# Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation using Deformable Shape Models

Leonid Kostrykin and Karl Rohr

**Abstract**—Cell nuclei segmentation is challenging due to shape variation and closely clustered or partially overlapping objects. Most previous methods are not globally optimal, limited to elliptical models, or are computationally expensive. In this work, we introduce a globally optimal approach based on deformable shape models and global energy minimization for cell nuclei segmentation and cluster splitting. We propose an implicit parameterization of deformable shape models and show that it leads to a *convex* energy. Convex energy minimization yields the *global* solution independently of the initialization, is fast, and robust. To *jointly* perform cell nuclei segmentation and cluster splitting, we developed a novel iterative global energy minimization method, which leverages the inherent property of *superadditivity* of the convex energy. This property exploits the lower bound of the energy of the union of the models and improves the computational efficiency. Our method provably determines a solution close to global optimality. In addition, we derive a closed-form solution of the proposed global minimization based on the superadditivity property for non-clustered cell nuclei. We evaluated our method using fluorescence microscopy images of five different cell types comprising various challenges, and performed a quantitative comparison with previous methods. Our method achieved state-of-the-art or improved performance.

**Index Terms**—Segmentation, cell cluster splitting, object representation, image models, surface fitting, global optimization

## 1 Introduction

A central task in many biological studies is the segmentation of cell nuclei in microscopy images, i.e. the identification of image regions corresponding to *individual* cell nuclei. Cell nuclei segmentation is important, for example, to study cellular movement or proliferation, but also to determine regions of interest to quantify other cellular structures. However, cell nuclei segmentation is challenging for many reasons such as irregular shapes, closely clustered or partially overlapping objects, and imaging artifacts (e.g., image noise, image blur, intensity inhomogeneities).

Fluorescence microscopy is one of the most important microscopy methods today (e.g., [1]). With this method, different fluorophores (chemical compounds) are used to label different types of cellular structures (e.g., cell nuclei, cytoplasms, mitochondria). Depending on the wavelength of the exciting light, the fluorescence intensities are captured in different image channels. Usually, the fluorescence of cell nuclei is captured in a single image channel. Cell nuclei typically appear as regions with bright intensity.

### 1.1 Related Work

In the recent years, *convolutional neural networks (CNN)* have been successfully used for cell segmentation in microscopy images (e.g., [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12]). U-Net [2] and Mask R-CNN [13] are the two predominant CNN architectures. To improve the performance, ensembles of multiple CNNs [5], [7], higher-dimensional embeddings [4], [8], focal loss [10], and CNNs in conjunction with graphical models [3], [5], [9] or gradient flow fields [11], [14] were proposed. In [9], Bayesian polytrees with learned deep features and globally optimal inference were used. Globally optimal approaches are intrinsically robust since they guarantee finding the best admissible solution. However, training deep CNNs generally requires large amount of manually annotated data, and manual annotation of pixel-based segmentation masks is tedious. To cope with this, synthetic training data [6] and higher-level annotations for classification of image regions [12] were used. However, training deep neural networks is computationally expensive in general. In addition, it was shown that neural networks are prone to adversarial perturbations (small intensity fluctuations in the input images) [15], universal perturbations (e.g., random noise, geometric transformations) [16], and misannotated image data (label noise) when using deep high-capacity models [17].

Another popular class of segmentation methods are *model-based approaches*, where a model of an object is fitted to the image data via *energy minimization*. Often, object contours are represented by control points (e.g., [18], [19], [20]) or *level sets* of functions in variational frameworks (e.g., [21], [22], [23]). However, the initialization of the control points or the level set functions is crucial due to local minima of the energies. To overcome this, [24] proposed a globally optimal approach for cell segmentation based on *convex* energy minimization [25], [26]. Due to the convexity property, only global energy minima exist and thus minimization does not depend on the initialization.

A central problem in cell segmentation is coping with closely clustered or partially overlapping objects. To address this, methods that exploit shape information have been

• *The authors are with the Biomedical Computer Vision Group, BioQuant, IPMB, Heidelberg University, Im Neuenheimer Feld 267, 69120 Heidelberg, Germany. E-mail: leonid.kostrykin@bioquant.uni-heidelberg.de, k.rohr@uni-heidelberg.de*

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

introduced. These *shape-based* methods often perform segmentation and separation of individual objects (cell cluster splitting) *consecutively*. There are two main schemes:

1) In the first scheme, *image binarization* and morphological analysis are performed to delineate individual objects (e.g., [27], [28], [29], [30], [31]). The approach in [27] uses thresholding [32] for binarization, followed by morphological analysis and clustering to delineate individual cell nuclei, and applies a random walker method [33] for the final segmentation. In [28], level sets were employed for binarization and a shortest-path method to delineate individual objects, which were identified in the binarized image. In [29], thresholding was used for binarization and a Markov point process (MPP) with circular models for cluster splitting. A related approach was introduced in [30], but using elliptical models and expectation maximization instead of MPP, which was later extended to better cope with partially overlapping objects [31]. A disadvantage of these approaches is that shape and intensity information are only used in *consecutive* steps, but not *jointly*. Individual objects are identified using only the initial binarized image.
2) In the second scheme, *prior object detection* is performed to identify individual cells before segmentation (e.g., [34], [35], [36], [37]). For example, [34] employed blob detection to localize individual cell nuclei, and subsequently used graph cuts [38] in conjunction with a shape prior for the final segmentation. Connected component analysis [35] and random decision forests [36] were also proposed for initial detection of individual objects, using level sets with shape priors for the final segmentation. Blob detection in conjunction with convex programming and elliptical models was used in [37]. These methods *jointly* exploit shape and intensity information, but they heavily depend on the result of the initial object detection, which determines how cell clusters are eventually split.

*Shape-based methods* without requiring prior object detection have also been proposed. These methods use variational level sets and *shape priors* (e.g., [39], [40], [41], [42]). Statistical shape priors have been used to segment overlapping cells in histological images, and were initialized by the watershed transform [39] or a variety of features including second-order image statistics [41], [42]. In [40], [42], the sparsity of shape representations in level set models was exploited for faster energy minimization. However, these methods do not yield a globally optimal solution.

Instead of using shape priors in variational level set methods, *shape models* based on either *explicit* (e.g., [43], [44], [45]) or *implicit* parameterizations (e.g., [46]) have been proposed, using stochastic [43], [45] or combinatorial optimization [44], [46] for energy minimization and model fitting. In [43], an overcomplete set of points was used, each corresponding to a possible object detection, and then *explicitly* parameterized elliptical models and an MPP were employed for the final segmentation. In [44], object detection was replaced by an MPP, and graph cuts [47] were used for optimization. A related approach using simulated annealing instead of graph cuts was described in [45]. In principle, this scheme converges to a global solution, but it is computationally expensive and requires careful calibration of the cooling parameters. Compared to explicitly parameterized models, *implicit* parameterizations have analytic and algorithmic advantages (e.g., easier representation of closed curves, convenient representation by matrix-vector multiplication). In particular, implicitly defined shape models are computationally advantageous since energy minimization can often be performed by convex optimization. In our previous work [46], we described a globally optimal approach based on implicit parameterizations and convex optimization for cell segmentation, which, however, is limited to *elliptical* models. None of the methods described above used implicitly defined deformable shape models and global optimization for cell segmentation.

### 1.2 Contributions

In this work, we introduce a new globally optimal approach for cell segmentation in microscopy images, which uses implicitly parameterized *deformable* shape models. Our approach intrinsically copes with non-elliptical shapes and jointly exploits shape and intensity information by convex energy minimization. Neither prior image binarization nor prior object detection are required. The approach is based on three main contributions:

1) We propose an *implicit* parameterization of *deformable shape models* for cell segmentation and show that this parameterization leads to a *convex* energy for model fitting. Minimization of the convex energy determines the global solution independently of the initialization, is fast, and robust.
2) We introduce a novel iterative global energy minimization method, which *jointly* performs cell segmentation and cluster splitting. The method exploits the inherent *superadditivity* property, simultaneously fits multiple models to the image data, and provably determines a solution close to global optimality. The superadditivity property leverages the lower bound of the energy of the union of models and improves the computational efficiency.
3) We also derive a closed-form solution of the global minimization for non-clustered cell nuclei, which is based on the superadditivity property. This further improves the efficiency since iterative minimization is not required.

The core idea of our approach is to consider the *infimum* of the convex energy for a deformable shape model as a *set energy function*, i.e. a function of the set of image regions where model fitting is performed. We determine optimal regions for fitting and show that for these regions the computation of set energy functions amounts to *convex energy minimization*. To perform joint cell segmentation and cluster splitting, we show that the set energy functions are *superadditive* for disjoint image regions. This structural property is established via the *set-packing polytope* and means that optimal image regions can be determined by only considering a subset of all possible image regions. We exploit the inherent property of superadditivity to develop a novel and computationally efficient global energy minimization method, which *iteratively* determines the optimal regions. In addition, we derive a *closed-form* solution of the proposed

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

global minimization, which directly determines optimal regions for non-clustered cell nuclei (without requiring iteration). Our energy minimization method does not suffer from local minima and scale-related hyperparameters are automatically determined to facilitate application to image data with different scales. The proposed approach intrinsically copes with intensity inhomogeneities and partial object overlap since shape and intensity information are used jointly. In contrast to our previous work [46], which was limited to elliptical models, the proposed approach copes with more general shapes by using deformable models, is more efficient since it exploits the property of superadditivity for energy minimization, and is scale invariant. None of the previous methods mentioned above exploited the superadditivity property. Our approach is denoted as SuperDSM since it leverages superadditivity and deformable shape models. To the best of our knowledge, the proposed approach is the first that combines convex optimization with deformable shape models for cell segmentation.

We have evaluated our approach using fluorescence microscopy datasets of five different cell types comprising various challenges, including publicly available benchmark datasets, and performed a quantitative comparison with previous methods. It turns out that the proposed approach generally yields competitive or improved results. In addition, we also demonstrate the applicability of our approach to another imaging modality, namely histopathology images with H&E-stained cell nuclei.

This paper is organized as follows. Section 2 introduces the implicitly parameterized deformable shape models, the corresponding convex energy, and the global energy minimization method which exploits the superadditivity property. Section 3 describes the proposed cell segmentation approach including the pre-processing scheme, the automatic choice of hyperparameters for scale invariance, and the post-processing scheme. Section 4 provides experimental results and a comparison with previous methods. We discuss the results of our work in Section 5.

## 2 Superadditivity and Convex Optimization for Deformable Shape Models

An overview of the proposed SuperDSM approach for cell nuclei segmentation using deformable shape models is shown in Fig. 1. The approach consists of four main steps: 1) Pre-processing (scale estimation, determination of intensity offsets, and detection of regions of possibly clustered objects), 2) coarse-to-fine region analysis (computation of the universe of image regions and the corresponding adjacency graph), 3) global energy minimization using deformable shape models, and 4) post-processing. Step 3 is most important and concerns our main contributions.

Below, we describe the proposed global energy minimization method. We first introduce the implicitly parameterized deformable models (Section 2.1) and the corresponding convex energy (Section 2.2). Then, we describe the superadditive set energy functions, the iterative method for cell segmentation and cluster splitting using global energy minimization, and the closed-form solution for non-clustered cell nuclei (Section 2.3).

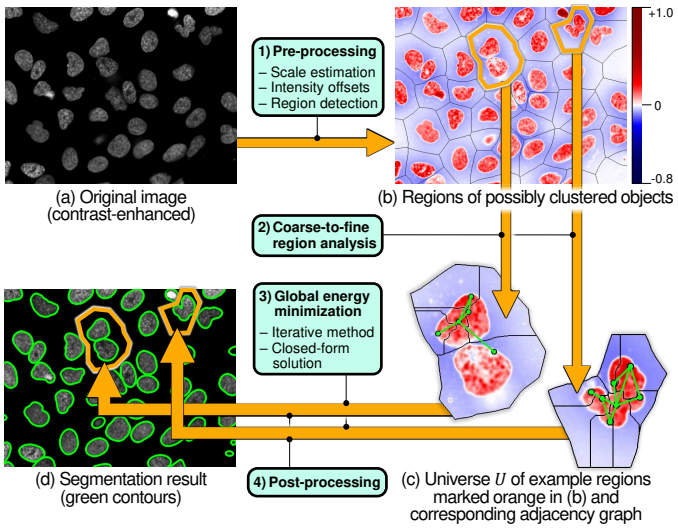


Fig. 1. Overview of our SuperDSM method for cell nuclei segmentation.

### 2.1 Implicit Shape Parameterization

We use the zero-level set of a model function to represent the shape of an object. Our model function consists of a polynomial and local deformations. We use a second-order polynomial $s\colon \Omega \to \mathbb{R}$, which maps each image point $x = (x_1, x_2)$ to a real value, where $x \in \Omega$ and $\Omega \subset \mathbb{R}^2$ are all points of an image. We employ the parameterization

$$\begin{aligned} s\left(x;\theta\right) &= \left\langle f_x, \theta \right\rangle, \qquad \theta \in \mathbb{R}^6, \\ f_x^\top &= \begin{bmatrix} x_1^2 & x_2^2 & 2x_1x_2 & x_1 & x_2 & 1 \end{bmatrix}, \end{aligned} \tag{1}$$

where $\theta$ are the polynomial parameters. The zero-level set of Eq. (1) corresponds to a conic section, which is limited to elliptical shapes and a few degenerated shapes (e.g., hyperbolic). We consider an image region $\omega \subseteq \Omega$, i.e. a non-empty subset of the image points $\Omega$ in an arbitrary but fixed order $\omega = \{x^{(1)}, \ldots, x^{(\#\omega)}\}$, where $\#$ denotes cardinality. Then,

$$S_\omega\left(\theta, \mathbb{0}\right) = F_\omega^\top \theta, \quad \text{where } F_\omega = \begin{bmatrix} f_{x^{(1)}} & \cdots & f_{x^{(\#\omega)}} \end{bmatrix}, \tag{2}$$

describes a *polynomial surface* within the image region $\omega$, where $\mathbb{0}$ is a vector of zeros with arbitrary dimension (used for notational consistency). The parameterization in Eq. (1) was used in our previous work [46] to describe *elliptical* shapes.

To represent more general non-elliptical shapes, we augment the polynomial surface by integrating local deformations. We represent the deformations by the smooth perturbation term $G_\omega \xi$ and define an *implicit deformable* shape model in an image region $\omega$ as

$$S_\omega\left(\theta, \xi\right) = F_\omega^\top \theta + G_\omega \xi, \qquad \xi \in \mathbb{R}^{\#\Omega}. \tag{3}$$

The $\#\omega \times \#\Omega$ matrix $G_\omega$ is a block Toeplitz matrix, where each row represents a Gaussian function with standard deviation $\sigma_G$ centered at the image points $x^{(1)}, \ldots, x^{(\#\omega)}$. The term $G_\omega \xi$ thus corresponds to a linear combination of Gaussian functions, and $\xi$ are the deformation parameters (weights of the Gaussian functions). The deformable model in Eq. (3) includes the elliptical model in Eq. (2) as a special case for $\xi = \mathbb{0}$. The implicit parameterization in Eq. (3)

L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.

has the advantage that it is linear in the model parameters $\theta, \xi$, which leads to a *convex* energy. Thus, minimization yields the global solution and can be performed efficiently (Section 2.2).

Any pair of model parameters $\theta, \xi$ induces two disjoint image regions, that are the zero-sublevel set $\mathcal{I}_S^- (\theta, \xi)$ of the deformable shape model $S_\omega (\theta, \xi) |_{\omega=\{x\}}$ as a function of $x \in \Omega$ and its corresponding zero-superlevel set $\mathcal{I}_S^+ (\theta, \xi)$,

$$\begin{aligned} \mathcal{I}_S^- (\theta, \xi) &= \left\{ x \in \Omega : S_\omega (\theta, \xi) |_{\omega=\{x\}} < 0 \right\}, \\ \mathcal{I}_S^+ (\theta, \xi) &= \left\{ x \in \Omega : S_\omega (\theta, \xi) |_{\omega=\{x\}} > 0 \right\}. \end{aligned} \tag{4}$$

These two regions correspond to the interior and exterior of the model, respectively.

### 2.2 Convex Energy Minimization

We use $g_x \in \mathbb{R}$ to denote the image intensity at an image point $x \in \Omega$. Given the image intensities $g_{x^{(1)}}, \ldots, g_{x^{(\#\omega)}}$, we assume local intensity offsets $\tau_{x^{(1)}}, \ldots, \tau_{x^{(\#\omega)}}$ so that

$$Y_\omega^\top = \begin{bmatrix} g_{x^{(1)}} - \tau_{x^{(1)}} & \ldots & g_{x^{(\#\omega)}} - \tau_{x^{(\#\omega)}} \end{bmatrix} \tag{5}$$

defines a coarse subdivision of the image into the two sets of points for the background $\Omega_{\text{bg}} = \left\{ x \in \Omega : Y_\omega |_{\omega=\{x\}} < 0 \right\}$ and the foreground $\Omega_{\text{fg}} = \left\{ x \in \Omega : Y_\omega |_{\omega=\{x\}} > 0 \right\}$. The intuition is that $g_x - \tau_x < 0$ (i.e. $x \in \Omega_{\text{bg}}$) indicates that an image point $x$ belongs to the background, and $g_x - \tau_x > 0$ (i.e. $x \in \Omega_{\text{fg}}$) indicates that an image point $x$ belongs to the foreground. Since cell nuclei in fluorescence microscopy images correspond to bright intensity regions compared to the background, the offsets $\tau_x$ can be determined, for example, by Gaussian filtering. In our implementation, we have developed a more sophisticated two-step scheme, which interpolates between Gaussian filtering of the original and clipped intensity values to better cope with boundary points, and is more robust to intensity inhomogeneities (for details see Supplemental Material 1).

To fit the implicitly parameterized deformable shape model to the image data, we seek to determine the model parameters $\theta$ and $\xi$ so that $\Omega_{\text{fg}} \approx \mathcal{I}_S^+ (\theta, \xi)$ and $\Omega_{\text{bg}} \approx \mathcal{I}_S^- (\theta, \xi)$. More formally, we minimize the cardinality of the intersections $\Omega_{\text{fg}} \cap \mathcal{I}_S^- (\theta, \xi)$ and $\Omega_{\text{bg}} \cap \mathcal{I}_S^+ (\theta, \xi)$,

$$\inf_{\theta, \xi} \psi_\omega (\theta, \xi), \qquad \psi_\omega (\theta, \xi) = \ell_\omega (\theta, \xi) + \alpha \cdot \|\xi\|_1, \tag{6}$$

where $\ell_\omega (\theta, \xi)$ is the cardinality of the intersections, using an $L_1$ regularization for the deformation parameters $\xi$. Direct minimization of the cardinality corresponds to the minimization of the $0/1$ loss. However, this is challenging since the $0/1$ loss is neither smooth nor convex. We thus use the *logistic loss* as a surrogate loss function (e.g., [48]) for the cardinality of the intersections, i.e.

$$\ell_\omega (\theta, \xi) = \mathbb{1}_{\#\omega}^\top \ln \left( 1 + \exp \left( -Y_\omega \cdot S_\omega (\theta, \xi) \right) \right), \tag{7}$$

where "ln", "exp", "$\cdot$" are defined component-wise, and $\mathbb{1}_{\#\omega}$ is a vector of ones with dimension $\#\omega$. Using the logistic loss is advantageous since Eq. (7) is *convex*. Another advantage of Eq. (7) is that the image intensities are directly exploited via $Y_\omega$ and image binarization is not required. The parameter $\alpha \geq 0$ in Eq. (6) governs the regularization of the deformations. Example segmentation results for different values of $\alpha$ for an image section of U2OS cells are shown in Fig. 2. The section shows a single cell nucleus (according to the ground truth from [49]) and has a size of $136 \times 108$ pixels. Increasing $\alpha$ leads to a smoother segmentation result.

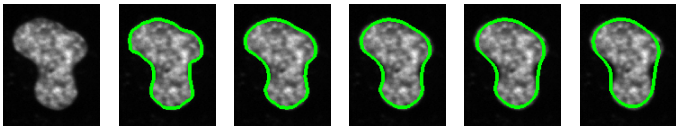

Fig. 2. Example segmentation results for different values of the regularization parameter $\alpha$. Left to right: Original image section, segmentation results using $\alpha = 0.001$, $\alpha = 0.002$, $\alpha = 0.003$, $\alpha = 0.004$, $\alpha = 0.005$.

We solve Eq. (6) which incorporates the convex loss function in Eq. (7) using a Newtonian solver [50] which is a fast numerical second-order method. Consecutive Newton steps determine the globally optimal solution independently of the initialization for $\theta$ and $\xi$ since the problem is unconstrained and convex (Property 1 in Supplemental Material 2).

### 2.3 Global Optimization based on Set Energy Functions and Superadditivity

The implicit deformable shape model introduced above represents a *single* object. For globally optimal model fitting for an entire image, we exploit that linearly parameterized single-object models such as Eq. (2) and Eq. (3) naturally generalize to the *multi-object* case [46]. Let the set $U$ be a *universe* of disjoint image regions, where each region comprises image points of at most one single object (and the image background). The objective then is to determine a *low-cardinality* and *minimal-energy family* $\mathscr{X}$ of sets of the regions $U$, subject to the constraint that $\bigcup \mathscr{X} = U$. The energy of a set $X \subseteq U$ of image regions is given by the solution of Eq. (6) for $\omega = \tilde{\omega} (X) \cup \Omega_{\text{bg}}$,

$$\inf_{\theta, \xi} \psi_{\tilde{\omega}(X) \cup \Omega_{\text{bg}}} (\theta, \xi), \qquad \text{where } \tilde{\omega} (X) = \bigcup X \tag{8}$$

is an image region defined by the set $X$ of disjoint image regions, and $\bigcup X$ are all image points of the set $X$. In [46], this result was used for elliptical models, but determining the optimal family $\mathscr{X}$ required the computation of *all* admissible sets using prior assumptions (e.g., maximum cardinality of $U$) to maintain computational tractability. In this work, we exploit the result for deformable shape models and for the property of *superadditivity*. Superadditivity denotes the property that the energy of any set $X$ is lower-bounded by the sum of energies of its disjoint subsets. This has the advantage that the optimal family $\mathscr{X}$ can be determined by only considering a *subset* of all possible image regions. The proposed global optimization method for deformable shape models is far more sophisticated than the one for elliptical models, since we automatically confine the computations to a meaningful subset of the admissible sets, using the analytical property of superadditivity instead of requiring prior assumptions.

The image regions $U$ are denoted *atomic image regions*, since any set of image regions $X \subseteq U$ is composed of these regions. Below, we first formally define the universe $U$ of atomic image regions, formulate Eq. (8) as a *set energy function*, and derive its property of superadditivity (Section 2.3.1). We then exploit this property to formally define

L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.

a suitable optimization objective for the optimal solution $\mathscr{X}$ (Section 2.3.2). Minimization of the obtained objective is NP-hard and thus computationally challenging. However, we further leverage superadditivity and convexity to decompose the challenging optimization problem into easily solvable sub-problems (Section 2.3.3).

#### 2.3.1 Set energy functions and superadditivity

Let $\mathcal{E} \subseteq U \times U$ represent adjacent image regions, i.e. $\{u, v\} \in \mathcal{E}$ if and only if $u$ is adjacent to $v$, using the following definition of adjacency:

**Definition 1.** Two atomic regions $u, v \in U$ are considered *adjacent* if and only if $\Omega_{\text{fg}} \cap (u \cup v)$ contains a path between $u$ and $v$. $\Pi \subseteq U \times U$ represents the *connected* regions (i.e. $\{u, v\} \in \Pi$ if and only if the adjacency graph $\mathcal{G} = (U, \mathcal{E})$ contains a path between the regions $u$ and $v$).

Fig. 1(c) shows an example universe of atomic image regions (black lines) and the corresponding adjacency graph (green lines) for two connected components.

In order to formally introduce set energy functions (see below), we first define the family of all connected subsets of $U$ with cardinality $k$ or less,

$$\mathbb{P}_k(U) = \{X \subseteq U | \#X \leq k, X \times X \subseteq \Pi\}. \tag{9}$$

In the following, we use $\mathbb{P}(U) = \mathbb{P}_{\#U}(U)$ as a short form.

The objective function in Eq. (8) is defined by the energy $\psi_\omega$ in Eq. (6) for the image region $\omega = \tilde{\omega}(X) \cup \Omega_{\text{bg}}$. Image points within large regions of image background generally yield low energy values and are thus negligible. Since the set $\Omega_{\text{bg}}$ mostly contains such image points, we have

$$\inf_{\theta,\xi} \psi_{\tilde{\omega}(X) \cup \Omega_{\text{bg}}}(\theta, \xi) \approx \inf_{\theta,\xi} \psi_{\tilde{\omega}(X)}(\theta, \xi). \tag{10}$$

This gives rise to the *set energy function* $c' \colon \mathbb{P}(U) \to \mathbb{R}_{\geq 0}$,

$$c'(X) = \inf_{\theta,\xi} \psi_{\tilde{\omega}(X)}(\theta, \xi). \tag{11}$$

Below, we describe a relation of Eq. (11) to the set-packing polytope, which we use to establish the property of superadditivity.

For any family of sets $X_1, \ldots, X_m \subseteq U$, the set $P(X_1, \ldots, X_m)$ of solutions $\eta \in \mathbb{R}^m_{\geq 0}$ for the inequality

$$\sum_{k \in [m]} [u \in X_k] \cdot \eta_k \leq 1 \quad \text{for all } u \in U, \tag{12}$$

is a *set-packing polytope* (the polytope associated with the *max set-packing* problem and its linear relaxation, e.g., [51]), using the Iverson brackets defined by $[\text{statement}] = \{1 \text{ if statement is true; } 0 \text{ else}\}$. Then, *any* family of sets $X_1, \ldots, X_m$ and associated weights $\eta \in P(X_1, \ldots, X_m)$ yields a lower bound of the set energy function in Eq. (11) for the set $X_1 \cup \cdots \cup X_m$ (Property 3 in Supplemental Material 2. The proof uses Property 2 which states a similar relation for the energy $\psi_\omega$.)

When the sets $X_1, \ldots, X_m$ are *disjoint*, a vector of ones with dimension $m$ is always contained in $P(X_1, \ldots, X_m)$ and thus the set energy function $c'$ is *superadditive*,

$$c'(X_1) + \cdots + c'(X_m) \leq c'(X_1 \cup \cdots \cup X_m). \tag{13}$$

Thus, for any two disjoint, non-empty sets $A, B \subset U$, the sum of their energies $c'(A) + c'(B)$ is a *lower bound* of the energy of their union, $c'(A \cup B)$. This means that the energy $c'$ of a set can be directly deduced (without optimization). Moreover, the *singleton* set $\{u\}$ (i.e. set with exactly one element) of any element $u \in U$ is the set with the *lowest* energy among all those containing the element $u$. In terms of energy minimization, this means that *any* image is best fitted by the singleton sets of $U$. This likely leads to over-segmentation and is thus not well-suited. Below, we describe an extension of the set energy functions which avoids over-segmentation.

#### 2.3.2 Extended Set Energy Functions and Optimization Objective

To obtain a meaningful segmentation result and avoid over-segmentation, we extend the set energy functions in Eq. (11). We add the constant term $\beta \geq 0$ and define the *extended set energy function* $c \colon \mathbb{P}(U) \to \mathbb{R}_{\geq 0}$,

$$c(X) = c'(X) + \beta. \tag{14}$$

In contrast to the original set energies $c'$, the extended energies $c(A) + c(B)$ of two disjoint, non-empty sets $A, B \subset U$ can actually be *higher* than the extended energy of their union, $c(A \cup B)$, since $c(A) + c(B) - c(A \cup B) \leq \beta$ due to Eq. (13). Thus, $\beta$ is the maximum allowed energy difference of merging $A$ and $B$. Only if the energy $c'(A \cup B)$ exceeds $c'(A) + c'(B)$ by less than $\beta$, merging $A$ and $B$ is beneficial in terms of energy minimization using the extended set energy $c$. Merging $A$ and $B$ corresponds to using a single deformable shape model for the union $A \cup B$ instead of two separate shape models for $A$ and $B$.

Using linear program (LP) relaxation of the *max-weight set-packing* problem (e.g., [51])

$$\begin{aligned} \text{MSP}_{\text{LP}}(\mathscr{S}) = \max_{\eta_1, \ldots, \eta_m} & \sum_{k \in [m]} \eta_k \cdot c'(X_k), \ \mathscr{S} = X_1, \ldots, X_m \\ \text{s.t.} \quad & \eta \in P(\mathscr{S}), \end{aligned} \tag{15}$$

we obtain the following lower bound of the extended set energy $c(X)$:

**Property 4.** *Given a set $X \subset U$ with cardinality $\#X = k+1 \geq 2$, the rhs of*

$$c(X) \geq \text{MSP}_{\text{LP}}(\mathbb{P}_k(X)) + \beta \tag{16}$$

*is a lower bound of its extended set energy (lhs).*

*Proof.* See Supplemental Material 2. □

To define the optimization objective for global energy minimization, we consider the overall minimal energy for a subset of regions $\mathscr{X} \subseteq \mathbb{P}(U)$, which covers the whole universe $U$. Using the extended set energy from Eq. (14), this formally corresponds to $\text{MSC}(\mathbb{P}(U))$ where

$$\text{MSC}(\mathscr{S}) = \min_{\mathscr{X} \subseteq \mathscr{S}} \sum_{X \in \mathscr{X}} c(X) \quad \text{s.t.} \bigcup \mathscr{S} = \bigcup \mathscr{X}, \tag{17}$$

which is an instance of the *min-weight set-cover* problem (e.g., [51]). Computation of $\text{MSC}(\mathbb{P}(U))$ is challenging for two reasons. First, Eq. (17) is NP-hard. To cope with this, we use an approximation algorithm which determines the global solution within a tight approximation ratio (see Section 3.3). Second, $\mathbb{P}(U)$ has a potentially large cardinality. We address this by avoiding the computation of the *whole* family $\mathbb{P}(U)$, as described below.

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

#### 2.3.3 Global Optimization Scheme

To cope with the potentially large cardinality of $\mathbb{P}(U)$, we are interested in a criterion for a set $X \subseteq U$ which guarantees that the set $X$ is *negligible*, i.e. $\mathrm{MSC}(\mathbb{P}(U)) = \mathrm{MSC}(\mathbb{P}(U) \setminus \{X\})$. Excluding such sets ultimately yields a subset $\mathscr{U} \subseteq \mathbb{P}(U)$, which suffices $\mathrm{MSC}(\mathscr{U}) = \mathrm{MSC}(\mathbb{P}(U))$ but is of lower cardinality than $\mathbb{P}(U)$. We derive such a criterion from the following *lower bound* of the global optimization objective:

**Property 5.** *Let $X \subseteq U$ be a set with cardinality $\#X = k+1 \geq 2$. If $X$ or a superset of $X$ are* not negligible, *i.e.*

$$\exists Y : X \subseteq Y \subseteq U \wedge \mathrm{MSC}(\mathbb{P}(U)) < \mathrm{MSC}(\mathbb{P}(U) \setminus \{Y\}), \tag{18a}$$

*then the lhs of*

$$c(X) + \sum_{u \in U \setminus X} c'(\{u\}) \leq \mathrm{MSC}(\mathbb{P}(U)) \tag{18b}$$

*is a lower bound of the global optimization objective (rhs).*

*Proof.* See Supplemental Material 2. □

In addition to the lower bound, the monotonicity $\mathbb{P}_k(U) \subseteq \mathbb{P}_{k+1}(U)$ yields *upper bounds* of $\mathrm{MSC}(\mathbb{P}(U))$,

$$\begin{aligned} &\mathrm{MSC}(\mathbb{P}(U)) = \\ &\mathrm{MSC}(\mathbb{P}_{\#U}(U)) \quad \leq \ldots \leq \quad \mathrm{MSC}(\mathbb{P}_1(U)) \\ &\qquad\qquad\qquad\qquad = \sum_{u \in U} c(\{u\}). \end{aligned} \tag{19}$$

By combining Property 5 for lower bounds with the upper bounds from Eq. (19), we obtain the following criterion to identify negligible sets:

**Criterion 1 (Negligible sets).** *Given a set $X \subseteq U$ with cardinality $\#X = k+1 \geq 2$, if*

$$c(X) > \mathrm{MSC}(\mathbb{P}_k(U)) - \sum_{u \in U \setminus X} c'(\{u\}), \tag{20a}$$

*then $X$ and its supersets are* negligible, *i.e.*

$$\mathrm{MSC}(\mathbb{P}(U)) = \mathrm{MSC}(\mathbb{P}(U) \setminus \{Y\}) \tag{20b}$$

*for all $Y : X \subseteq Y \subseteq U$.*

*Proof.* See Supplemental Material 2. □

To compute $\mathrm{MSC}(\mathbb{P}(U))$, we consider the sequence $\mathscr{U}_1, \ldots, \mathscr{U}_{\#U}$, where each subset $\mathscr{U}_k \subseteq \mathbb{P}_k(U)$ is obtained by excluding sets according to Criterion 1. This procedure guarantees $\mathrm{MSC}(\mathscr{U}_{\#U}) = \mathrm{MSC}(\mathbb{P}(U))$ due to Eq. (20b) and is formally described in Algorithm 1. First, Property 4 and Property 5 are used to determine a lower bound $c_{\min}$ (line 9) and an upper bound $c_{\max}$ (line 8) of the extended set energy $c(X)$. The set is excluded if the lower bound exceeds the upper bound (line 10). Otherwise, $c(X)$ is computed (line 11) and the set $X$ is excluded if $c(X)$ exceeds the upper bound $c_{\max}$ (line 12). The lower and upper bounds $c_{\min}$ and $c_{\max}$ are tightened from iteration to iteration (due to monotonic increase of the family $\mathscr{U}$ and monotonic decrease of the variable `value`). Thus, if a set $X$ is excluded, all supersets $Y \supset X$ are also excluded in subsequent iterations and computation of $c(Y)$ is not required due to Property 4. The number of iterations is upper-bounded by the cardinality of the universe. An example run-through of Algorithm 1 is given in Supplemental Material 3.

**Algorithm 1:** Iterative solution of $\mathrm{MSC}(\mathbb{P}(U))$.

**input:** Adjacency graph $\mathcal{G} = (U, \mathcal{E})$

1 `iter1` $\leftarrow \{\{u\} | u \in U\}$; *// initialize $\mathscr{U}_1 = \mathbb{P}_1(U)$*
2 $\mathscr{U} \leftarrow$ `iter1`; *// sets for which $c(X)$ was computed*
3 **do** *// iterate $\mathscr{U} = \mathscr{U}_1, \mathscr{U}_2, \ldots$*
4 `value` $\leftarrow \mathrm{MSC}(\mathscr{U})$; *// optimization objective*
5 `iter0` $\leftarrow$ family of all sets $X \cup \{u\}$ where $X \in$ `iter1`, $u \in U \setminus X$, and $\exists v \in X, \{u, v\} \in \mathcal{E}$;
6 `iter1` $\leftarrow \{\}$;
7 **for** $X \in$ `iter0` **do**
8 $c_{\max} \leftarrow$ `value` $- \sum_{u \in U \setminus X} c'(\{u\})$; *// Property 5*
9 $c_{\min} \leftarrow \beta + \mathrm{MSP}_{\mathrm{LP}}\{Y \in \mathscr{U} | Y \subset X\}$; *// Property 4*
10 **if** $c_{\min} \leq c_{\max}$ **then** *// Criterion 1*
11 compute $c(X)$ and insert $X$ into $\mathscr{U}$;
12 **if** $c(X) \leq c_{\max}$ **then** *// Criterion 1*
13 `iter1` $\leftarrow$ `iter1` $\cup \{X\}$;
14 **until** $\#$`iter1` $= 0$;
15 **return** family $\mathscr{X}$ corresponding to "`value`", cf. Eq. (17)

Compared to classically tree-based branch-and-bound schemes, Algorithm 1 builds *multiple* trees $\mathcal{H}$ along the edges of the adjacency graph $\mathcal{G} = (U, \mathcal{E})$, each rooted in the singleton sets of $U$. This corresponds to the directed acyclic graph (DAG)

$$\mathcal{H} = (\mathscr{U}, \mathcal{E}'), \qquad \text{where } (X, Y) \in \mathcal{E}' \text{ iff } X \subset Y. \tag{21}$$

The graph $\mathcal{H}$ comprises only the subset $\mathscr{U}$ of the admissible nodes $\mathbb{P}(U)$. Also, in contrast to previous DAG-based approaches (e.g., [9], [52]), the segmentation is *not* encoded in our graph structure. These two properties naturally lead to comparably *shallow* graphs, thus, neither heuristic pruning [52] nor prior assumptions [46] are required to obtain graphs of computationally tractable size.

Algorithm 1 can be interpreted that it excludes sets corresponding to falsely merged objects (see the example in Supplemental Material 3). However, if the *whole* universe $U$ corresponds to a *single* object, i.e. if $c(U) = \mathrm{MSC}(\mathbb{P}(U))$, then no falsely merged objects can possibly occur. In this case, Eq. (18b) is fulfilled for any $X \subseteq U$ due to Eq. (13), and the computational cost (cardinality of $\mathscr{U}$) grows to $\#\mathbb{P}(U)$. To avoid this, we introduce Criterion 2, which identifies this case a priori and provides a closed-form solution:

**Criterion 2 (Closed-form solution of $\mathrm{MSC}(\mathbb{P}(U))$).** *If $c(U) \leq 2\beta + \sum_{u \in U} c'(\{u\})$, then* $\mathrm{MSC}(\mathbb{P}(U)) = c(U)$.

*Proof.* See Supplemental Material 2. □

Criterion 2 is applied for direct segmentation of non-clustered cell nuclei without using the iterative Algorithm 1 (see Section 3).

The iterative Algorithm 1 is computationally more efficient for *small* values of $\beta$ (Criterion 1 then excludes more sets). On the other hand, the closed-form solution (Criterion 2) is more efficient for *large* values of $\beta$ (since in this case the margin of the inequality in Criterion 2 is larger).

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

## 3 Cell Nuclei Segmentation using Superadditivity, Convex Optimization, and Deformable Shape Models

The proposed SuperDSM method for cell nuclei segmentation consists of four main steps (cf. Fig. 1). First, an image is pre-processed to estimate the scale of the objects, determine the image intensity offsets, and detect image regions corresponding to possibly clustered objects. Second, a coarse-to-fine region analysis is performed to compute the universe $U$ of atomic image regions and the corresponding adjacency graph. Third, global energy minimization is performed by iterative (Algorithm 1) and direct (Criterion 2) solution of $\mathrm{MSC}\left(\mathbb{P}\left(U\right)\right)$ in Eq. (17) above. This includes the automatic choice of hyperparameters for scale invariance, comprising the weight $\alpha$ of the regularization of the deformations in Eq. (6) and the constant term $\beta$ of the extended set energy functions in Eq. (14). Both parameters are determined based on the scale $\sigma$, which is computed automatically in the first step. Fourth, post-processing is performed. The steps are detailed below.

### 3.1 Scale Estimation, Intensity Offsets, and Detection of Regions of Possibly Clustered Objects

To estimate the scale of cell nuclei, we compute the Hessian matrix of the image intensities [53] and determine local maxima of the determinant of the Hessian in scale space [54]. False-positive detections are discarded if *(i)* the detection corresponds to a non-negative response of the Laplacian of Gaussian filter or *(ii)* the scale of the detection is an outlier (determined based on the mean absolute difference to the median scale). The mean of the remaining inliers is then associated with the scale $\sigma$ of an image.

The intensity offsets $\tau_x$ for all image points $x \in \Omega$ are computed using modified Gaussian filtering with standard deviation according to the estimated scale $\sigma$ (see Supplemental Material 1). The denoised image intensities $g_x$ (obtained by a Gaussian filter with standard deviation $\sqrt{2}$) and the intensity offsets $\tau_x$ are then used to compute $Y_\Omega$ (image intensities with $\tau_x$ offset) for all image points $x \in \Omega$ according to Eq. (5) by setting $\omega = \Omega$.

We also determine the universe $U$ of atomic image regions and the corresponding adjacency graph $\mathcal{G} = (U, \mathcal{E})$. In general, $\mathcal{G}$ is disconnected due to Definition 1 (Section 2.3.1). Each *connected* component corresponds to a *region of possibly clustered objects*. For computational efficiency, we first determine these regions, and then the corresponding *connected* adjacency graph $\mathcal{G}$ for each region of possibly clustered objects (formally this is the same as considering the disconnected adjacency graph of the whole image). We consider the connected components of the foreground region $\Omega_{\mathrm{fg}}$ (defined as in Section 2.2). This generally yields several components which correspond to the background due to image noise. We identify such components by considering the perimeter-to-area ratio (P/A ratio). The intuition is that components due to image noise have strongly irregular contours (jagged or wavelike), while isolated and clustered cell nuclei have rather smooth contours. Components with strongly irregular contours can be identified by a large perimeter compared to the area. We discarded such image regions when the P/A ratio is larger than a threshold of 0.2, which was chosen empirically. Our method is not sensitive to the choice of this value (see Supplemental Material 4). We then obtain the regions of possibly clustered objects as the regions of the Voronoi diagram of the remaining connected components (cf. Fig. 1(b)).

### 3.2 Coarse-to-Fine Region Analysis

We next determine the universe $U$ of atomic image regions and the corresponding *connected* adjacency graph $\mathcal{G}$ separately for each region of possibly clustered objects (cf. Fig. 1(c)). The main requirement for the universe $U$ is that each atomic image region overlaps with at most one object, while not generating a universe $U$ of an unnecessarily large cardinality (which would increase the run time of Algorithm 1). We thus start with a region of possibly clustered objects as a whole, determine the *irregularity* of an object in that region, and split the region into smaller parts as long as the irregularity is large, ultimately obtaining the atomic regions (which are not split further). To determine the irregularity of an image region $\omega$, we minimize the energy function in Eq. (6) for $\xi = \mathbb{0}$ and consider the *normalized energy*

$$r\left(\omega\right) = \inf_{\theta} \psi_\omega\left(\theta, \mathbb{0}\right) / \#\omega. \tag{22}$$

Eq. (22) corresponds to fitting an elliptical model to the region $\omega$. It is beneficial to use elliptical models here, since the energy of these models is more sensitive to shape irregularities of objects than the energy of deformable shape models. Since the analysis is performed by splitting large image regions into smaller parts, a *coarse-to-fine region analysis* scheme is induced (in contrast to Algorithm 1 which uses a fine-to-coarse scheme). See Supplemental Material 5 for details.

### 3.3 Scale Invariant Global Energy Minimization

Each graph $\mathcal{G} = (U, \mathcal{E})$ determined as described in Section 3.2 is processed as follows. First, the extended set energies $c\left(U\right)$ and $c\left(\{u\}\right)$ are computed for all atomic image regions $u \in U$ by solving Eq. (6). Second, it is checked whether $U$ corresponds to a non-clustered cell nucleus using the inequality in Criterion 2. If it does, the closed-form solution $\mathscr{X} = \{U\}$ is applied. Otherwise, the iterative Algorithm 1 is used to determine the solution $\mathscr{X}$ of $\mathrm{MSC}(\mathbb{P}\left(U\right))$ using $\mathcal{G}$. Below, we describe the automatic choice of the hyperparameters $\alpha$ and $\beta$ for Criterion 2 and Algorithm 1 to establish scale invariance, and introduce efficient implementations of Algorithm 1 and Eq. (6).

The extended set energy function $c\left(X\right) = c'\left(X\right) + \beta$ in Eq. (14) depends on the hyperparameter $\beta \geq 0$. To properly choose a value for $\beta$, we need to understand how $c'\left(X\right) = \inf_{\theta,\xi} \psi_{\tilde{\omega}(X)}\left(\theta, \xi\right)$ in Eq. (11) depends on the *scale* $\sigma$ of an image. For an arbitrary image region $\omega = \tilde{\omega}\left(X\right)$, recall that the energy $\inf_{\theta,\xi} \psi_\omega\left(\theta, \xi\right)$ approximates the cardinality of a set of image points (Section 2.2). Since the number of image points corresponds to the area, the number changes quadratically with respect to the scale (for 2D images). Thus, it is reasonable to assume that the energy $\inf_{\theta,\xi} \psi_\omega\left(\theta, \xi\right)$ depends quadratically on the scale. Fig. 3(b) shows the energy for image sections of different cell types in different

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

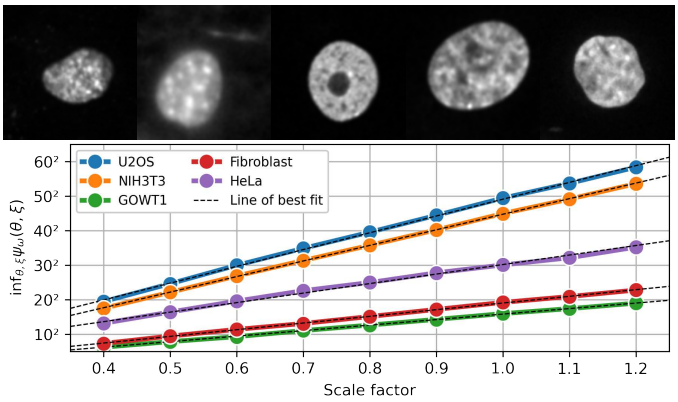


Fig. 3. Relation of the energy $\inf_{\theta,\xi}\psi_\omega(\theta,\xi)$ and the scale $\sigma$. Top: Example image regions of different cell types, from left to right: U2OS, NIH3T3, GOWT1, Fibroblast, HeLa. Bottom: Corresponding energy $\inf_{\theta,\xi}\psi_\omega(\theta,\xi)$ as a function of the scale.

datasets generated by Gaussian filtering and sub-sampling at different scales $\sigma$. It can be seen that the energy depends quadratically on the scale (note that the vertical axis is scaled quadratically). Thus, it is reasonable to choose $\beta = \beta_{\texttt{factor}} \cdot \sigma^2$ as a quadratic function of the scale to achieve scale invariance. In our experiments using image data of different scales and cell types, $\beta_{\texttt{factor}} = 0.33$ turned out to be a reasonable choice (cf. Section 4).

For Algorithm 1, the extended set energies $c(X) = \inf_{\theta,\xi}\psi_{\tilde{\omega}(X)}(\theta,\xi)+\beta$ have to be computed often. Each computation amounts to solving the convex problem in Eq. (6). Efficient implementation of convex energy minimization is thus crucial. Eq. (6) includes the term $G_\omega\xi$ via Eq. (7) and Eq. (3), which can be interpreted as radial basis function interpolation [55] due to the block Toeplitz structure of the matrix $G_\omega$. Reducing the number of points (summands of the interpolant) motivates the approximation $G_\omega\xi \approx \tilde{G}_\omega\tilde{\xi}$, where the matrix $\tilde{G}_\omega$ is constructed from a subset of the columns of $G_\omega$ (the rows are normalized to the sum of 1) and $\tilde{\xi}$ is a vector of lower dimension than $\xi$ (the dimension of $\xi$ corresponds to the overall number of points in an image). The columns of $G_\omega$ correspond to the regular grid of all image points $\Omega$, and the columns of $\tilde{G}_\omega$ correspond to the sub-sampled regular grid of image points within the region $\omega$ spaced by $2\sigma_G$ (where $\sigma_G$ is the standard deviation of the Gaussian function used for the matrix $G_\omega$, cf. Section 2.1). We used $\sigma_G = 0.2\,\sigma$ (where $\sigma$ is the scale of an image). Due to sub-sampling, the number of points of the regular grid and thus also the dimension of $\tilde{\xi}$ scales inverse-quadratically with $\sigma$. Since the regularization parameter $\alpha$ of the energy in Eq. (6) is a factor of $\|\tilde{\xi}\|_1$ (using $\tilde{\xi}$ instead of $\xi$), inverse-quadratic scaling of the dimension of $\tilde{\xi}$ needs to be compensated by quadratically scaling $\alpha$. We used $\alpha = \alpha_{\texttt{factor}} \cdot \sigma^2$ and $\alpha_{\texttt{factor}} = 5 \cdot 10^{-4}$ in our experiments (Section 4). The approximation $G_\omega\xi \approx \tilde{G}_\omega\tilde{\xi}$ concerns only the deformations of the shape representation in Eq. (3). The approximated shape representation has a somewhat lower expressive power (since the matrix $\tilde{G}_\omega$ is of lower rank than $G_\omega$), but substantially increases the computational efficiency, since less parameters need to be determined (the dimension of $\tilde{\xi}$ is lower than $\xi$). Another interpretation of the perturbation term $G_\omega\xi$ is that it corresponds to a low-pass filter of $\xi$, which is due to the block Toeplitz structure of the matrix $G_\omega$ and since each row represents a Gaussian function (see above). High frequencies in $\xi$ are suppressed and thus sub-sampling $\xi$ introduces only minor errors. Therefore, minimization using the approximation yields a solution close to the globally optimal solution for the whole segmentation task with the original minimization. Further implementation details of Algorithm 1 and an initialization scheme for convex programming with fast convergence are described in Supplemental Material 6.

**Algorithm 2:** Approximative solution of MSC ($\mathscr{S}$).

**input:** Family $\mathscr{S} \subseteq \mathbb{P}(U)$; number of iterations `max_iter`; decay factor $\gamma \in (0,1)$

let $\tilde{c}(X,\beta') = c'(X) + \beta'$;
**for** `iter` $= 1,\ldots,$ `max_iter` **do**
  $\mathscr{X}' \leftarrow \{\}$; $V \leftarrow U$; $\mathscr{Z} \leftarrow \mathscr{S} = \{X_1,\ldots,X_n\}$;
  $\beta' \leftarrow \beta \cdot \gamma^{(\texttt{iter}-1)}$;
  **while** $\#V > 0$ **do** *// greedy step*
    let $N_k = \#(X_k \cap V)$;
    let $T_k = \{\tilde{c}(X_i,\beta')/N_k$ if $N_k > 0$; $\infty$ else$\}$;
    set $k' \in [n]$ so that $T_{k'} = \min_k T_k$;
    insert $X_{k'}$ into $\mathscr{X}'$;
    $V \leftarrow V \setminus X_{k'}$;
  **while** $\#\mathscr{Z} > 0$ **do** *// merge step*
    set $k' \in [n]$ so that $X_{k'} \in \mathscr{Z} \wedge \tilde{c}(X_{k'},\beta') = \min_{k:X_k\in\mathscr{Z}} \tilde{c}(X_{k'},\beta')$;
    **if** $X_{k'} \notin \mathscr{X}'$ and $\exists \mathscr{Y} \subseteq \mathscr{X}' : \bigcup \mathscr{Y} = X_{k'} \wedge \tilde{c}(X_{k'},\beta') < \sum_{X\in\mathscr{Y}} \tilde{c}(X,\beta')$ **then**
      $\mathscr{X}' \leftarrow (\mathscr{X}' \setminus \mathscr{Y}) \cup \{X_{k'}\}$;
    $\mathscr{Z} \leftarrow \mathscr{Z} \setminus \{X_{k'}\}$;
  **if** `iter` $= 1$ or $\sum_{X\in\mathscr{X}'} c(X) < \sum_{X\in\mathscr{X}} c(X)$ **then**
    $\mathscr{X} \leftarrow \mathscr{X}'$;
**return** $\mathscr{X}$;

To solve the NP-hard min-weight set-cover problem MSC in Algorithm 1, we use the approximative Algorithm 2, which iteratively performs a two-step scheme. First, a *greedy* step [56] determines the family $\mathscr{X} \subseteq \mathscr{S}$ for Eq. (17) so that the value $C = \sum_{X\in\mathscr{X}} c(X)$ of the global optimization objective $\text{MSC}(\mathbb{P}(U))$ is at most factor $h(\#U) = \sum_{k=1}^{\#U} 1/k$ higher than the globally optimal solution (due to false merges/splits). Any false splits that may arise are coped with by the subsequent *merge* step, which merges subsets of $\mathscr{X}$ if this decreases the value $C$ of the global optimization objective. To cope with possible occurrences of false merges, both steps are repeated using a more conservative merging strategy, i.e. decreasing $\beta$ by a decay factor $\gamma$, and the family $\mathscr{X}$ which yields the overall *lowest* value $C$ is considered as final solution. We used 5 iterations and $\gamma = 0.8$. The approximation ratio of Algorithm 2 is $h(\#U)$ or better, since modifications of the greedy solution are only permitted if the value $C$ is lowered.

This conservative upper bound of the approximation ratio can be tightened *a posteriori*. Let $C^*$ be the unknown exact solution of Eq. (17) and $C^*_{\text{LP}} \leq C^*$ the exact solution of the LP relaxation. The approximation ratio $C^*/C$ is thus at

L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.

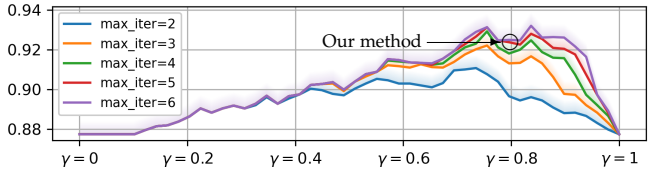


Fig. 4. A posteriori assessment of global optimality of Algorithm 2 for different numbers of iterations (`max_iter`) and values of $\gamma$ based on all instances of MSC($\mathscr{S}$) in Algorithm 1 in our experiments. Each curve shows a *lower bound* (indicated by the shading) of the ratio of de-facto exact solutions.

worst $C^*_{\mathrm{LP}}/C$. In our experiments, the average approximation ratio was at least $99.7\,\%$ (median: $100.0\,\%$). Moreover, we found that a de-facto exact solution ($C^*_{\mathrm{LP}}/C \geq 0.99$) was determined in at least $92.1\,\%$ of the cases. Fig. 4 shows lower bounds of the ratio of de-facto exact solutions for different values of $\gamma$ and iteration numbers. Note that the special case $\gamma = 1$ corresponds to our previous approach [46]. It can be seen that for $\gamma < 1$, higher lower bounds are obtained, thus, our new method improves the global optimality (we used $\gamma = 0.8$ and 5 iterations as indicated in the figure).

Algorithm 1 also requires solving $\mathrm{MSP}_{\mathrm{LP}}$ defined in Eq. (15). An exact solution can be obtained in almost-quadratic time using linear programming [57]. However, the lower bound computed by $\mathrm{MSP}_{\mathrm{LP}}$ serves the only purpose to determine whether computation of the exact set energy $c(X)$ in line 11 is necessary. Relaxing this bound thus, at worst, leads to more frequent computations of the exact set energy than necessary, but does not affect the segmentation result. This means that for $\mathrm{MSP}_{\mathrm{LP}}$, low run time is more important than accuracy. We thus approximated $\mathrm{MSP}_{\mathrm{LP}}$ by packing disjoint elements of $\mathscr{S}$ in decreasing order of their respective set energies $c'$, which performs in linear-logarithmic run time, and found that it overall yields the lowest run time.

### 3.4 Post-Processing

Post-processing of the segmentation result from the globally optimal solution $\mathscr{X}$ is performed by refining the segmentation masks (e.g., hole filling using morphological operations) and rejecting falsely detected objects (e.g., imaging artifacts or debris objects). Details are given in Supplemental Material 7. Note that objects are neither split nor merged in the post-processing.

## 4 Experimental Results

We have applied the proposed approach to 2D fluorescence microscopy image data. For performance evaluation, we have used six image datasets of five different cell types comprising various challenges, including publicly available benchmark datasets. We studied the segmentation accuracy and the cluster splitting performance, and carried out a quantitative comparison with previous methods. We used region-based and contour-based performance measures, as well as detection-based measures. We also provide an analysis of the run time performance and describe the application of our approach to another imaging modality.

### 4.1 Performance Measures

The used *region-based* performance measures are defined based on the set of all ground truth objects $\mathscr{R}_{\mathrm{gt}}$ within an image and the set of all segmented objects $\mathscr{R}_{\mathrm{seg}}$:

**Dice similarity coefficient (Dice).** The Dice coefficient is defined as

$$\mathrm{Dice}\left(\mathscr{R}_{\mathrm{gt}}, \mathscr{R}_{\mathrm{seg}}\right) = \frac{2 \cdot \left|\left(\cup \mathscr{R}_{\mathrm{gt}}\right) \cap \left(\cup \mathscr{R}_{\mathrm{seg}}\right)\right|}{\left|\cup \mathscr{R}_{\mathrm{gt}}\right| + \left|\cup \mathscr{R}_{\mathrm{seg}}\right|} \tag{23}$$

and measures the overlap of the ground truth and the segmentation result, where 0 means no overlap and 1 means perfect agreement. Dice corresponds to the pixel-based $F_1$ score (harmonic mean of precision and recall).

**Rand index (Rand).** The Rand index measures the similarity of the ground truth and the segmentation result [49]. Rand corresponds to the pixel-based accuracy score and, in contrast to Dice, is not biased towards positive or negative detections. A Rand value of 0 means no overlap, and a Rand value of 1 means perfect agreement.

**Object-based Jaccard index (SEG).** In contrast to Dice and Rand, which only consider the union of all objects within an image, the SEG measure [58] takes into account the performance for individual objects. For each ground truth object $G \in \mathscr{R}_{\mathrm{gt}}$, the measure is defined as

$$\mathrm{SEG}\left(G, \mathscr{R}_{\mathrm{seg}}\right) = \begin{cases} \frac{|G \cap S|}{|G \cup S|} & \text{if } \exists S \in \mathscr{R}_{\mathrm{seg}} : |G \cap S| > 0.5 \cdot |G|\,, \\ 0 & \text{else} \end{cases} \tag{24}$$

and attains values between 0 and 1. The SEG value is 0 if no segmented object overlaps the ground truth objects by at least $50\,\%$ (e.g., due to very inaccurate segmentations, falsely split/merged, or undetected objects). The SEG value is 1 if a ground truth object is perfectly segmented.

Notably, Dice and Rand are sensitive to false-positive detections, but invariant to falsely split/merged objects. On the other hand, SEG is sensitive to falsely split/merged objects and false-negative detections (but invariant to false-positive detections). Overall, SEG is the most comprehensive and best suited measure for region-based segmentation performance since it incorporates both detection and object-based segmentation performance.

We also used two *contour-based* performance measures. Both measures are based on the Euclidean distance $\mathrm{dist}_{\partial G}(x) = \min_{x' \in \partial G} \|x - x'\|$ of an image point $x$ to the contour $\partial G$ of the ground truth object $G \in \mathscr{R}_{\mathrm{gt}}$ and the corresponding segmented object $S \in \mathscr{R}_{\mathrm{seg}}$:

**Object-based Hausdorff distance (HSD).** The Hausdorff distance

$$\mathrm{HSD}\left(\mathrm{G}, \mathrm{S}\right) = \max_{x \in \partial S} \mathrm{dist}_{\partial G}(x) \tag{25}$$

is the maximum distance of the object contour $\partial G$ to the contour $\partial S$ of the segmented object [59]. The HSD is not upper-bounded and attains 0 if the two objects are identical.

L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.

**Object-based normalized sum of distances (NSD).**
The NSD measure is defined by

$$\mathrm{NSD}\left(\mathrm{G},\mathrm{S}\right) = \sum_{x \in G \triangle S} \mathrm{dist}_{\partial G}\left(x\right) / \sum_{x \in G \cup S} \mathrm{dist}_{\partial G}\left(x\right), \quad (26)$$

where $G \triangle S = (G \setminus S) \cup (S \setminus G)$ is the symmetric difference of $G$ and $S$. The NSD is the ratio of the number of image points, which are either only in $G$ or only in $S$, where each image point is weighted by its distance to the ground truth object contour [49]. The measure ranges from 0 ($G$ and $S$ are identical) to 1 (no overlap of $G$ and $S$).

We computed HSD and NSD for all segmented objects of an image. In case of ambiguities, the correspondences between the ground truth objects $\mathscr{R}_{\mathrm{gt}}$ and the segmented objects $\mathscr{R}_{\mathrm{seg}}$ were established by choosing the segmented object $S \in \mathscr{R}_{\mathrm{gt}}$ which is *closest* to the ground truth object $G \in \mathscr{R}_{\mathrm{gt}}$ using the respective distance function.

In addition, we used two *detection-based* performance measures to better assess the segmentation performance of closely clustered or overlapping objects. We computed the average number of falsely merged (Merge) and split (Split) objects per image. These events were identified using the definition proposed in [49].

### 4.2 Datasets

We used fluorescence microscopy image datasets of five different cell types. The data is challenging for a variety of reasons, including autofluorescence artifacts, low signal-to-noise ratio, closely clustered and partially overlapping objects, strongly non-elliptical shapes, and different object scales.

**NIH3T3 dataset.** This dataset [49] consists of 49 Hoechst-stained images with a size of $1344 \times 1024$ pixels. The dataset includes 2212 cell nuclei in total and is challenging because of strong autofluorescence artifacts, visible debris, significant intensity inhomogeneities, and many closely clustered or partially overlapping cell nuclei.

**U2OS dataset.** The dataset [49] consists of 48 images of U2OS cells that were stained with Hoechst 33342. The images have a size of $1349 \times 1030$ pixels and the dataset includes 1836 cell nuclei. The dataset is challenging due to frequent occurrence of strongly non-elliptical and closely clustered cell nuclei.

**GOWT1 datasets.** We used two datasets of GFP-transfected mouse embryonic stem cells from the IEEE ISBI Cell Tracking Challenge training data [58]. The datasets are temporal image sequences, where each image has a size of $1024 \times 1024$ pixels. GOWT1 datasets 1 and 2 consist of 31 images (with 150 cell nuclei) and 20 images (128 cell nuclei), respectively. GOWT1 is challenging due to low signal-to-noise ratio, low image contrast, and cell nucleoli (distinct dark regions within cell nuclei). The ground truth consists of eight fully annotated images and partial annotations for the other images. To avoid wrong false-positive detections due to partial annotation, only performance measures which are invariant to false-positive detections can be used for the whole dataset (SEG, Merge, Split). In addition, we used Dice, Rand, HSD, and NSD for the fully-annotated images.

**Fibroblast dataset.** This dataset [60] contains 175 3-D stacks of 35–50 DAPI-stained images of human Fibroblast cells. Each image has a size of $1024 \times 1024$ pixels. The dataset includes 985 cell nuclei. From each of the 175 stacks, we used the image slice with the highest object density for the evaluation. The ground truth does not include objects at the image boundaries, causing wrong false-positive detections. To avoid misleading results, objects detected at image boundaries were ignored for Dice and Rand. The dataset is challenging due to partially strong intensity inhomogeneities of the cell nuclei.

**HeLa dataset.** We also used 25 DAPI-stained HeLa images. The images contain 282 cell nuclei and each image has a size of $1200 \times 1620$ pixels.

Overall, 348 images were used, including 5593 annotated cell nuclei in total.

### 4.3 Evaluation

We applied the SuperDSM approach using the same set of hyperparameters for all six datasets described in Section 4.2. In addition, we applied our approach using dataset-specific adaptations (SuperDSM*). Details on the hyperparameters are given in Supplemental Material 8.

We performed a comparison with previous methods, including nine state-of-the-art methods comprising those which were reported to achieve the best results on the respective datasets. We also used three standard methods (Otsu, BLOB-LS, and BLOB-RW) for the Fibroblast and HeLa datasets. These methods were optimized for each dataset using a grid search scheme, which maximizes the equally weighted sum of the Dice and SEG values using two randomly chosen images per dataset. These images are not included in the test set. Note that we did not use hyperparameter validation for our method or the other methods, which is more realistic in practical applications. Also, we are on the safe side that we do not give an advantage to our method in the comparison.

**Global intensity thresholding (Otsu).** A global intensity threshold is computed from the image histogram [32]. Connected component analysis is employed to identify individual objects.

**Blob detection-based level sets (BLOB-LS).** A variational level set model [61] and multi-scale Laplacian of Gaussian filtering [54] for initialization are used. Individual objects are determined using connected component analysis.

**Blob detection-based random walker (BLOB-RW).** This algorithm [33] expands object and background markers based on the image intensities. For initialization, a multi-scale Laplacian of Gaussian blob detector [54] and a watershed transform are used.

**Convex variational level sets (CVX-LS).** The two-step approach [24] exploits the convexity of level set functionals. First, an image is segmented and then semi-local refinement is performed. Shape information is not used.

**Region-based progressive localization (RPL).** Progressive contrast enhancement and pre-trained classifiers to detect salient image regions are used [62]. Cell clusters are split using a binary classifier.

L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.

**Blob detection-based approach (BLOB).** Elliptical filter banks for nuclei detection and a watershed transform to obtain segmentation candidates are used [63]. Temporal information from the image sequence is exploited to select or reject candidates.

**Bayesian risk-based level sets (BR-LS).** A level set functional based on the Bayesian classification risk is used, followed by morphological analysis to separate individual cell nuclei [28].

**Cell proposal network (CPN).** The method [64] uses two convolutional neural networks (CNN). First, a CNN based on [65] is employed to determine candidates for cell nuclei bounding boxes. Then, another CNN based on the U-Net [2] is used for segmentation of each candidate. Finally, multiple hypothesis tracking is performed and the candidates corresponding to the most plausible trajectories are selected.

**Ellipse-based shape decomposition (SEG-SELF, RFOVE).** First, locally adaptive thresholding is used to binarize an image. Then, cell cluster splitting is performed by approximation of the binary image by a low-cardinality set of ellipses, either using hard (SEG-SELF [30]) or soft (RFOVE [31]) constraints for the overlap of the ellipses and the binary image.

**Cellpose.** First, a modified U-Net [2] with residual blocks and global average pooling is used to predict a vector field. Then, individual objects are identified by grouping image points whose vectors point to the same location [11]. The authors trained the model using a wide spectrum of different microscopy images including the NIH3T3 and U2OS datasets (Section 4.2).

**Globally optimal collaborative ellipses (GOCELL).** Our previous approach [46] uses elliptical models and does not exploit superadditivity. Additional pre-processing is required for NIH3T3 (local background subtraction) and GOWT1 (nucleoli removal).

For the U2OS and NIH3T3 datasets, we computed the performance values for SEG-SELF and RFOVE based on the *original* segmentation results published by the authors. Some results for HSD and NSD are somewhat different from previously reported values [46] since object correspondences are established differently (cf. Section 4.1). Results for HSD somewhat differ from [30], [31], possibly due to different computation of object correspondences. For the two GOWT1 datasets, we applied the original SEG-SELF and RFOVE implementations. In addition, we applied Cellpose to all six datasets. Note that the authors had used the NIH3T3 and U2OS datasets for training the network. Thus, the results of Cellpose for these two datasets should be treated with caution. For the NIH3T3 and the two GOWT1 datasets, we had to manually adapt one input parameter of Cellpose to achieve useful results (the optimal nuclei diameter, which we determined based on the ground truth, and which by default is computed automatically). For CVX-LS, RPL, BLOB, BR-LS, and CPN, we have used the performance values provided in the respective publications. For CPN, only performance values for the SEG measure were reported.

The results for all approaches and all datasets are provided in Table 1. Below, we discuss the datasets individually.

The NIH3T3 dataset contains many closely clustered cell nuclei. SuperDSM yields better results than SEG-SELF and RFOVE regarding SEG, Dice, Rand, HSD, NSD, and Split. Compared to RPL, SuperDSM yields a slightly worse Dice value, but improved NSD and strongly improved HSD values. The lowest number of falsely merged/split objects is obtained by Cellpose and the second-lowest by GOCELL and SuperDSM. However, for Cellpose this dataset was used for training, and GOCELL employs dataset-specific parameters and pre-processing. For SEG, Dice, HSD, and NSD, SuperDSM performs substantially better than Cellpose. Besides using fixed parameters for all datasets, we also employed dataset-specific adaptations for our approach (SuperDSM*) and discarded objects by post-processing which

TABLE 1
Segmentation performance of different approaches. For SEG, Dice, Rand, higher is better. For HSD, NSD, Merge, Split, lower is better. Not available results are indicated by "—". Best results are highlighted.

| | SEG | Dice | Rand | HSD | NSD | Merge | Split |
|---|---|---|---|---|---|---|---|
| **NIH3T3 cells** | | | | | | | |
| RPL | — | 0.91 | 0.93 | 14.1 | 0.09 | — | — |
| SEG-SELF | 0.80 | 0.89 | 0.92 | 12.9 | 0.11 | 0.8 | 1.3 |
| GOCELL | 0.84 | **0.92** | **0.94** | **8.3** | **0.06** | 0.7 | 0.5 |
| RFOVE | 0.80 | 0.89 | 0.92 | 13.3 | 0.12 | 0.9 | 0.9 |
| Cellpose | 0.75 | 0.86 | 0.92 | 33.4 | 0.17 | **0.0** | **0.4** |
| SuperDSM | 0.82 | 0.90 | 0.93 | 8.8 | 0.08 | 0.8 | **0.4** |
| SuperDSM* | **0.85** | **0.92** | **0.94** | **8.3** | 0.07 | 0.5 | 0.6 |
| **U2OS cells** | | | | | | | |
| CVX-LS | — | 0.94 | — | 12.8 | 0.05 | — | — |
| RPL | — | **0.96** | **0.96** | 10 | **0.02** | — | — |
| BR-LS | — | **0.96** | — | 12.7 | — | — | — |
| SEG-SELF | 0.85 | 0.94 | 0.95 | 13.7 | 0.08 | 0.3 | 4.6 |
| GOCELL | 0.75 | 0.92 | 0.93 | 15.5 | 0.09 | 0.4 | 3.3 |
| RFOVE | 0.77 | 0.92 | 0.93 | 15.8 | 0.16 | 1.8 | 1.9 |
| Cellpose | 0.89 | **0.96** | **0.96** | 11.6 | 0.06 | **0.1** | **0.2** |
| SuperDSM | 0.86 | 0.93 | 0.94 | 8.8 | 0.05 | 0.6 | 0.5 |
| SuperDSM* | **0.90** | **0.96** | **0.96** | **7.3** | 0.06 | 0.9 | 0.4 |
| **GOWT1 dataset 1** | | | | | | | |
| BLOB | 0.74 | — | — | — | — | — | — |
| CPN | 0.85 | — | — | — | — | — | — |
| SEG-SELF | 0.52 | 0.89 | 0.97 | 30.3 | 0.17 | **0.0** | 0.1 |
| GOCELL | 0.85 | **0.94** | **0.98** | 4.7 | 0.02 | **0.0** | **0.0** |
| RFOVE | 0.60 | 0.89 | 0.97 | 25.6 | 0.15 | **0.0** | 0.1 |
| Cellpose | 0.72 | 0.86 | 0.97 | 28.2 | 0.16 | 0.1 | **0.0** |
| SuperDSM | 0.84 | **0.94** | **0.98** | 4.3 | **0.01** | **0.0** | **0.0** |
| SuperDSM* | **0.87** | **0.94** | **0.98** | **4.2** | **0.01** | **0.0** | **0.0** |
| **GOWT1 dataset 2** | | | | | | | |
| BLOB | 0.91 | — | — | — | — | — | — |
| CPN | 0.87 | — | — | — | — | — | — |
| SEG-SELF | 0.82 | 0.92 | 0.97 | 18.4 | 0.12 | **0.0** | 1.1 |
| GOCELL | 0.91 | **0.95** | **0.98** | 3.9 | **0.01** | **0.0** | **0.0** |
| RFOVE | 0.79 | 0.90 | 0.97 | 18.9 | 0.13 | **0.0** | 0.7 |
| Cellpose | 0.73 | 0.92 | 0.97 | 18.2 | 0.13 | **0.0** | 0.6 |
| SuperDSM | 0.89 | 0.94 | **0.98** | 4.3 | 0.02 | **0.0** | 0.5 |
| SuperDSM* | **0.92** | 0.94 | **0.98** | **3.6** | **0.01** | **0.0** | **0.0** |
| **Fibroblasts** | | | | | | | |
| Otsu | 0.78 | 0.86 | 0.97 | 19.6 | 0.17 | 0.1 | 0.9 |
| BLOB-LS | 0.70 | 0.77 | 0.95 | 27.1 | 0.22 | 0.1 | 1.2 |
| BLOB-RW | 0.63 | 0.71 | 0.93 | 32.7 | 0.22 | **0.0** | 0.1 |
| GOCELL | 0.93 | **0.90** | **0.98** | 6.5 | **0.01** | **0.0** | **0.0** |
| Cellpose | 0.54 | 0.56 | 0.94 | 115.0 | 0.39 | **0.0** | 0.3 |
| SuperDSM | 0.94 | 0.89 | **0.98** | 5.8 | 0.02 | **0.0** | 0.1 |
| SuperDSM* | **0.95** | **0.90** | **0.98** | **5.1** | **0.01** | **0.0** | **0.0** |
| **HeLa cells** | | | | | | | |
| Otsu | 0.85 | **0.94** | **0.98** | **10.7** | 0.08 | 0.2 | 2.8 |
| BLOB-LS | 0.85 | 0.93 | **0.98** | 13.5 | 0.06 | 0.2 | 0.1 |
| BLOB-RW | 0.67 | 0.80 | 0.93 | 32.7 | 0.18 | **0.0** | 0.4 |
| GOCELL | 0.89 | **0.94** | **0.98** | 15.9 | 0.03 | **0.0** | 0.3 |
| Cellpose | 0.69 | 0.76 | 0.95 | 106.7 | 0.24 | **0.0** | 0.3 |
| SuperDSM | **0.90** | 0.93 | **0.98** | 13.9 | **0.02** | 0.1 | **0.0** |
| SuperDSM* | **0.90** | **0.94** | **0.98** | 13.2 | 0.03 | 0.1 | **0.0** |

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

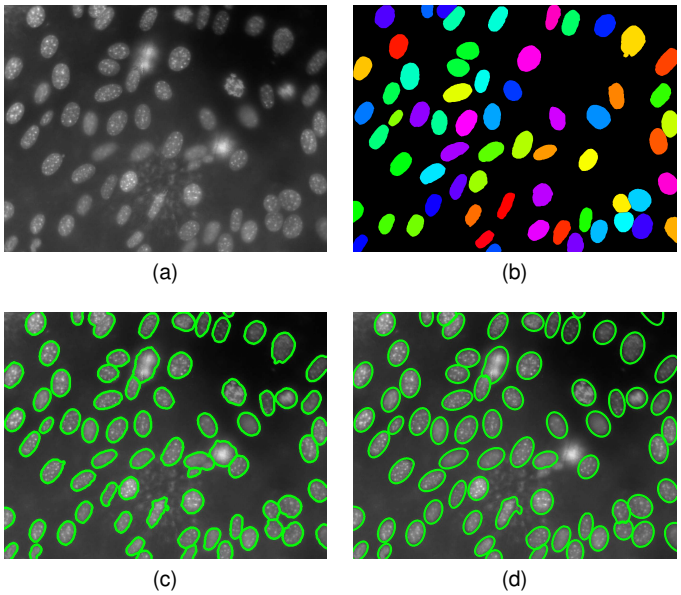


Fig. 5. Example segmentation results (green contours) for the NIH3T3 dataset. (a) Original image. (b) Ground truth. (c) Result of RFOVE. (d) Result of SuperDSM*.

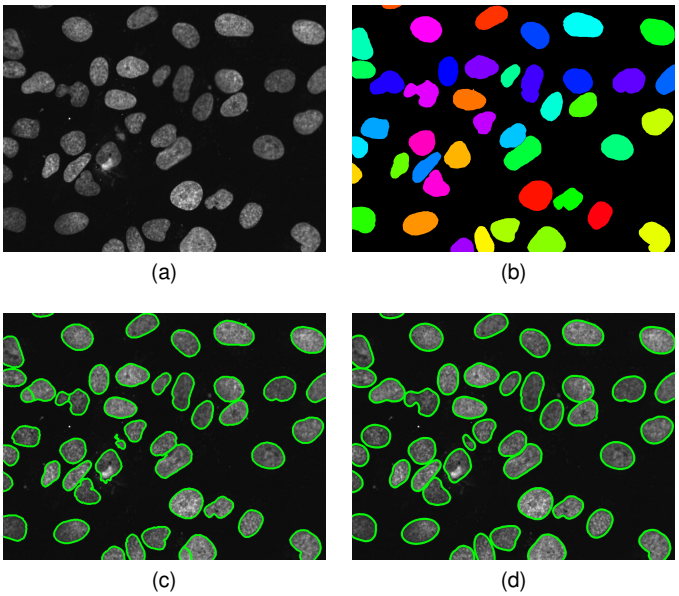


Fig. 6. Example segmentation results (green contours) for the U2OS dataset. (a) Original image (contrast-enhanced). (b) Ground truth. (c) Result of SEG-SELF. (d) Result of SuperDSM*.

likely correspond to autofluorescence artifacts (based on the radius and connected component analysis). SuperDSM* yields the best results for SEG, Dice, Rand, and HSD. An example segmentation result is shown in Fig. 5. It can be seen that clustered and non-clustered cell nuclei are well segmented.

The U2OS dataset is difficult due to strongly non-elliptical shape of the cell nuclei, which is challenging for the merging/splitting schemes of the segmentation methods. Cellpose yields the lowest number of falsely merged/split objects, however, this dataset was used for training the network. SEG-SELF achieves the second-lowest false merging rate (0.3 per image), but has a strong tendency to over-segmentation (4.6 falsely split objects per image). SuperDSM yields only 0.6 falsely merged and only 0.5 falsely split cell nuclei per image. This good merging/splitting performance is in agreement with an improved SEG performance (SEG is sensitive to false merges/splits). Compared to RPL, SuperDSM performs worse for Dice, Rand, and NSD. However, Dice and Rand are invariant to false merges/splits, which were not reported for RPL. SuperDSM* yields the best results for SEG, Dice, Rand, and HSD. SEG-SELF and GOCELL yield fewer false merges, but significantly more false splits. Thus, overall, SuperDSM* performs best. Fig. 6 shows example segmentation results. Our approach yields no false merges/splits, whereas SEG-SELF yields four falsely split cell nuclei. The object contours of all cell nuclei are accurately segmented.

For GOWT1 dataset 1, SuperDSM yields overall very good results. Compared to SEG-SELF and RFOVE, strong improvements can be observed for SEG (0.60 to 0.84), Dice (0.89 to 0.94), HSD (25.6 to 4.3), and NSD (0.15 to 0.01). Improvements are also large compared to Cellpose and concern mostly SEG (0.72 to 0.84), Dice (0.86 to 0.94), HSD (28.2 to 4.3), and NSD (0.16 to 0.01). GOCELL and CPN perform slightly better (SEG is 0.85), however, only SEG was reported for CPN (which is invariant to false-positive detections) and GOCELL used dataset-adapted parameters and pre-processing. Using our method with dataset-specific adaptations (SuperDSM*) yields the best results for all performance measures. For GOWT1 dataset 2, SuperDSM performs competitively. SEG is improved compared to CPN, SEG-SELF, RFOVE, and Cellpose, but not as good as BLOB (which did not perform well for GOWT1 dataset 1). SuperDSM* yields the best results regarding all measures. Fig. 7 shows example segmentation results. All objects are accurately segmented, including the low-intensity cell nuclei and the irregularly shaped nucleus in GOWT1 dataset 2.

For the Fibroblast dataset, SuperDSM yields the best results for SEG, Rand, HSD, and Merge, but is slightly worse for Dice, NSD, Split compared to GOCELL. Compared to Cellpose, the results are strongly improved for SEG (0.54 to 0.94), Dice (0.56 to 0.89), Rand (0.94 to 0.98), HSD (115.0 to 5.8), NSD (0.39 to 0.02), and Split (0.3 to 0.1). The dataset contains multiple images with only few cell nuclei and it turns out that the pre-processing somewhat underestimates the scale $\sigma$. Adapting this parameter (SuperDSM*) yields the best results for all measures. Example segmentation results are shown in Fig. 8. Our approach reliably segments the contours of both elliptical and non-elliptical cell nuclei. GOCELL performs worse since elliptical shape models are insufficient. BLOB-LS and BLOB-RW perform worse since their initialization is prone to the non-elliptical shapes and closely clustered objects.

For the HeLa dataset, SuperDSM performs overall better than Otsu, BLOB-LS, BLOB-RW, and Cellpose. Compared to GOCELL, SuperDSM yields better results for SEG, HSD, NSD, and Split. Overall, SuperDSM yields the best result for SEG, Rand, NSD, and Split. SuperDSM* yields the overall best result.

Considering all datasets, it turns out that our SuperDSM approach achieves better results than previous methods for the Fibroblast and HeLa datasets and competitive results for the U2OS, NIH3T3, GOWT1 datasets using the same set of

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

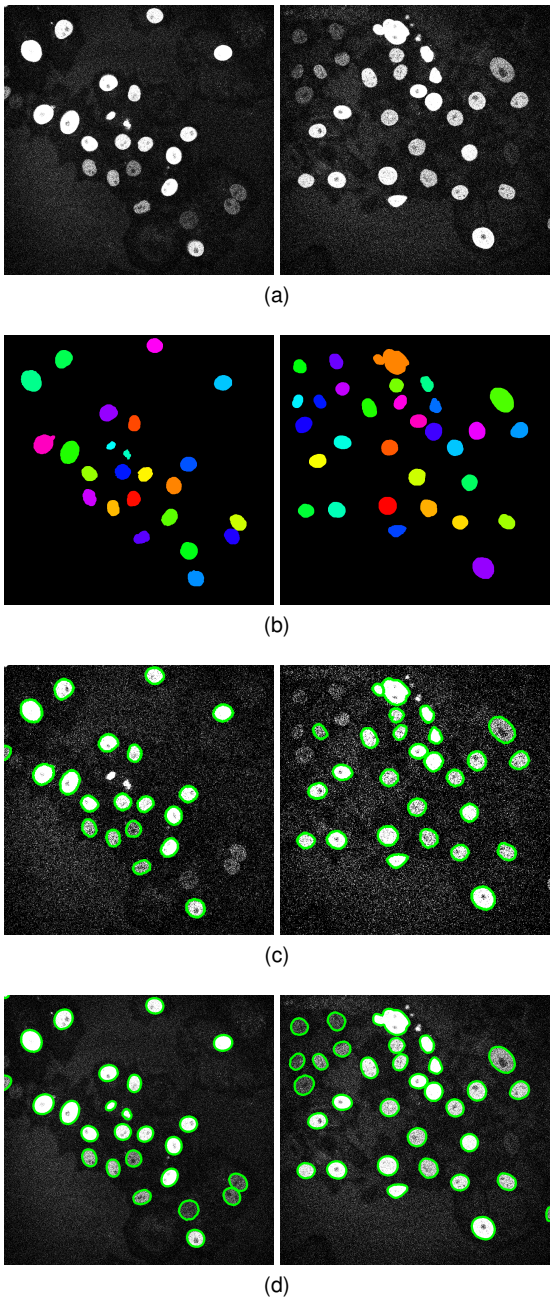


Fig. 7. Example segmentation results (green contours) for GOWT1 dataset 1 (left column) and GOWT1 dataset 2 (right column). (a) Original images (contrast-enhanced). (b) Ground truth. (c) Result of RFOVE. (d) Result of SuperDSM*.

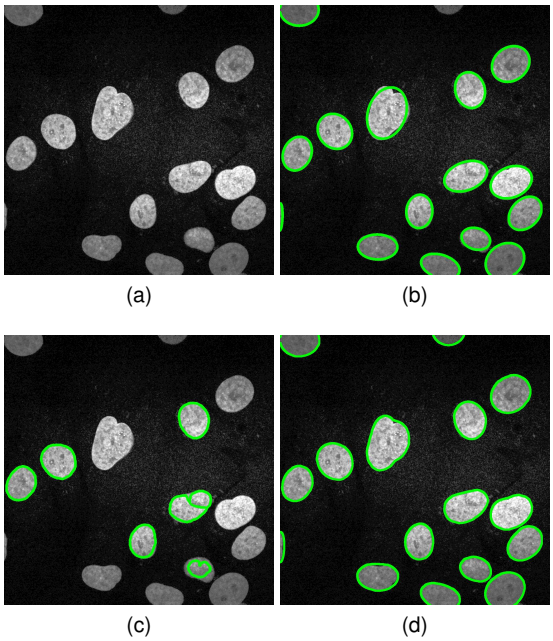


Fig. 8. Example segmentation results (green contours) for the Fibroblast dataset. (a) Original image (contrast-enhanced). (b) Result of GOCELL. (c) Result of Cellpose. (d) Result of SuperDSM*.

hyperparameters for all six datasets. The results are generally slightly worse than using dataset-specific adaptations (SuperDSM*). Using such adaptations, our approach generally yields best results for all datasets for the region-based measure SEG. For Dice and Rand, the results are competitive. Concerning contour-based measures, the performance of our method is best for four out of six datasets, and second-best for two datasets. Regarding cluster splitting, our method generally yields the best results for all datasets for the number of falsely merged/split objects (sum of Merge and Split in Table 1), and achieves very low Merge and Split counts (less than one falsely merged/split objects per image). For the NIH3T3 and U2OS datasets, Cellpose yields fewer falsely merged/split objects, however, both datasets were used for training the network. For the other datasets, our method achieves strongly improved results compared to Cellpose for all region-based and contour-based measures. The best performing methods besides our approach are RPL and GOCELL. However, for RPL, published results for SEG, Merge, and Split are not available. GOCELL yields worse results than SuperDSM for non-elliptical cell nuclei (e.g., U2OS dataset). We thus conclude that our method performs overall best in this study.

In practical applications, the hyperparameters of the global energy minimization of our SuperDSM* method can be adapted as follows starting from the default values (SuperDSM). For example, when the computed cell contours are too smooth, one could reduce the weight $\alpha$ of the regularization of the deformations in Eq. (6). When objects are falsely split (e.g., due to irregular shapes) one could increase the constant term $\beta$ of the extended set energy functions in Eq. (14). The scale $\sigma$ can be adapted by using $\sigma = R/\sqrt{2}$, where $R$ is the average radius of cell nuclei which could be determined coarsely.

### 4.4 Run Time Performance

In addition, we studied the run time performance of our approach. As described in Section 3, our approach separately processes regions of possibly clustered objects. This involves the computation of the extended set energies $c(X)$ in Eq. (14) for different sets $X \in \mathbb{P}(U)$ to determine the solution $\mathscr{X} \subseteq \mathbb{P}(U)$ of MSC $(\mathbb{P}(U))$, see Algorithm 1. The computational cost is the number of sets (image regions), for which the set energies must be computed (see Section 2.3.3). This number is upper-bounded by the cardinality of the universe of a region of possibly clustered objects. Regions of possibly clustered objects with universe cardinality $\#U \leq 2$ are computationally cheap, since $\#\mathbb{P}(U) \leq 3$ and thus at

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

most three sets must be computed. We found that such regions occurred in $82.7\,\%$ of the cases. The other cases are computationally more challenging, and our method reduces the computational cost by excluding sets *before* computing the corresponding energies using Criterion 1 and Criterion 2. The more sets are excluded, the more efficient the method is. As a measure of efficiency, we use the *set exclusion rate (SER)*, which is the ratio of excluded sets compared to the cardinality of $\mathbb{P}(U)$. To quantify the computational cost of our approach for the computationally challenging cases, below we consider SER for regions of possibly clustered objects with universe cardinality $\#U > 2$. We also study the run time of our method.

Fig. 9(a) shows a histogram of SER values for the computationally challenging cases ($\#U > 2$) for all six datasets. The median SER is $0.33$ and the maximum is $0.97$, thus, the computational cost was typically reduced by $33\,\%$ and at best by $97\,\%$. For the U2OS, GOWT1, Fibroblast, and HeLa datasets, relatively high SER values were obtained. The median SER is $0.62$ for GOWT1 dataset 1 and $0.43$ for the other datasets. Thus, the computational cost was typically reduced by $43\,\%$ or $62\,\%$, respectively. Only for the NIH3T3 dataset lower SER values were obtained. The reason is probably that this dataset is most difficult in terms of clustered objects. Overall, we can conclude that the two criteria effectively reduce the computational cost for the computationally challenging cases.

Fig. 9(b) shows the average computation time per image of the individual processing steps of our approach. It can be seen that both pre-processing and post-processing performed in almost constant time (see the scattering indicated by the error bars). The computationally most expensive task in pre-processing is scale estimation (cf. Section 3.1). The strongest scattering in the average run time can be observed for the coarse-to-fine region analysis, which is probably due to the linear search to determine seeds for splitting regions with high normalized energies (cf. Section 3.2). Regarding the global energy minimization (cf. Section 3.3), we observe that images with many closely clustered objects correlate with longer run times (U2OS and NIH3T3 datasets). This confirms that our approach efficiently copes with non-clustered cell nuclei. The overall average run time was $45.3$ seconds per image. For comparison, the average run time of our previous globally optimal approach [46] was 1:23 minutes per image using the same hardware (see below). Thus, we achieved a speed-up of $183\,\%$ compared to our previous method [46]. This is remarkable, since deformable shape models used in the proposed approach are computationally much more challenging than elliptical models due to the higher dimension of the parameter space. The average run time of Cellpose [11] was $18.7$ seconds per image. However, additional time is required for training, while for our approach training is not necessary. For the globally optimal approach using circular models [45], a run time of $38$ minutes was reported in [46]. This emphasizes the computational efficiency of the proposed approach. In biological applications, most time is required for preparation of the specimen such as staining and for image acquisition (e.g., DAPI staining takes at least 2–4 minutes [1]). Thus, the run time of our method is suited for practical applications.

Fig. 9(c) shows the overall run time of our method per image. From the mode of this histogram it can be seen that for the vast majority of the images, the typical computation time was less than 1 minute per image ($84.1\,\%$ of the images). For $95.8\,\%$ of the images, the overall run time was less than 2 minutes. We observed only two cases, where the computation lasted unexpectedly long (a single image from the Fibroblast dataset, which took 14 minutes for SuperDSM and 12 minutes for SuperDSM*). This is due to strong image noise leading to a large number of detected intensity peaks used as seed points by the coarse-to-fine region analysis for splitting image regions (cf. Supplemental Material 5).

All experiments were performed using an AMD Ryzen Threadripper 3970X CPU and 32 GB of RAM. We used Intel Math Kernel Library 20.0 for fast sparse and dense linear algebra. For the coarse-to-fine region analysis, up to 16 regions of possibly clustered objects were processed in parallel. For global energy minimization, the energies for up to 16 sets were computed in parallel. For post-processing, up to 16 objects were processed in parallel. Note that no GPU acceleration was used for the experiments. Faster run times can be achieved by increasing the degree of parallelization, which is straightforward.

### 4.5 Application to Different Imaging Modality

So far, we have studied the performance of our method using a wide range of fluorescence microscopy image data. In a final experiment, we investigated the applicability of our approach to another imaging modality, namely histopathology images stained with haematoxylin and eosin (H&E). These images are color images, where cell nuclei appear in blue or dark purple. We used the training dataset of the MICCAI 2018 MoNuSeg challenge [66]. The image size is $1000 \times 1000$ pixels and the data contains 1390 cell nuclei from the histological section of a human stomach in adenocarcinoma disease condition (see Fig. 10(a)). The data is challenging due to very densely clustered cell nuclei, a wide variety of nuclei shapes, and strongly inhomogeneous background.

We used minor methodological adaptations to account for the very different imaging modality. This concerns only pre-processing and the computation of $Y_\Omega$ (image intensities with $\tau_x$ offset). The idea is to transform the image intensities so that cell nuclei (dark purple regions) correspond to bright intensities. To this end, we first average the image intensities over the three color channels $g_x^{\mathrm{r}}, g_x^{\mathrm{g}}, g_x^{\mathrm{b}}$ using $g_x = 1 - \frac{1}{3}\left(g_x^{\mathrm{r}} + g_x^{\mathrm{g}} + g_x^{\mathrm{b}}\right)$. Second, we apply Gaussian filtering (standard deviation $\sigma$) followed by local maximum filtering ($2\sigma \times 2\sigma$ neighborhood) to determine the locally maximal responses $g_x^{\max}$. Third, we compute $Y_\Omega$ (image intensities with $\tau_x$ offset, see Eq. (5)) by $\tau_x = \max\{g_x^{\max}, \mathrm{mean}_{x\in\Omega}\, g_x\}$ using the mean intensity over all image points.

Our method performed the segmentation within 3:06 minutes and the result is shown in Fig. 10(c). We found that $97.1\,\%$ of the cell nuclei were detected (using the detection measure in [49]), and only $4.5\,\%$ were falsely merged or split, respectively. Given that our approach is designed for fluorescence microscopy images rather than H&E-stained histopathology images, the result is promising. In Fig. 10(c), right bottom, few small tissue regions are segmented since they are slightly darker than their neighborhood (as for

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

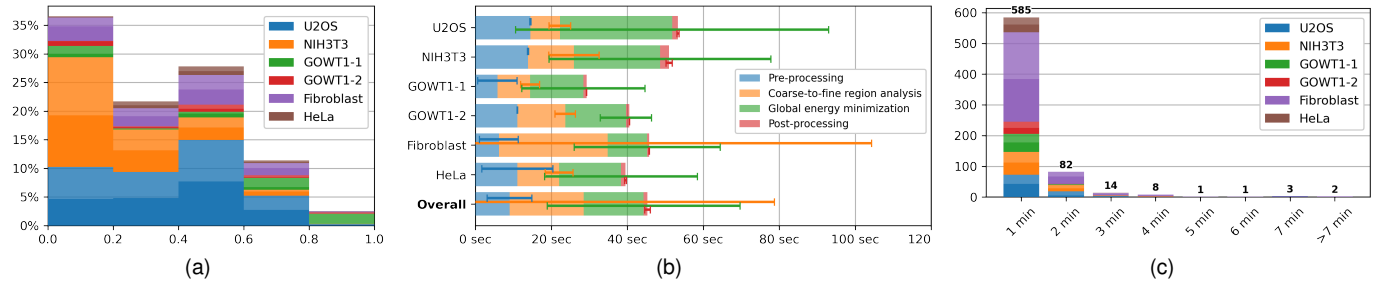


Fig. 9. Run time performance of our approach for the six datasets (darker color shades correspond to SuperDSM, brighter color shades correspond to SuperDSM*). (a) Histogram of the set exclusion rate SER (relative frequencies). (b) Computation time of the individual processing steps of our approach (mean and standard deviation). (c) Histogram of the total run time per image (absolute frequencies). The labels on the horizontal axis denote intervals (e.g., 0 to 1 minutes, 1 to 2 minutes).

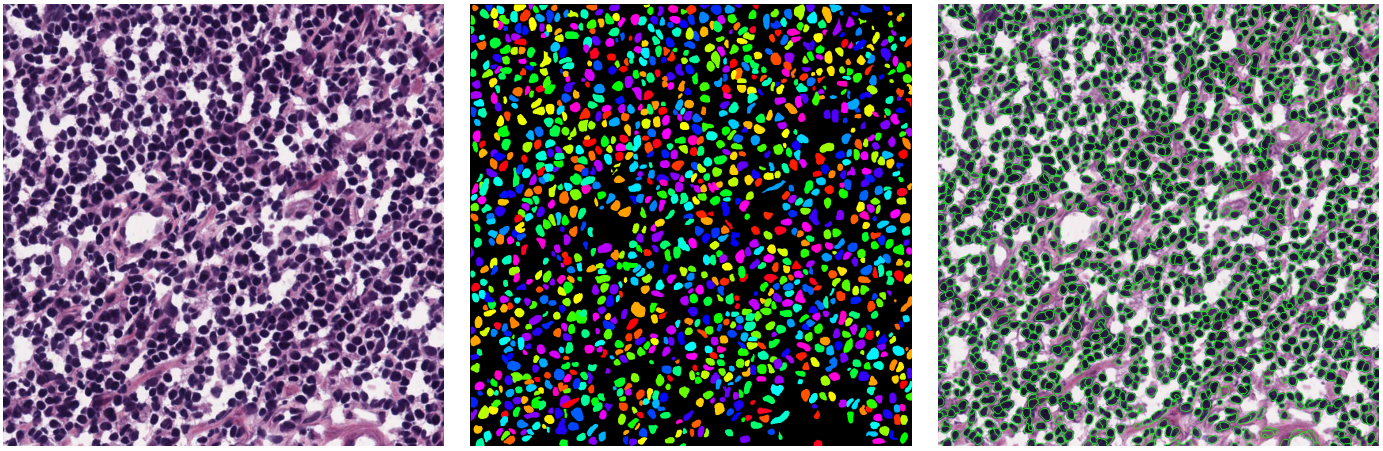


Fig. 10. (a) Original H&E-stained histopathology image. (b) Ground truth. (c) Segmentation result (green contours) of SuperDSM*.

cells). The result could be improved by not only taking into account the brightness, but also the color hue and saturation for computing $Y_\Omega$. Overall, this experiment shows that our approach can be generalized to other imaging modalities.

## 5 Conclusion

We have introduced a new *globally optimal* approach based on *deformable* shape models and global energy minimization for cell nuclei segmentation in microscopy images. The approach intrinsically copes with non-elliptical shapes, jointly exploits shape and intensity information, and is based on an *implicit* parameterization, which leads to a *convex* energy. Thus, energy minimization is independent of the initialization, fast, and robust. To jointly perform cell segmentation and cluster splitting, we have considered the infimum of the convex energy as a *set energy function*, i.e. a function of the set of image regions where model fitting is performed. We have proposed a novel iterative global energy minimization method, which provably determines the optimal image regions close to global optimality. The method exploits the inherent property of *superadditivity* of the set energy function, which is established via the *set-packing polytope*. Intuitively, the property of superadditivity means that a deformable shape model cannot fit better to an image region than it fits to any of its sub-regions. Thus, it is not necessary to consider all possible image regions for optimization. Instead, the proposed energy minimization method considers image regions in order of increasing size and leverages superadditivity to exclude regions corresponding to falsely merged objects using a fine-to-coarse scheme. This improves the computational efficiency, since when excluding a region, all its supersets are also excluded. We have also described a coarse-to-fine region analysis scheme, which determines the universe of atomic image regions used as input for global energy minimization. In addition, we have derived a closed-form solution of the proposed global energy minimization based on the superadditivity property for non-clustered cell nuclei, which further accelerates the computation.

The regularization parameter $\alpha$ of the convex energy is used to control the shape variability of the deformable shape models. An *extended set energy function* has been introduced to avoid over-segmentation, which uses the hyperparameter $\beta$ defining the maximum allowed energy difference for merging two image regions (i.e. two deformable shape models that are fitted in these regions). Our approach automatically determines scale-related hyperparameters based on scale estimation. The objective function of our global energy minimization method corresponds to a min-weight set-cover problem, which is NP-hard to compute. We have thus used a fast approximation algorithm, which determines a solution close to global optimality. In addition, the design of the algorithm directly addresses the false splits and false merges possibly introduced by using an approximation. We have performed an analysis of global optimality and

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

found that the global solution was exactly determined in at least 92.1 % of our experiments, the average approximation ratio of the solution was at least 99.7 %, and the median was 100.0 %. To compute the set energy function, we have used a fast numerical second-order method which directly determines global solutions by convex energy minimization.

We have applied our approach to a wide range of 348 fluorescence microscopy images of five different cell types comprising 5593 cell nuclei, and performed a quantitative comparison with previous methods. It turned out that our approach generally demonstrates the best or second-best cluster splitting performance. The segmentation accuracy is better compared to previous methods according to region-based measures, and is competitive according to contour-based measures. For the region-based SEG performance measure used in the cell segmentation benchmark [58], which is the best suited measure for overall segmentation performance (since it incorporates both detection and object-based segmentation performance), our approach generally yields superior results for all datasets. Our approach is robust since it achieves competitive or improved results even when using a fixed set of hyperparameters for all datasets, compared to nine state-of-the-art methods comprising those which previously achieved best results on the respective datasets. In addition, we have demonstrated that our approach can be generalized to other imaging modalities.

Future work will be concerned with the development of a modality-agnostic segmentation method. Other work will be on improved scale estimation and incorporation of higher-level image features (e.g., information from other image channels). In addition, the derivation of even tighter energy bounds to further accelerate the global energy minimization and a study of other optimization methods will be subject of future work.

## Acknowledgments

This work has been funded by the German Research Foundation (DFG) within the RTG 1653 and by the BMBF within the projects de.NBI and CancerTelSys. We thank Christian Ritter for helpful discussions and Ronja Rappold for providing the HeLa dataset.

## References

[1] M. Mulisch and U. Welsch, Eds., *Romeis - Mikroskopische Technik*. Springer-Verlag, 2015.

[2] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 234–241.

[3] Y. Song, E.-L. Tan, X. Jiang, J.-Z. Cheng, D. Ni, S. Chen, B. Lei, and T. Wang, "Accurate cervical cell segmentation from overlapping clumps in pap smear images," *IEEE Trans. Med. Imag.*, vol. 36, no. 1, pp. 288–300, Jan. 2017.

[4] A. Böhm, A. Ücker, T. Jäger, O. Ronneberger, and T. Falk, "ISOO DL: Instance segmentation of overlapping biological objects using deep learning," in *Proc. Int. Symp. Biomed. Imag.*, 2018, pp. 1225–1229.

[5] M. Fan and J. Rittscher, "Global probabilistic models for enhancing segmentation with convolutional networks," in *Proc. Int. Symp. Biomed. Imag.*, 2018, pp. 1234–1238.

[6] W. Xie, J. A. Noble, and A. Zisserman, "Microscopy cell counting and detection with fully convolutional regression networks," *Comput. Methods Biomech. Biomed. Eng. Imaging. Vis.*, vol. 6, no. 3, pp. 283–292, 2018.

[7] A. O. Vuola, S. U. Akram, and J. Kannala, "Mask-RCNN and U-Net ensembled for nuclei segmentation," in *Proc. Int. Symp. Biomed. Imag.*, 2019, pp. 208–212.

[8] C. Payer, D. Štern, M. Feiner, H. Bischof, and M. Urschler, "Segmenting and tracking cell instances with cosine embeddings and recurrent hourglass networks," *Med. Image Anal.*, vol. 57, pp. 106–119, 2019.

[9] H. Fehri, A. Gooya, Y. Lu, E. Meijering, S. A. Johnston, and A. F. Frangi, "Bayesian polytrees with learned deep features for multi-class cell segmentation," *IEEE Trans. Image Process.*, vol. 28, no. 7, pp. 3246–3260, Jan. 2019.

[10] T. Wollmann, M. Gunkel, I. Chung, H. Erfle, K. Rippe, and K. Rohr, "GRUU-net: Integrated convolutional and gated recurrent neural network for cell segmentation," *Med. Image Anal.*, vol. 56, pp. 68–79, 2019.

[11] C. Stringer, T. Wang, M. Michaelos, and M. Pachitariu, "Cellpose: a generalist algorithm for cellular segmentation," *Nat. Methods*, vol. 18, no. 1, pp. 100–106, 2020.

[12] O. Ciga and A. L. Martel, "Learning to segment images with classification labels," *Med. Image Anal.*, vol. 68, p. 101912, 2021.

[13] K. He, G. Gkioxari, P. Dollár, and R. Girshick, "Mask R-CNN," in *Proc. Int. Conf. Comput. Vis.*, 2017, pp. 2961–2969.

[14] G. Li, T. Liu, J. Nie, L. Guo, J. Chen, J. Zhu, W. Xia, A. Mara, S. Holley, and S. Wong, "Segmentation of touching cell nuclei using gradient flow tracking," *J. Microsc.*, vol. 231, no. 1, pp. 47–58, 2008.

[15] I. J. Goodfellow, J. Shlens, and C. Szegedy, "Explaining and harnessing adversarial examples," in *Proc. Int. Conf. Learn. Representations*, 2015.

[16] A. Fawzi, S. M. Moosavi Dezfooli, and P. Frossard, "The robustness of deep networks – a geometric perspective," *Signal Process. Mag.*, vol. 34, no. 6, pp. 50–62, 2017.

[17] C. G. Northcutt, A. Athalye, and J. Mueller, "Pervasive label errors in test sets destabilize machine learning benchmarks," *arXiv preprint arXiv:2103.14749*, 2021.

[18] C. Zimmer and J.-C. Olivo-Marin, "Coupled parametric active contours," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 11, pp. 1838–1842, Nov. 2005.

[19] P. Thevenaz, R. Delgado-Gonzalo, and M. Unser, "The Ovuscule," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 2, pp. 382–393, Feb. 2011.

[20] R. Delgado-Gonzalo, N. Chenouard, and M. Unser, "Spline-based deforming ellipsoids for interactive 3d bioimage segmentation," *IEEE Trans. Image Process.*, vol. 22, no. 10, pp. 3926–3940, Oct. 2013.

[21] C. de Solorzano, R. Malladi, S. Lelievre, and S. Lockett, "Segmentation of nuclei and cells using membrane related protein markers," *J. Microsc.*, vol. 201, no. 3, pp. 404–415, 2001.

[22] A. Dufour, V. Shinin, S. Tajbakhsh, N. Guillén-Aghion, J.-C. Olivo-Marin, and C. Zimmer, "Segmenting and tracking fluorescent cells in dynamic 3-D microscopy with coupled active surfaces," *IEEE Trans. Image Process.*, vol. 14, no. 9, pp. 1396–1410, Sep. 2005.

[23] C. Li, C.-Y. Kao, J. C. Gore, and Z. Ding, "Minimization of region-scalable fitting energy for image segmentation," *IEEE Trans. Image Process.*, vol. 17, no. 10, pp. 1940–1949, Oct. 2008.

[24] J.-P. Bergeest and K. Rohr, "Efficient globally optimal segmentation of cells in fluorescence microscopy images using level sets and convex energy functionals," *Med. Image Anal.*, vol. 16, no. 7, pp. 1436–1444, 2012.

[25] T. F. Chan, S. Esedoglu, and M. Nikolova, "Algorithms for finding global minimizers of image segmentation and denoising models," *SIAM J. Appl. Math.*, vol. 66, no. 5, pp. 1632–1648, 2006.

[26] T. Goldstein, X. Bresson, and S. Osher, "Geometric applications of the split Bregman method: Segmentation and surface reconstruction," *J. Sci. Comput.*, vol. 45, no. 1, pp. 272–293, 2010.

[27] Y. He, Y. Meng, H. Gong, S. Chen, B. Zhang, W. Ding, Q. Luo, and A. Li, "An automated three-dimensional detection and segmentation method for touching cells by integrating concave points clustering and random walker algorithm," *PloS One*, vol. 9, no. 8, p. e104437, 2014.

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

[29] P. Markowsky, S. Reith, T. Zuber, R. König, K. Rohr, and C. Schnörr, "Segmentation of cell structures using model-based set covering with iterative reweighting," in *Proc. Int. Symp. Biomed. Imag.*, 2017, pp. 392–396.

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**

[30] C. Panagiotakis and A. Argyros, "Cell segmentation via region-based ellipse fitting," in *Proc. Int. Conf. Image Proc.*, 2018, pp. 2426–2430.

[31] ——, "Region-based fitting of overlapping ellipses and its application to cells segmentation," *Image Vis. Comput.*, vol. 93, p. 103810, 2020.

[32] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[33] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Sep. 2006.

[34] X. Lou, U. Köthe, J. Wittbrodt, and F. A. Hamprecht, "Learning to segment dense cell nuclei with shape prior," in *Proc. Conf. Comp. Vis. Pattern Recognit.*, 2012, pp. 1012–1018.

[35] Z. Lu, G. Carneiro, and A. P. Bradley, "An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells," *IEEE Trans. Image Process.*, vol. 24, no. 4, pp. 1261–1272, Apr. 2015.

[36] M. Nosrati and G. Hamarneh, "Segmentation of overlapping cervical cells: a variational method with star-shape prior," in *Proc. Int. Symp. Biomed. Imag.*, 2015, pp. 186–189.

[37] L. Kostrykin, C. Schnörr, and K. Rohr, "Segmentation of cell nuclei using intensity-based model fitting and sequential convex programming," in *Proc. Int. Symp. Biomed. Imag.*, 2018, pp. 654–657.

[38] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient ND image segmentation," *Internat. J. Computer Vis.*, vol. 70, no. 2, pp. 109–131, 2006.

[39] S. Ali and A. Madabhushi, "An integrated region-, boundary-, shape-based active contour for multiple object overlap resolution in histological imagery," *IEEE Trans. Med. Imag.*, vol. 31, no. 7, pp. 1448–1460, Jul. 2012.

[40] F. Xing and L. Yang, "Fast cell segmentation using scalable sparse manifold learning and affine transform-approximated active contour," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2015, pp. 332–339.

[41] J. Kong, P. Zhang, Y. Liang, G. Teodoro, D. Brat, and F. Wang, "Robust cell segmentation for histological images of Glioblastoma," in *Proc. Int. Symp. Biomed. Imag.*, 2016, pp. 1041–1045.

[42] P. Zhang, F. Wang, G. Teodoro, Y. Liang, M. Roy, D. Brat, and J. Kong, "Effective nuclei segmentation with sparse shape prior and dynamic occlusion constraint for glioblastoma pathology images," *J. Med. Imag.*, vol. 6, no. 1, p. 017502, 2019.

[43] G. Dong and S. Acton, "Detection of rolling leukocytes by marked point processes," *J. Electron. Imag.*, vol. 16(3), p. 033013, 2007.

[44] E. Soubies, P. Weiss, and X. Descombes, "Graph cut based segmentation of predefined shapes: Applications to biological imaging," in *Pattern Recognit. Appl. Methods*, 2015, pp. 153–170.

[45] X. Descombes, "Multiple objects detection in biological images using a marked point process framework," *Methods*, vol. 115, pp. 2–8, 2017.

[46] L. Kostrykin, C. Schnörr, and K. Rohr, "Globally optimal segmentation of cell nuclei in fluorescence microscopy images using shape and intensity information," *Med. Image Anal.*, vol. 58, p. 101536, 2019.

[47] A. G. Eldin, X. Descombes, G. Charpiat, and J. Zerubia, "Multiple birth and cut algorithm for multiple object detection," *J. Multimed. Processing Technol.*, 2012.

[48] P. L. Bartlett, M. I. Jordan, and J. D. McAuliffe, "Convexity, classification, and risk bounds," *J. Am. Stat. Assoc.*, vol. 101, no. 473, pp. 138–156, 2006.

[49] L. Coelho, A. Shariff, and R. Murphy, "Nuclear segmentation in microscope cell images: A hand-segmented dataset and comparison of algorithms," in *Proc. Int. Symp. Biomed. Imag.*, 2009, pp. 518–521.

[50] M. Andersen, J. Dahl, and L. Vandenberghe, "CVXOPT: A Python package for convex optimization, version 1.2.5," 2020. [Online]. Available: http://cvxopt.org

[51] V. Vazirani, *Approximation algorithms*. Springer, 2001.

[52] J. Funke, C. Zhang, T. Pietzsch, M. A. G. Ballester, and S. Saalfeld, "The candidate multi-cut for cell segmentation," in *Proc. Int. Symp. Biomed. Imag.*, 2018, pp. 649–653.

[53] H. Bay, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis.*, 2006, pp. 404–417.

[54] T. Lindeberg, "Feature detection with automatic scale selection," *Int. J. Comput. Vis.*, vol. 30, no. 2, pp. 79–116, 1998.

[55] R. Franke, "Scattered data interpolation: tests of some methods," *Math. Comput.*, vol. 38, no. 157, pp. 181–200, 1982.

[56] D. S. Johnson, "Approximation algorithms for combinatorial problems," *J. Comput. Syst. Sci.*, vol. 9, no. 3, pp. 256–278, 1974.

[57] S. Jiang, Z. Song, O. Weinstein, and H. Zhang, "Faster dynamic matrix inverse for faster LPs," *arXiv preprint arXiv:2004.07470*, 2020.

[58] M. Maška, V. Ulman, D. Svoboda, P. Matula, P. Matula, C. Ederra, A. Urbiola, T. España, S. Venkatesan, D. Balak *et al.*, "A benchmark for comparison of cell tracking algorithms," *Bioinform.*, vol. 30, no. 11, pp. 1609–1617, 2014.

[59] P. Bamford, "Empirical comparison of cell segmentation algorithms using an annotated dataset," in *Proc. Int. Conf. Image Proc.*, vol. 2, 2003, pp. II–1073–1076.

[60] A. A. Kalinin, A. Allyn-Feuer, A. Ade, G.-V. Fon, W. Meixner, D. Dilworth, R. Jeffrey, G. A. Higgins, G. Zheng, A. Creekmore *et al.*, "3D cell nuclear morphology: Microscopy imaging dataset and voxel-based morphometry classification results," in *Proc. Conf. Comp. Vis. Pattern Recognit. Workshops*, 2018, pp. 2272–2280.

[61] T. F. Chan and L. A. Vese, "Active contours without edges," *IEEE Trans. Image Process.*, vol. 10, no. 2, pp. 266–277, Feb. 2001.

[62] Y. Song, W. Cai, H. Huang, Y. Wang, D. D. Feng, and M. Chen, "Region-based progressive localization of cell nuclei in microscopic images with data adaptive modeling," *BMC Bioinform.*, vol. 14:173, no. 1, pp. 1–16, 2013.

[63] S. Akram, J. Kannala, L. Eklund, and J. Heikkilä, "Joint cell segmentation and tracking using cell proposals," in *Proc. Int. Symp. Biomed. Imag.*, 2016, pp. 920–924.

[64] ——, "Cell tracking via proposal generation and selection," *arXiv preprint*, vol. arXiv:1705.03386, 2017.

[65] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 818–833.

[66] N. Kumar, R. Verma, D. Anand, Y. Zhou, O. F. Onder, E. Tsougenis, H. Chen, P.-A. Heng, J. Li, Z. Hu *et al.*, "A multi-organ nucleus segmentation challenge," *IEEE Trans. Med. Imag.*, vol. 39, no. 5, pp. 1380–1391, May 2019.


**Leonid Kostrykin** received the B.Sc. degree in scientific programming from the FH Aachen, Germany, in 2013, and the M.Sc. degree in computer science from the University of Düsseldorf, Germany, in 2016. Since 2016 he has been with the Biomedical Computer Vision Group (BMCV) at Heidelberg University, Germany, where he is currently pursuing the Ph.D. degree. His research interests include image analysis methods and biomedical applications.

**Karl Rohr** received the Dipl.-Ing. degree in Electrical Engineering from the University of Karlsruhe (KIT), Germany, and the Ph.D. and Habilitation degrees in Computer Science from the University of Hamburg, Germany, in 1994 and 1999, respectively. He is currently an Associate Professor and the Head of the Biomedical Computer Vision Group (BMCV) at Heidelberg University. From 2000-2004 he was an Associate Professor at the International University in Germany. In 1999 he spent a research stay at the Surgical Planning Laboratory, Harvard Medical School, Boston/MA, USA. His research interests are in biomedical image analysis with focus on cell segmentation and tracking, particle tracking, non-rigid image registration, vessel segmentation, and landmark localization. He has written a book entitled Landmark-Based Image Analysis (Kluwer Academic Publishers, 2001) and published more than 300 peer-reviewed scientific articles. He was an Associate Editor of IEEE Transactions on Biomedical Engineering (2007–2012) and served the editorial board of the journal Pattern Recognition (2000–2006). He was Program Chair of the IEEE International Symposium on Biomedical Imaging (ISBI) 2016.

**L. Kostrykin and K. Rohr, "Superadditivity and Convex Optimization for Globally Optimal Cell Segmentation Using Deformable Shape Models," in IEEE Transactions on Pattern Analysis and Machine Intelligence, 2022, doi: 10.1109/TPAMI.2022.3185583.**