MODELING STRUCTURAL DISSIMILARITY BASED ON SHAPE EMBODIMENT FOR CELL SEGMENTATION

Hyun-Gyu Lee, Adiba Orzikulova, Bo-Gyu Park, and Sang-Chul Lee

Department of Computer Science and Engineering, INHA University

ABSTRACT

Accurate cell segmentation is one of the critical, yet challenging problems in microscopy images due to ambiguous boundaries as well as a wide variation of shapes and sizes of cells. Even though a number of existing methods have achieved decent results for cell segmentation, boundary vagueness between adjoining cells tended to cause generation of perceptually inaccurate segmentation of stained nuclei. We propose a segmentation method of cells based on structural dissimilarity between embodied and imaged cells. From assumption that the shape of the region of adjoining cells follows a 2D Gaussian mixture model, the cell region is divided by an expectation-maximization method. The lowest structural dissimilarity using embodied cells decides on the number of components of the 2D Gaussian mixture model. The region of interest is extracted by implementation of both global and local thresholdings, which performs binarization of the local image with a seed at the center, where the seed is obtained by the maximally stable extremal regions. Our approach presented considerably higher evaluation scores compared with other five existing methods in terms of both accuracies of region of interest (ROI) detection and boundary discrimination.

Index Terms— Cell segmentation, cell division, embodied cell, Gaussian mixture model, ROI detection

1. INTRODUCTION

Detection and quantification of cells using automated approaches have a significant role to play in medical diagnosis and treatment, since it provides useful information in analysis of cell structures. There are different types of techniques that can be applied to cell detection and segmentation. However, these cell-based researches can pose a challenging problem due to diversity in microscopy images caused by noises, artifacts, as well as a wide range of cell patterns. Especially, organization of cells in overlapping clusters tends to present complex and irregular segmentation, therefore it can be considered as the primary hindrance to maintain ideal and accurate performance in cell segmentation.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2015R1D1A1A01057968).

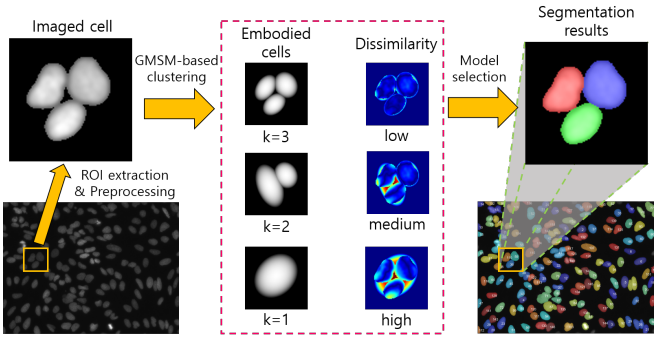


Fig. 1. Overview of the proposed segmentation based on structural dissimilarity.

Some classical techniques involving thresholding intensity of the image such as watershed [10] and Otsu's [18] segmentations, even when combined together [4], might not achieve desired accurate results due to their sensitivity to noise and sophisticated cell structures or characteristics. Segmentation approaches using region-based [8] or edge-based [7] active contour model (ACM) [11] that share energy minimization concept are robust to the problem caused by sensitivity. However, these ACM-based segmentation techniques are time-consuming, and reliable boundary initialization is prerequisite for accurate segmentation in these methods, whereas the initialization is not required for more progressive energy minimization methods such as level sets [13], gradient flow tracking [14][15], modular-interactive nuclear segmentation (MINS) [17] and graph cut [1][19].

Recently, in order to enhance the quality and efficiency of cell segmentation results, machine learning techniques with a variety of features and classifiers were introduced. Maximum a posteriori probability (MAP) and Kalman filter [2] are used to detect and track cells in a high-throughput microscopy video. A support vector machine, which has good generation ability, is also useful for a variety of datasets, such as fluorescence images of human embryonic kidney (HEK) cells, phase-contrast images of HeLa cells, and histopathological images of breast cancers [3]. In another supervised learning approach [12], color-texture feature was designed using local Fourier transform. Simple statistical features, such as mean, variance, and histogram of a pixel's neighborhood, and

a random forest classifier are also performed well in different types of images. However, such progressive approaches [2][3][12][20] can perform well only if ground truth sample images are provided.

We aim to propose a novel method based on structural dissimilarity between a cell embodied by a Gaussian mixture based shape model (GMSM) and an imaged cell, as shown in Fig. 1., to tackle above mentioned challenges including misleading (jagged, straight) boundaries between overlapped cells as well as a high variety of intensities in images which can be detrimental in obtaining expected accurate results. Our approach contributes to the accuracy in cell segmentation in terms of the following aspects:

- A high performance in quantification of cells can be obtained without accurate seed detection. This is a significant progress when compared with other cell segmentation methods which require high accuracy of seed detection for accurate segmentation [1][21].
- Boundaries estimated using the proposed method is similar to what human perceives as boundaries between aggregated cells.

The seed detection using maximally stable extremal regions (MSER) [5][22], which is effective in detection of elliptical objects, is the initial step towards obtaining accurate segmentation. The derived seeds are used in local thresholding. A combination of local and global thresholdings extracts cell regions in spite of large variation of intensities of cells. Expectation-maximization (EM) [6] is applied to divide the extracted regions, and it can overcome difficulties concerning boundary estimation between adjoining cells. In our experiments, we compared performance of our method with other state-of-the-arts methods in Broad Bioimage Benchmark Collection (BBBC) dataset [16], and achieved improved results especially in terms of identifying the boundaries of cells.

2. COMBINATION OF THRESHOLDING USING SEED DETECTION BASED ON MSER

A key idea of our segmentation method is to find the number of components, which corresponds to the number of actual cells, of 2D Gaussian mixture model. Our method, using structural dissimilarity between embodied and actual cells, consists of two parts: ROI detection to be used as an input of the Gaussian mixture model, and cell division using the dissimilarity.

We assume that the shape of adjoining cells follows a 2D Gaussian mixture model, and each component of the model relates to a cell. Identification of cell position is an important stage for the estimation of the model in the adjoining cells. Otsu's method [18] can be successful at the discrimination of foreground and background, but it yields poor results in regions where variations of intensities are locally large depending on cells. To extract the reliable ROI under the large variation of intensities, we define a ROI detector combined with local and global thresholding, as follows:

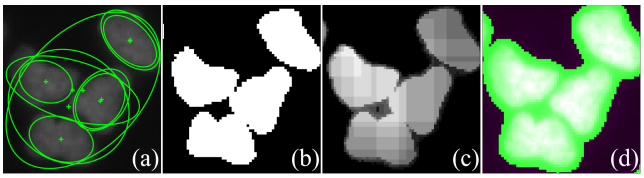


Fig. 2. Examples of (a) seed detection using MSER, (b) a binary image by global thresholding (Eq. (2)), (c) an accumulated binary image by local thresholding (Eq. (3)), and (d) final ROI (green area) using the combination of these thresholdings (Eq. (1)).

$$\Psi(x,y) = \begin{cases} 1, \text{if } \Psi_{\text{blur}}(x,y) > \delta \\ 0, \text{otherwise} \end{cases}, \qquad \Psi_{\text{blur}} = (\Psi_{\text{global}} + \frac{1}{\alpha}\Psi_{\text{local}}) * g(\sigma), \tag{1}$$

where, $g(\sigma)$ is the Gaussian kernel, which produces a smooth boundary, by a combination of local and global thresholdings. To normalize Ψ_{local}, we introduce α, the maximum intensity of an accumulated binary image Ψ_{local} produced by local thresholding. The threshold δ controls the range of the ROI, and a low threshold generates a larger ROI.

The binary image Ψ_{global} produced by global thresholding and the accumulated binary image Ψ_{local} are defined using Otsu's method and MSER [5][22], as follows:

$$\Psi_{\text{global}}(x,y) = \begin{cases} 1, \text{if } G(x,y,\sigma) > \delta_{\text{global}} \\ 0, \text{otherwise} \end{cases}, \tag{2}$$

$$\Psi_{\text{local}}(x,y) = \sum_{i=1}^{n} \begin{cases} 1, \text{if } G(x,y,\sigma) > \delta_{\text{local}}^{(i)} \\ 0, \text{otherwise} \end{cases}, \quad \text{s.t. } \hat{x}_i - r_i \le x \le \hat{x}_i + r_i, \; \hat{y}_i - r_i \le y \le \hat{y}_i + r_i, (\hat{x}_i, \hat{y}_i) \in \hat{\mathbf{x}}_i, \tag{3}$$

where, δ_{global} is the global threshold obtained by applying Otsu's method to a Gaussian blurred image ($G = I * g(\sigma)$) produced from an original image I. Each local threshold $\delta_{\text{local}}^{(i)}$ is obtained by applying Otsu's method to a subset image, centered on the position $\hat{\mathbf{x}}_i$ of the seed, of size r_i^2, and r_i is a scaled length of the major axis of ith extremal region. $\hat{\mathbf{x}}$ is a seed position detected by MSER that can find extremal region where all pixels have higher or lower intensity values compared to other pixel values on outer boundary.

Since MSER may detect more than two seeds in a cell, we select the largest length of the major axis in seeds grouped by mutual proximity manner, and eliminate other seeds for low computation in the local thresholding. The global thresholding can detect an approximate ROI of the whole image,

but parts of the ROI may be classified as background, due to the low contrast between a cell and background. Our local thresholding near detected seeds compensates the defect of the global thresholding. Fig. 2 shows process of the combination of these two thresholdings.

3. CELL DIVISION USING STRUCTURAL DISSIMILARITY

According to the assumption that the shape of ROI follows Gaussian mixture model, we split the ROI based on Gaussian distribution and define this divided distribution model as the Gaussian mixture based shape model (GMSM). Thefore, we estimate a GMSM using EM, and the classified cell consists of pixels that contribute to generate each component of the GMSM.

If the positions of ROI are used only as features in EM, regions near boundaries between adjoining cells may be classified as other cells. To classify the regions as their own cell, the ROI position and its local maximum which was obtained by gradient ascent from each ROI position of Gaussian blurred image G were used as features in EM. In case that the number of components of a GMSM is the same as the number of the cells in the ROI, selecting the number of components of the GMSM is important in cell division using EM, since the EM method can adequately divide ROI consisted of aggregated cells. To find the number of components of GMSM in the ROI, we define structural dissimilarity $D(\mathbf{X}, \Theta_K)$ between cells embodied by parameters of a GMSM and actual cells, as follows:

$$\begin{aligned} D(\mathbf{X}, \Theta_K) &= \sum_{i=1} g(\mathbf{x}_i, \rho\Theta_1)|\mathbf{V}(\mathbf{x}_i, \Theta_K) - G(\mathbf{x}_i, \sigma)|, \\ \mathbf{V}(\mathbf{x}_i, \Theta_K) &= \max_{\theta_k} \big(V(\mathbf{x}_i, \theta_k)\big), \mathbf{x}_i \in \mathbf{X}, \theta_k \in \Theta_K, \\ &\quad \text{s.t.}\, 1 \le k \le K \le A/A_{\min}, \end{aligned} \tag{4}$$

where, $\mathbf{V}(\mathbf{x}_i, \Theta_K)$ is the maximum intensity at each position $\mathbf{x}_i$ of ROI in k embodied cells, produced using parameter set Θ_K of K components of the GMSM for the ROI, and θ_k are parameters, consisted of mean μ_k and covariance Σ_k, of each component of the GMSM. Θ_1 are the parameters of the GMSM with one component, and ρ controls the size of the Gaussian kernel. A is the size of the segment of ROI, and $A_{\min}$ is the minimum size of cell regions detected by MSER.

The dissimilarity represents Gaussian-weighted difference of intensities of the embodied cells $\mathbf{V}(\mathbf{x}_i, \Theta_K)$ and blurred actual cells. Hence, we solve the problem of finding the number of seeds by estimating the GMSM with the lowest dissimilarity. In order to solve $\text{argmin}_K\, D(\mathbf{X}, \Theta_K)$, we measure the dissimilarities for all candidate GMSMs with component ranged from one to $A/A_{\min}$, and select the best model that similarly embodies original cells. Finally, the pixels, which contributed to generating a component of the selected model, are classified as one segment. The embodied cell is defined using exponents for each pixel distance and whitening transform [6] as following:

$$\begin{aligned} V(\mathbf{x}, \theta) &= \frac{\max_{\mathbf{x}_c}\left(e^{2\|\mathbf{x}_c\mathbf{W}\|}\right) - e^{2\|\mathbf{x}_c\mathbf{W}\|}}{\max_{\mathbf{x}_c}\left(e^{2\|\mathbf{x}_c\mathbf{W}\|}\right) - \min_{\mathbf{x}_c}\left(e^{2\|\mathbf{x}_c\mathbf{W}\|}\right)} + \delta_t, \\ &\quad \text{s.t.}\, P(\mathbf{x}_c|\theta) > \delta_t, \\ \delta_t &= \underset{0<t\le 1}{\text{argmin}} \left| \Big(\sum_{\mathbf{x}_c} P(\mathbf{x}_c|\theta) \Big) - \gamma \right|, \text{s.t.}\, P(\mathbf{x}_c|\theta) < t, \end{aligned} \tag{5}$$

where, $P(\mathbf{x}_c|\theta)$ is a multivariate Gaussian distribution, and $\mathbf{x}_c$ is the distance in pixels between the given $\mathbf{x}$ and μ in θ. The term δ_t is the minimum probability of $P(\mathbf{x}_c|\theta)$ obtained from γ that determines the size of the embodied cell, and $\mathbf{W}$ is a 2×2 whitening matrix used in the whitening transformation for representation of intensities similar to those on the surface of the original cell.

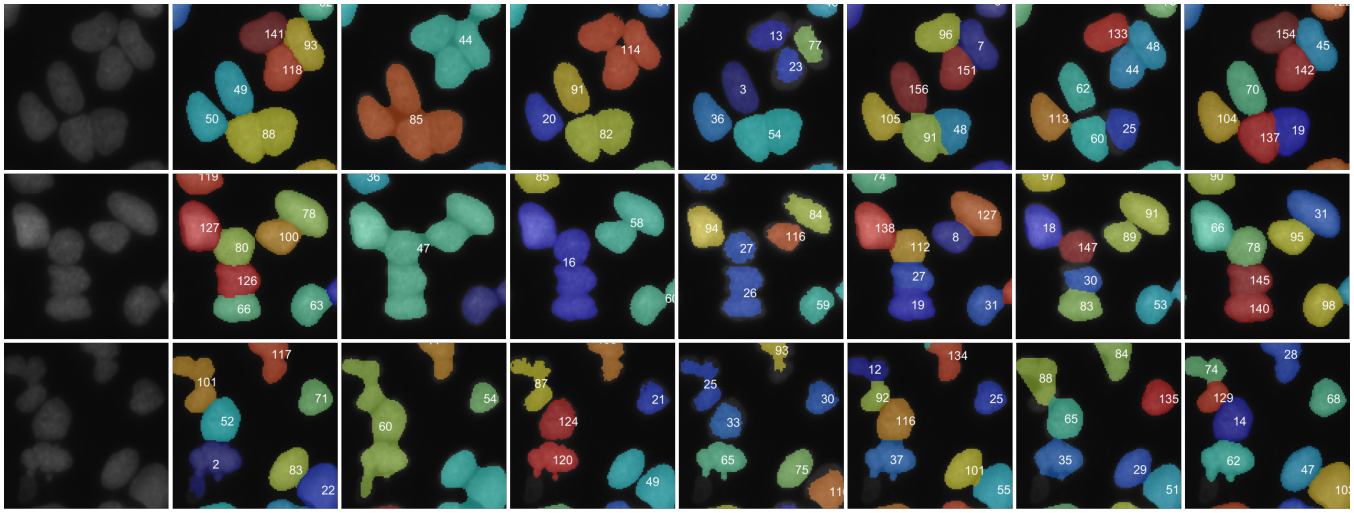
Using Eq. (4), we obtained the cell where brightness characteristics are similar to the ones of the original cell as shown in Fig. 1.

4. EXPERIMENTAL RESULTS

In order to evaluate the proposed method, our experiments were carried out on the BBBC006v1 dataset, publicly available from Broad Bioimage Benchmark Collection [16]. This dataset was produced to test image analysis methods for the life sciences, and it comprises 768 microscopy images along with corresponding ground-truth images. For experimental evaluation, we used a total of 758 16-bit grayscale images of 696×520 resolution in the BBBC dataset except the images without ground truth.

4.1. Comparison methods

Evaluation of our approach was performed by comparing it with other five existing methods such as non-overlapping extremal regions based (NOER) [3], graph cut-based (GC) [1], interactive learning and segmentation toolkit (ilastik) [20], MINS [17], and Otsu [18]. Parameters for techniques using machine learning for cell segmentation such as NOER and ilastik were chosen empirically. Since GC and Otsu are working fully automatically, we did not consider any parameters. The ilastik method provides a feature list, and we selected mean, variance, and histogram of each pixel's neighborhood intensities in the list. Parameter selection for NOER was decided by relying on its authors recommended values, and then we trained and tested samples using 10-fold cross-validation.



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

Fig. 3. Comparison with existing cell segmentation methods. (a) Original image, (b) ground-truth image, (c) ilastik [20], (d) Otsu [18], (e) NOER [3], (f) GC [1], (g) MINS [17], and (h) the proposed method. As a result of segmentation, each cell was randomly assigned a unique color and a number.

4.2. Evaluation metrics

We evaluated the results obtained from proposed and five comparison methods using three metrics [9] for segmentation, as follows:

$$Sensitivity = \frac{|R \cap S|}{|R|}, J = \frac{|R \cap S|}{|R \cup S|}, DSC = \frac{2|R \cap S|}{|R| + |S|}, \quad (6)$$

where, $|R|$ is the number of pixels in a reference cell in a ground-truth image, and $|S|$ is the number of pixels in a cell in a segmented image, which is closest to the reference cell.

$Sensitivity$ measures the proportion of pixels of a segmented cell, which is correctly identified in the region of a reference cell. Jaccard similarity coefficient J measures the proportion of an intersected region between two cells R and S in a union region of the cells. Dice similarity coefficient DSC is known as F1 score, and it measures the similarity of the two cells.

4.3. Segmentation results

We selected two parameters ρ and δ, where the sum of the three evaluation metrics was highest in randomly selected 10% of images, and the selected parameters were applied to the remaining images in the dataset. This procedure was repeated ten times to perform reliable validation.

The segmentation result of an image is shown in Fig. 3. Most of the cells were correctly classified. Especially, using our method, the boundaries of adjoining cells were identified close to the boundaries that are recognized by a human, such as 45, 142, and 154 in Fig. 3h. Moreover, the ROI corresponding to cell regions was adequately detected by our combination of thresholdings, except the regions with extremely low contrast. Proposed method outperformed the other comparison methods in all evaluation metrics, as shown in Table 1.

Table 1. Comparison of our method with five cell segmentation methods.

Methods	$Sensitivity$	J	DSC
ilastik [20]	0.540	0.494	0.604
Otsu [18]	0.592	0.568	0.634
NOER [3]	0.684	0.679	0.781
GC [1]	0.728	0.722	0.794
MINS [17]	0.767	0.758	0.829
Proposed	**0.876**	**0.845**	**0.886**

5. CONCLUSION

We have proposed a new method of cell segmentation based on structural dissimilarity between embodied and actual cells. The expectation-maximization method adequately divided regions of overlapped cells, which is in the form of 2D Gaussian mixture model. Identification of the number of components in the Gaussian mixture model was carried out by the implementation of lowest structural dissimilarity between embodied and imaged cells. We showed the efficiency of using local and global thresholding in order to deal with the issues regarding intensity difference between cells in the microscopy images. We showed that our method has a remarkable performance in terms of evaluation measurements in cell segmentation, without provision of seed detection. Moreover, our method managed to achieve desirable results while dealing with boundary identification of adjoining cells. Our future work is to improve limitations in over segmentation that might occur when the surface of a cell is not smooth and irregular.

6. REFERENCES

[1] Y. Al-Kofahi, W. Lassoued, W. Lee, and B. Roysam, "Improved automatic detection and segmentation of cell nuclei in histopathology images," *Biomedical Engineering, IEEE Transactions on*, vol. 57, no. 4, pp. 841–852, 2010.

[2] A. Arbelle, N. Drayman, M. Bray, U. Alon, A. Carpenter, and T. R. Raviv, "Analysis of high-throughput microscopy videos: Catching up with cell dynamics," in *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*. Springer, 2015, pp. 218–225.

[3] C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman, "Learning to detect cells using non-overlapping extremal regions," in *Medical image computing and computer-assisted intervention–MICCAI 2012*. Springer, 2012, pp. 348–356.

[4] R. Beare and G. Lehmann, "The watershed transform in itk-discussion and new developments," *The Insight Journal*, vol. 1, pp. 1–24, 2006.

[5] W. Berchtold, M. Schäfer, and M. Steinebach, "Maximal stable extremal regions for robust video watermarking," *Electronic Imaging*, vol. 2016, no. 8, pp. 1–8, 2016.

[6] C. M. Bishop, "Pattern recognition and machine learning." New York, USA: Springer, 2006.

[7] V. Caselles, R. Kimmel, and G. Sapiro, "Geodesic active contours," *International journal of computer vision*, vol. 22, no. 1, pp. 61–79, 1997.

[8] T. F. Chan and L. A. Vese, "Active contours without edges," *Image processing, IEEE transactions on*, vol. 10, no. 2, pp. 266–277, 2001.

[9] H.-H. Chang, A. H. Zhuang, D. J. Valentino, and W.-C. Chu, "Performance measure characterization for evaluating neuroimage segmentation algorithms." *NeuroImage*, vol. 47, no. 1, pp. 122–135, 2009.

[10] K. Jiang, Q.-M. Liao, and S.-Y. Dai, "A novel white blood cell segmentation scheme using scale-space filtering and watershed clustering," in *Machine Learning and Cybernetics, 2003 International Conference on*, vol. 5. IEEE, 2003, pp. 2820–2825.

[11] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International journal of computer vision*, vol. 1, no. 4, pp. 321–331, 1988.

[12] H. Kong, M. Gurcan, and K. Belkacem-Boussaid, "Partitioning histopathological images: an integrated framework for supervised color-texture segmentation and cell splitting," *Medical Imaging, IEEE Transactions on*, vol. 30, no. 9, pp. 1661–1677, 2011.

[13] C. Li, C. Xu, C. Gui, and M. D. Fox, "Distance regularized level set evolution and its application to image segmentation," *Image Processing, IEEE Transactions on*, vol. 19, no. 12, pp. 3243–3254, 2010.

[14] G. Li, T. Liu, A. Tarokh, J. Nie, L. Guo, A. Mara, S. Holley, and S. T. Wong, "3d cell nuclei segmentation based on gradient flow tracking," *BMC cell biology*, vol. 8, no. 1, p. 1, 2007.

[15] T. Liu, J. Nie, G. Li, L. Guo, and S. T. Wong, "Zfiq: a software package for zebrafish biology," *Bioinformatics*, vol. 24, no. 3, pp. 438–439, 2008.

[16] V. Ljosa, K. L. Sokolnicki, and A. E. Carpenter, "Annotated high-throughput microscopy image sets for validation," *Nat Methods*, vol. 9, no. 7, p. 637, 2012.

[17] X. Lou, M. Kang, P. Xenopoulos, S. Munoz-Descalzo, and A.-K. Hadjantonakis, "A rapid and efficient 2d/3d nuclear segmentation method for analysis of early mouse embryo and stem cell image data," *Stem cell reports*, vol. 2, no. 3, pp. 382–397, 2014.

[18] N. Otsu, "A threshold selection method from gray-level histograms," *Automatica*, vol. 11, no. 285-296, pp. 23–27, 1975.

[19] H. Sharma, A. Alekseychuk, P. Leskovsky, O. Hellwich, R. Anand, N. Zerbe, and P. Hufnagl, "Determining similarity in histological images using graph-theoretic description and matching methods for content-based image retrieval in medical diagnostics," *Diagnostic pathology*, vol. 7, no. 1, p. 134, 2012.

[20] C. Sommer, C. Straehle, U. Koethe, and F. A. Hamprecht, "ilastik: Interactive learning and segmentation toolkit," in *Biomedical Imaging: From Nano to Macro, 2011 IEEE International Symposium on*. IEEE, 2011, pp. 230–233.

[21] J. Stegmaier, J. C. Otte, A. Kobitski, A. Bartschat, A. Garcia, G. U. Nienhaus, U. Strhle, and R. Mikut, "Fast segmentation of stained nuclei in terabyte-scale, time resolved 3d microscopy image stacks," *PLOS ONE*, vol. 9, no. 2, pp. 1–11, 02 2014.

[22] H. Zhu, J. Sheng, F. Zhang, J. Zhou, and J. Wang, "Improved maximally stable extremal regions based method for the segmentation of ultrasonic liver images," *Multimedia Tools and Applications*, vol. 75, no. 18, pp. 10 979–10 997, 2016.