

# Interactive Cell Segmentation Based on Active and Semi-Supervised Learning

Hang Su\*, Zhaozheng Yin, Seungil Huh, Takeo Kanade, and Jun Zhu

**Abstract**—Automatic cell segmentation can hardly be flawless due to the complexity of image data particularly when time-lapse experiments last for a long time without biomarkers. To address this issue, we propose an interactive cell segmentation method by classifying feature-homogeneous superpixels into specific classes, which is guided by human interventions. Specifically, we propose to actively select the most informative superpixels by minimizing the expected prediction error which is upper bounded by the transductive Rademacher complexity, and then query for human annotations. After propagating the user-specified labels to the remaining unlabeled superpixels via an affinity graph, the error-prone superpixels are selected automatically and request for human verification on them; once erroneous segmentation is detected and subsequently corrected, the information is propagated efficiently over a gradually-augmented graph to un-labeled superpixels such that the analogous errors are fixed meanwhile. The correction propagation step is efficiently conducted by introducing a verification propagation matrix rather than rebuilding the affinity graph and re-performing the label propagation from the beginning. We repeat this procedure until most superpixels are classified into a specific category with high confidence. Experimental results performed on three types of cell populations validate that our interactive cell segmentation algorithm quickly reaches high quality results with minimal human interventions and is significantly more efficient than alternative methods, since the most informative samples are selected for human annotation/verification early.

**Index Terms**—Active learning, cell segmentation, label propagation, microscopy image analysis, verification propagation.

## I. INTRODUCTION

**L**IGHT microscope has been manipulated to help biologists understand underlying principles of specimens' life at cellular level ever since its development. Although analysis

of microscopy images can be conducted manually, it is always time-consuming and labor-intensive due to large volumes of image data captured from the long-term high-throughput biological experiments. Therefore, computerized techniques become essential such that the tasks can be performed automatically. Among the components in the computer-aided microscopy imaging system, cell segmentation referring to locate and detect cells in microscopy images is a central problem, since it is a cornerstone for lots of subsequent analysis, e.g., cell tracking, intercellular processes, etc. Although great efforts have been made in developing reliable cell segmentation techniques, it still offers a great challenge to develop automated computer vision and pattern recognition algorithms aiming at perfect accuracy, especially when staining methods are not involved in monitoring live cells. The challenge is ascribed to the heterogeneity and complexity of the data generated in modern imaging experiments. Therefore, it motivates us to allow users to participate and guide the segmentation procedure, which is called *interactive cell segmentation* [1], [2].

### A. Related Work

1) *Interactive Image Segmentation*: Interactive image segmentation has been extensively studied during the past decades. Different types of user intervention are expected to incorporate with photometric features of images to generate an improved result with minimal amounts of human efforts. One category of algorithms is recognized as boundary-based approaches. It requires users to specify an initial contour to approximate the target's boundary that will be evolved towards the desirable results. Rother [3] proposed a GrabCut method that models the color distribution within a box surrounding the intended segmentation targets, which are consequently segmented by iterative graph cut. However, its performance heavily relies on the global color model, and is often unsatisfactory for the object regions that share some similar color distributions with background. Recently, bounding box is further used to impose a powerful prior, e.g., tight shape prior [4] or high-level (semantic) knowledge of the intended target [5]. Nevertheless, the box-driven approaches require great attentions to the boundary specifications, e.g., the areas and shapes of the boxes. The performance undergoes a significant degeneration if the boundary box is rough, especially for targets with complex shapes. As another category, the region-based algorithms are also studied intensively, e.g., the graph-based segmentation [6], [7]. In particular, images are treated as a weighted graph, and identities of the unlabeled pixels are propagated from the neighboring labeled pixels on the affinity graph, e.g., graph cut [8] and random walk [9], [10]. Yilmaz *et al.* [11] further improved the graph-based

Manuscript received September 27, 2015; accepted October 15, 2015. Date of publication October 26, 2015; date of current version March 01, 2016. This research is partly supported by National Key Foundation R&D Projects No. 2013CB329403, National Natural Science Foundation of China (Nos. 61571261, 61322308, 61332007), NSF EPSCoR grant IIA-1355406 and NSF CAREER award IIS-1351049, and also supported by CPSF 2015M580099. *Asterisk indicates corresponding author.*

\*H. Su is with the Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China, and also with the Robotics Institute, Carnegie Mellon University (e-mail: suhangss@gmail.com).

Z. Yin is with the Department of Computer Science, Missouri University of Science and Technology, Rolla, MO 65401 USA.

S. Huh is with Google, Pittsburgh, PA 15213 USA.

T. Kanade is with the Robotics Institute, Carnegie Mellon University, Pittsburgh, PA 15201 USA.

J. Zhu is with the Department of Computer Science and Technology, Tsinghua University, Beijing 100084, China.

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2015.2494582

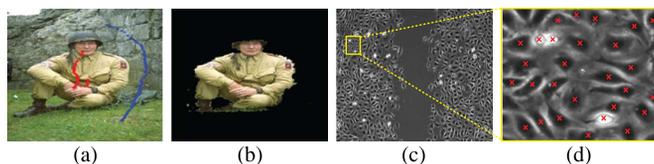


Fig. 1. Sample results of interactive segmentation. (a) An input image with user-supplied seeds for label propagation; (b) Foreground segmentation based on random walk [9]; (c) An input microscopy image; (d) Seed annotation (red cross markers on each cell) for cell segmentation.

approaches by partitioning an image into superpixels [12] in advance and constructing an affinity graph by measuring the pairwise similarity among the superpixels rather than among pixels. Label propagation can also be implemented from the initiated seeds along the gradients of an image in watershed [13]. Recently, Couprie *et al.* [14], [15] proposed a framework to unify the aforementioned region-based algorithms in terms of power watershed, and also pointed out the limitation of these techniques, namely, small over-segmented regions due to the small cut phenomenon. Addressing the detection of massive objects, Lin *et al.* [16] propose to decompose scenes into meaningful semantic parts with coherent motions and facilitate the analysis of complex and dense scenarios.

Although the aforementioned interactive approaches have gained success in lots of scenarios, there still exist limitations. Firstly, most of them can be recognized as foreground/background image segmentation, i.e., the images contain a dominant or a small number of foreground objects on the background (See Figs. 1(a) and (b)). If a large number of targets are scattered widely in an image, e.g.,  $10^2 \sim 10^3$  cells in microscopy images, specifying the boundary or labeling seeds for each object is labor-intensive (See Fig. 1(c) and (d)). Moreover, although the segmentation performance is greatly dependent on the information provided by users, it is not easy for an user to decide the most informative interventions under the present passive strategies, i.e., the users specify the intervention subjectively. An active strategy that can suggest a more informative intervention is desirable, since it will improve the performance and reduce the user's efforts at the same time. Finally, most of the current approaches do not provide a verification paradigm, namely, if there exist false segmentation results, the users have to *re-initialize* the annotation, and *re-perform* the segmentation from the scratch. Nevertheless, it is expensive or impossible to recollect the initial labeling set  $\mathcal{L}$ , because it requires a lot of human annotations or special devices/experiments to obtain the ground truth.

2) *Cell Segmentation*: The society of cell image analysis has witnessed the growing efforts in automatic cell segmentation during the past years. Thresholding [17], which segments cells out of background via a pre-defined threshold, may fail to separate cells of low contrast compared to the culturing medium. House *et al.* [18] proposed to segment cells by locating edges of them, but this work also suffers from low contrast or fuzzy boundaries. Cell segmentation is also challenged by the dissimilarity between different subcellular structures. Watershed-based method [19] segments cells based on the assumption that the intensity gradient inside cells is small while large around boundaries between cells and the background, but it often results in over segmentation due to regional minima. The de-

formable model, including active contour [20], [21] and level sets [22]–[24], attains the boundary by minimizing an energy function using intensity variance inside and outside contours. However, the result is sensitive to the initialization since the contour energy function is non-convex and there exists local minima. Therefore, it requires a sufficiently perfect initialization to obtain a high quality result. To cope with cells that appear almost transparent in the images, Zhang *et al.* [25] propose to train a classifier to detect (partial) cell boundaries, which is applicable to different modalities and cell types. In [26], the authors introduce a tree-structured discrete graphical model that is used to select and label a set of non-overlapping regions in the image such that overlapping and clustered cells are detected automatically. In [27], [28], Yin *et al.* propose to estimate a cell-sensitive camera response function by taking multiple exposures of phase contrast microscopy images on the same petri dish, which responds to cells' irradiance signals but generates a constant on non-cell background signal. To address the issue of mitosis event segmentation in large-scale time-lapse phase contrast microscopy image sequences of stem cell populations, the authors propose to recognize the mitosis based on semi-Markov model [29], [30] and sparse coding [31], [32]. In [33], [34], the authors propose a general algebraic framework for preconditioning microscopy images, which transforms an image that is unsuitable for direct analysis into an image that can be effortlessly segmented using global thresholding. More recently, Yin *et al.* [35] and Su *et al.* [36] studied the optical properties of phase contrast microscope, and developed imaging models to approximate the image formation process of microscopy images, based on which the optics-related features are obtained. In [48], the authors propose to co-restore phase contrast and differential interference contrast (DIC) microscopy images captured on the same cell dish simultaneously. However, it still needs to further study how to utilize these features to realize a more effective cell segmentation.

To surpass the limits of fully automated cell segmentation methods, we recently proposed a semi-supervised learning based cell segmentation method [2]. In the work, we presented cell segmentation as a superpixel classification problem, and performed it by propagating the labels of seed superpixels annotated by a human to unlabeled superpixels. Nevertheless, in this framework, the interventions are specified by users subjectively, but it is not easy for the users to decide the most informative interventions, thereby may require more human efforts in annotation. Moreover, although this approach was shown to be effective in cell segmentation, it was still not error-free during the long-term cell proliferation. Therefore, it is worth to consider how to further incorporate human guidance to boost the results, which leads to the following two questions:

- How to determine the most informative superpixel(s) that should be annotated as seed(s) for consequent cell segmentation, i.e., the query strategy during the interactive scenario?
- How to fix the analogous errors efficiently by verifying or correcting a small amount of error-prone superpixel(s)?

## B. Our Proposal

In this paper, we propose an interactive cell segmentation scheme by answering the aforementioned questions based

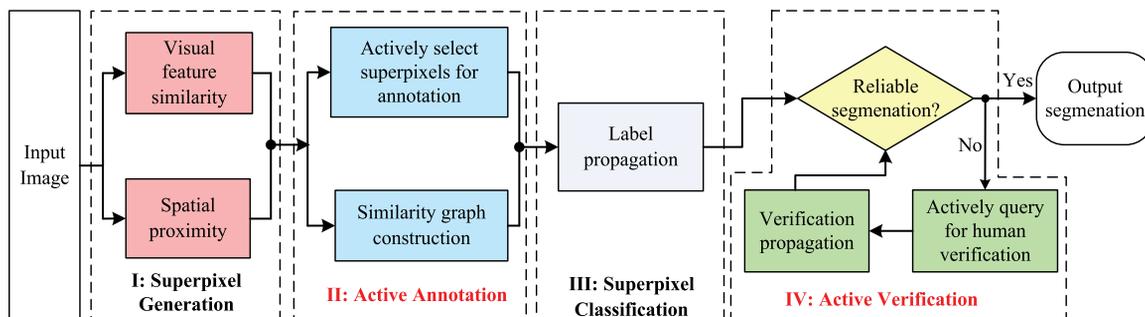


Fig. 2. A general framework of our proposed interactive cell segmentation scheme. For specific image sequences, we first partition an image into feature-homogeneous superpixels [2] in module I, and then select the most informative examples for annotations actively in module II. The step is conducted by minimizing the expected prediction error on the unlabeled superpixels, which is upper bounded by the transductive Rademacher complexity [37]. The human specified labels are then propagated to the unlabeled superpixels via an affinity graph in module III, which characterizes the pairwise similarity among the superpixels. In case that the segmentation is not reliable, we measure the uncertainty of each superpixel with entropy, and choose the most uncertain superpixels to query for verification. Human interventions are then systematically propagated to other unlabeled superpixels to fix the analogous misclassified state via a gradually-augmented affinity graph in module IV. Cell segmentation is realized until most of the unlabeled superpixels are classified into specific classes with high confidence.

on active and semi-supervised learning. Specifically, (1) the most informative superpixels are selected automatically by minimizing the expected prediction error, and (2) the human verifications are systematically propagated over a gradually-augmented affinity graph by introducing a verification propagation matrix, such that superpixels undergoing analogous misclassification are fixed meanwhile. The general framework of our proposed approach is illustrated in Fig. 2 with four major modules:

- 1) **Superpixel generation by clustering neighboring pixels** [2]. As the first step, we partition microscopy images into feature-homogeneous superpixels by clustering neighboring pixels with similar visual features, thereby eliminating the local redundancy of images.
- 2) **Active annotation by drawing samples that minimize the expected prediction error.** In order to reduce the human efforts in annotation, we seek to select the most informative superpixels early and annotate them first such that the expected prediction error is minimized after label propagation. Specifically, we use the tool of transductive Rademacher complexity [37] as a surrogate to guarantee a small expected prediction error. Consequently, human efforts are greatly reduced without impairing the performance of segmentation.
- 3) **Superpixel classification by propagating labels over the affinity graph.** In order to implement the label propagation, we construct an affinity graph that characterizes the pairwise similarity between the superpixels. Afterwards, the superpixels are classified into specified classes by propagating the human specified-labels to the unlabeled superpixels. Cell segmentation is consequently performed by grouping neighboring superpixels with similar identities [2].
- 4) **Active verification to fix misclassifications.** In case that the segmentation result is not reliable, we seek to measure the uncertainty of each unlabeled superpixel using entropy [38], and select a batch of error-prone superpixels to query for verification. Once errors are detected and corrected, the information is systematically propagated over an augmented affinity graph to the unlabeled superpixels. This

step is efficiently conducted by introducing a verification correction propagation matrix rather than re-performing the label propagation from the beginning. We repeat this procedure until most superpixels are classified into a specific category with high confidence.

Compared with the previous approaches for cell segmentation based on label propagation on *random annotation* [2], we propose to *actively select* the most informative samples such that the human efforts in annotation are reduced significantly (module II in Fig. 2). Moreover, most of the previous algorithms do not provide a strategy that boosts the results incrementally when misclassification is detected, which is common in many real-world scenarios when the characteristics of samples change gradually. In this paper, we propose a verification paradigm such that a batch of uncertain samples are *actively selected* for human verifications, whose correction will be propagated to the other unlabeled samples to correct similar misclassifications (module IV in Fig. 2).

As an empirical study, we conduct cell segmentation in phase contrast microscopy images in this paper, which still remains as a challenging task in computer-aided cell image analysis due to the low contrast and artifacts resulted from the particular optics of phase contrast microscope [35]. Specifically, we restore the phase feature vector of each pixel and then partition the images into phase-homogeneous superpixels [2], which are used as elementary units for subsequent cell segmentation. Note that other superpixel segmentation methods can also be used for other types of microscopy. Afterwards, cell segmentation is realized by classifying the superpixels into specific classes via our interactive procedure. Experimental results validate that our proposed interactive framework is effective and efficient for cell segmentation.

The remainder of this paper is organized as follows. In Section II, we elaborate the active annotation algorithm for interactive image segmentation, and we propose the verification propagation algorithm that fixes analogous errors in Section III. The experimental setup and results with discussions are presented in Section IV. Finally, the paper is summarized in Section V.

## II. CELL SEGMENTATION WITH ACTIVE ANNOTATION

In this section, we first present the graph model of our previous cell segmentation based on label propagation; then we elaborate the active annotation criterion by minimizing the expected prediction error using the tool of transductive Rademacher complexity; finally, we propose a sequential optimization algorithm to solve the problem.

### A. Overview of Cell Segmentation Through Label Propagation

In our previous work [2], segmentation is performed by partitioning a microscopy image into superpixels and classifying them into several predefined categories, such as bright cell, dark cell, and background regions. Mathematically, we define the set of annotated superpixels as  $\mathcal{L} \triangleq \{(\mathbf{x}_l, \mathbf{y}_l)\}_{l=1}^{N_l}$  where  $N_l$  is the number of annotated superpixels;  $\mathbf{x}_l$  is a feature vector that describes the visual characteristics of the  $l$ th labeled superpixel; and  $\mathbf{y}_l$  is a binary indicator corresponding to the identity of the  $l$ th superpixel, i.e.,  $\mathbf{y}_l(k) = 1$  means that the  $l$ th superpixel is labeled to belong to the  $k$ th class. In the same way, we define  $\mathcal{U} \triangleq \{(\mathbf{x}_u, \mathbf{y}_u)\}_{u=1}^{N_u}$  with  $N_u$  being the number of unlabeled superpixels and  $\{\mathbf{x}_u\}$  being the feature vector of the unlabeled superpixels. The label vectors of unlabeled superpixels  $\{\mathbf{y}_u\}$  are going to be inferred via the label propagation procedure.

Firstly, we construct an  $\epsilon$ -nearest neighboring graph [39] to characterize the pairwise similarity between superpixels. The adjacency matrix of the weight graph is denoted as  $\mathbf{W} = [w_{ij}]$ , which is an  $N \times N$  nonnegative and symmetric matrix with  $N = N_l + N_u$  being the number of total superpixels. To implement label propagation, the Laplacian matrix  $\mathbf{L}$  that corresponds to  $\mathbf{W}$ , i.e.,  $\mathbf{L} = \mathbf{D} - \mathbf{W}$  with  $\mathbf{D}$  being the diagonal degree matrix, is rearranged by splitting it into labeled and unlabeled sub-matrices as

$$\mathbf{L} \leftarrow \begin{bmatrix} \mathbf{L}_{ll} & \mathbf{L}_{lu} \\ \mathbf{L}_{ul} & \mathbf{L}_{uu} \end{bmatrix}, \quad (1)$$

where  $\mathbf{L}_{ll}$  is the Laplacian sub-matrix that characterizes the relationship between labeled superpixels;  $\mathbf{L}_{uu}$  denotes the sub-matrix corresponding to the unlabeled superpixels;  $\mathbf{L}_{lu}$  is a sub-matrix which interrelates the labeled and unlabeled superpixels; and  $\mathbf{L}_{ul} = \mathbf{L}_{lu}^T$ . Then, the problem of label propagation is formulated in a matrix form [2] as

$$\mathbf{Y}_u^* = \arg \min_{\mathbf{Y}_u} \text{tr} \left( [\mathbf{Y}_l; \mathbf{Y}_u]^T \begin{bmatrix} \mathbf{L}_{ll} & \mathbf{L}_{lu} \\ \mathbf{L}_{ul} & \mathbf{L}_{uu} \end{bmatrix} [\mathbf{Y}_l; \mathbf{Y}_u] \right), \quad (2)$$

where  $\text{tr}$  is the trace operator;  $\mathbf{Y}_l$  and  $\mathbf{Y}_u$  are binary indicator matrices corresponding to the labeled and unlabeled superpixels, respectively. Specifically, the indicator matrices  $\mathbf{Y}_l$  and  $\mathbf{Y}_u$  are constructed by stacking up the binary row indicators  $\{\mathbf{y}_n\}_{n=1}^N$  in rows correspondingly, namely, for the labeled superpixels  $\mathbf{Y}_l = [\mathbf{y}_1; \dots; \mathbf{y}_{N_l}]$ , and for unlabeled superpixels  $\mathbf{Y}_u = [\mathbf{y}_{N_l+1}; \dots; \mathbf{y}_u; \dots; \mathbf{y}_{N_l+N_u}]$ .

Based on the work in [39], label propagation is solved by Gaussian fields harmonic Functions (GFHF), and the solution is given by a hypotheses  $\mathcal{H}$  as

$$\begin{aligned} \mathcal{H} : \mathbf{Y}_u^* &= -(\mathbf{D}_{uu} - \mathbf{W}_{uu})^{-1} \mathbf{L}_{ul} \mathbf{Y}_l \\ &= -\mathbf{L}_{uu}^{-1} \mathbf{L}_{ul} \mathbf{Y}_l \\ &\triangleq -\mathbf{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{Y}_l, \end{aligned} \quad (3)$$

where  $\mathbf{D}_{uu}$  is a diagonal degree matrix with  $\mathbf{D}_{uu}(i, i) = \sum_{j=1}^{N_l+N_u} \mathbf{W}(i, j)$ ; and  $\mathbf{\Gamma}_{uu}$  is the inverse of the Laplacian matrix  $\mathbf{L}_{uu}$ , i.e.,  $\mathbf{\Gamma}_{uu} = \mathbf{L}_{uu}^{-1}$ . Note that the elements of  $\mathbf{Y}_u^*$  are not binary but real numbers, thus  $\mathbf{Y}_u^*$  can be considered as soft label results. The hard label vector can be obtained simply by converting the maximum value in each  $\mathbf{y}_u^*$  into 1 and the others into 0.

### B. Active Annotation by Minimizing Expected Prediction Error

In interactive cell segmentation, it is always the case that data are abundant but labels are scarce or expensive to acquire. This motivates us to select the most informative superpixels for annotation to reduce the labeling efforts. In this section, we propose an active annotation criterion that minimizes the expected prediction error on the unlabeled superpixels using the tool of transductive Rademacher complexity [37], since it outperforms other bounds based on combinatorial dimensions [40] and can reflect the properties of the particular probability distribution [41]. As the first step, we generalize the transductive Rademacher complexity [37] to a multi-class version as

*Definition (Transductive Rademacher Complexity):* For a sample set  $\mathcal{D} \triangleq \mathcal{L} \cup \mathcal{U} = \{\mathbf{x}_n\}_{n=1}^N$  with  $N = N_l + N_u$ , if  $\mathcal{H}$  is a class of real-valued function on  $\mathcal{D}$ , the transductive Rademacher complexity of  $\mathcal{H}$  is defined as

$$\hat{\mathcal{R}}(\mathcal{H}, \mathcal{L}) = \left( \frac{1}{N_l} + \frac{1}{N_u} \right) \mathbb{E}_{\sigma} \left[ \sup_{\mathbf{h} \in \mathcal{H}} \sum_{i=1}^{N_c} \sigma^T \mathbf{h}_i(\mathbf{X}) \right], \quad (4)$$

where  $N_c$  is the number of classes,  $\mathbf{X}$  is a matrix by collecting the feature vectors  $\mathbf{X} = [\mathbf{x}_1; \dots; \mathbf{x}_N]$ ;  $\mathbf{h}_i(\mathbf{X}) = [h_i(\mathbf{x}_1); \dots; h_i(\mathbf{x}_N)]$  is a column vector of hypothesis functions for the  $i$ th class, and  $\sigma^T = [\sigma_1, \dots, \sigma_N]$  is a row vector of i.i.d. random variables such that  $\sigma_n$  is equal to 1 or -1 with the probability  $p \in [0, 1/2]$  for each, or 0 with the probability  $1 - 2p$ .

It has been proved that minimizing the bound of transductive Rademacher complexity is a proxy for minimizing a prediction error [42]. In the following, we derive the upper bound of the transductive Rademacher complexity in Theorem 1. It has been empirically proved that the upper bound of Rademacher complexity is close to the true Rademacher complexity as more samples are drawn from the dataset [42] and serves as the theoretical foundation of our proposed active annotation algorithm.

*Theorem 1. (Bound of Transductive Rademacher Complexity):* The transductive Rademacher complexity of label propagation based on  $\mathcal{H}$  in (3) is upper bounded by

$$\hat{\mathcal{R}}(\mathcal{H}; \mathcal{L}) \leq C \sqrt{\text{tr}(\mathbf{P}\mathbf{P}^T)}, \quad (5)$$

where  $\mathbf{P}$  is the label propagation matrix  $\mathbf{P} = -\mathbf{\Gamma}_{uu} \mathbf{L}_{ul}$ , and  $C$  is a constant.

*Proof:* In (3), the label propagation is given by

$$\mathbf{Y}_u = -\mathbf{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{Y}_l \triangleq \mathbf{P} \mathbf{Y}_l. \quad (6)$$

By definition 1, the transductive Rademacher complexity for label propagation is computed as

$$\hat{\mathcal{R}}(\mathcal{H}; \mathcal{L}) = \left( \frac{1}{N_l} + \frac{1}{N_u} \right) \mathbb{E}_{\sigma} \left[ \sup_{\mathbf{h} \in \mathcal{H}} \sum_{i=1}^{N_c} \sigma^T \mathbf{P} \mathbf{Y}_l^{(i)} \right], \quad (7)$$

where  $\mathbf{Y}_l^{(i)}$  is the  $i$ th column of  $\mathbf{Y}_l$ , i.e.,  $\mathbf{Y}_l^{(i)}(n) = 1$  indicates that the  $n$ th sample belongs to the  $i$ th class, otherwise  $\mathbf{Y}_l^{(i)}(n) = 0$ . Therefore,  $\mathbf{P}\mathbf{Y}_l^{(i)}$  indicates the probability of corresponding samples belonging to the  $i$ th class. By applying the Cauchy-Schwarz inequality, we obtain

$$\begin{aligned} \hat{\mathcal{R}}(\mathcal{H}; \mathcal{L}) &= \left( \frac{1}{N_l} + \frac{1}{N_u} \right) \mathbb{E}_{\boldsymbol{\sigma}} \left[ \sup_{\mathbf{h} \in \mathcal{H}} \boldsymbol{\sigma}^T \mathbf{P} \sum_{i=1}^{N_c} \mathbf{Y}_l^{(i)} \right] \\ &\leq \left( \frac{1}{N_l} + \frac{1}{N_u} \right) \mathbb{E}_{\boldsymbol{\sigma}} \left[ \sup_{\mathbf{h} \in \mathcal{H}} \|\boldsymbol{\sigma}^T \mathbf{P}\|_2 \left\| \sum_{i=1}^{N_c} \mathbf{Y}_l^{(i)} \right\|_2 \right] \\ &= \left( \frac{1}{N_l} + \frac{1}{N_u} \right) \mathbb{E}_{\boldsymbol{\sigma}} \left[ \sup_{\mathbf{h} \in \mathcal{H}} \|\boldsymbol{\sigma}^T \mathbf{P}\|_2 \|\mathbf{Y}_l\|_2 \right]. \quad (8) \end{aligned}$$

Since there are  $N_l$  labeled samples,  $\|\mathbf{Y}_l\|_2 = \sqrt{N_l}$ . Using the property of inner product, we have

$$\begin{aligned} \hat{\mathcal{R}}(\mathcal{H}; \mathcal{L}) &\leq \left( \frac{1}{N_l} + \frac{1}{N_u} \right) \sqrt{N_l} \mathbb{E}_{\boldsymbol{\sigma}} \left[ \sqrt{\boldsymbol{\sigma}^T \mathbf{P} \mathbf{P}^T \boldsymbol{\sigma}} \right] \\ &= \left( \frac{1}{N_l} + \frac{1}{N_u} \right) \sqrt{N_l} \mathbb{E}_{\boldsymbol{\sigma}} \\ &\quad \times \left[ \sqrt{\sum_{i,j=1}^{N_l+N_u} \sigma_i \sigma_j \langle \mathbf{P}(i, :), \mathbf{P}(j, :)\rangle} \right]. \quad (9) \end{aligned}$$

Using the Jensen inequality, (9) is upper bounded by

$$\begin{aligned} \hat{\mathcal{R}}(\mathcal{H}; \mathcal{L}) &\leq \left( \frac{\sqrt{N_l}}{N_l} + \frac{\sqrt{N_l}}{N_u} \right) \\ &\quad \times \left[ \sqrt{\sum_{i,j=1}^{N_l+N_u} \mathbb{E}_{\boldsymbol{\sigma}} \{ \sigma_i \sigma_j \langle \mathbf{P}(i, :), \mathbf{P}(j, :)\rangle \}} \right] \\ &= \left( \frac{\sqrt{N_l}}{N_l} + \frac{\sqrt{N_l}}{N_u} \right) \\ &\quad \times \left[ \sqrt{\sum_{i=1}^{N_l+N_u} \frac{2N_l N_u}{(N_l + N_u)^2} \langle \mathbf{P}(i, :), \mathbf{P}(i, :)\rangle} \right] \quad (10) \end{aligned}$$

$$= C \sqrt{\text{tr}(\mathbf{P}\mathbf{P}^T)}, \quad (11)$$

where  $C = \sqrt{2/N_u}$ . Note that (10) is obtained by the expectation of Rademacher variables in definition 1. Hereby, we obtain the upper bound of the transductive Rademacher complexity as

$$\hat{\mathcal{R}}(\mathcal{H}; \mathcal{L}) \leq C \sqrt{\text{tr}(\mathbf{P}\mathbf{P}^T)}. \quad (12)$$

By ignoring the constant scalar and square root function in Theorem 1, the active annotation can be implemented equivalently by minimizing the upper bound of transductive Rademacher complexity as

$$\min_{\mathcal{L} \subset \mathcal{D}} \text{tr}(\mathbf{P}\mathbf{P}^T). \quad (13)$$

After some mathematical derivations, we propose the following active annotation criterion in Theorem 2.

*Theorem 2 (Active Annotation Criterion):* The active annotation can be implemented by solving the problem as

$$\min_{\mathcal{L} \subset \mathcal{D}} \text{tr} \left( (\mathbf{L}_{uu})^{-2} (\mathbf{L}^2)_{uu} \right), \quad (14)$$

where  $\mathbf{L}$  is the Laplacian matrix.

*Proof:* According to the cyclic property of trace, it can be proved that

$$\begin{aligned} \text{tr}(\mathbf{P}\mathbf{P}^T) &= \text{tr} \left( \boldsymbol{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{L}_{ul}^T \boldsymbol{\Gamma}_{uu}^T \right) \\ &= \text{tr} \left( \boldsymbol{\Gamma}_{uu}^T \boldsymbol{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{L}_{lu} \right) \\ &= \text{tr} \left( \boldsymbol{\Gamma}_{uu} \boldsymbol{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{L}_{lu} \right). \quad (15) \end{aligned}$$

Eq. (15) holds due to the fact that  $\boldsymbol{\Gamma}_{uu} = \boldsymbol{\Gamma}_{uu}^T$ . Additionally, we can obtain that

$$\begin{aligned} (\mathbf{L}^2)_{uu} &= \left( \begin{bmatrix} \mathbf{L}_{ll} & \mathbf{L}_{lu} \\ \mathbf{L}_{ul} & \mathbf{L}_{uu} \end{bmatrix} \begin{bmatrix} \mathbf{L}_{ll} & \mathbf{L}_{lu} \\ \mathbf{L}_{ul} & \mathbf{L}_{uu} \end{bmatrix} \right)_{uu} \\ &= \mathbf{L}_{ul} \mathbf{L}_{lu} + \mathbf{L}_{uu} \mathbf{L}_{uu}. \quad (16) \end{aligned}$$

By substituting (16) into (15), we obtain

$$\begin{aligned} \text{tr}(\mathbf{P}\mathbf{P}^T) &= \text{tr} \left( \boldsymbol{\Gamma}_{uu}^2 \left( (\mathbf{L}^2)_{uu} - \mathbf{L}_{uu} \mathbf{L}_{uu} \right) \right) \\ &= \text{tr} \left( \boldsymbol{\Gamma}_{uu}^2 (\mathbf{L}^2)_{uu} \right) - \text{tr} \left( \boldsymbol{\Gamma}_{uu}^2 \mathbf{L}_{uu}^2 \right), \\ &= \text{tr} \left( (\mathbf{L}_{uu})^{-2} (\mathbf{L}^2)_{uu} \right) - N_u. \quad (17) \end{aligned}$$

Therefore, the problem in (13) is equivalent to

$$\min_{\mathcal{L} \subset \mathcal{D}} \text{tr} \left( (\mathbf{L}_{uu})^{-2} (\mathbf{L}^2)_{uu} \right). \quad \blacksquare$$

### C. Sequential Annotation of Samples

Since the active annotation criterion in Theorem 2 is a combinatorial optimization problem, it is a NP-hard problem to find the global optimal solution. In order to solve it, we propose a sequential annotation scheme as follows. First, we introduce a matrix  $\tilde{\mathbf{S}}(N \times N_u)$  to indicate the unselected samples, which is defined as  $\tilde{\mathbf{S}} = [s_{ij}]$  with

$$s_{ij} = \begin{cases} 1, & \text{if the } i_{\text{th}} \text{ sample in } \mathcal{D} \text{ is corresponding} \\ & \text{to the } j_{\text{th}} \text{ sample in } \mathcal{U} \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

Therefore, the problem in Theorem 2 can be rewritten using matrix  $\tilde{\mathbf{S}}$  as

$$\begin{aligned} \min_{\tilde{\mathbf{S}}} \text{tr} \left( (\tilde{\mathbf{S}}^T \mathbf{L} \tilde{\mathbf{S}})^{-2} (\tilde{\mathbf{S}}^T \mathbf{L}^2 \tilde{\mathbf{S}}) \right). \\ \text{s.t. } \begin{cases} \tilde{\mathbf{S}} \in \{0, 1\}^{N \times N_u}, \\ \tilde{\mathbf{S}}^T \tilde{\mathbf{S}} = \mathbf{I}_{N_u}. \end{cases} \quad (19) \end{aligned}$$

Using eigenvalue analysis, the Laplacian matrix is decomposed as

$$\mathbf{L} = \mathbf{Q}\boldsymbol{\Lambda}\mathbf{Q}^T, \quad (20)$$

where  $\mathbf{Q}$  and  $\mathbf{\Lambda}$  are corresponding to the eigenvector and eigenvalue matrices, respectively. Therefore, the problem in (19) can be rewritten as

$$\min_{\bar{\mathbf{S}}} \text{tr} \left( (\bar{\mathbf{S}}^T \mathbf{Q} \mathbf{\Lambda} \mathbf{Q}^T \bar{\mathbf{S}})^{-2} (\bar{\mathbf{S}}^T \mathbf{Q} \mathbf{\Lambda}^2 \mathbf{Q}^T \bar{\mathbf{S}}) \right). \quad (21)$$

Once  $\bar{\mathbf{S}}$  is determined, the selected samples for annotation are obtained by excluding the samples corresponding to  $\bar{\mathbf{S}}$  in  $\mathcal{U}$ . By denoting  $\mathbf{R} = \bar{\mathbf{S}}^T \mathbf{Q}$ , the problem can be reformulated as

$$\min_{\mathbf{R}} \text{tr} \left( (\mathbf{R} \mathbf{\Lambda} \mathbf{R}^T)^{-2} (\mathbf{R} \mathbf{\Lambda}^2 \mathbf{R}^T) \right). \quad (22)$$

Since the eigenvectors are orthogonal, it can be proved that  $\mathbf{R} \mathbf{R}^T = \mathbf{I}_{N_u}$ . Thus,

$$\begin{aligned} (\mathbf{R} \mathbf{\Lambda} \mathbf{R}^T)^{-1} &= (\mathbf{R} (\mathbf{\Lambda} + \mathbf{I}) \mathbf{R}^T - \mathbf{R} \mathbf{R}^T)^{-1} \\ &= (-\mathbf{I} + \mathbf{R} (\mathbf{\Lambda} + \mathbf{I}) \mathbf{R}^T)^{-1} \\ &= -\mathbf{I} - \mathbf{R} ((\mathbf{\Lambda} + \mathbf{I})^{-1} - \mathbf{R}^T \mathbf{R})^{-1} \mathbf{R}^T \\ &\triangleq -\mathbf{I} - \mathbf{R} \Phi \mathbf{R}^T, \end{aligned} \quad (23)$$

where  $\Phi = ((\mathbf{\Lambda} + \mathbf{I})^{-1} - \mathbf{R}^T \mathbf{R})^{-1}$ . Therefore, we can obtain

$$\begin{aligned} &\text{tr} \left( (\mathbf{R} \mathbf{\Lambda} \mathbf{R}^T)^{-2} (\mathbf{R} \mathbf{\Lambda}^2 \mathbf{R}^T) \right) \\ &= \text{tr} \left( (-\mathbf{I} - \mathbf{R} \Phi \mathbf{R}^T)^2 (\mathbf{R} \mathbf{\Lambda}^2 \mathbf{R}^T) \right) \\ &= \text{tr} \left( (\mathbf{I} + 2\mathbf{R} \Phi \mathbf{R}^T + \mathbf{R} \Phi \mathbf{R}^T \mathbf{R} \Phi \mathbf{R}^T (\mathbf{R} \mathbf{\Lambda}^2 \mathbf{R}^T)) \right) \\ &= \text{tr}(\mathbf{R} \mathbf{\Lambda}^2 \mathbf{R}^T) + 2\text{tr}(\mathbf{R} \Phi \mathbf{R}^T \mathbf{R} \mathbf{\Lambda}^2 \mathbf{R}^T) \\ &\quad + \text{tr}(\mathbf{R} \Phi \mathbf{R}^T \mathbf{R} \Phi \mathbf{R}^T \mathbf{R} \mathbf{\Lambda}^2 \mathbf{R}^T) \\ &= \text{tr}(\mathbf{\Lambda}^2 \mathbf{R}^T \mathbf{R}) + 2\text{tr}(\Phi \mathbf{R}^T \mathbf{\Lambda}^2 \mathbf{R}^T \mathbf{R}) \\ &\quad + \text{tr}(\Phi \mathbf{R}^T \mathbf{R} \Phi \mathbf{R}^T \mathbf{R} \mathbf{\Lambda}^2 \mathbf{R}^T \mathbf{R}). \end{aligned} \quad (24)$$

In order to solve the NP-hard problem, we proposed a sequential minimization algorithm to find the global optimization of  $\mathbf{R}$ . Specifically, suppose  $k-1$  samples have been selected, then the subsequent  $k$ th sample can be selected by solving the following optimization problem

$$\min_{\mathbf{R}_k} \left\{ \begin{aligned} &\text{tr}(\mathbf{\Lambda}^2 \mathbf{R}_k^T \mathbf{R}_k) + 2\text{tr}(\Phi_k \mathbf{R}_k^T \mathbf{R}_k \mathbf{\Lambda}^2 \mathbf{R}_k^T \mathbf{R}_k) \\ &+ \text{tr}((\Phi_k \mathbf{R}_k^T \mathbf{R}_k)^2 \mathbf{\Lambda}^2 \mathbf{R}_k^T \mathbf{R}_k) \end{aligned} \right\}, \quad (25)$$

where  $\mathbf{R}_k = [\mathbf{q}_1, \dots, \mathbf{q}_k]$  with  $\mathbf{q}_k$  being the column vector, which is obtained by transposing the row vector of  $\mathbf{Q}$  corresponding to the sample selected in the  $k$ th iteration; and  $\Phi_{k-1} = ((\mathbf{\Lambda} + \mathbf{I})^{-1} - \mathbf{R}_{k-1}^T \mathbf{R}_{k-1})^{-1}$ . Since  $\mathbf{R} = \bar{\mathbf{S}}^T \mathbf{Q}$  is to select the corresponding column vectors in the  $\mathbf{Q}$  matrix that is removed in  $k$ th iteration, we initialize  $\mathbf{R}_k$  as  $\mathbf{R}_0 = \mathbf{Q}$ . The problem can be solved sequentially by searching a sample in  $\mathcal{U}$  which results in the minimum increment of the trace once  $k-1$  samples are determined (i.e.,  $\mathbf{R}_{k-1}$  is determined). The corresponding vector  $\mathbf{q}_k$  is removed from  $\mathbf{R}_{k-1}$  when the  $k$ th sample is selected for annotation.

It is noted that only the diagonal entries contribute to the final result when conducting a trace calculation, hence the problem

in (25) can be solved in a more efficient manner as follows. According to the definition of  $\mathbf{R}_k$ , we have

$$\begin{aligned} \mathbf{R}_{k-1}^T \mathbf{R}_{k-1} &= [\mathbf{q}_1, \dots, \mathbf{q}_{N+1-k}]^T [\mathbf{q}_1, \dots, \mathbf{q}_{N+1-k}] \\ &= \sum_{i=1}^{N+1-k} \sum_{j=1}^{N+1-k} \mathbf{q}_i \mathbf{q}_j^T, \\ &= \sum_{i=1}^{N-k} \sum_{j=1}^{N-k} \mathbf{q}_i \mathbf{q}_j^T + \mathbf{q}_k \sum_{j=1}^{N-k} \mathbf{q}_j^T + \sum_{j=1}^{N-k} \mathbf{q}_j (\mathbf{q}_k)^T \\ &\quad + \mathbf{q}_k \mathbf{q}_k^T \\ &= \mathbf{R}_k^T \mathbf{R}_k + \mathbf{q}_k \sum_{j=1}^{N-k} \mathbf{q}_j^T + \sum_{j=1}^{N-k} \mathbf{q}_j (\mathbf{q}_k)^T \\ &\quad + \mathbf{q}_k \mathbf{q}_k^T. \end{aligned} \quad (26)$$

Note that  $\mathbf{q}_i^T \mathbf{q}_j = 0$  for  $i \neq j$  due to the orthogonality of row vectors in the normalized eigenvector matrix, (26) can be rewritten as

$$\mathbf{R}_k^T \mathbf{R}_k = \mathbf{R}_{k-1}^T \mathbf{R}_{k-1} - \mathbf{q}_k \mathbf{q}_k^T. \quad (27)$$

By substituting (27) into (25) iteratively, we transform the trace calculation into a multiplication of vector-matrix-vector problem. By denoting  $\mathbf{A}_k = \mathbf{R}_k^T \mathbf{R}_k$ , the optimization problem can be solved by sequentially selecting the sample that results in the minimal trace increment as

$$\min_k \left( \mathbf{q}_k^T \mathbf{X}_k \mathbf{q}_k + \frac{2}{c_k} \mathbf{q}_k^T \mathbf{Y}_k \mathbf{q}_k + \frac{1}{c_k^2} \mathbf{q}_k^T \mathbf{D}_k \mathbf{q}_k \mathbf{q}_k^T \mathbf{F}_k \mathbf{q}_k \right), \quad (28)$$

where  $c_k = 1 - \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k$ ;  $\mathbf{X}_k = -(\mathbf{\Lambda}^2 + \mathbf{B}_k + \mathbf{C}_k + \mathbf{E}_k)$  with  $\mathbf{B}_k = (2\Phi_k + \Phi_k \mathbf{A}_k \Phi_k) \mathbf{A}_k \mathbf{\Lambda}^2$ ,  $\mathbf{C}_k = \mathbf{\Lambda}^2 \mathbf{A}_{k-1} (2\Phi_k + \Phi_k \mathbf{A}_k \Phi_k)$  and  $\mathbf{E}_k = \Phi_k \mathbf{A}_{k-1} \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \Phi_k$ ;  $\mathbf{Y}_k = -\mathbf{D}_k (\mathbf{I} + \mathbf{A}_{k-1} \Phi_{k-1})$  with  $\mathbf{D}_k = \Phi_{k-1} \mathbf{A}_{k-1} \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \Phi_{k-1}$ ; and  $\mathbf{F}_k = \Phi_{k-1} \mathbf{A}_{k-1} \Phi_{k-1}$  (We detail the proof in Appendix A). Note that by transforming the trace calculation in (25) to a vector-matrix-vector multiplication in (28), only the trace-related rather than all the entries in (25) are calculated to solve the problem, and thus the computation cost is reduced significantly.

Additionally, once the  $k$ th sample is selected,  $\Phi_k$  can be updated based on  $\Phi_{k-1}$  using the Binomial inverse theorem as

$$\begin{aligned} \Phi_k &= ((\mathbf{\Lambda} + \mathbf{I})^{-1} - \mathbf{R}_k^T \mathbf{R}_k)^{-1} \\ &= ((\mathbf{\Lambda} + \mathbf{I})^{-1} - \mathbf{R}_{k-1}^T \mathbf{R}_{k-1} + \mathbf{q}_k \mathbf{q}_k^T)^{-1} \\ &= (\Phi_{k-1}^{-1} + \mathbf{q}_k \mathbf{q}_k^T)^{-1} \\ &= \Phi_{k-1} - \frac{\Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k}. \end{aligned} \quad (29)$$

Hence,  $\Phi_k^{-1}$  can be computed by matrix (vector) multiplication and addition, which is much more efficient than matrix inverse. In summary, we detail the active sample selection algorithm in Algorithm 1.

### III. FIXING ERRONEOUS SEGMENTATION BY ACTIVE VERIFICATION PROPAGATION

Although we have actively selected informative superpixels for annotation, it is not realistic to expect the result is error-free

---

**Algorithm 1** Active Sample Selection by Minimizing Transductive Rademacher Complexity
 

---

- 1: **Input:** Laplacian matrix  $\mathbf{L}$ , sample set  $\mathcal{D}$ , number of samples to select  $N_l$
- 2: **Output:** Actively selected sample set  $\mathcal{L}$
- 3: Perform eigenvalue decomposition:  $\mathbf{L} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ .
- 4: **Initialize:**  $\mathbf{R}_0 = \mathbf{Q}$ ;  $\mathbf{A}_0 = \mathbf{R}_0^T \mathbf{R}_0$ ;  $\mathbf{\Phi}_0 = \mathbf{\Lambda} + \mathbf{I}$ .
- 5: **for**  $k = 0 \rightarrow N_l - 1$  **do**
- 6:    //Calculate the trace increment if the corresponding superpixel in  $\mathcal{U}_k$  is selected in the  $k^{\text{th}}$  iteration.
- 7:    **for all** columns  $\{\mathbf{q}_i\}_{i=1}^{N-k}$  in  $\mathbf{R}_k$  **do**
- 8:        // Compute the corresponding factor matrices as

$$\begin{aligned} \mathbf{A}_k^i &= \mathbf{A}_{k-1} - \mathbf{q}_i \mathbf{q}_i^T, \\ \mathbf{\Phi}_k^i &= \mathbf{\Phi}_{k-1} - \frac{\mathbf{\Phi}_{k-1} \mathbf{q}_i \mathbf{q}_i^T \mathbf{\Phi}_{k-1}}{1 + \mathbf{q}_i^T \mathbf{\Phi}_{k-1} \mathbf{q}_i} \\ \mathbf{B}_k^i &= (2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k^i \mathbf{\Phi}_k) \mathbf{A}_k^i \mathbf{\Lambda}^2, \\ \mathbf{C}_k^i &= \mathbf{\Lambda}^2 \mathbf{A}_{k-1}^i (2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k^i \mathbf{\Phi}_k), \\ \mathbf{D}_k^i &= \mathbf{\Phi}_{k-1} \mathbf{A}_{k-1} \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \mathbf{\Phi}_{k-1}, \\ \mathbf{E}_k^i &= \mathbf{\Phi}_k \mathbf{A}_{k-1} \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \mathbf{\Phi}_k \\ \mathbf{F}_k^i &= \mathbf{\Phi}_{k-1} \mathbf{A}_{k-1} \mathbf{\Phi}_{k-1}, \\ c_i &= 1 + \mathbf{q}_i^T \mathbf{\Phi}_{k-1} \mathbf{q}_i \\ \mathbf{X}_k^i &= -(\mathbf{\Lambda}^2 + \mathbf{B}_k^i + \mathbf{C}_k^i + \mathbf{E}_k^i), \\ \mathbf{Y}_k^i &= -\mathbf{D}_k^i (\mathbf{I} + \mathbf{A}_{k-1}^i \mathbf{\Phi}_{k-1}) \end{aligned}$$

- 9:    **end for**
- 10:    Select and annotate the superpixel that results in the minimum trace increment by solving

$$\min_k \left( \mathbf{q}_k^T \mathbf{X}_k \mathbf{q}_k + \frac{2}{c_k} \mathbf{q}_k^T \mathbf{Y}_k \mathbf{q}_k + \frac{1}{c_k^2} \mathbf{q}_k^T \mathbf{D}_k \mathbf{q}_k \mathbf{q}_k^T \mathbf{F}_k \mathbf{q}_k \right).$$

- 11:    Update the selected sample set:  $\mathcal{L}_k \leftarrow \mathcal{L}_{k-1} \cup \{k\}$ .
- 12:    Update the matrices as

$$\begin{aligned} \mathbf{A}_k &\leftarrow \mathbf{A}_{k-1} - \mathbf{q}_k \mathbf{q}_k^T \\ \mathbf{\Phi}_k &\leftarrow \mathbf{\Phi}_{k-1} - \frac{\mathbf{\Phi}_{k-1} \mathbf{q}_k \mathbf{q}_k^T \mathbf{\Phi}_{k-1}}{1 + \mathbf{q}_k^T \mathbf{\Phi}_{k-1} \mathbf{q}_k} \end{aligned}$$

- 13: **end for**
- 

especially for the long-time high-throughput images. Therefore, it is worth to consider how to further incorporate human guidance to achieve better results. In this case, it needs to figure out which samples are error-prone and should be validated; and propagate the human verifications to unlabeled samples to fix analogous errors.

In this section, we propose a verification scheme such that the false segmentations are corrected efficiently. First, we measure the uncertainty of the predicted label on each unlabeled superpixel using entropy [38], and actively select the most uncertain superpixels to query for verification. Once errors are detected and corrected, the information is systematically propagated over a gradually-augmented affinity graph to the unlabeled superpixels. This step is efficiently conducted by introducing a verification propagation matrix rather than re-performing the label propagation from the beginning. We repeat this procedure until most of superpixels are classified into a specific category with high confidence. The procedure is detailed in the following section.

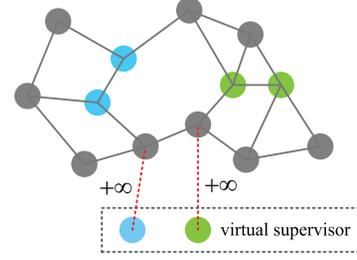


Fig. 3. An augmented graph example where blue and green nodes indicate differently labeled samples and gray nodes are unlabeled samples. Human labels are given on the most confusing samples by adding auxiliary nodes named virtual supervisors and connecting them to the samples with an infinite weight.

#### A. Active Sample Selection for Verification

Although it is expected that there exist misclassifications after label propagation, it is not easy for the user to determine the most effective intervention such that human interventions are minimized and the resultant performance is maximized. Recently, sample selection based on active learning [43] has provided a promising direction to help the user select the most informative samples to be corrected.

Given the segmentation results acquired through label propagation, we measure the uncertainty of the predicted label of each unlabeled sample with entropy [38]. After normalizing the indicator of the unlabeled sample to  $[0, 1]$  as  $\hat{\mathbf{y}}_u^* = (1/\|\mathbf{y}_u^*\|_2) \mathbf{y}_u^*$ , the entropy of prediction is thus

$$H(u) = -\hat{\mathbf{y}}_u^* \log \hat{\mathbf{y}}_u^{*T}. \quad (30)$$

A larger entropy implies that the label is more uncertain. During each iteration of the interactive segmentation, we actively select  $K$  samples with the  $K$  largest entropies out of  $\mathcal{U}$ , the set of which is denoted by  $\mathcal{K}$ , and request human validation on them.

#### B. Verification Propagation Over Graph

Once a user recognizes some errors during verification and corrects them, it is desirable to search for similar errors that can be fixed based on the given human intervention. The simplest way to conduct this task would be to rebuild or modify the affinity matrix in (1) and re-perform label propagation. However, this scheme is too inefficient for the verification to be interactive because the inverse of the Laplacian matrix in (3) needs to be recalculated.

Instead, we propose a verification propagation scheme by extending the augmented graph method introduced by [44] to handle a batch of samples at each round for more effective and efficient interaction. We build an augmented graph by adding auxiliary nodes, which are called *virtual supervisors* and denoted by  $\mathcal{S} = \{\mathbf{y}_s\}_{s=1}^{N_s}$  with  $N_s$  being the number of virtual supervisors and  $\mathbf{y}_s$  being the human label given on the  $s$ th sample, and connect them to the corrected samples with weight  $w$  such that  $w \rightarrow +\infty$ , as shown in Fig. 3. Then the adjacency matrix of the augmented graph can be built by adding some rows and columns to the original adjacency matrix as follows:

$$\mathbf{W}^+ \leftarrow \begin{bmatrix} \mathbf{W}_{ll} & \mathbf{W}_{lu} & \mathbf{O}_{ls} \\ \mathbf{W}_{ul} & \mathbf{W}_{uu} & \mathbf{W}_{us} \\ \mathbf{O}_{sl} & \mathbf{W}_{su} & \mathbf{O}_{ss} \end{bmatrix} \quad (31)$$

where  $\mathbf{O}_{ls}$ ,  $\mathbf{O}_{sl}$  and  $\mathbf{O}_{ss}$  are zero submatrices with appropriate sizes, and  $\mathbf{W}_{us} = w\mathbf{Z}_{us}$  and  $\mathbf{W}_{su} = w\mathbf{Z}_{su} = w\mathbf{Z}_{us}^T$  where  $\mathbf{Z}_{us}$  and  $\mathbf{Z}_{su}$  are binary indicator matrices indicating which virtual supervisor is connected to which unlabeled sample.

The label propagation over this augmented graph can be obtained by minimizing the following objective function [45]:

$$f(\mathbf{Y}_u) = \text{tr} \left( [\mathbf{Y}_l; \mathbf{Y}_u; \mathbf{Y}_s]^T \mathbf{L}^+ \begin{bmatrix} \mathbf{Y}_l \\ \mathbf{Y}_u \\ \mathbf{Y}_s \end{bmatrix} \right), \quad (32)$$

where  $\mathbf{L}^+$  is the Laplacian matrix of  $\mathbf{W}^+$ ; and  $\mathbf{Y}_s$  is the label matrix of virtual supervisors. Minimizing this objective yields a closed form solution:

$$\begin{aligned} \mathbf{Y}_u^+ &= -(\mathbf{L}_{uu}^+)^{-1} [\mathbf{L}_{ul} \quad \mathbf{L}_{us}] \begin{bmatrix} \mathbf{Y}_l \\ \mathbf{Y}_s \end{bmatrix} \\ &= -\mathbf{\Gamma}_{uu}^+ (\mathbf{L}_{ul} \mathbf{Y}_l - w \mathbf{Z}_{us} \mathbf{Y}_s) \end{aligned} \quad (33)$$

where  $\mathbf{L}_{uu}^+$  denotes the Laplacian submatrix of the augmented graph corresponding to the unlabeled samples;  $\mathbf{\Gamma}_{uu}^+$  is its inverse; and  $\mathbf{L}_{us} = -\mathbf{W}_{us} = -w \mathbf{Z}_{us}$ .

In the rest of this section, we will show how to efficiently compute  $\mathbf{\Gamma}_{uu}^+$  and  $\mathbf{Y}_u^+$ . The diagonal submatrix of the augmented graph corresponding to the unlabeled samples  $\mathbf{D}_{uu}^+$  is computed as

$$\begin{aligned} \mathbf{D}_{uu}^+(i, i) &= \sum_{j=1}^{N_l} \mathbf{W}_{ul}(i, j) + \sum_{j=1}^{N_u} \mathbf{W}_{uu}(i, j) + \sum_{j=1}^{N_s} \mathbf{W}_{us}(i, j) \\ &= \mathbf{D}_{uu}(i, i) + w(\mathbf{Z}_{us} \mathbf{Z}_{su})(i, i). \end{aligned} \quad (34)$$

We can easily prove that  $\sum_{j=1}^{N_s} \mathbf{Z}_{us}(i, j) = (\mathbf{Z}_{us} \mathbf{Z}_{su})(i, i)$  with the fact that  $\mathbf{Z}_{us}$  is a binary indicator matrix. Note that  $\mathbf{Z}_{us} \mathbf{Z}_{su}$  is a diagonal matrix.

Then,  $\mathbf{L}_{uu}^+$  can be computed as

$$\begin{aligned} \mathbf{L}_{uu}^+ &= \mathbf{D}_{uu}^+ - \mathbf{W}_{uu}^+ \\ &= \mathbf{L}_{uu} + w \mathbf{Z}_{us} \mathbf{Z}_{su}, \end{aligned} \quad (35)$$

and accordingly  $\mathbf{\Gamma}_{uu}^+$  can be calculated using the Binomial inverse theorem as

$$\begin{aligned} \mathbf{\Gamma}_{uu}^+ &= (\mathbf{L}_{uu}^+)^{-1}, \\ &= (\mathbf{L}_{uu} + w \mathbf{Z}_{us} \mathbf{Z}_{su})^{-1}, \\ &= \mathbf{L}_{uu}^{-1} - w \mathbf{L}_{uu}^{-1} \mathbf{Z}_{us} (\mathbf{I}_{N_s} + w \mathbf{Z}_{su} \mathbf{L}_{uu}^{-1} \mathbf{Z}_{us})^{-1} \mathbf{Z}_{su} \mathbf{L}_{uu}^{-1} \\ &= \mathbf{\Gamma}_{uu} - \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{ku} \end{aligned} \quad (36)$$

where  $\mathbf{I}_{N_s}$  is an  $N_s \times N_s$  identity matrix, and  $\mathbf{\Gamma}_{uk}$ ,  $\mathbf{\Gamma}_{kk}$ , and  $\mathbf{\Gamma}_{ku}$  are submatrices of  $\mathbf{\Gamma}_{uu}$  such that  $\mathbf{\Gamma}_{uk} = \mathbf{\Gamma}_{uu} \mathbf{Z}_{us} = [\mathbf{\Gamma}_{uu}(i, j)]$ ,  $\forall i \in \mathcal{U}, j \in \mathcal{K}$ ;  $\mathbf{\Gamma}_{kk} = [\mathbf{\Gamma}_{uu}(i, j)]$ ,  $\forall i \in \mathcal{K}, j \in \mathcal{K}$ ; and  $\mathbf{\Gamma}_{ku} = \mathbf{Z}_{su} \mathbf{\Gamma}_{uu} = [\mathbf{\Gamma}_{uu}(i, j)]$ ,  $\forall i \in \mathcal{K}, j \in \mathcal{U}$  where  $\mathcal{K}$  is a set of selected samples among  $\mathcal{U}$  for human verification.

By substituting (36) into (33), we obtain

$$\begin{aligned} \mathbf{Y}_u^+ &= -\mathbf{\Gamma}_{uu}^+ (\mathbf{L}_{ul} \mathbf{Y}_l - w \mathbf{Z}_{us} \mathbf{Y}_s) \\ &= -\left( \mathbf{\Gamma}_{uu} - \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{ku} \right) \mathbf{L}_{ul} \mathbf{Y}_l \\ &\quad + w \left( \mathbf{\Gamma}_{uu} - \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{ku} \right) \mathbf{Z}_{us} \mathbf{Y}_s. \end{aligned} \quad (37)$$

Since  $w \rightarrow +\infty$ , the first term in (37) can be computed as

$$\begin{aligned} & - \left( \mathbf{\Gamma}_{uu} - \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{ku} \right) \mathbf{L}_{ul} \mathbf{Y}_l, \\ &= -\mathbf{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{Y}_l + \mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1} \mathbf{\Gamma}_{ku} \mathbf{L}_{ul} \mathbf{Y}_l, \\ &= -\mathbf{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{Y}_l + \mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1} \mathbf{Z}_{su} \mathbf{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{Y}_l (\because \mathbf{\Gamma}_{ku} = \mathbf{Z}_{su} \mathbf{\Gamma}_{uu}), \\ &= \mathbf{Y}_u - \mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1} \mathbf{Z}_{su} \mathbf{Y}_u (\because -\mathbf{\Gamma}_{uu} \mathbf{L}_{ul} \mathbf{Y}_l = \mathbf{Y}_u \text{ in Eq (3)}), \\ &= \mathbf{Y}_u - \mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1} \mathbf{Y}_k. \end{aligned} \quad (38)$$

The second term in (37) can be calculated as

$$\begin{aligned} & w \left( \mathbf{\Gamma}_{uu} - \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{ku} \right) \mathbf{Z}_{us} \mathbf{Y}_s \\ &= w \mathbf{\Gamma}_{uu} \mathbf{Z}_{us} \mathbf{Y}_s - w \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{ku} \mathbf{Z}_{us} \mathbf{Y}_s \\ &= w \mathbf{\Gamma}_{uk} \mathbf{Y}_s - w \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{kk} \mathbf{Y}_s \\ &= w \mathbf{\Gamma}_{uk} \left( \mathbf{I}_{N_s} - (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{kk} \right) \mathbf{Y}_s \\ &= w \mathbf{\Gamma}_{uk} \left( (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk}) - (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{\Gamma}_{kk} \right) \mathbf{Y}_s \\ &= w \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk} - \mathbf{\Gamma}_{kk}) \mathbf{Y}_s \\ &= \mathbf{\Gamma}_{uk} (\mathbf{I}_{N_s} / w + \mathbf{\Gamma}_{kk})^{-1} \mathbf{Y}_s \\ &= \mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1} \mathbf{Y}_s. \quad (\because w \rightarrow +\infty) \end{aligned} \quad (39)$$

Substituting (38) and (39) into (37) yields

$$\begin{aligned} \mathbf{Y}_u^+ &= \mathbf{Y}_u - \mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1} \mathbf{Y}_k + \mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1} \mathbf{Y}_s \\ &= \mathbf{Y}_u + \mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1} (\mathbf{Y}_s - \mathbf{Y}_k), \end{aligned} \quad (40)$$

where  $\mathbf{Y}_k$  is a submatrix of  $\mathbf{Y}_u$  that is constructed by stacking the rows of  $\mathbf{Y}_u$  that correspond to queried samples.

We denote  $\mathbf{\Gamma}_{uk} \mathbf{\Gamma}_{kk}^{-1}$  as the *verification propagation matrix*. By propagating the human verification information through this matrix, the segmentation results can be incrementally improved. Note that we compute the inverse of  $\mathbf{\Gamma}_{kk}$  which is an  $N_s \times N_s$  matrix, rather than  $\mathbf{L}_{uu}^+$  which is an  $N_u \times N_u$  matrix. As a result, verification propagation can be efficiently performed, leading to the efficient interactive segmentation.

Since the parameter  $K$  is essential, we select an empirical  $K$  which should be efficient and user-friendly in this paper. On the one hand, if  $K$  is set to be small, e.g., only one query is selected at a time as in [44] and the classifier has to be retrained accordingly, the user has to wait until the classifier is updated which is apparently very slow and making the method impractical.

On the other hand, if  $K$  is set to be large, the user has to verify/correct too many instances simultaneously, which is not user-friendly. Moreover, it will also require user to correct some analogous misclassification, which is inefficient and should be fixed with correction propagation.

#### IV. EXPERIMENTAL RESULTS

We validate our approach on three time-lapse phase contrast microscopy images that capture different types of cells. The detailed specifications of each sequence is illustrated in Table I.

TABLE I  
IMAGE DATA SPECIFICATIONS

	cell type	# of frames	# of cells per image	# of manually labeled cells	resolution	cell stages
Seq1	muscle stem cell of a progeroid mouse	532	50+ to 300+	1459	696 × 520	intermitosis/mitosis
Seq2	bovine aortic endothelial cell	489	500+ to 800+	7020	1040 × 1392	intermitosis/mitosis
Seq3	C2C12 myoblastic stem cell	383	300+	5830	1040 × 1392	intermitosis/apoptosis(dead)

From each sequence, we select image frames with a certain time interval and manually label the cells in the frames. The number of cells in each frame varies from 50 to 800.<sup>1</sup> Although the number of images is relatively small, each image contains a large number of cells (from tens to hundreds) yielding more than  $10^4$  cells with more than  $10^7$  cell pixels in total. Moreover, we annotated various cells undergoing different cell stages, e.g., inter-mitosis, mitosis, apoptosis. In this case, the results obtained in our experiments are therefore statistically significant. In this section, we show the visual results of cell segmentation based on active annotation and verification propagation algorithm, and also quantitatively evaluate the performance in terms of rate of convergence and Tanimoto coefficient, which is defined as

$$TC = \frac{\text{area}(D_{dt} \cap D_{gt}) + \text{area}(B_{dt} \cap B_{gt})}{\text{area}(D_{dt} \cup D_{gt}) + \text{area}(B_{dt} \cup B_{gt})} \quad (41)$$

where  $D_{dt}$  and  $D_{gt}$  are the detected and annotated dark cell regions, respectively, and  $B_{dt}$  and  $B_{gt}$  are the detected and annotated bright cell regions, respectively.

#### A. Cell Segmentation Based on Active Annotation

1) *Qualitative Results*: Fig. 4 shows sample results based on active annotation, which demonstrates that high quality cell segmentation is realized with only a small amount of human annotations. In order to eliminate the local redundancy of an image and reduce the computational cost, we first partition the phase contrast microscopy images into phase-homogeneous superpixels by aggregating the neighboring pixels with similar phase retardation features [2], which are used as elementary units for cell segmentation and shown in column (c). In this paper, we investigate the variance of intra-superpixel features to determine the number of superpixels to partition as in [2]. It is desirable to end up with a small number of superpixels to reduce the complexity of classification problem while preserve the local structures, i.e., visual features in a superpixel should remain homogeneous.

Afterwards, we actively select the most informative phase-homogeneous superpixels by minimizing the prediction error in Algorithm 1, and annotate them as seeds for the subsequent label propagation. In column (c), we mark the superpixels as red, green and blue regions if the corresponding superpixels are annotated as bright cells, dark cells and background, respectively. Column (d) shows the soft segmentation results by propagating the human annotation to the unlabeled superpixels, in which the

red, green and blue channel values indicate the likelihoods of the superpixel belonging to bright cell, dark cell and background regions, respectively. Each soft segmentation result was converted into a hard segmentation result by finding the label with the maximum likelihood, and grouping neighboring superpixels with the same labels, as is shown in column (e). As the results demonstrate, high quality segmentation is realized based on the informative superpixel annotation and label propagation. As a comparison, we also implemented cell segmentation based on unsupervised spectral clustering [46], as is shown in column (f). The results demonstrate that there exist false cells detections for some challenging cases, as marked with yellow circles. Halos are recognized as cell regions in Seq1 when cells form as clusters, and some bright cells with low contrast to the background are also missed in Seq2. It is observed that some bright cells or bright regions in apoptotic cells are also missed in Seq3, which is due to the inconsistency of phase retardation feature when cells undergo apoptotic process [47]. Oppositely, the results are improved by incorporating guided human annotations to handle the challenging cases.

2) *Quantitatively Evaluation*: In this section, we quantitatively evaluate the performance against the unsupervised spectral clustering [46] in terms of TC, which is reported in Table II. The results demonstrate that our proposed algorithm outperforms the unsupervised spectral clustering, since human interventions can indeed facilitate cell segmentation to handle challenging cases when the characteristic of visual features is complex, e.g., cells form clusters, or the inconsistency of visual characteristics within apoptotic cells. Additionally, it is not easy to determine the optimal cluster number in advance for the unsupervised spectral clustering which greatly affects the final results.

Obviously, the supervised method is time-consuming or tedious since a user has to annotate lots of cell regions manually. In this case, we propose to implement cell segmentation by propagating human annotation to the unlabeled regions. In addition to comparing with unsupervised clustering, we also compare our active annotation with random annotation in [2] which randomly selects and annotates superpixels as seeds for label propagation. Fig. 5 demonstrates the Tanimoto coefficient curves with respect to the percentage of different annotation samples among all the superpixels, in which the curves are averaged over 10 trials. As is seen and expected, TC generally increases as more labeled information is provided, and our proposed active annotation algorithm converges to high quality results much more quickly than the random annotation. The main reason is because the most informative superpixels are actively selected early and annotated first. In this case, human efforts are

<sup>1</sup>We build a large dataset of phase contrast microscopy images to evaluate our algorithm, which is a public benchmark. All data can be downloaded at <http://www.celltracking.ricmu.edu/downloads.html>.

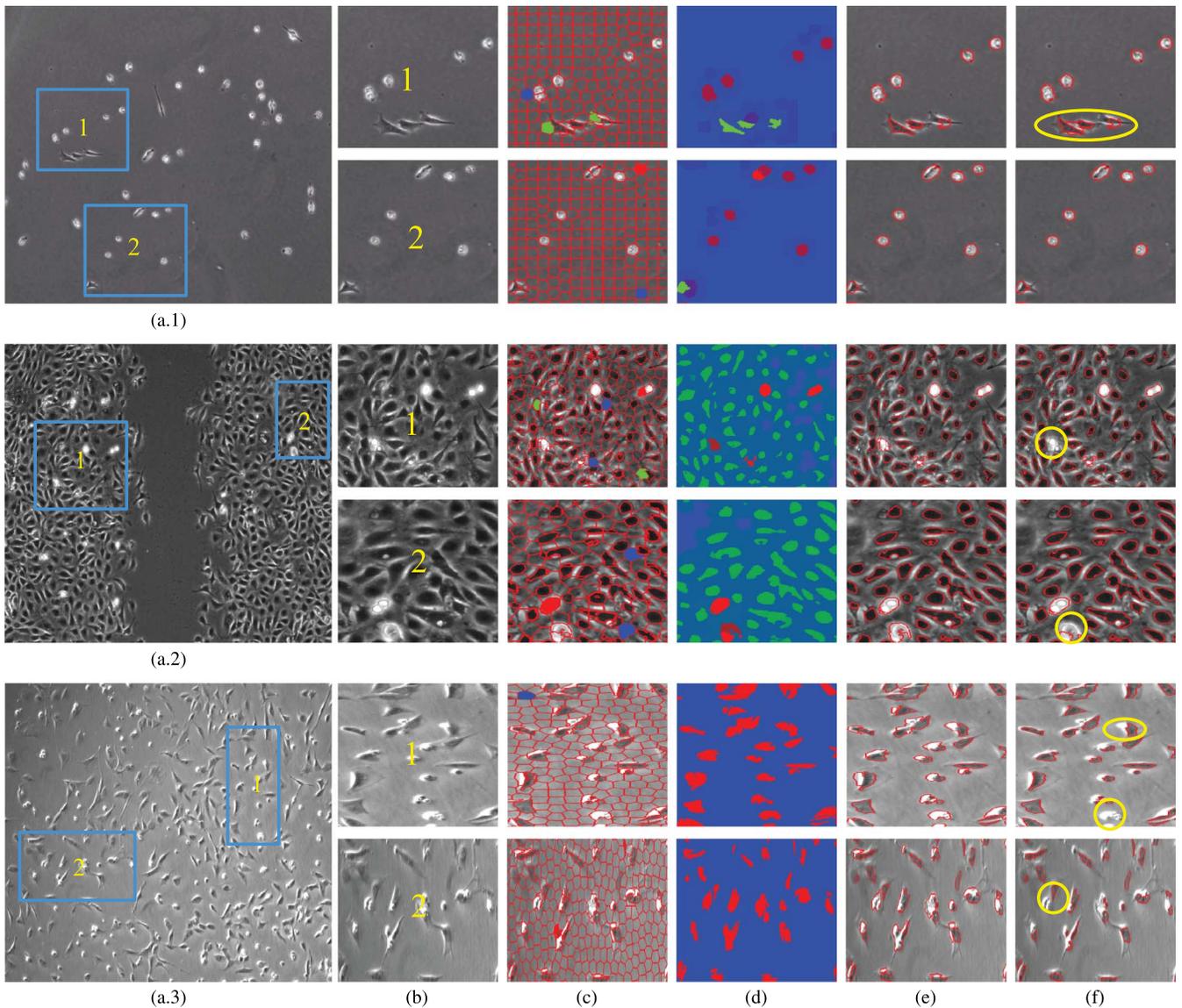


Fig. 4. Sample results of cell segmentation based on active annotation in Seq1 (top), Seq2 (middle), and Seq3 (bottom). (a) Input phase contrast microscopy images; (b) Zoom-in sub-images; (c) Active sample selection and annotation over the phase-homogeneous superpixels, where the red, green and blue annotations indicate the corresponding superpixels are annotated as bright cells, dark cells, and background, respectively; (d) Soft classification results based on label propagation with human annotations. The red, green and blue channel values in the soft classification results indicate the likelihoods of the superpixels belonging to bright cell, dark cell, and background regions, respectively. (e) Cell segmentation by finding the labels with the maximum likelihood, and grouping the neighboring superpixels with the same labels. (f) Cell segmentation based on unsupervised spectral clustering [46] as comparison.

TABLE II  
COMPARISON WITH UNSUPERVISED CLUSTERING

	Seq1	Seq2	Seq3
Active Annotation	0.9701	0.9597	0.9813
Unsupervised Spectral Clustering [46]	0.9266	0.9301	0.9012

reduced significantly by guiding human interventions with active paradigm.

**B. Cell Segmentation Based on Active Verification Propagation**

In practice, a classifier learned from initial labeling tends to result in more and more misclassifications when the datasets expand over time with more and more unseen data. However, it is expensive to rebuild a classifier from scratch using newly collected training data, because it requires a lot of human annotations. It would be nice to re-use the previous labeled data and

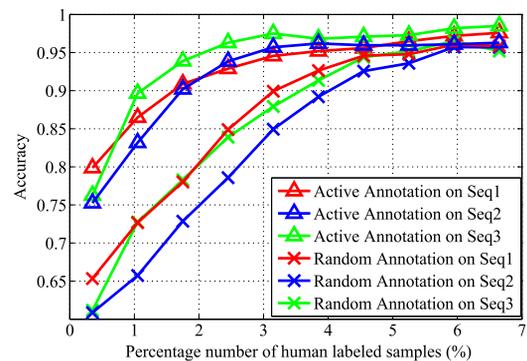


Fig. 5. Comparison with random annotation [2].

incorporate subsequent human verifications or corrections. In order to validate our proposed algorithm, we actively annotate

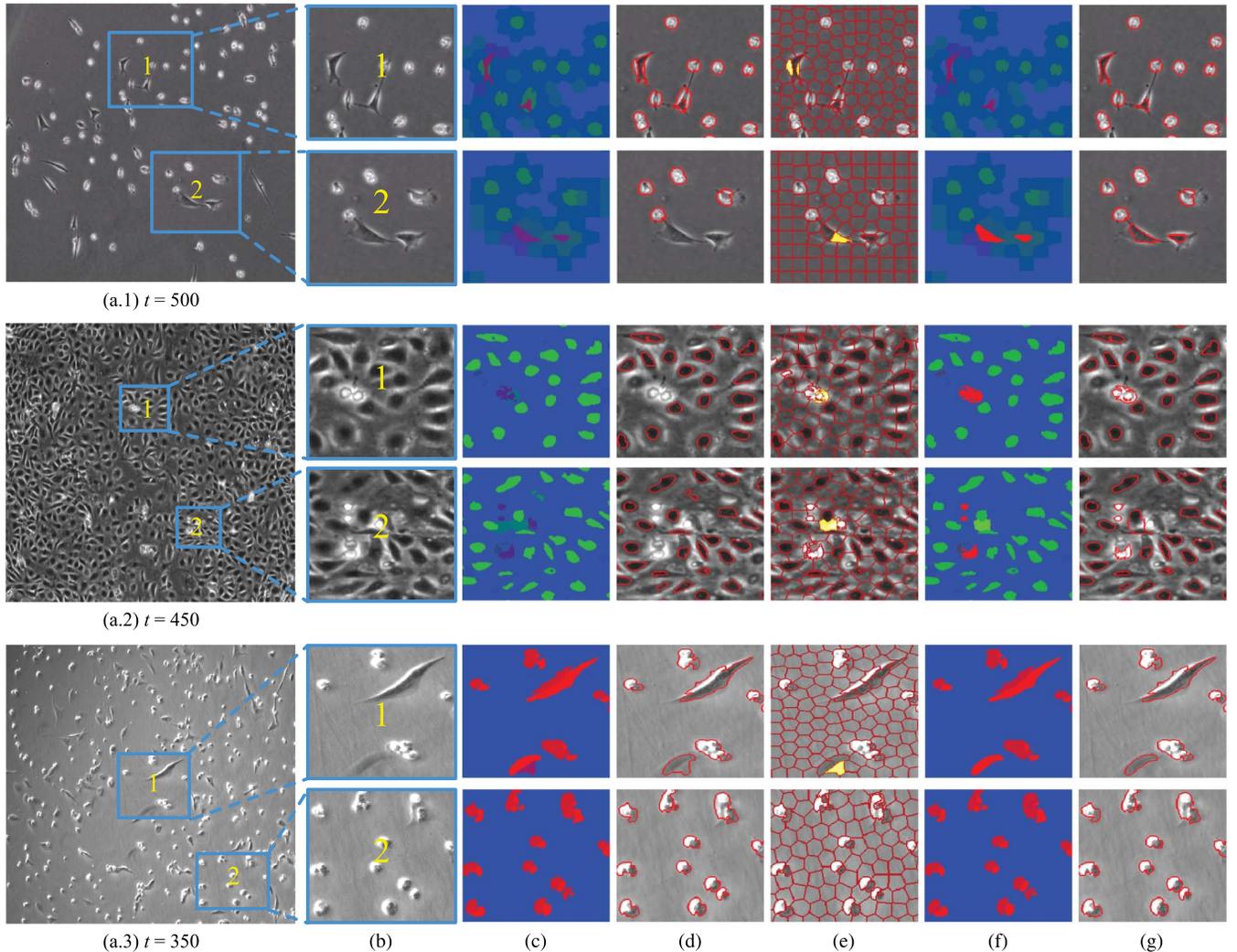


Fig. 6. Erroneous segmentation fixing examples with verification propagation on three images from Seq1 (top), Seq2 (middle), and Seq3 (bottom). (a) Input phase contrast microscopy images; (b) Zoom-in sub-imgs; (c) Soft segmentation based on label propagation using annotation acquired in the first frame; (d) Hard based on the corresponding soft segmentation; (e) Actively selected superpixels for verification marked by yellow color. (f) Improved soft segmentation after verification propagation; (g) Improved hard segmentation based on the improved soft segmentation.

the most informative superpixels in the first frame of each sequence, which are used as seeds during label propagation. Afterwards, we implement cell segmentation in each image of a sequence, and examine the results frame-by-frame. Once the reliability of segmentation is below a threshold or obvious erroneous segmentations are detected, the system queries for active verification actively. Then, the user will examine the results, and correct misclassifications by clicking the corresponding regions with a mouse. Human correction will be propagated to the unlabeled samples to fix analogous misclassifications.

1) *Qualitative Results*: In the following, we conduct image segmentation by propagating the labels obtained in the first image of each sequence to the sub-sequential frames. Although the images are acquired at different moments, the annotated samples can be reused for others since the cells show similar visual properties and thus with similar phase retardations. Nevertheless, there also exist cells with different phase retardation features from the initial annotations due to proliferation processes of live cells, and thus may lead to erroneous segmentation. In this case, we fix these misclassifications based on our

proposed verification propagation method. In this paper, we set  $K = 4$  which is an empirically efficient and user-friendly in practice.

Examples of the cell segmentation results for the subsequent frames are shown in Fig. 6. In columns (c) and (d), we demonstrate the soft and hard segmentation results based on label propagation of the initial human annotation from the first image. As can be seen in column (d), there contain several errors in the segmentation result, i.e., some dark cells are missed and bright halos are misclassified into cells in Seq1; bright cells with low contrast are missed in Seq2; and some dark-adjacent superpixels are classified into dark cells in Seq3. These errors are ascribed to inadequate or unbalanced human labels, since the phase retardation features of some cells are different from the first frame in the sequence. Some of the superpixels containing errors were automatically selected by our algorithm to be verified and corrected by human, as is shown in column (e). In columns (f) and (g), we demonstrate the improved soft and hard segmentation results after human correction and its propagation. As shown in column (g), in addition to the errors of sub-

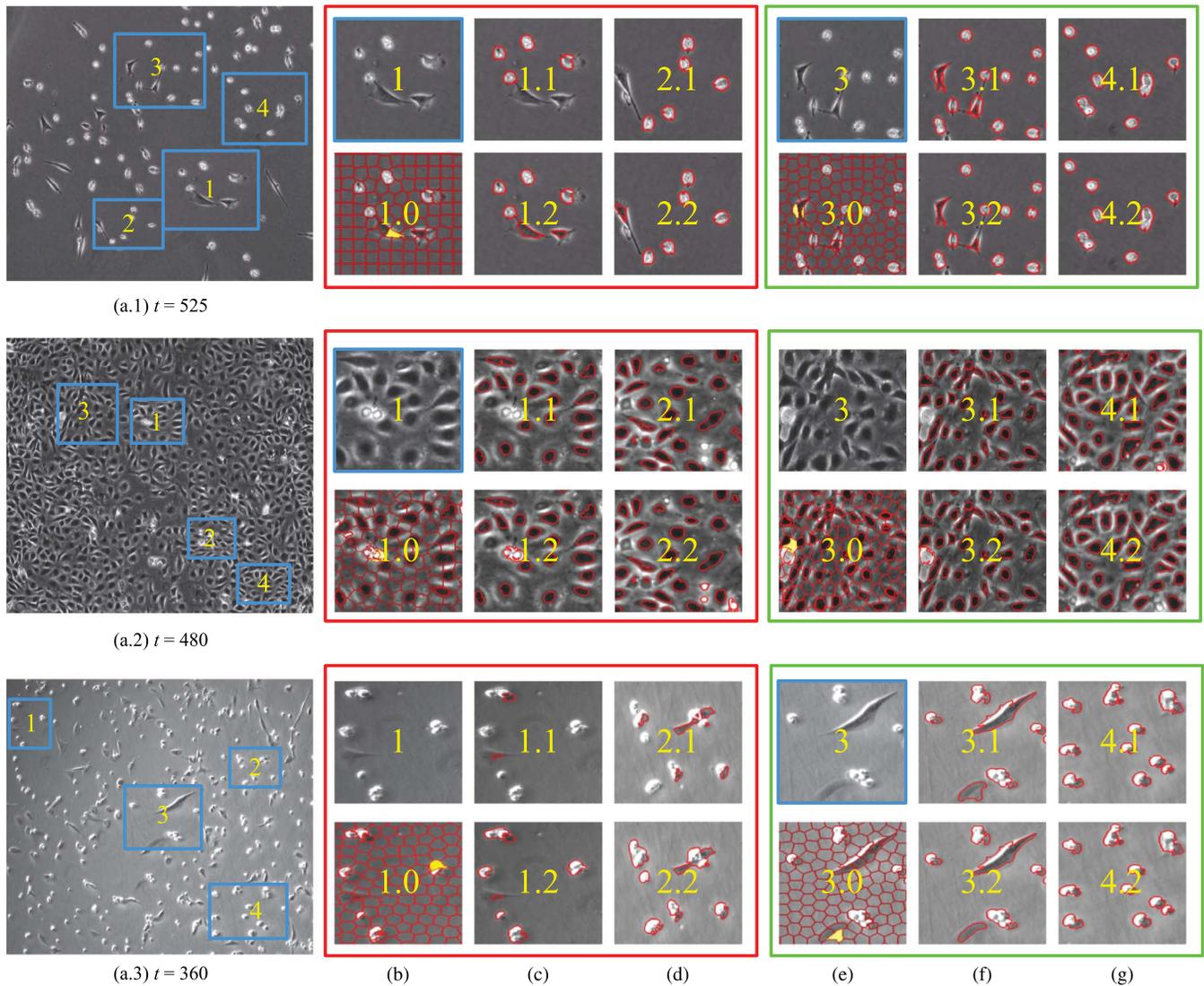


Fig. 7. Sample results of verification propagation on three images from Seq1 (top), Seq2 (middle), and Seq3 (bottom). (a) Input phase contrast microscopy images. (b, c, d) The sub-images in red rectangles are corresponding to the zoom-in images that are corrected after the first verification propagation. Fig. (1) Zoom-in sub-images; (1.0) Actively selected superpixels for verification marked by yellow color; (1.1) Segmentation result based on initial label propagation; (1.2) Improved cell segmentation after the 1st verification propagation iteration; (2.1) and (2.2) are corresponding to the initial cell segmentation and the improved results, which demonstrates that similar misclassification are corrected via verification propagation. (e, f, g) Similarly, sub-images in the green rectangles are corresponding to results based on human verification (3.0, 3.1, 3.2) and verification propagation (4.1, 4.2). The results in Fig. 4(a) and Fig. 4(b) demonstrate that segmentation results are not influenced by subsequent human correction if there is no misclassification.

perpixels selected by active superpixel selection, similar errors of other superpixels are also effectively fixed. It is also noted that if there is no misclassification, the previous label propagation results are not influenced by subsequent human correction, as is shown in columns (d) and (g) corresponding to the sub-image (2) of Fig. 6(a).

In Fig. 7, we demonstrate the step-by-step results of active verification propagation procedure, in which the sub-images in the red and green rectangles are corresponding visual results after the first and second verification propagation procedures. After human verifications in Fig. 7(b), the erroneous segmentations in Fig. 7(c) are corrected in Fig. 7(c). Simultaneously, the analogous misclassifications in Fig. 7(d) are corrected by propagating the human verification. Similarly, it is observed that erroneous segmentations in Fig. 7(f) are corrected after human verification in Fig. 7(e), resulting in an improved result in Fig. 7(f).

However, the valid results in Fig. 7(g) are not influenced by human verification, as is shown in Fig. 7(g).

2) *Quantitative Evaluation*: To verify the effectiveness of our active verification propagation in cell segmentation, we compare our method against alternative learning algorithms quantitatively. First, we implemented two cell segmentation algorithms solely based on the initially annotated samples without subsequent human verification:

- **Random annotation and no verification** [2]: Random annotation performs uniform sampling on the first frame of each sequence to select samples for initial labeling. Cell segmentation is implemented by (3) without the verification step, which acts as the baseline method.
- **Active annotation and no verification**: We apply the same label propagation algorithm by using the most informative samples as seeds that are drawn actively in

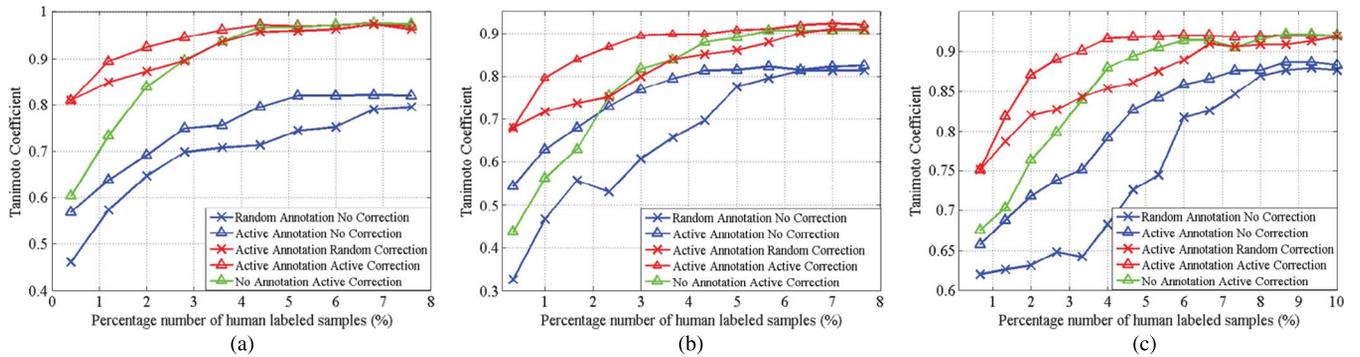


Fig. 8. Comparison between different segmentation algorithms. Figs. 8(a), (b) and (c) are corresponding to the results of Seq1, Seq2 and Seq3, respectively. The blue, red and green curves are corresponding to the schemes of annotation without verification, annotation along with verification and re-annotation on the unseen frames, respectively. The curves with triangle and cross markers are corresponding to the active sample selection and random sample selection scheme, respectively.

Algorithm 1 in the first frame, but without any human verification.

We also implemented two algorithms that apply the verification mechanism on top of active annotation at the beginning:

- **Active annotation and random verification:** We select a portion of the most informative samples via the active annotation algorithm in the first frame of each sequence as initial labeling, then randomly select samples for human verification in each sub-sequential frame. Manual interventions are propagated to the unlabeled samples following (40).
- **Active annotation and active verification:** After obtaining a portion of samples via active annotation, we actively check the results based on our proposed sample selection and verification propagation algorithms.

Additionally, we implement the segmentation task by ignoring the labeled data in the first frame as

- **Active re-annotation:** We re-annotated each frame actively, and then utilize them to conduct cell segmentation during the label propagation.

In order to reduce the bias, the result is averaged over 10 trials on the entire dataset by randomly selecting samples, either applied during annotation or correction.

The quantitative evaluation is reported in Fig. 8. As the figure shows, no matter whether the initially annotated samples are selected randomly or actively, cell segmentation without correction (blue curves) is not comparable to the paradigm when subsequent correction is involved. The main reason is that the characteristics or phase retardation features of cells change gradually over time due to the mitosis and apoptosis of cells or interaction of them. If human verification and correction are performed after the label propagation from the initial annotation, the performance is improved greatly as more information is provided by a user, since cells with different features that are not included in the first frame are identified, and well segmented. The human examination of the first 2% of samples results in more than ten percentage points of accuracy improvement, as shown in the very early stage of human correction (beginning of red curves with triangle markers). This implies that the samples initially selected have typical errors, so correction on them can fix a lot of similar cases, thereby significantly reducing human efforts in refining the results. In particular, verification guided

by active sample selection converges more quickly than random verification, since more error-prone samples are selected early and corrected first.

As the experiments show, high quality segmentation can be obtained via both annotation and verification. One may be interested in that why we need verification since we could annotate each frame in a sequence. We implement the scheme by ignoring the previous human labeling and re-annotating each frame (Active Re-Annotation in Fig. 7). As is observed, the performance converges to a comparable performance with active annotation and active verification eventually. However, human need to annotate more samples to achieve high quality results since it does not leverage the information from the previous labeled examples. Therefore, re-using the annotated sample reduces human efforts in annotation or verification greatly.

Base on the aforementioned results, we propose to actively annotate the most informative superpixels in the first frame and re-use them as seeds during label propagation for the subsequent frames. Once false segmentation is detected, we proposed to select the error-prone superpixels automatically, and propagate the human verification to the remaining superpixels to fix analogous misclassifications for the subsequent frames in each sequence. The scheme is effective and more efficient compared with alternative algorithms in practice.

*Cell-Wise Evaluation:* The pixel-wise evaluation based on Tanimoto coefficient is useful for some biomedical applications, e.g., cellular structure and morphology analysis. In lineage analysis, users are only interested in the centroid of cells to conduct cell tracking, in which a cell-wise evaluation is preferable. In the following, we also include a cell-wise evaluation to demonstrate the effectiveness of the proposed algorithm for cell tracking as

$$acc = \frac{N_{suc}}{N_{cell}}, \quad (42)$$

where  $N_{cell}$  is the total number of cells, and  $N_{suc}$  is the successful detected cells, in which the distance of centroid between the detected cells and ground truth satisfy  $d < th$ , as is shown in Fig. 9. The quantitative evaluation is demonstrated in Table III, which demonstrates that more than 98% cells are detected successfully with proper human verification harnessed, which can be applied for lineage analysis.

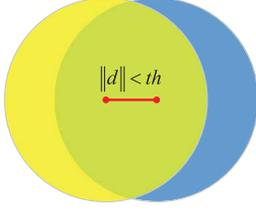


Fig. 9. Illustration for cell-wise evaluation. A segmentation is defined as a successful detection when the distance of centroid between the detected cells (blue circle) and ground truth (yellow circle) is below threshold  $th$ .

TABLE III  
CELL-WISE EVALUATION BASED ON VERIFICATION PROPAGATION

	Seq1	Seq2	Seq3
With Correction	0.9931	0.9907	0.9893
Without Correction	0.9036	0.8901	0.9127

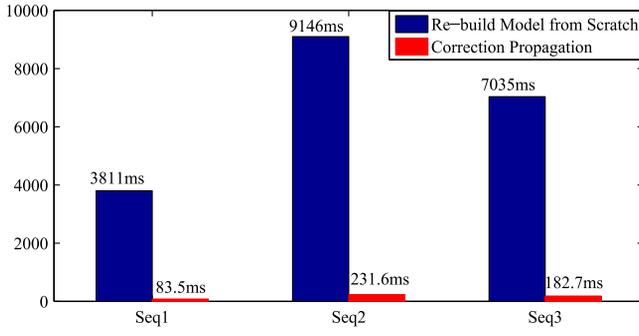


Fig. 10. Time cost evaluation.

*Time Cost Analysis:* The bottleneck of time cost in label propagation is the inverse matrix computation with a rough time complexity  $O(N^3)$ . For building everything from scratch, we need to compute the inverse of a huge matrix, the size of which is the number of unlabeled samples (1000 to 10000 in our experiments), whereas our incremental correction propagation only computes the inverse of a small matrix, the size of which is the number of queries (approximately 10 in our experiments). All experiments are implemented in MATLAB R2014 and run on a 3.2 GHz, 8 GB RAM Core 2Duo PC. We show the result is illustrated in Fig. 10, which demonstrates the computation cost saving is orders of magnitude.

## V. CONCLUSION

In this paper, we propose an interactive cell segmentation algorithm based on active annotation and verification propagation. After partitioning the frames into feature-homogeneous superpixels, we actively select the most informative samples that minimize the expected prediction error on the unlabeled superpixels. After performing the initial classification through label propagation, we actively check the most error-prone superpixels and correct the false segmentation. The correction is propagated to the remaining unlabeled samples through our proposed verification propagation method, which is efficient since it does not involve reconstruction of the affinity graph, resulting in effective corrections on similar errors. Experimental results demon-

strate that high quality cell segmentation is realized via the proposed active annotation and verification propagation algorithms with less human efforts.

## APPENDIX A

### DERIVATION OF SAMPLE SELECTION IN (28)

*Solution:* Based on the update of the corresponding matrix of  $\mathbf{R}_k$  in (27), we have

$$\mathbf{R}_k^T \mathbf{R}_k = \mathbf{R}_{k-1}^T \mathbf{R}_{k-1} - \mathbf{q}_k \mathbf{q}_k^T, \quad (43)$$

where  $\mathbf{q}_k$  is the column vector, which is obtained by transposing the row vector in  $\mathbf{Q}$  corresponding to the samples selected in the  $k$ th iteration. By denoting  $\mathbf{A}_{k-1} = \mathbf{R}_{k-1}^T \mathbf{R}_{k-1}$  and  $\mathbf{A}_k = \mathbf{A}_{k-1} - \mathbf{q}_k \mathbf{q}_k^T$ , and substituting the corresponding items in (25), we have

$$\begin{aligned} & \text{tr} \left( \mathbf{\Lambda}^2 \mathbf{R}_k^T \mathbf{R}_k \right) + 2\text{tr} \left( \mathbf{\Phi}_k \mathbf{R}_k^T \mathbf{R}_k \mathbf{\Lambda}^2 \mathbf{R}_k^T \mathbf{R}_k \right) \\ & + \text{tr} \left( \left( \mathbf{\Phi}_k \mathbf{R}_k^T \mathbf{R}_k \right)^2 \mathbf{\Lambda}^2 \mathbf{R}_k^T \mathbf{R}_k \right) \\ & = \text{tr} \left( \mathbf{\Lambda}^2 \mathbf{A}_k \right) + \text{tr} \left( \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_k \mathbf{\Lambda}^2 \mathbf{A}_k \right) \\ & = \text{tr} \left( \mathbf{\Lambda}^2 \left( \mathbf{A}_{k-1} - \mathbf{q}_k \mathbf{q}_k^T \right) \right) \\ & + \text{tr} \left( \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_k \mathbf{\Lambda}^2 \left( \mathbf{A}_{k-1} - \mathbf{q}_k \mathbf{q}_k^T \right) \right) \\ & = \text{tr} \left( \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \right) - \mathbf{q}_k^T \mathbf{\Lambda}^2 \mathbf{q}_k - \mathbf{q}_k^T \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_k \mathbf{\Lambda}^2 \mathbf{q}_k \\ & + \text{tr} \left( \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_k \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \right) \\ & \triangleq \text{tr} \left( \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \right) - \mathbf{q}_k^T \mathbf{\Lambda}^2 \mathbf{q}_k - \mathbf{q}_k^T \mathbf{B}_k \mathbf{q}_k \\ & + \text{tr} \left( \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_k \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \right), \end{aligned} \quad (44)$$

where  $\mathbf{B}_k = \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_k \mathbf{\Lambda}^2$ . Note that the first terms is from the  $(k-1)$ th iteration, which is a constant when solving the optimization regarding to  $\mathbf{q}_k$ , and therefore we focus on the rest terms. Using (43), the fourth term in (44) can be rewritten as

$$\begin{aligned} & \text{tr} \left( \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_k \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \right) \\ & = \text{tr} \left( \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_k \right) \\ & = \text{tr} \left( \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \left( \mathbf{A}_{k-1} - \mathbf{q}_k \mathbf{q}_k^T \right) \right) \\ & = -\mathbf{q}_k^T \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{q}_k \\ & + \text{tr} \left( \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_{k-1} \right) \\ & \triangleq -\mathbf{q}_k^T \mathbf{C}_k \mathbf{q}_k + \text{tr} \left( \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right) \mathbf{A}_{k-1} \right), \end{aligned} \quad (45)$$

where  $\mathbf{C}_k = \mathbf{\Lambda}^2 \mathbf{A}_{k-1} \left( 2\mathbf{\Phi}_k + \mathbf{\Phi}_k \mathbf{A}_k \mathbf{\Phi}_k \right)$ . Using the

$$\begin{aligned} \mathbf{\Phi}_k & = \left( \left( \mathbf{\Lambda} + \mathbf{I} \right)^{-1} - \mathbf{R}_k^T \mathbf{R}_k \right)^{-1} \\ & = \left( \left( \mathbf{\Lambda} + \mathbf{I} \right)^{-1} - \mathbf{R}_{k-1}^T \mathbf{R}_{k-1} + \mathbf{q}_k \mathbf{q}_k^T \right)^{-1} \\ & = \left( \mathbf{\Phi}_{k-1}^{-1} + \mathbf{q}_k \mathbf{q}_k^T \right)^{-1} \\ & = \mathbf{\Phi}_{k-1} - \frac{\mathbf{\Phi}_{k-1} \mathbf{q}_k \mathbf{q}_k^T \mathbf{\Phi}_{k-1}}{1 + \mathbf{q}_k^T \mathbf{\Phi}_{k-1} \mathbf{q}_k}. \end{aligned} \quad (46)$$

By substituting (46) into (47), the second term in (45) can be and rewritten as

$$\begin{aligned}
& \text{tr} \left( \Lambda^2 \mathbf{A}_{k-1} (2\Phi_k + \Phi_k \mathbf{A}_k \Phi_k) \mathbf{A}_{k-1} \right) \\
&= \text{tr} \left( \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} (2\mathbf{I} + \Phi_k \mathbf{A}_k) \Phi_k \right) \\
&= 2\text{tr}(\mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1}) \\
&\quad - \frac{2\mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \\
&\quad + \text{tr}(\mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k \mathbf{A}_k \Phi_k) \\
&\triangleq 2\text{tr}(\mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1}) - \frac{2}{c_k} \mathbf{q}_k^T \mathbf{D}_k \mathbf{q}_k \\
&\quad + \text{tr}(\mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k \mathbf{A}_k \Phi_k), \tag{47}
\end{aligned}$$

where  $\mathbf{D}_k = \Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1}$ , and  $c_k = 1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k$ . We omit the first term since it is from the  $(k-1)$ th iteration and is a regarding.

The third term in (47) can be rewritten as

$$\begin{aligned}
& \text{tr}(\mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k \mathbf{A}_k \Phi_k) \\
&= \text{tr} \left( \Phi_k \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k (\mathbf{A}_{k-1} - \mathbf{q}_k \mathbf{q}_k^T) \right) \\
&= -\mathbf{q}_k^T \Phi_k \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k \mathbf{q}_k \\
&\quad + \text{tr}(\Phi_k \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k \mathbf{A}_{k-1}) \\
&\triangleq -\mathbf{q}_k^T \mathbf{E}_k \mathbf{q}_k + \text{tr}(\Phi_k \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k \mathbf{A}_{k-1}), \tag{48}
\end{aligned}$$

where  $\mathbf{E}_k = \Phi_k \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k$ .

Using (46), the second term in (48) can be rewritten as

$$\begin{aligned}
& \text{tr}(\Phi_k \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k \mathbf{A}_{k-1}) \\
&= \text{tr} \left( \left( \Phi_{k-1} - \frac{\Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \right) \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \right) \\
&\quad \left( \Phi_{k-1} - \frac{\Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \right) \mathbf{A}_{k-1} \\
&= \text{tr}(\Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{A}_{k-1}) \\
&\quad - \text{tr} \left( \frac{\Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{A}_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \right) \\
&\quad - \text{tr} \left( \frac{\Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \right) \\
&\quad + \text{tr} \left( \frac{\Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \right) \\
&\quad \times \frac{\Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \mathbf{A}_{k-1}. \tag{49}
\end{aligned}$$

It is easy to prove that both  $\mathbf{A}_{k-1}$  and  $\Phi_{k-1}$  are symmetric matrix, since

$$\begin{aligned}
\mathbf{A}_{k-1}^T &= (\mathbf{R}_{k-1}^T \mathbf{R}_{k-1})^T \\
&= \mathbf{R}_{k-1}^T \mathbf{R}_{k-1} \\
&= \mathbf{A}_{k-1}, \tag{50}
\end{aligned}$$

$$\begin{aligned}
\Phi_{k-1}^T &= \left( ((\Lambda + \mathbf{I})^{-1} - \mathbf{R}_{k-1}^T \mathbf{R}_{k-1})^{-1} \right)^T \\
&= \left( ((\Lambda + \mathbf{I})^{-1} - \mathbf{R}_{k-1}^T \mathbf{R}_{k-1})^T \right)^{-1} \\
&= ((\Lambda + \mathbf{I})^{-1} - \mathbf{R}_{k-1}^T \mathbf{R}_{k-1})^{-1} \\
&= \Phi_{k-1}. \tag{51}
\end{aligned}$$

Using (50) and (51), we can obtain that the second and third terms in (49) are equal, since

$$\begin{aligned}
& \text{tr} \left( \frac{\Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{A}_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \right) \\
&= \frac{\mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \\
&= \frac{\mathbf{q}_k^T \Phi_{k-1}^T \mathbf{A}_{k-1}^T \Phi_{k-1}^T \Phi_{k-1}^T \mathbf{A}_{k-1}^T \Lambda^2 \mathbf{A}_{k-1}^T \Phi_{k-1}^T \mathbf{q}_k}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \\
&= \frac{\mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \\
&= \text{tr} \left( \frac{\Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1}}{1 + \mathbf{q}_k^T \Phi_{k-1} \mathbf{q}_k} \right). \tag{52}
\end{aligned}$$

Therefore, (49) can be rewritten as

$$\begin{aligned}
& \text{tr}(\Phi_k \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_k \mathbf{A}_{k-1}) \\
&= \text{tr}(\Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{A}_{k-1}) \\
&\quad - \frac{2}{c_k} \mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k \\
&\quad + \frac{1}{c_k^2} \mathbf{q}_k^T \Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1} \\
&\quad \times \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k \\
&\triangleq \text{tr}(\Phi_{k-1} \mathbf{A}_{k-1} \Lambda^2 \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{A}_{k-1}) \\
&\quad - \frac{2}{c_k} \mathbf{q}_k^T \mathbf{D}_k \mathbf{A}_{k-1} \Phi_{k-1} \mathbf{q}_k + \frac{1}{c_k^2} \mathbf{q}_k^T \mathbf{D}_k \mathbf{q}_k \mathbf{q}_k^T \mathbf{F}_k \mathbf{q}_k, \tag{53}
\end{aligned}$$

where  $\mathbf{F}_k = \Phi_{k-1} \mathbf{A}_{k-1} \Phi_{k-1}$ .

And thus the (25) can be summed up as

$$\begin{aligned}
& \min_{\mathbf{R}_k} \left\{ \text{tr}(\Lambda^2 \mathbf{R}_k^T \mathbf{R}_k) + 2\text{tr}(\Phi_k \mathbf{R}_k^T \mathbf{R}_k \Lambda^2 \mathbf{R}_k^T \mathbf{R}_k) \right\} \rightarrow \\
& \min_k \left\{ \begin{aligned} & + \text{tr} \left( (\Phi_k \mathbf{R}_k^T \mathbf{R}_k)^2 \Lambda^2 \mathbf{R}_k^T \mathbf{R}_k \right) \\ & - \mathbf{q}_k^T (\Lambda^2 + \mathbf{B}_k + \mathbf{C}_k + \mathbf{E}_k) \mathbf{q}_k \\ & - \frac{2}{c_k} \mathbf{q}_k^T \mathbf{D}_k (\mathbf{I} + \mathbf{A}_{k-1} \Phi_{k-1}) \mathbf{q}_k + \frac{1}{c_k^2} \mathbf{q}_k^T \mathbf{D}_k \mathbf{q}_k \mathbf{q}_k^T \mathbf{F}_k \mathbf{q}_k \end{aligned} \right\}. \tag{54}
\end{aligned}$$

Let  $\mathbf{X}_k = -(\Lambda^2 + \mathbf{B}_k + \mathbf{C}_k + \mathbf{E}_k)$ ,  $\mathbf{Y}_k = -\mathbf{D}_k (\mathbf{I} + \mathbf{A}_{k-1} \Phi_{k-1})$  and  $\mathbf{Z}_k = \mathbf{D}_k \mathbf{q}_k \mathbf{q}_k^T \Phi_{k-1} \Phi_{k-1}$ , the problem can be rewritten as

$$\min_k \left( \mathbf{q}_k^T \mathbf{X}_k \mathbf{q}_k + \frac{2}{c_k} \mathbf{q}_k^T \mathbf{Y}_k \mathbf{q}_k + \frac{1}{c_k^2} \mathbf{q}_k^T \mathbf{D}_k \mathbf{q}_k \mathbf{q}_k^T \mathbf{F}_k \mathbf{q}_k \right). \tag{55}$$

Hereby, we finish the solution.

## REFERENCES

- [1] H. Su, Z. Yin, T. Kanade, and S. Huh, "Interactive cell segmentation based on correction propagation," in *Proc. IEEE Int. Symp. Biomed. Imag.*, Apr. 2014, pp. 712–716.
- [2] H. Su, Z. Yin, S. Huh, and T. Kanade, "Cell segmentation in phase contrast microscopy images via semi-supervised classification over optics-related features," *Med. Image Anal.*, vol. 17, pp. 746–765, 2013.
- [3] C. Rother, V. Kolmogorov, and A. Blake, "Grabcut": Interactive foreground extraction using iterated graph cuts," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 309–314, Aug. 2004.
- [4] V. Lempitsky, P. Kohli, C. Rother, and T. Sharp, "Image segmentation with a bounding box prior," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 277–284.
- [5] L. Grady, M. P. Jolly, and A. Seitz, "Segmentation from a box," in *Proc. IEEE Int. Conf. Comput. Vis.*, Nov. 2011, pp. 367–374.
- [6] Y. Boykov and G. Funka-Lea, "Graph cuts and efficient n-d image segmentation," *Int. J. Comput. Vis.*, vol. 70, no. 2, pp. 109–131, Nov. 2006.
- [7] X. Bai and G. Sapiro, "A geodesic framework for fast interactive image and video segmentation and matting," in *Proc. IEEE 11th Int. Conf. Comput. Vis.*, Oct. 2007, pp. 1–8.
- [8] B. Price, B. Morse, and S. Cohen, "Geodesic graph cut for interactive image segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2010, pp. 3161–3168.
- [9] L. Grady, "Random walks for image segmentation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 11, pp. 1768–1783, Nov. 2006.
- [10] W. Yang, J. Cai, J. Zheng, and J. Luo, "User-friendly interactive image segmentation through unified combinatorial user inputs," *IEEE Trans. Image Process.*, vol. 19, no. 9, pp. 2470–2479, Sep. 2010.
- [11] D. Lei and A. Yilmaz, "Image segmentation as learning on hyper-graphs," in *Proc. 7th Int. Conf. Mach. Learn. Appl.*, Dec. 2008, pp. 247–252.
- [12] X. Ren and J. Malik, "Learning a classification model for segmentation," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Oct. 2003, vol. 1, pp. 10–17.
- [13] C. Straehle *et al.*, "Seeded watershed cut uncertainty estimators for guided interactive segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 765–772.
- [14] C. Couprie, L. Grady, L. Najman, and H. Talbot, "Power watersheds: A new image segmentation framework extending graph cuts, random walker and optimal spanning forest," in *Proc. IEEE 12th Int. Conf. Comput. Vis.*, Oct. 2009, pp. 731–738.
- [15] C. Couprie, L. Grady, L. Najman, and H. Talbot, "Power watershed: A unifying graph-based optimization framework," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 7, pp. 1384–1399, Jul. 2011.
- [16] W. Wang *et al.*, "Finding coherent motions and semantic regions in crowd scenes: A diffusion and clustering approach," in *Proc. 11th Eur. Conf. Comput. Vis.*, 2014, pp. 756–771.
- [17] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE Trans. Syst., Man Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.
- [18] W. House and B. M. , "Tracking of cell populations to understand their spatio-temporal behavior in response to physical stimuli," in *Proc. Workshop Math. Model. Biomed. Image Anal.*, Miami, FL, 2009, pp. 186–193.
- [19] N. N. Kachouie, P. Fieguth, and E. Jervis, "Watershed deconvolution for cell segmentation," in *Proc. 30th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2008, pp. 375–378.
- [20] H. P. Grimm, A. B. Verkhnovsky, A. Mogilner, and J. J. Meister, "Analysis of actin dynamics at the leading edge of crawling cells: Implications for the shape of keratocyte lamellipodia," *Eur. Biophys. J.*, vol. 32, no. 6, pp. 563–577, Sep. 2003.
- [21] F. Li, X. Zhou, H. Zhao, and S. T. C. Wong, "Cell segmentation using front vector flow guided active contours," in *12th Int. Conf. MICCAI*, 2009, pp. 609–616.
- [22] G. Xiong *et al.*, "Segmentation of drosophila rnai fluorescence images using level sets," in *IEEE Int. Conf. Image Process.*, Oct. 2006, pp. 73–76.
- [23] D. Padfield, J. Rittscher, N. Thomas, and B. Roysam, "Spatio-temporal cell cycle phase analysis using level sets and fast marching methods," *Med. Image Anal.*, vol. 13, no. 1, pp. 143–155, 2009.
- [24] M. Ambuhl, C. Brepsant, J.-J. Meister, A. Verkhnovsky, and I. F. Szalzarini, "High-resolution cell outline segmentation and tracking from phase-contrast microscopy images," *J. Microsc.*, vol. 245, pp. 161–170, 2012.
- [25] C. Zhang, J. Yarkony, and F. A. Hamprecht, "Cell detection and segmentation using correlation clustering," *Proc. MICCAI*, pp. 9–16, 2014.
- [26] C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman, "Detecting overlapping instances in microscopy images using extremal region trees," *Med. Image Anal.*, 2015.
- [27] Z. Yin, H. Su, E. Ker, M. Li, and H. Li, "Cell-sensitive microscopy imaging for cell image segmentation," *Proc. MICCAI*, pp. 41–48, 2014.
- [28] Z. Yin, H. Su, E. Ker, M. Li, and H. Li, "Cell-sensitive phase contrast microscopy imaging by multiple exposures," *Med. Image Anal.*, vol. 25, no. 1, pp. 111–121, 2015.
- [29] A.-A. Liu, K. Li, and T. Kanade, "A semi-Markov model for mitosis segmentation in time-lapse phase contrast microscopy image sequences of stem cell populations," *IEEE Trans. Med. Imag.*, vol. 31, no. 2, pp. 359–369, Feb. 2012.
- [30] A. Liu, T. Hao, Z. Gao, Y. Su, and Z. Yang, "Nonnegative mixed-norm convex optimization for mitotic cell detection in phase contrast microscopy," *Computat. Math. Methods Med.*, vol. 2013, 2013.
- [31] A. Liu, T. Hao, Z. Gao, Y. T. Su, and Z. Yang, "Sequential sparse representation for mitotic event recognition," *Electron. Lett.*, vol. 49, no. 14, pp. 869–871, 2013.
- [32] A. Liu, Z. Gao, H. Tong, Y. Su, and Z. Yang, "Sparse coding induced transfer learning for hep-2 cell classification," *Bio-Med. Mater. Eng.*, vol. 24, no. 1, pp. 237–243, 2013.
- [33] K. Li and T. Kanade, "Nonnegative mixed-norm preconditioning for microscopy image segmentation," *Inf. Process. Med. Imag.*, pp. 362–373, 2009.
- [34] Z. Yin and T. Kanade *et al.*, "Restoring DIC microscopy images from multiple shear directions," *Inf. Process. Med. Imag.*, pp. 384–397, 2011.
- [35] Z. Yin, T. Kanade, and M. Chen, "Understanding the phase contrast optics to restore artifact-free microscopy images for segmentation," *Med. Image Anal.*, vol. 16, no. 5, pp. 1047–1062, 2012.
- [36] H. Su, Z. Yin, T. Kanade, and S. Huh, "Phase contrast image restoration via dictionary representation of diffraction patterns," in *Proc. MICCAI*, 2012, vol. 3, pp. 615–622.
- [37] R. El-Yaniv and D. Pechyony, "Transductive rademacher complexity and its applications," in *Proc. 20th Annu. Conf. Learn. Theory*, 2007, pp. 157–171.
- [38] T. M. Cover and J. A. Thomas, *Elements of Information Theory*. New York, USA: Wiley-Interscience, 1991.
- [39] X. Zhu, Z. Ghahramani, and J. Lafferty, "Semi-supervised learning using gaussian fields and harmonic functions," in *Proc. 20th Int. Conf. Mach. Learn.*, 2003, pp. 912–919.
- [40] M. Ji and J. Han, "A variance minimization criterion to active learning on graphs," in *Proc. 15th Int. Conf. Artif. Intell. Stat.*, 2012, pp. 556–564.
- [41] P. L. Bartlett and S. Mendelson, "Rademacher and Gaussian complexities: Risk bounds and structural results," *J. Mach. Learn. Res.*, vol. 3, pp. 463–482, 2003.
- [42] R. El-Yaniv and D. Pechyony, "Transductive rademacher complexity and its applications," *J. Artif. Intell. Res.*, vol. 35, pp. 193–234, 2009.
- [43] L. Zhang and Q. Ji, "A Bayesian network model for automatic and interactive image segmentation," *IEEE Trans. Image Process.*, vol. 20, no. 9, pp. 2582–2593, Sep. 2011.
- [44] X. Zhu, J. Lafferty, and Z. Ghahramani, "Combining active learning and semi-supervised learning using Gaussian fields and harmonic functions," in *ICML 2003 Workshop on the Continuum From Labeled to Unlabeled Data in Machine Learning and Data Mining*, 2003, pp. 58–65.
- [45] W. Liu and S.-F. Chang, "Robust multi-class transductive learning with graphs," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2009, pp. 381–388.
- [46] H. Su, Z. Yin, S. Huh, and T. Kanade, "Cell segmentation via spectral analysis on phase retardation features," in *Proc. Int. Symp. Biomed. Imag.*, Apr. 2012.
- [47] S. Huh, D. F. E. Ker, H. Su, and T. Kanade, "Apoptosis detection for adherent cell populations in time-lapse phase-contrast microscopy images," in *Proc. MICCAI*, 2012, pp. 331–339.
- [48] M. Li and Z. Yin, "Co-restoring multi-modal microscopy images," in *Proc. MICCAI*, 2015.