicki, K.L., Carpenter, A.E., 2012. Annotated high-throughput microscopy image sets for validation. Nat. Methods 9, 637.

Lou, X., Köthe, U., Wittbrodt, J., Hamprecht, F.A., 2012. Learning to segment dense cell nuclei with shape prior, in: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), IEEE. pp. 1012–1018.

Magnusson, K., Jaldén, J., 2012. A batch algorithm using iterative application of the viterbi algorithm to track cells and construct cell lineages, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 382–385.

Markowsky, P., Reith, S., Zuber, T., König, R., Rohr, K., Schnörr, C., 2017. Segmentation of cell structures using model-based set covering with iterative reweighting, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 392–396.

43

Maška, M., Daněk, O., Garasa, S., Rouzaut, A., Muñoz-Barrutia, A., Ortiz-de Solorzano, C., 2013. Segmentation and shape tracking of whole fluorescent cells based on the Chan–Vese model. IEEE Trans. Med. Imag. 32, 995–1006.

Maška, M., Ulman, V., Svoboda, D., Matula, P., Matula, P., Ederra, C., Urbiola, A., España, T., Venkatesan, S., Balak, D., et al., 2014. A benchmark for comparison of cell tracking algorithms. Bioinform. 30, 1609–1617.

Mollenhoff, T., Laude, E., Moeller, M., Lellmann, J., Cremers, D., 2016. Sublabel-accurate relaxation of nonconvex energies, in: Proceedings of the Conference on Computer Vision and Pattern Recognition (CVPR), IEEE. pp. 3948–3956.

Murphy, K.P., 2012. Machine Learning: A Probabilistic Perspective. MIT press.

Nosrati, M., Hamarneh, G., 2015. Segmentation of overlapping cervical cells: a variational method with star-shape prior, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 186–189.

Otsu, N., 1979. A threshold selection method from gray-level histograms. IEEE Trans. Syst., Man, Cybern. 9, 62–66.

Panagiotakis, C., Argyros, A.A., 2018. Cell segmentation via region-based ellipse fitting, in: Proceedings of the International Conference on Image Processing (ICIP), IEEE.

Parzen, E., 1962. On estimation of a probability density function and mode. Ann. Math. Statistics 33, 1065–1076.

Plissiti, M.E., Nikou, C., Charchanti, A., 2011. Automated detection of cell nuclei in pap smear images using morphological reconstruction and clustering. IEEE Trans. Inf. Technol. Biomed. 15, 233–241.

Poulain, E., Prigent, S., Soubies, E., Descombes, X., 2015. Cells detection using segmentation competition, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 1208–1211.

44

Ronneberger, O., Fischer, P., Brox, T., 2015. U-Net: Convolutional networks for biomedical image segmentation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 234–241.

Singh, R., Sharma, M., Bhattacharya, M., 2017. Segmentation of heavily clustered cell nuclei in histopathological images, in: European Congress on Computational Methods in Applied Sciences and Engineering, Springer. pp. 244–254.

Singh, S., Janoos, F., Pécot, T., Caserta, E., Huang, K., Rittscher, J., Leone, G., Machiraju, R., 2011. Non-parametric population analysis of cellular phenotypes, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 343–351.

de Solorzano, C., Malladi, R., Lelievre, S., Lockett, S., 2001. Segmentation of nuclei and cells using membrane related protein markers. J. Microsc. 201, 404–415.

Song, Y., Cai, W., Huang, H., Wang, Y., Feng, D.D., Chen, M., 2013. Region-based progressive localization of cell nuclei in microscopic images with data adaptive modeling. BMC Bioinform. 14:173, 1–16.

Soubies, E., Weiss, P., Descombes, X., 2015. Graph cut based segmentation of predefined shapes: Applications to biological imaging, in: Pattern Recognit. Applic. and Methods. Springer, pp. 153–170.

Türetken, E., Wang, X., Becker, C., Haubold, C., Fua, P., 2017. Network flow integer programming to track elliptical cells in time-lapse sequences. IEEE Trans. Med. Imag. 36, 942–951.

Vazirani, V., 2001. Approximation algorithms. Springer, Berlin New York.

Wählby, C., Lindblad, J., Vondrus, M., Bengtsson, E., Björkesten, L., 2002. Algorithms for cytoplasm segmentation of fluorescence labelled cells. Anal. Cell. Pathol. 24, 101–111.

45

Xie, W., Noble, J.A., Zisserman, A., 2018. Microscopy cell counting and detection with fully convolutional regression networks. Computer Methods in Biomechanics and Biomed. Eng.: Imag. & Vis. 6, 283–292.

Xing, F., Yang, L., 2015. Fast cell segmentation using scalable sparse manifold learning and affine transform-approximated active contour, in: International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI), Springer. pp. 332–339.

Zhang, P., Wang, F., Teodoro, G., Liang, Y., Brat, D., Kong, J., 2017. Automated level set segmentation of histopathologic cells with sparse shape prior support and dynamic occlusion constraint, in: Proceedings of the International Symposium on Biomedical Imaging (ISBI), IEEE. pp. 718–722.