# JOINT CELL SEGMENTATION AND TRACKING USING CELL PROPOSALS

Saad Ullah Akram<sup>1,2</sup>, Juho Kannala<sup>3</sup>, Lauri Eklund<sup>2,4</sup>, and Janne Heikkilä<sup>1</sup>

<sup>1</sup> Center for Machine Vision Research, <sup>2</sup> Biocenter Oulu, <sup>4</sup> Oulu Center for Cell-Matrix Research, and <sup>4</sup> Faculty of Biochemistry and Molecular Medicine, University of Oulu, Finland <sup>3</sup> Department of Computer Science, Aalto University, Finland

### ABSTRACT

Time-lapse microscopy imaging has advanced rapidly in last few decades and is producing large volume of data in cell and developmental biology. This has increased the importance of automated analyses, which depend heavily on cell segmentation and tracking as these are the initial stages when computing most biologically important cell properties. In this paper, we propose a novel joint cell segmentation and tracking method for fluorescence microscopy sequences, which generates a large set of cell proposals, creates a graph representing different cell events and then iteratively finds the most probable path within this graph providing cell segmentations and tracks. We evaluate our method on three datasets from ISBI Cell Tracking Challenge and show that our greedy nonoptimal joint solution results in improved performance compared with state of the art methods.

*Index Terms*— joint segmentation and tracking, cell tracking, cell segmentation, cell proposals

## 1. INTRODUCTION

In last couple of decades, advances in microscopy techniques have enabled the investigation of dynamic processes of cells at increasing temporal and spatial resolution. To a large extent, microscopy imaging can be automated resulting in a huge amount of data with single imaging experiments generating up-to TBs of data [1]. Manual analysis of these huge datasets is highly inefficient, not easily reproducible, often only qualitative and limits the hypotheses which can be tested. In cell and developmental biology, to better understand cell functions and tissue development, it is often important to analyze cell behavior at individual cell level. Robust cell segmentation and tracking is necessary for automating these detailed analyses.

Microscopic images often suffer from a low signal to noise ratio, poor staining, variable fluorescence in cells or cell organelles and high cell density. This results in images which have some very ambiguous regions, often leading to mistakes. Utilizing temporal constraints can be very helpful in these situations as most ambiguities do not last for more than few frames.

### 2. RELATED WORK

Most cell tracking methods can be grouped into two main categories, tracking by assignment and tracking by model evolution. In tracking by model evolution, a mathematical model of cells, frequently based on level-sets [2], is evolved from frame to frame to jointly segment and track cells. These methods require high frame rate, high resolution and usually need special handling of cells entering or leaving the imaged region.

Tracking by assignment is by far the most popular approach within cell tracking field, with five out of six methods in ISBI 2013 Cell Tracking Challenge [3] belonging to this category. It decouples cell segmentation and tracking, making it more generalizable and easier to manage. In the first stage, cells are segmented and some features are extracted. In the second stage, these features are used to link segmentations in neighboring frames. The tracks can then be obtained by considering the whole graph with all frames linked together [4, 5] or just neighboring frames or a combination of both of these [6].

In recent years, object proposals have gained prominence in computer vision field when it comes to object detection, segmentation and tracking. They have also recently been used in cell detection [7] and tracking [8, 9, 10]. Proposal based joint segmentation and tracking methods [8, 9, 10] first generate a large number of cell proposals, connect them temporally, and then use integer linear programming for inference. These methods heavily restrict possible transitions between frames to speed up inference for long sequences [9] which is not always suitable especially if sequences have low frame rate and high cell density.

We propose a novel cell segmentation and tracking method, which extends the idea of iteratively finding the lowest cost path in a graph [4] by using cell proposals instead of cell segmentations. Previously a similar idea [11], using the lowest cost path in a graph representing proposals, has been used to extract a single object from a video sequence. However, we extend this idea to extract multiple cells from sequences with many different challenges including mitosis and a low frame rate.

Our method consists of three main stages: cell proposal generation (Section 3), graph construction (Section 4) and



**Fig. 1**: Cell Proposal Generation: a) Original Image from HeLa-01, b) Binary Segmentation, c) Individual filter responses for a filter bank with filters shown in top left corner, d) Accumulated filter response, e) Detected cell seeds, and f) A set of segmentation proposals.

greedy approximate inference (Section 5).

### 3. PROPOSAL GENERATION

The first stage in our method is proposal generation, which aims to generate a large number of segmentation proposals such that they have a high recall. Fig. 1 shows main steps in generation of one set of segmentation proposals: first, cells are segmented from background; second, blob detection is used to detect individual cells; finally, individual cells are segmented using watershed. Initial binary segmentation can be obtained using any method that can separate cells from background. We use either graph-cuts [12] or thresholding depending on the sequence.

To split cell clusters in binary segmented image, we use N generalized Laplacian of Gaussian (gLOG) filter banks [13] covering common cell sizes and aspect ratios. All filters within a filter bank are rotated version of a base filter having scale ( $\sigma_x, \sigma_y$ ) and aspect ratio ( $\frac{\sigma_x}{\sigma_y}$ ). Each filter bank is applied to the image and response image  $R_n$  is obtained by summing the log scale normalized [13] response  $R_{nm}$  of all M filters within it using:

$$R_n = \sum_{m=1}^{M} (1 + \log(\sigma_x)^{\alpha})(1 + \log(\sigma_y)^{\alpha})R_{nm} \qquad (1)$$

where  $\alpha$  is the parameter which controls the eccentricities of the detected blobs [13]. This normalization is necessary to ensure that responses of all filter banks have same scale and can be compared. Local peaks, above a threshold, in the response image  $R_n$  are detected and used as cell markers by watershed transform to split cell clusters and obtain individual cell segmentations. This results in N sets of cell segmentation proposals, which contain many duplicate cell proposals.



**Fig. 2**: a) Segmentation proposal tree for a connected region from binary segmented image. b) Super-node structure. c) Simplified graph structure (Mitosis and Apoptosis edges are not shown).

Non-maxima suppression is used to remove these duplicate proposals. Remaining proposals are arranged in hierarchical proposal trees as shown in Fig. 2a (Fig. 2b and 2c will be explained in Section 4), where proposals along a branch conflict with each other and only one proposal can be selected from each branch during inference.

#### 4. GRAPHICAL MODEL

In the second stage, our method creates a directed acyclic graph for the whole sequence which includes links representing all probabilities necessary for selecting good cell proposals and their correct associations between frames.

## 4.1. Graph Structure

Fig. 2c shows a simplified graph structure for a sequence with three frames and six cell proposals. Each cell track, starting at  $t_{start}$  and ending at  $t_{end}$ , in this graph begins from *Start* node, passes through an entry node  $E_{t_{start}}$ , some cell proposal super-nodes,  $P_{t,i}$ , an exit node  $L_{t_{end}}$  and terminates at End node. A high gating threshold is used to restrict connections between proposals in adjacent frames which are too far away from each other. In many cell tracking applications, cells frequently enter or leave the imaged region; we use entry nodes  $(E_t)$  and exit nodes  $(L_t)$  to allow new tracks to begin and end respectively at each time instant t for these cells.

### 4.2. Super Node Structure

Fig. 2b shows the structure of a super-node,  $P_{t,i}$ , which represents each cell segmentation proposal. It consists of five nodes, so that four of these nodes,  $D_{t,i}$  (move),  $M_{t,i}$  (mitosis),  $A_{t,i}$  (apoptosis) and  $V_{t,i}$  (leave), represent the events that the cell can go through and fifth node,  $F_{t,i}$ , connects this proposal to entry node  $E_t$  and proposals in previous frames.

#### 4.3. Model Probabilities

The green edges in Fig. 2b connecting  $F_{t,i}$  with event nodes represent the probability of that proposal being a cell. This probability is computed using an SVM classifier which is fed basic shape and appearance features. These features include area, perimeter, solidity, eccentricity, extent, filter scale, hu moments, mean and standard deviation of the gLOG response and the intensity image. It ensures that proposals which resemble actual cells are more likely to be selected over proposals with atypical shape or appearance.

Without specific intrinsic or external guidance cells frequently and abruptly change their movement speed and direction so our method uses Brownian motion model. The probabilities of blue edges in Fig. 2b and Fig. 2c connecting proposals in adjacent frames are computed using the distance between centroids of both proposals and output of SVM classifier which uses above mentioned features from both proposals in addition to their overlap to predict if they belong to same cell. Our model handles missed detections by allowing edges, which have an additional cost term depending on the number of skipped frames, between proposals in non-adjacent frames.

Our method assumes that when a cell goes through mitosis, it results in appearance of two daughter cells very close to the parent's centroid in next frame. It uses the overlap of daughter proposal pair with parent proposal, distances between them and above mentioned features from all three proposals to compute the probability of a parent proposal dividing into the daughter proposal pair. It then connects, using red edges in Fig. 2b, the parent proposal  $P_{t,i}$  with the daughter proposal pair,  $P_{t+1,j}$  and  $P_{t+1,k}$ , which has the highest mitosis probability among all daughter proposal pairs.

Cell disappearance also poses some challenges as any failure to detect disappearance of a cell due to either cell death or it leaving the imaged region can cause its track to continue through a nearby proposal, which can interfere with the paths of future tracks. Since in normal cell culture conditions cell death is a relatively rare event and there are usually not enough training samples, our method uses maximum overlap of a proposal with proposals in next frame to compute this probability, brown edges in Fig. 2b. In a sequence with high death rate, shape and appearance features can be used to learn this probability.

Cells frequently enter and leave the field of view at the image boundaries and from axial direction. In some sequences, cells entering/leaving the imaged region from axial direction often have smaller size and lower intensity compared with other cells. Our method uses a proposal's size, intensity and distance from nearest image border to compute its probability of entering/leaving the imaged region, pink/orange edges in Fig. 2b and Fig. 2c.

### 5. GREEDY APPROXIMATE INFERENCE

The edges, e, in the directed acyclic graph represent the probability, Pr(e), of different cellular events (move, mitosis, death, entering, leaving) which each cell proposal can go through. Each path, k, through this graph has a cost,  $C_k$ , which is the negative logarithm of the probability of the cell track that goes through the proposals and transitions (cell events) traversed by the set of edges, S, in that path.

$$C_k = \sum_{e \in S} -\log(Pr(e)) \tag{2}$$

Our method tries to cover all proposals with cell tracks so that the combined cost of all K tracks is minimized.

$$cost = \min \sum_{k=1}^{K} C_k \tag{3}$$

Globally optimal solution of (3) can be found using Integer Linear Programming [8, 9, 10] but it can be computationally expensive and may not always be feasible for very dense and long sequences. So we use a greedy iterative shortest path algorithm for performing inference on this graph. Our method uses Dijkstra's algorithm to iteratively find the lowest cost path S from *Start* node to *End* node in this directed graph passing through a set of cell proposals  $P_S$ . Then, it removes all proposals which lie along the branches containing any proposal in  $P_S$  and creates swap nodes [4], which allow future tracks to make modifications to previously found tracks.

The first track found by our greedy method is globally optimal since all possible proposals and transitions are considered. However, since there is no way to recover deleted proposals, the subsequent tracks are no longer globally optimal. Nevertheless, the sub-optimal paths still consider a very large number of proposals and transitions and result in selection of very good tracks.

If a track passes through a mitosis node,  $M_{t,i}$ , of a proposal, indicating that a mitosis event has taken place, then the next track is initiated from the second daughter cell node,  $F_{t+1,j}$ , instead of *Start* node. If a track jumps over few frames, our method uses linear motion to predict the locations of that cell in skipped frames and places the last detected cell at those locations. Once all the proposals have been exhausted, search is terminated.

### 6. RESULTS

We evaluate our proposed method on three fluorescence microscopy sequences, Fluo-N2DH-GOWT1-01, Fluo-N2DH-

**Table 1**: Tracking and segmentation results for our method,KTH [4] and LEID [2].

		TRA	SEG	TP	FN	FP	NS	EA	EC	ED2
GOWT1	Ours	0.9628	0.7678	2006	46	222	11	81	17	1
	KTH	0.9348	0.5834	1927	125	81	0	137	1	0
	LEID	0.9462	0.6849	1955	97	147	0	100	0	0
HeLa	Ours	0.9782	0.7707	8560	79	601	15	364	90	58
	KTH	0.9775	0.8018	8496	143	318	17	222	38	27
	LEID	0.9082	0.7568	7960	679	636	109	727	35	15
SIM+	Ours	0.9878	0.8528	2586	21	48	4	43	23	1
	KTH	0.9752	0.8377	2554	53	66	1	83	16	5



(a) Frame from Fluo-N2DH-Hela-01

(b) Our Method



**Fig. 3**: Segmentation results for a frame (segmented cells are labeled using 7 repeating colors). Contrast has been enhanced in (a) to show dark cells. • in (a) mark the centroids of ground truth cells. **FN**, **FP**, NS and **Over-segmentation** errors are highlighted in (b), (c) and (d).

HeLa-01 and Fluo-N2DH-SIM+-01, from ISBI 2015 Cell Tracking Challenge. We obtain initial segmentation for Fluo-N2DH-SIM+-01 using thresholding, and for other two using graph cuts [12]. We use the tracking performance measure (TRA) and Jaccard similarity index (SEG) used in the ISBI 2015 Cell Tracking Challenge for comparison. TRA is designed to mirror the manual effort needed to correct the errors in the tracks generated by tracking algorithms by differently penalizing following errors: FN (False Negatives), FP (False Positives), NS (Under-segmentations), EA (Missing edges), EC (Miss-labeled edges) and ED2 (Extra edges). Both TRA and SEG values range between 0 and 1 (perfect result), making comparison easier.

We compare our method with two methods from ISBI 2013 Cell tracking challenge [3]. KTH method [4] creates a graph by connecting cell segmentations in adjacent frames and then iteratively finds lowest cost paths in this graph using Viterbi algorithm. It had the best overall performance

in ISBI 2013 Cell Tracking Challenge. LEID method [2] is based on model (level sets) evolution approach. We have used the parameters provided with these methods for HeLa and GOWT1 sequences. Fluo-N2DH-SIM+ is a new dataset introduced in 2014 and the parameters for it are not provided with these methods, so we only compare our method with the KTH method and have tuned its parameters ourselves.

Table 1 lists the results for all three methods. Our method has better tracking score, higher TRA, than both these methods for all three sequences. It is able to detect more cells accurately (higher TP). For HeLa-01, even though our method detects more cells, it has lower SEG score due to worse pixel level segmentation. One shortcoming of our method is the higher false positive rate for most sequences, this is due to the fact that we try to include enough proposals so that recall among proposals would be high. Including too many proposals sometimes also leads to a higher number of errors in the links connecting cells in adjacent frames. We tried stronger pruning in the proposal generation stage to reduce the number of proposals, which led to lower number of FP, EA, EC and ED errors but increased FN and NS errors and had a lower overall tracking score. Using a better proposal generation method or cell classifier should alleviate this issue.

Fig. 3 shows one frame from the results produced by all methods; errors are highlighted using colored boxes. All methods suffer from false negative errors (red boxes), but our method has only one such error compared with five for KTH and six for LEID. Our method and LEID fail to split two cells which are in contact with each other leading to one undersegmentation error (green box). All methods suffer from few false positive errors especially near the image border as cells which are within 25 pixel wide band around image border are not marked in the ground truth.

## 7. CONCLUSION

We have presented a novel joint cell segmentation and tracking method, which utilizes cell segmentation proposals to create a directed acyclic graph and then iteratively finds the shortest path in this graph, which provides segmentations, tracks and events for individual cells. Experimental results show that our method achieves better performance than state of the art cell tracking methods. They also indicate that even when making greedy joint segmentation and tracking decisions, it improves performance compared with when segmentation and tracking are performed separately.

#### 8. REFERENCES

[1] J. Stegmaier, J. C. Otte, A. Kobitski, A. Bartschat, A. Garcia, G. U. Nienhaus, U. Strähle, and R. Mikut, "Fast Segmentation of Stained Nuclei in Terabyte-Scale, Time Resolved 3D Microscopy Image Stacks," *PLoS ONE*, 2014. 1

- [2] O. Dzyubachyk, W. van Cappellen, J. Essers, W. Niessen, and E. Meijering, "Advanced Level-Set-Based Cell Tracking in Time-Lapse Fluorescence Microscopy," *IEEE Transactions on Medical Imaging*, 2010. 1, 4
- [3] M. Maška, V. Ulman, D. Svoboda, et al., "A Benchmark for Comparison of Cell Tracking Algorithms," *Bioinformatics*, 2014. 1, 4
- [4] K. E. G. Magnusson and J. Jalden, "A Batch Algorithm using Iterative Application of the Viterbi Algorithm to Track Cells and Construct Cell Lineages," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, 2012. 1, 3, 4
- [5] X. Lou, M. Schiegg, and F. A. Hamprecht, "Active Structured Learning for Cell Tracking: Algorithm, Framework, and Usability," *IEEE Transactions on Medical Imaging*, 2014. 1
- [6] T. Kanade, Z. Yin, R. Bise, S. Huh, S. Eom, M. F. Sandbothe, and M. Chen, "Cell Image Analysis: Algorithms, System and Applications," in *IEEE Workshop on Applications of Computer Vision (WACV)*, 2011. 1
- [7] C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman, "Learning to Detect Cells Using Non-overlapping Extremal Regions," in *International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI)*. 2012. 1
- [8] F. Jug, T. Pietzsch, D. Kainmüller, J. Funke, M. Kaiser, E. van Nimwegen, C. Rother, and G. Myers, "Optimal Joint Segmentation and Tracking of Escherichia Coli in the Mother Machine," in *Bayesian and grAphical Models for Biomedical Imaging (MICCAI Workshop)*. 2014. 1, 3
- [9] M. Schiegg, P. Hanslovsky, C. Haubold, U. Koethe, L. Hufnagel, and F. A. Hamprecht, "Graphical Model for Joint Segmentation and Tracking of Multiple Dividing Cells," *Bioinformatics*, 2015. 1, 3
- [10] E. Türetken, X. Wang, C. Becker, C. Haubold, and P. Fua, "Globally Optimal Cell Tracking using Integer Programming," *arXiv:1501.05499*, 2015. 1, 3
- [11] D. Zhang, O. Javed, and M. Shah, "Video Object Segmentation through Spatially Accurate and Temporally Dense Extraction of Primary Object Regions," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013. 1
- [12] S. U. Akram, J. Kannala, M. Kaakinen, L. Eklund, and J. Heikkilä, "Segmentation of Cells from Spinning Disk Confocal Images Using a Multi-stage Approach," in *Asian Conference on Computer Vision (ACCV)*, 2014. 2, 4

[13] H. Kong, H. C. Akakin, and S. E. Sarma, "A Generalized Laplacian of Gaussian Filter for Blob Detection and Its Applications," *IEEE Transactions on Cybernetics*, 2013. 2