

The Candidate Multi-Cut for Cell Segmentation

Jan Funke¹, Chong Zhang², Tobias Pietzsch³, Stephan Saalfeld⁴

¹ Institut de Robotica i Informàtica Industrial, UPC, Barcelona

² Universitat Pompeu Fabra, Barcelona

³ MPI-CBG, Dresden

⁴ Janelia Research Campus, Ashburn VA

Abstract. Two successful approaches for the segmentation of biomedical images are (1) the selection of segment candidates from a merge-tree, and (2) the clustering of small superpixels by solving a Multi-Cut problem. In this paper, we introduce a model that unifies both approaches. Our model, the *Candidate Multi-Cut* (CMC), allows joint selection and clustering of segment candidates from a merge-tree. This way, we overcome the respective limitations of the individual methods: (1) the space of possible segmentations is not constrained to candidates of a merge-tree, and (2) the decision for clustering can be made on candidates larger than superpixels, using features over larger contexts. We solve the optimization problem of selecting and clustering of candidates using an integer linear program. On datasets of 2D light microscopy of cell populations and 3D electron microscopy of neurons, we show that our method generalizes well and generates more accurate segmentations than merge-tree or Multi-Cut methods alone.

1 Introduction

In this paper, we are addressing the problem of segmenting multiple objects in biomedical images, possibly against background. For this problem, *merge-tree methods* and *Multi-Cut methods* are amongst the best performing for a range of data modalities.

Both methods start with an initial set of superpixels, that is assumed to provide an oversegmentation. In **merge-tree methods**, these superpixels are iteratively merged to obtain a hierarchy of candidate segments. Amongst all candidates, a cost-minimal and non-overlapping subset is selected to yield a segmentation. The advantage of these methods is that they can consider candidates larger than the initial superpixels and thus use more meaningful features. A clear disadvantage is the limited expressiveness, as these methods require that each object is correctly segmented by one candidate in the merge-tree. Merge-tree methods demonstrate state-of-the-art performance for the segmentation of cells in 2D light microscopy [4,5,8]. Variations of this approach have also been successfully applied to the segmentation of neurons in volumes of electron microscopy [7,9], but were ultimately outperformed by Multi-Cut methods [12]. **Multi-Cut methods** consider finding a segmentation as an instance of a clustering problem on superpixels [1,10]. For that, edges in an adjacency graph of

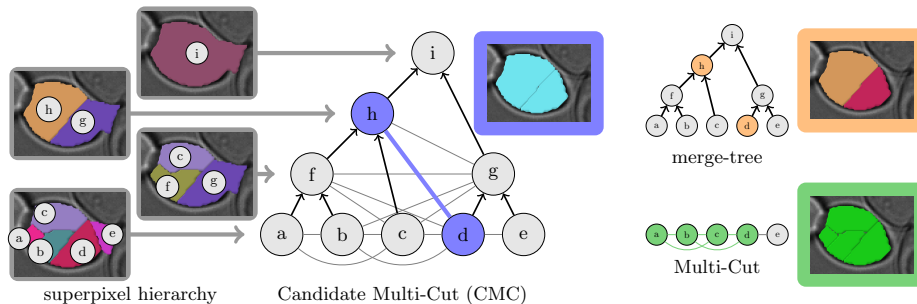


Fig. 1: Illustration of CMC compared to a merge-tree and Multi-Cut segmentation. In the superpixel hierarchy shown on the left, there is no candidate producing the desired segmentation of the cell. Consequently, the best merge-tree segmentation (which does not use the adjacency edges between candidates) is the selection of candidates d and h as two separate objects (highlighted in orange) which introduces a split error. CMC, in contrast, introduces adjacency edges between all neighboring candidates. Therefore, CMC can pick candidates d and h , and the adjacency edge (d, h) to merge these candidates (highlighted in blue). This way, a segment can be assembled that is not limited by the heuristic superpixel hierarchy. The Multi-Cut on the same initial superpixels could possibly achieve the same segmentation, but has to resort to smaller superpixels that lack structural information (highlighted in green).

superpixels are cut. A segmentation is obtained as the connected components of a cost-minimal cut, where constraints ensure that there is no path connecting two separated superpixels. In contrast to merge-tree methods, a correct segmentation can theoretically always be obtained, if the initial superpixels are oversegmenting. However, small superpixels carry the risk of not capturing meaningful features, like the local orientation of a cellular structure or the diameter of a cell, which could help to resolve ambiguities during inference. Multi-Cut methods are the current state of the art for the segmentation of neurons in electron microscopy volumes [6,10,1,12], but are outperformed by merge-tree methods on the segmentation of cells in light microscopy images [8].

The specific advantages and disadvantages of the two methods make them perform differently depending on the characteristics of a given dataset. In particular, none of the two methods performs well on *both* 2D segmentation of cells in light microscopy and 3D segmentation of neurons in electron microscopy.

To combine the advantages of both methods (larger feature context of merge-tree methods and the expressiveness of Multi-Cut methods), we introduce a segmentation model that jointly selects *and* clusters segment candidates from a merge tree. First, we obtain a merge-tree of segment candidates following greedy merging on initial superpixels. We then introduce adjacency edges between all adjacent candidates across all levels of the tree and train a classifier on ground-truth to obtain a cost for the selection of each candidate and each merge of an adjacency edge. Finally, we find the globally cost-minimal selection and cluster-

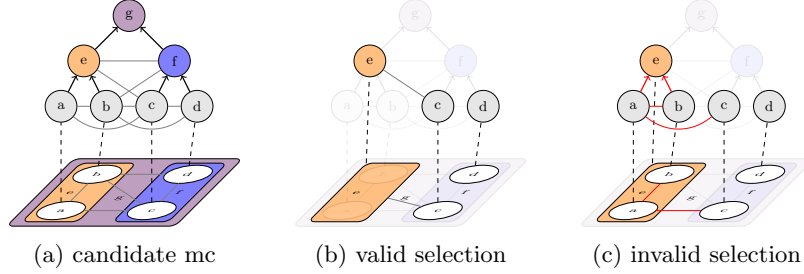


Fig. 2: Illustration of the proposed model. (a) Lower part: Given initial superpixels (a, b, c , and d), a heuristic region merging strategy is used to generate larger candidates (e, f , and g). Upper part: The *candidate region adjacency graph* (CRAG) is used to represent all candidates, their adjacencies across *all* levels (solid gray lines), and subset relations (black arrows). (b) A selection of candidates and adjacency edges producing a valid segmentation. Note that this segmentation can not be produced by the merge-tree method alone, since it can not merge e and c . A Multi-Cut method could generate the same segmentation by merging a , b , and c , but would not be able to exploit the features extracted on the larger candidate e . (c) An invalid selection of candidates and adjacency edges: Candidates a , b overlap with e and can not be selected with e at the same time. Further, a is merged both with b and c , but b and c are not merged, thus violating the transitivity of equivalence.

ing of candidates by solving an integer linear program (ILP) with a cutting-plane method. Compared to the standard formulations for merge-tree and Multi-Cut segmentations, our model has two advantages:

First, our model unifies the two method families including both as special cases (see Fig. 1). A merge-tree segmentation in the style of [4,8] can be obtained by simply omitting the adjacency edges. Similarly, the Multi-Cut formulation [1,10] can be obtained by omitting candidates other than the initial superpixels. In our model, however, a valid solution allows to select higher-level candidates and merge them with lower-level candidates. This allows us to train a classifier on more meaningful features that are only available for larger candidates. In contrast to merge-tree methods, however, we are not limited by the choice of the extracted candidates. Every possible segmentation given the initial superpixels can still be realized.

Second, by allowing candidates to not be selected at all, our formulation is particularly well suited to segment several foreground objects against background. This is required in 2D cell segmentation from light microscopy images where foreground objects are not tiling the plane. Our model has a dedicated cost contribution for the selection of candidates which depends on features of foreground objects. This is in contrast to previous Multi-Cut methods that required a post-processing step to filter background segments from foreground segments [14,13].

We evaluate our model in two different and dissimilar setups: First, on the segmentation of cells in 2D light microscopy (involving three datasets of different resolution, microscopy modality, and cell types), and second, on the reconstruction of neurons from electron microscopy volumes of neural tissue. Our model shows a consistent improvement over both merge-tree methods and Multi-Cut methods. This is of particular interest since neither merge-tree methods nor Multi-Cut methods deliver state-of-the-art performance on *both* datasets jointly.

2 Method

In order to combine the advantages of merge-tree methods and Multi-Cut methods, we introduce a model generalizing both: the *Candidate Multi-Cut* (CMC). In our model, the standard Multi-Cut formulation [1] is extended by considering additional candidate segments formed by merging initial superpixels to obtain a merge-tree. Various merging strategies can be used to obtain a merge-tree. In Section 3, we show two strategies for the datasets used here.

Given a merge-tree, we are addressing a segmentation problem in terms of the selection of candidate regions and adjacency edges (see Fig. 2a). Let $G = (V, E, S, f, g)$ be the *candidate region adjacency graph* (CRAG), where V is the set of all candidate regions (including the original superpixels), $E \subset V \times V$ the set of undirected edges indicating region adjacencies across all levels of the merge-tree, $S \subset V \times V$ the set of directed edges indicating subset relations of the candidate regions of the merge-tree, and $f : V \mapsto \mathbb{R}$ and $g : E \mapsto \mathbb{R}$ are cost functions for the selection of candidates and adjacency edges to merge, respectively. These costs are trained on features extracted for candidates and adjacency edges, see Section 3 for details. We encode a selection and merging of candidates with binary indicator variables $\mathbf{y} = (y_i \in \{0, 1\} \mid i \in V)$ and $\mathbf{m} = (m_e \in \{0, 1\} \mid e \in E)$. Setting $y_i = 1$ means that the candidate represented by node i is part of an object (as opposed to being considered background). Setting $m_{(i,j)} = 1$ states that the adjacent candidates i and j are part of the same object. By rewriting the costs $f(i)$ and $g(e)$ as vectors $\mathbf{f} = (f_i \in \mathbb{R} \mid i \in V)$ and $\mathbf{g} = (g_e \in \mathbb{R} \mid e \in E)$, such that they are congruent to \mathbf{y} and \mathbf{m} , we find a cost-minimal segmentation by minimizing

$$\langle \mathbf{y}, \mathbf{f} \rangle + \langle \mathbf{m}, \mathbf{g} \rangle. \quad (1)$$

However, not every assignment of the indicators \mathbf{y} and \mathbf{m} results in a valid segmentation. We ensure consistency with the introduction of three types of constraints: *overlap constraints* ensure that no overlapping candidates are selected at the same time, *incidence constraints* force incident candidates of selected adjacency edges to be selected as well, and *path constraints* state that for every adjacent pair of candidates that are not merged, there is no path of selected adjacency edges connecting them indirectly. See Fig. 2c for an illustration of these

constraints. More formally, we require:

$$\sum_{i \in C} y_i \leq 1 \quad \forall C \in \mathcal{C} \quad (2)$$

$$2m_{(i,j)} - y_i - y_j \leq 0 \quad \forall (i,j) \in E \quad (3)$$

$$\sum_{e \in P} m_e - m_{(i,j)} \leq |P| - 1 \quad \forall (i,j) \in E, \forall P \in \mathcal{P}_{(i,j)}. \quad (4)$$

For the overlap constraints in Eq. 2, $\mathcal{C} \subset 2^V$ denotes the set of all *conflict cliques*, i.e., a set of candidates that are mutually overlapping. For tree-shaped candidate subset graphs like those used here, the set \mathcal{C} simply contains all candidates of all paths from a leaf node to the root node. In the example in Fig. 2a, $\mathcal{C} = \{\{a, e, g\}, \{b, e, g\}, \{c, f, g\}, \{d, f, g\}\}$. The incidence constraints in Eq. 3 force the indicators y_i and y_j to be selected, if $m_{(i,j)}$ is selected. Finally, the path constraints in Eq. 4 ensure that if an adjacency edge was not selected (i.e., $m_{(i,j)} = 0$), there is no path of selected adjacency edges connecting them otherwise (i.e., the sum of selected edges along the path is strictly smaller than the length of the path). $\mathcal{P}_{(i,j)} \subset 2^E$ denotes the set of all paths on adjacency edges connecting candidates i and j . Since there are in general exponentially many paths connecting two candidates in a CRAG, we do not add these constraints initially. Following a cutting plane strategy, we solve an ILP without those constraints, and add violated constraints as needed and resolve until a consistent solution is found.

3 Experiments and Results

Datasets. We validate our method on four datasets (see examples in Fig. 3) which differ greatly in image resolution, microscopy modality, and cell type (shape, appearance and size): (A) phase contrast images of cervical cancer Hela cells [3] (B) bright field images of in-focus Diploid yeast cells [14] (C) bright field images of Fission yeast cells [11], and (D) 3D serial section electron microscopy volume of neural tissue [2].

Merge-Tree Generation. Our model requires us to generate a merge-tree to build the CRAG. For datasets A, B, and C, we ran a seeded watershed algorithm to obtain an initial oversegmentation on the pixel-wise boundary predictions from [8]. We obtained a merge-tree by iteratively merging neighboring superpixels with minimal merge score. The merge score we used is the product of the smaller region’s size and the median intensity of the boundary pixels separating the neighbors. From this merge-tree, we included all candidates that are the result of 5 or less merges in the CRAG, and added adjacency edges between each pair of touching candidates (considering a 4-neighborhood) across all levels. For dataset D, we generated merge-trees in a similar fashion for each 2D image of the stack individually, using the boundary predictions from [2]. In the CRAG, we only included the leaf nodes of the merge-tree and the largest candidates below a merge score that still favors oversegmentation.

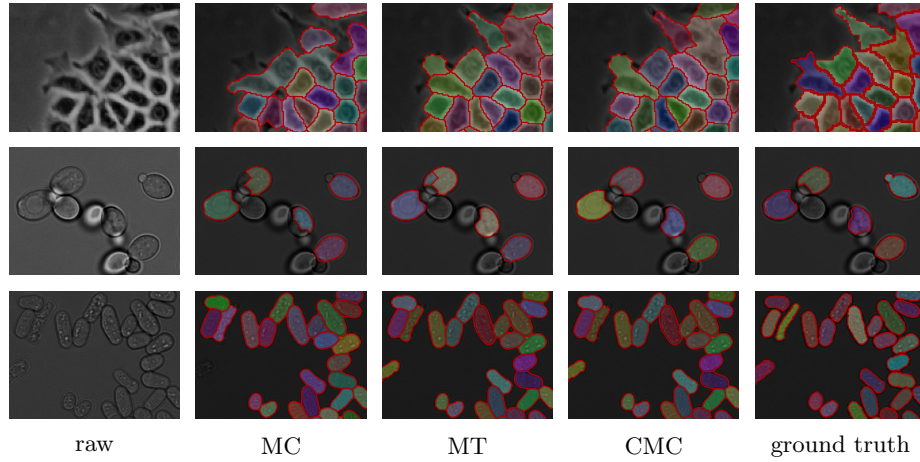


Fig. 3: Samples of the segmentations obtained on datasets A, B, and C (from top to bottom), using the Multi-Cut formulation (MC) [1], the merge-tree segmentation (MT) [8], and our model (CMC).

Features. For all experiments, we extracted the same features for each candidate: size, circularity, eccentricity, and a contour angle histogram with 16 bins. In addition to that, we extracted intensity features (sum, mean, variance, skewness, kurtosis, histogram with 20 bins, 7 histogram quantiles) from the raw and boundary prediction images for the whole candidate and for the contour pixels. For edge features, we used the size of the contact area between neighboring candidates, as well as mean, variance, and skewness of the intensities across the boundary pixels separating the neighbors. In addition to that, we added the absolute difference, min, max, and sum of each feature of the two adjacent candidates.

Training. The training consists of finding suitable costs for candidates and edges, given a training dataset with a CRAG and a ground truth image or volume. To learn the costs, we trained a random forest classifier by first generating a consistent CMC solution that is closest to the ground truth. For that, we assign each leaf node to the ground truth region with largest overlap. We then selected all candidates that cover leaf nodes with the same label as positive candidate instances. Similarly, we selected edges linking selected candidates with the same label as positive edge instances. All other candidates or edges were considered negative instances. Note that this training setup includes edges spanning different levels of the CRAG, and thus allows the classifier to relate features of candidates of different sizes. For datasets A, B, and C, we only trained the edge costs as described, and learned the node costs as suggested in [8].

Comparison. We compare the performance of our model on datasets A, B, and C against a Multi-Cut (MC) [1] and the merge-tree (MT) method described in [8]. For both, we use exactly the same candidates (for MT), edges of leaf nodes (for MC), features, and learning method as described above. On dataset D, we compare against a Multi-Cut (MC) and the assignment model (AM) proposed

	VOI			DS		
	split	merge	total	prec.	rec.	F-sc.
MC	0.404	0.644	1.048	0.987	0.775	0.868
MT	0.488	0.297	0.785	0.903	0.997	0.948
CMC	0.487	0.295	0.782	0.907	0.997	0.950
MT*	0.493	0.222	0.715	0.688	1.000	0.815
CMC*	0.433	0.226	0.659	0.995	0.997	0.996
DATASET A						

	VOI			DS		
	split	merge	total	prec.	rec.	F-sc.
MC	0.249	0.567	0.816	0.977	0.703	0.817
MT	0.423	0.219	0.643	0.895	0.945	0.919
CMC	0.424	0.215	0.638	0.896	0.945	0.920
MT*	0.349	0.108	0.457	0.792	1.000	0.884
CMC*	0.312	0.109	0.421	1.000	1.000	1.000
DATASET B						

	VOI			DS		
	split	merge	total	prec.	rec.	F-sc.
MC	0.225	0.238	0.463	0.991	0.821	0.898
MT	0.308	0.095	0.403	0.959	0.956	0.958
CMC	0.308	0.095	0.403	0.959	0.956	0.958
MT*	0.274	0.072	0.346	0.922	1.000	0.959
CMC*	0.268	0.072	0.340	1.000	1.000	1.000
DATASET C						

	VOI			RAND
	split	merge	total	
AM	3.644	0.772	4.416	0.950
MC	1.477	0.604	2.081	0.964
CMC	1.477	0.601	2.078	0.965
AM*	2.914	0.544	3.458	0.960
CMC*	1.041	0.569	1.610	0.966
DATASET D				

Table 1: Segmentation and detection results on datasets A, B, and C (2D light microscopy) and segmentation results on dataset D (3D electron microscopy). The best values of the used measures are highlighted: variation of information (VOI), detection score (DS) [8], and Rand index (RAND). We compare the standard Multi-Cut formulation (MC) [1] and the merge-tree segmentation (MT) [8] against our model (CMC) on datasets A, B, and C. For the neuron reconstruction in dataset D, we compare against the assignment model (AM) [7] and the Multi-Cut (MC). Entries with * have been found by matching the ground truth as closely as possible. They show the best achievable result given the initial superpixels on the respective dataset and thus show the performance limit of each method. Note that MT and AM have little room for improvement (using better features, for instance), whereas the CMC could benefit a lot from better features or different learning methods.

in [7], which is a variation of a merge-tree method. Again, we used the same candidate hierarchy for AM, the same adjacency edges for MC, and the same learning method as used for CMC. We trained each method on the same training subset of each dataset, and report results on the remaining images or volume.

Evaluation. For datasets A, B, and C, we report two measures, as in [8]: variation of information (VOI) to measure segmentation accuracy and detection score (DS) to measure the detection accuracy (see Table 1). The Candidate Multi-Cut improves accuracy on two of the three datasets, and produces the exact same result as MT on dataset B. In the same table, we also report the results of the best achievable segmentation of each method, denoted as MT* and CMC*. We found these segmentations in the same way we generated the training samples as described above. It can clearly be seen that CMC has a higher expressiveness, although it uses the same initial superpixels as MT. For dataset D, we report VOI and the Rand index (RAND) to measure segmentation accuracy (see Table 1). The Candidate Multi-Cut improves accuracy compared to both MC and

AM. Again, we report the best achievable segmentation as AM* and CMC* and find CMC to be more expressive than AM.

4 Conclusions

We presented a generalization of two successful segmentation methods for biomedical images and demonstrated that our model combines the advantages of both methods. This is a worthy contribution on its own as it broadens the range of applications that a single method can be used for. We find this especially remarkable as our method does not introduce new hyperparameters, and thus its advantages come basically “for free”. On top of that, we could report a modest improvement in segmentation and detection accuracy which shows that our generalized model is more than competitive with its specialized variants. Considering that our method has a higher expressiveness compared to merge-tree methods and larger context than Multi-Cut methods, we believe that further improvements are possible by learning node and edge costs in a more principled way. We believe that this, together with the design or learning of features that are more discriminative, is a promising direction for further research.

References

1. B. Andres, T. Kröger, K. L. Briggman, W. Denk, N. Korogod, G. Knott, U. Köthe, and F. A. Hamprecht. Globally Optimal Closed-surface Segmentation for Conectomics. In *ECCV*, 2012.
2. I. Arganda-Carreras, S. H. Seung, A. Vishwanathan, and D. R. Berger. SNEMI 3D: 3D Segmentation of Neurites in EM Images, 2013.
3. C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman. Learning to Detect Cells Using Non-overlapping Extremal Regions. In *MICCAI*, 2012.
4. C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman. Learning to detect partially overlapping instances. In *CVPR*, 2013.
5. C. Arteta, V. Lempitsky, J. A. Noble, and A. Zisserman. Detecting Overlapping Instances in Microscopy Images using Extremal Region Trees. *MIA*, 2016.
6. T. Beier, F. A. Hamprecht, and J. H. Kappes. Fusion Moves for Correlation Clustering. In *CVPR*, 2015.
7. J. Funke, B. Andres, F. A. Hamprecht, A. Cardona, and M. Cook. Efficient Automatic 3D-Reconstruction of Branching Neurons from EM Data. In *CVPR*, 2012.
8. J. Funke, F. A. Hamprecht, and C. Zhang. Learning to Segment: Training Hierarchical Segmentation under a Topological Loss. In *MICCAI*, 2015.
9. V. Kaynig, A. Vazquez-Reina, S. Knowles-Barley, M. Roberts, T. R. Jones, N. Kasthuri, E. Miller, J. Lichtman, and H. Pfister. Large-Scale Automatic Reconstruction of Neuronal Processes from Electron Microscopy Images. *TMI*, 2015.
10. T. Kröger, S. Mikula, W. Denk, U. Köthe, and F. A. Hamprecht. Learning to Segment Neurons with Non-Local Quality Measures. In *MICCAI*, 2013.
11. J. Y. Peng, Y. J. Chen, M. D. Green, S. A. Sabatino, S. L. Forsburg, and C. N. Hsu. PombeX: Robust Cell Segmentation for Fission Yeast Transillumination Images. *PLoS One*, 2013.

12. Snemi3d challenge, leader board, <http://brainiac2.mit.edu/SNEMI3D/leaders-board>, 03/2016.
13. J. Yarkony, C. Zhang, and C. C. Fowlkes. Hierarchical Planar Correlation Clustering for Cell Segmentation. In *EMMCVPR*, 2015.
14. C. Zhang, J. Yarkony, and F. A. Hamprecht. Cell Detection and Segmentation using Correlation Clustering. In *MICCAI*, 2014.