Introducing the Boundary-Aware loss for deep image segmentation

Minh Ôn Vû Ngoc, Yizi Chen, Nicolas Boutry, Joseph Chazalon, Edwin Carlinet, Jonathan Fabrizio, Clément Mallet, Thierry Géraud

Introducing the Boundary-Aware loss for deep image segmentation

Minh Ôn Vũ Ngọc *[1]
minh.on.vu.ngoc@lrde.epita.fr
Yizi Chen *[12]
yizi.chen@ign.fr
Nicolas Boutry[1]
Joseph Chazalon[1]
Edwin Carlinet[1]
Jonathan Fabrizio[1]
Clément Mallet[2]
Thierry Géraud[1]

[1] EPITA Research and Development Lab. LRDE
Le Kremlin-Bicêtre, France

[2] Univ. Gustave Eiffel, IGN-ENSG, LaSTIG
Saint-Mandé, France

Abstract

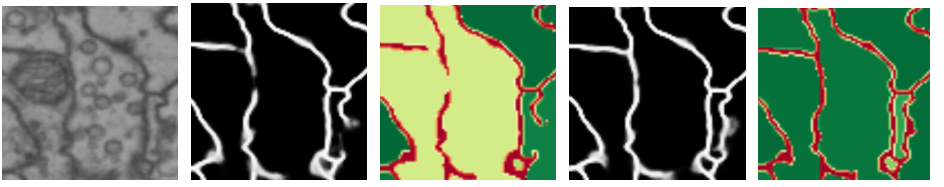
Most contemporary supervised image segmentation methods do not preserve the initial topology of the given input (like the closeness of the contours). One can generally remark that edge points have been inserted or removed when the binary prediction and the ground truth are compared. This can be critical when accurate localization of multiple interconnected objects is required. In this paper, we present a new loss function, called, Boundary-Aware loss (BALoss), based on the Minimum Barrier Distance (MBD) cut algorithm. It is able to locate what we call the *leakage pixels* and to encode the boundary information coming from the given ground truth. Thanks to this adapted loss, we are able to significantly refine the quality of the predicted boundaries during the learning procedure. Furthermore, our loss function is differentiable and can be applied to any kind of neural network used in image processing. We apply this loss function on the standard U-Net and DC U-Net on Electron Microscopy datasets. They are well-known to be challenging due to their high noise level and to the close or even connected objects covering the image space. Our segmentation performance, in terms of Variation of Information (VOI) and Adapted Rank Index (ARI), are very promising and lead to ∼15% better scores of VOI and ∼5% better scores of ARI than the state-of-the-art. The code of boundary-awareness loss is freely available at https://github.com/onvungocminh/MBD_BAL

1 Introduction

Image segmentation is one of the most fundamental building blocks in the field of computer vision and pattern recognition [36, 42, 45, 48]. From a graph theory perspective, image segmentation corresponds to the simultaneous partitioning and labeling of all pixels in the image [9]. Learning the global structure of the image is in particular mandatory to maintain the connection between pixels on the boundaries of the regions. Such borders are of superior

*These authors contributed equally to this work.

*This work has been conducted in the context of the SoDuCo project, grant ANR-18-CE38-0013.



(a) Original image (b) Output [48] (c) Segmentation [48] (d) Our output (e) Our segmentation

Figure 1: Importance of broken connection restoration for image segmentation and the topological correction of our method. (a) Neuron image. (b) U-Net boundary prediction [48]. (c) Resulting segmentation [48]. A leakage position leads to a confusion between regions. (d) Boundary prediction with our topological correction method. (e) Final segmentation.

importance in the case of dense image segmentation and multiple instances [46]: (i) objects share the information of contours, and (ii) erroneous or incomplete borders are detrimental for instance retrieval. In biomedical images, correctly delineated instances (membranes, vessels) are essential to provide accurate structure qualification in order to count the right number of instances for the medical usage [5, 12, 17, 60, 62].

Contemporary deep image segmentation methods can efficiently extract and combine deep features to achieve satisfactory per-pixel level accuracy [39]. However, they are prone to fail in recovering thin connections, finer details of structures, accurate location of borders [13], and subsequently the correct topology of images. The issue of broken pixels in the borders does not significantly influence the correctness of detected pixels but might cause catastrophic results, according to the number of instances (see Fig. 1(b) and 1(c)). While a significant amount of literature has focused on boundary refinement [35, 63, 72], often supported by semantic information [1, 10, 65, 70], few papers are dedicated to broken pixel restoration. To tackle such an issue, [40] considers the continuity of selected features as a loss function to connect pixels. However, this approach does not guarantee to get the correct topology of the detected objects. Topology preservation and broken pixel refinement in the boundary of the detected instances is alternatively addressed with the persistent homology framework [21], again in a supervised way. However, their method is sensitive to noise, all the broken pixels have not been fully located and the training procedure is slow. It requires prior knowledge on the number of instances, impossible in object counting challenges.

In order to better localize the broken pixels while maintaining the close properties of the objects, we propose a seeded deep-based segmentation approach where the object boundaries can be extracted during training and eventually encourage the network to focus on boundary structures. We do not make any assumption on the class, shape, and number of objects. The seeded approach alleviates this common limitation. The overview of our method is illustrated in Fig. 2. Intuitively speaking, we treat the boundary prediction of the network as a terrain function from which we extract all landscape ridges. We respectively consider each seeded pixel inside or outside (deduced automatically from the ground truth) as the foreground or background of its region. One common region-based boundary extraction strategy consists in relying on graph-cut methods [11, 49]. However, they have a limitation that produces short-cutting segmentation that cuts across the interior of an object due to the boundary term's bias. Lately, a new distance metric, the Minimum Barrier Distance (MBD) [51, 68], has shown to exhibit a rather limited sensitivity to noise, blur, and seed positions. It has been successfully applied to salient object detection and object segmentation [14, 41, 55, 68]. In this paper, we leverage the MBD information for identifying the boundary of the regions.

To alleviate current limitations, we introduce the Boundary-Aware loss function (*BALoss*), computed from the extracted boundaries, that penalizes the topologically-wrong pixels in the initial boundary prediction image. In our implementation, the *BALoss* is coupled with the binary cross-entropy loss to further refine the boundary details. The uniqueness of our *BALoss* is that it allows propagating the boundary information for regulating the network behavior while keeping a similar network structure. To the best of our knowledge, our method is the first one that addresses the boundary localization of instances for learning the structure of the image. We test our loss function on biomedical image (Electron Microscopy) datasets, which are highly challenging with the high level of noise and the multiplicity of interconnected objects. Training with the *BALoss*, the network performance is significantly improved, especially on the region boundaries (see Fig. 1(d) and 1(e)).

Our contributions in this paper are two-fold: (i) We propose a new Boundary-Aware loss function that comes along with a new seeded approach to correctly localize the boundaries of regions. (ii) We prove that our loss function can be adapted in deep segmentation networks and systematically improve the results of the network, with low sensitivity.

2 Related work

Image segmentation. Recently, various CNN-based segmentation methods have been proposed and achieved high pixel-level accuracy [7, 8, 16, 19, 27, 36, 53]. In the application of biomedical images, U-Net-based frameworks [37, 48, 73] consist of a contracting path and an expansive path to integrate multi-level features to improve the accuracy of localization yet without guarantee on topology preservation. Moreover, to increase the quality of detected segmentation, the current research is mainly focusing on processing boundaries as separated information, E.g., [34, 53, 66, 71] proposed a joint task framework combining both semantic segmentation and semantic boundary detection. [54] extracted and refined boundaries of instances by using a series of small boundary patches with higher image resolution. In [26, 29, 57], a boundary loss is proposed to measure the difference between the segmented and the ground truth boundaries. Though additional border pixels can boost the performance of segmentation results, those methods do not provide any information of the pixels that are required to refine and maintain the topological properties of objects [21].

Pixel-connectivity-preserving segmentation. Conditional Random Field (CRF) and Markov Random Field (MRF) are straightforward formalisms that involve neighboring pixel relationships in the training procedure [43, 44, 58, 67]. Specific functions can be formulated in order to preserve pixel connections or to include high-order cues (lines, object proposals, [38]). In [25, 64], the ConnNet proposed to learn and predict the pixel-pair connectivity to group or connect pixels. The MALIS and MALA methods [17, 56] focus on the affinity prediction based on maximin edges (it is a local comparison) so that it will ensure sufficiently low maximin edges to obtain a good segmentation after threshold whereas our method is to compute distance based on dynamics (which is global information). Furthermore, the MALIS is computed on small patches due to the limited speed of the network and noise level of the image, while our method can be applied to full-size images, which simplifies the procedure. Moreover, the Mosin [40] and IterNet [33] proposed to use a multiple-iterations framework to gradually improve the pixel connections in each iteration. Yet, these methods do not have a strong guarantee to achieve closed contours for images.

Topology-preserving segmentation. However, pixel-wise loss or pixel connectivity alone is not sufficient to capture the whole topological structure in the image [21]. To maintain the topological correctness of objects, [20] proposed to integrate topological signatures

to deep neural networks and to learn a task-optimal representation during training. [21, 22] designed topology-preserving losses based on persistent homology, which is differentiable and can be applied to any end-to-end deep neural networks in segmentation tasks. Still, training a neural network with these two losses does not solve the issues of border pixel localization, sensitivity to image noise, and slow training speed.

Seeded image segmentation. Another strategy that maintains the topology properties while partitioning the image is seeded image segmentation [12, 32, 50, 69]. This method is first used in the classical version of the watershed algorithm [9]. In [11, 18, 49], several multi-labeling image segmentation methods have been proposed to partition the image by using a random walker and graph-cut algorithms. However, these methods penalize segmented boundary length, thereby leading to a shrinkage bias problem. [47] proposed alternatively to combine the long-range information from the geodesic distance with edge information in a graph-cut optimization framework. Still, this method only used low-level features that do not correctly capture object boundaries. [61] proposed a learned watershed algorithm that trains the altitude function together with subsequent region assignment decisions in a reinforcement learning [59] style. This strategy is however greedy in terms of training data and computing time. In this paper, we keep the idea of seeded graph-cut based segmentation and push the work further by using a differentiable loss function that can propagate the properties through the whole training process. We rely on the Minimum Barrier Distance for that purpose.

3 Background

This section recalls the definition of the Minimum Barrier Distance (MBD) [51], which is the cornerstone of our paper. An image is associated with a graph in which vertices represent pixels on the image. Let $\pi = \langle ..., \pi_i, \pi_{i+1} ... \rangle$ denote the path of pixels on the graph. Also, the set of paths going from the vertex x' to the vertex x is denoted by $\Pi(x', x)$. The *barrier strength* τ of a path π in the given gray-level image u is defined as *the dynamic distance between the highest and lowest pixel values along the path*:

$$\tau_u(\pi) = \max_{p_i \in \pi} u(p_i) - \min_{p_i \in \pi} u(p_i). \tag{1}$$

The *Minimum Barrier Distance* d^{MB} between two vertices x' and x in u is then defined as the minimum of the barrier strengths of all the paths between two given vertices:

$$d_u^{\text{MB}}(x', x) = \min_{\pi \in \Pi(x', x)} \tau_u(\pi). \tag{2}$$

It is common to derive a distance map from the MBD. Given a minimum barrier strength function and a set X of seed points, a distance map S^{MBD} from every point x' of the image u to the set X of seed points can be computed by:

$$S_u^{\text{MBD}}(x', X) = \min_{x \in X} d_u^{\text{MB}}(x', x). \tag{3}$$

In the next section, we propose to use this distance to recover the object boundaries in the image, as a basis to compute our Boundary-Aware loss (*BALoss*) function.

4 Proposed method

4.1 Overview of the method

Our method is a seeded two-step approach (Fig. 2), in which the object boundaries can be extracted during training and eventually encourage the network to focus in boundary struc-

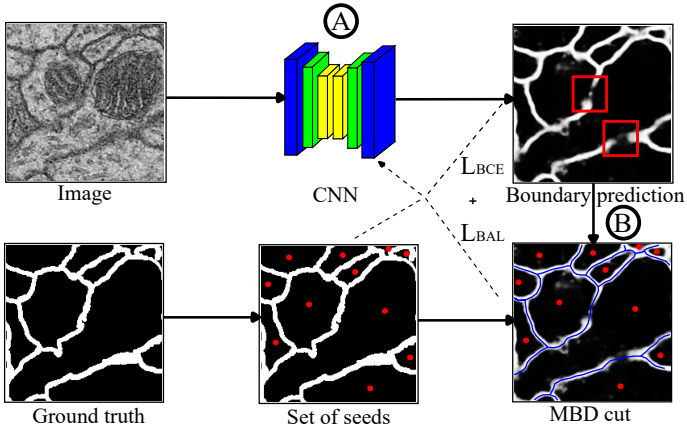


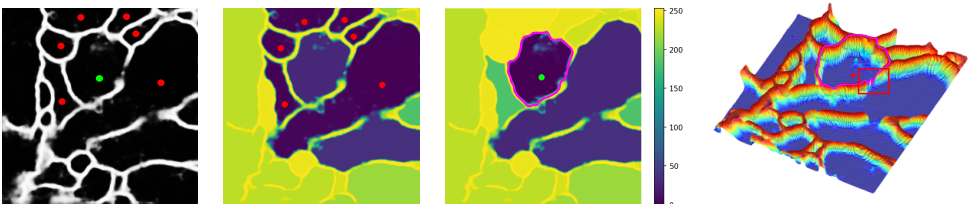
Figure 2: Overview of our approach. A set of seed points (red points) are placed from the ground truth. A boundary prediction image is the output of the CNN network (step A). Leakage positions are highlighted inside the red squares. We locate the boundaries of regions by using the MBD cut algorithm (blue lines), then compute the *BALoss* function (step B).

tures. To train the neural network, we are using the combination of the binary cross-entropy (BCE) and the *BALoss*. First, a convolutional neural network (CNN) predicts a boundary likelihood map from the original image. Secondly, the *BALoss* is computed from the extracted boundaries of regions by using the *MBD-cut* with a seed node provided inside each connected component of the label ground truth image. These seeds are the points that have the maximum Euclidean distance value w.r.t the boundary of each connected component. We respectively consider the inside/outside seeds of the region as foreground/background seeds, then compute a foreground/background distance map on the prediction image thanks to the MBD distance. Through comparing the values on these two distance maps, the boundaries of the regions are identified and the *BALoss* is derived by computing the pixel-wise error between the extracted boundaries and the ground truth image. The advantage of our *BALoss* function is that it helps the network to focus on important broken missing pixels on each region, thus preserving the topological structure of the image, without any class information.

4.2 MBD-cut

The boundaries of the regions are extracted using the *MBD-cut*. This preserves the topology of the image and measures the quality of the segmentation. MBD distance has a low sensitivity to noise, blur, and seed positions [51, 63]. In our approach, we push the idea of seeded graph-cut based segmentation further by using the high-level features computed from a convolutional neural network.

We denote the likelihood prediction map as u, and the ground truth label image as S. We respectively consider the seed point x_i inside the region S_i as the *foreground* seed and all the seeds x_j of the neighbor connected components S_j as the *background* seeds. We respectively compute the MBD distance map from the background/foreground seeds by using the front propagation approach [23] (see Eq. 3). The idea behind is that we consider the seed pixels as sources of water, the water can flow from source pixels to other pixels with a different priority which is determined by the MBD cost. We use the priority queue to keep track of the order of pixels to propagate the distance value to every pixel in the image (lower cost means earlier flow). The algorithm stops when all pixels in the image were scrutinized. The



(a) Likelihood image from the CNN network (b) Background MBD distance map (c) Foreground MBD distance map (d) The density map

Figure 3: (a) Prediction image from CNN network (green/red points are foreground/background seeds). (b) Background MBD distance map from the seed of neighbor regions. (c) Foreground MBD distance map and the MBD cut (pink contour). (d) The density map represents the values of the prediction image. Z-axis represents the value of pixels the prediction image. Leakage position is shown inside the highlight square.

complexity of our front propagation algorithm is $\mathcal{O}(n \log n)$, where n is the number of pixels in the image. Our method is efficiently computed, so that we can get the MBD distance map immediately from the set of the foreground and background seed points. The background/foreground MBD distance maps are illustrated in Fig. 3(b) and 3(c). After computing these maps, we are able to label the pixels as *background* or *foreground* based on their distances to the seed set. We also recover the boundary of the region C_i (pink contour in Fig. 3(d)). The segmented boundary is pivotal in computing the Boundary-Aware loss function.

4.3 Training using the Boundary-Aware loss

Most CNN-based segmentation networks use the binary cross-entropy (BCE) as a loss function. It is defined as a measure of the difference between two probability distributions for a given random variable or set of events [24]. BCE is known to be adapted to measure boundary shifts [10, 31]. Here, we present a new *BAL* function to enhance segmentation results and detail how to implement it. The *BAL* function is computed from the values of the binary extracted contour C_i of the region S_i using the *MBD-cut*. The total loss is the sum of the *BALoss* for every region:

$$L_{BAL}(u, GT) = \sum_{i \in N} BCE(u \circ C_i, GT \circ C_i), \tag{4}$$

where u represents the likelihood prediction map, GT is the boundary ground truth image, and $\circ$ is the Hadamard product.

Our loss function measures the segmentation quality for each region. We target to check if there are leakage positions on the boundaries, thereby ensuring the topological structure in the image. A high value of the Boundary-Aware loss corresponds to many broken connections. When the loss function L_{BAL} is zero, the prediction image is exactly the same as the ground truth image. The pixel-wise binary cross-entropy remains crucial to maintain the global information of every pixel in the image.

$$L_{total} = L_{BCE}(u, GT) + \alpha\, L_{BAL}(u, GT), \tag{5}$$

where α tunes the trade-off between both losses.

	Method	ARI ↑	VOI ↓	CREMI score ↓
CREMI dataset	MALIS [56]	0.6136 (±0.0185)	2.6983 (±0.1474)	1.0210 (±0.1033)
	MALA [17]	0.7663 (±0.0093)	2.4138 (±0.0843)	0.7510 (±0.0684)
	Mosin [40]	0.9185 (±0.0125)	0.6278 (±0.0434)	0.2262 (±0.0406)
	Iternet [63]	0.9240 (±0.0390)	0.7147 (±0.0208)	0.2330 (±0.0339)
	Segfix [68]	0.7416 (±0.0098)	1.9595 (±0.1122)	0.7115 (±0.0854)
	U-Net + Dice [62]	0.8929 (±0.0094)	0.4904 (±0.0650)	0.2291 (±0.0582)
	U-Net + Lovasz [6]	0.8935 (±0.0065)	0.4569 (±0.0175)	0.2205 (±0.0159)
	U-Net + BL [28]	0.8466 (±0.0285)	0.5279 (±0.0543)	0.2845 (±0.0484)
	U-Net + Topoloss [21]	0.9257 (±0.0012)	0.8021 (±0.0168)	0.2441 (±0.0156)
	DC U-Net [30]	0.9184 (±0.0212)	0.7215 (±0.0621)	0.2426 (±0.0591)
	DC U-Net + BAL	**0.9321 (±0.0026)**	**0.6647 (±0.0552)**	**0.2124 (±0.0515)**
	U-Net + wBCE [43]	0.9246 (±0.0023)	0.7562 (±0.0238)	0.2388 (±0.0221)
	U-Net + BAL	**0.9366 (±0.0064)**	**0.6893 (±0.0102)**	**0.2090 (±0.0105)**
ISBI12 dataset	MALIS [56]	0.6399 (±0.0178)	2.385 (±0.119)	0.9267 (±0.0872)
	MALA [17]	0.6239 (±0.0954)	3.0663 (±0.3966)	1.0738 (±0.3850)
	Mosin [40]	0.7833 (±0.0128)	1.1332 (±0.0799)	0.4955 (±0.0643)
	Iternet [63]	0.751 (±0.0379)	1.4614 (±0.1332)	0.6032 (±0.1144)
	Segfix [68]	0.7461 (±0.0217)	1.5555 (±0.0747)	0.6284 (±0.0652)
	U-Net + Dice [62]	0.7804 (±0.0348)	1.1325 (±0.0897)	0.4986 (±0.0804)
	U-Net + Lovasz [6]	0.7961 (±0.0088)	1.3239 (±0.0640)	0.5195 (±0.0523)
	U-Net + BL [28]	0.7621 (±0.0336)	1.4576 (±0.2064)	0.5888 (±0.1649)
	U-Net + Topoloss [21]	0.7944 (±0.0246)	1.1438 (±0.1860)	0.4849 (±0.1505)
	DC U-Net [30]	0.7518 (±0.0404)	1.3736 (±0.2432)	0.5839 (±0.1913)
	DC U-Net +BAL	**0.7983 (±0.0312)**	**1.0101 (±0.0567)**	**0.4514 (±0.0552)**
	U-Net+ wBCE [43]	0.7203 (±0.0417)	1.5644 (±0.0423)	0.6615 (±0.0720)
	U-Net + BAL	**0.8138 (±0.0191)**	**1.0557 (±0.0257)**	**0.4433 (±0.0290)**
ISBI13 dataset	MALIS [56]	0.5355 (±0.0187)	3.5324 (±0.0995)	1.2809 (±0.0848)
	MALA [17]	0.7713 (± 0.046)	3.4541 (±0.5073)	0.8887 (±0.4223)
	Mosin [40]	0.7504 (±0.0403)	1.5018 (±0.1379)	0.6122 (±0.1200)
	Iternet [63]	0.8686 (±0.0014)	1.5856 (±0.1412)	0.4564 (±0.1226)
	Segfix [68]	0.7995 (±0.0356)	2.4555 (±0.1452)	0.7016 (±0.1454)
	U-Net + Dice [62]	0.8078 (±0.0268)	1.2030 (±0.1436)	0.4808 (±0.1205)
	U-Net + Lovasz [6]	0.7993 (±0.0184)	1.3103 (±0.0953)	0.5128 (±0.0799)
	U-Net + BL [28]	0.6805 (±0.1048)	1.6832 (±0.4705)	0.7333 (±0.3688)
	U-Net + Topoloss [21]	0.8864 (±0.0265)	1.4623 (±0.0493)	0.4076 (±0.0584)
	DC U-Net [30]	0.7336 (±0.0137)	2.0900 (±0.1748)	0.7462 (±0.1314)
	DC U-Net + BAL	**0.8061 (±0.0096)**	**1.5523 (±0.2235)**	**0.5486 (±0.1808)**
	U-Net + wBCE [43]	0.8919 (±0.0164)	1.4270 (±0.1284)	0.3927 (±0.1169)
	U-Net + BAL	**0.9023 (±0.0237)**	**1.3761 (±0.0753)**	**0.3666 (±0.0754)**

Table 1: Quantitative results for different models on ISBI 13, ISBI 12 and CREMI datasets. Best scores are underlined while performance improvements brought by BAL are in bold.

5 Experiments

Our evaluation of the BAL performance is two-fold: an ablation study to assess its relevance and a comparison with current state-of-the-art segmentation methods.

Datasets. In our experiments, we use three highly challenging neuron Electron Microscopy Images: ISBI12 [3], ISBI13 [2], and CREMI [15]. The ISBI12 EM Segmentation Challenge [3] is a neuron segmentation challenge that contains 30 images which have size 512×512, ISBI13 [2] consists of 100 images with larger size 1024×1024. For the CREMI dataset [15], we test on volume A which has 125 slices with size 1250×1250.

Setting. We use a 3-fold cross-validation. To train the neural network, Adam [30] optimization algorithm is used with a learning rate of 10^{-4}, and an early-stop mechanism to stop

(a) BCE likelihood map

(b) BCE+BAL map

(c) Improvement

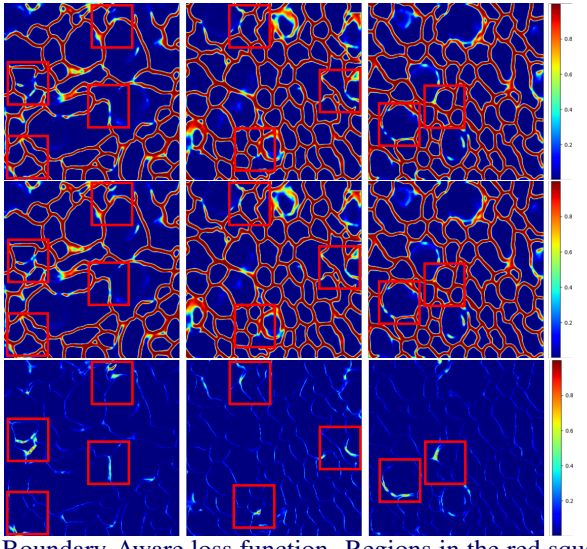


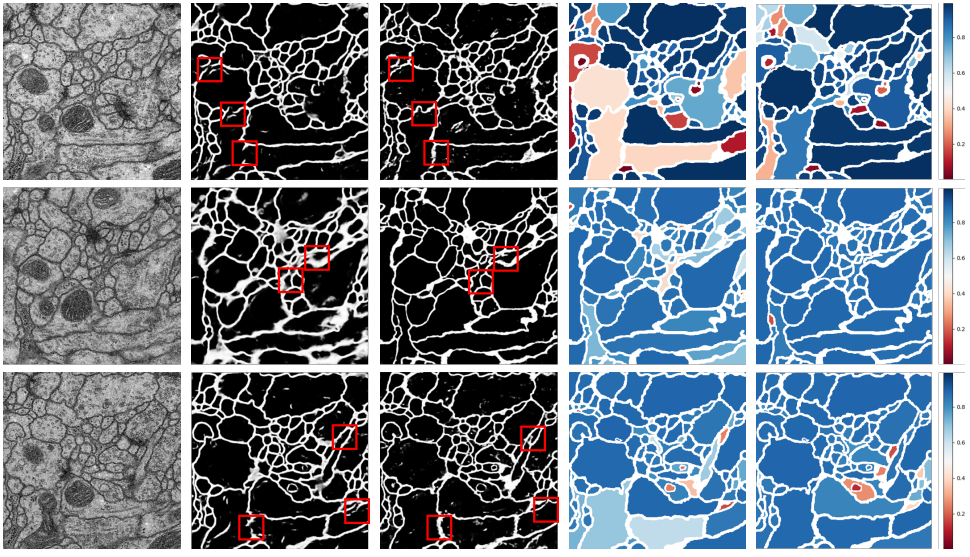
Figure 4: Positive impact of our Boundary-Aware loss function. Regions in the red squares demonstrate the closing action on the boundaries of the regions. The magnitude of the improvement is illustrated as from blue to red color.

the training when the validation loss is no longer decreasing during 15 epochs. α (Eq. 5) is set to a value of 0.1. We noted a limited sensitivity.

Evaluation metrics. The final goal of the work is to perform (dense) instance segmentation of cells, enabling to count them accurately for instances. The partition-based measures for instance segmentation provide a clear insight into the performance of the target applications. The creators of these three datasets suggest to use of the Variation of Information (VOI), Adapted Rand Index (ARI), and CREMI-score to facilitate the comparison with existing and future approaches. In detail, **ARI** is the maximal F-score of the foreground-restricted Rand index (a measure of similarity between two clusters). This version of the Rand index excludes the zero component of the original labels (background pixels of the ground truth). **VOI** is a measure of the distance between two clusters, closely related to mutual information. VOI is used to measure split and merge errors of the segmentation results. **CREMI-score** is the geometric mean of Adapted Rand Error (**ARAND**) and **VOI** scores, where **ARAND = 1-ARI**.

5.1 Ablation study

We show the impact of our *BALoss*, based on the improvement of the likelihood map predicted with a standard U-Net on the CREMI dataset [15]. We respectively predict the likelihood maps with and without the *BALoss* from the same pre-trained weights, then compute the difference between both. The impacts of the *BALoss* are illustrated in Fig. 4. We can see that the improvements concentrate mostly on the boundaries of regions. Since it is clear that our *BALoss* function penalizes the broken connection on the boundaries, the quality of boundaries increases in every training epoch. Further, the positive results are demonstrated in Tab. 1. It shows that *BALoss* can encourage to refine the boundary of regions, which leads to a significant improvement of the evaluation scores (ARI and VOI scores).



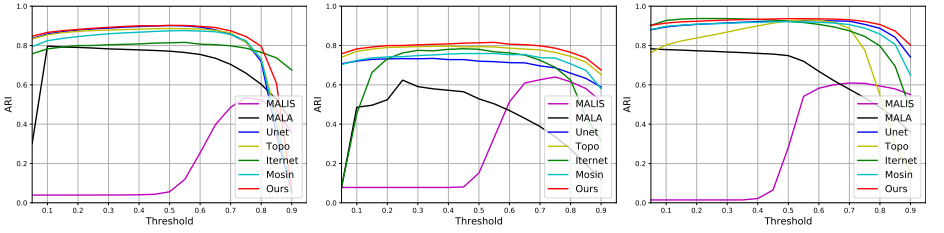
(a) Image (b) U-Net (c) Ours (d) Recall (U-Net) (e) Recall (Ours)

Figure 5: Qualitative results of our proposed method compared to a standard U-Net. (a) The original image. (b), (c) The segmentation results of U-Net and U-Net + BAL. Regions in the red squares demonstrate the closing action on the boundaries of the regions. (d), (e) The recall maps of U-Net and U-Net + BAL. The scale from bad to good segmentation is illustrated as from red to blue color.

5.2 Comparison with state-of-the-art methods

We compare our method with various open-sources state-of-the-art segmentation models (MALIS [56], MALA [17], U-Net [48], DC U-Net [37], Mosin [40], Iternet [53] and different loss functions (Dice [52], Lovasz Softmax [6], Boundary Loss [28] and Topoloss [21]).

Tab.1 shows the quantitative results for three different neuron image datasets, ISBI12, ISBI13, and CREMI. Our first observation is, our ***BALoss*** **outperforms the current state-of-the-art models for maintaining the topology-correctness in segmentation tasks** (the first block in the table). Our method (U-Net + BAL) outperforms the best score of these models, improving **+0.0172** (compared to Mosin [40] in CREMI), **+0.0522** (compared to Mosin [40] in ISBI2012) and **+0.0898** (compared to Iternet [53] in ISBI2013) in CREMI-score. According to **MALIS** and **MALA** results, it is worth noting that the high noise levels usually exist in biomedical images, the maximum affinity models for predicting every pair of pixels required to be highly tuned. Comparatively, our method correctly localizes pixels of the weak edges and enforces the network to correct pixels in the boundary (due to the implicit denoting which occurs during the training procedure). Nonetheless, the complexity of the affinity matrix between pixels is expensive $O(n^2)$ where our method is in $O(nlog(n))$. We also test the semantic segmentation method Segfix (which achieves state-of-the-art in natural images application), however, it is not designed to recover thin connections, finer details of structures such as the location of boundaries (membranes and tiny neurons). Our second observation is that ***BALoss*** **has a better performance compared to existing loss functions for topology-preserving purposes** (the second block in the table). The U-Net + BAL has the highest performance in CREMI-score, boosting up **+0.0115** (compared to



(a) ARI ISBI13 (b) ARI ISBI12 (c) ARI CREMI

Figure 6: ARI w.r.t segmentation thresholds. Our method achieves better the results across all threshold values.

U-Net + Lovasz [6] in CREMI), **+0.0416** (compared to U-Net + Topoloss [21] in ISBI2012) and **+0.0410** (compared to U-Net + Topoloss [21] in ISBI2013). ***BALoss* is model-agnostic for any neural networks** (the third block in the table). By applying *BALoss* into U-Net and DC U-Net, the CREMI-score increases **+0.0298** and **+0.0302** in CREMI, **+0.2182** and **+0.1325** in ISBI2012, **+0.0261** and **+0.1987** in ISBI2013, in original U-Net and DC U-Net respectively. This confirms that the *BALoss* can potentially be applied in different neural network architectures. Based on these observations, the common conclusion that can be drawn for the three datasets is that our *BALoss* function improves the segmentation results in the matter of the two topological correctness metrics (**∼5%** in average for ARI and **∼15%** in average for VOI). It is noteworthy that, especially for the ISBI12 dataset, our loss function significantly enhances the segmentation results. To conclude, our *BALoss* function is able to detect the leakage positions and refine boundaries, thereby improving segmentation results.

Fig. 5 shows qualitative results with and without the *BAL* function. We can see that our method combined with U-Net is able to enhance the broken connections and leads to more correct regions than with U-Net only. In particular, our *BALoss* successfully retrieves more regions all over the images in each iteration. Our method exhibits more consistency in terms of structure and topology and it is able to close contour regions in the image, hence improving the segmentation results. More results are shown in the supplementary material. We also investigate the relationships between ARI index and threshold values for likelihood images. For each image, we binarize the corresponding likelihood map with nine different values from 0.1 to 0.9 and compute the ARI index with different thresholds between the prediction and the ground truth images (Fig. 6). We note that our methods get the better area under the curve (AUC) compare to other methods. This is an advantage since the "best" threshold remains unknown and is image-dependent.

6 Conclusions

In this paper, we propose a new Boundary-Aware loss function embedded into a seeded approach for the purpose of image segmentation. Our loss function, based on the Minimum Barrier Distance cut algorithm is able to locate the object boundaries in the image and integrate the boundary information into the neural network. Moreover, our loss function focuses on penalizing the broken connections for each connected component, thereby recovering the closed contours in the image. Furthermore, the loss can be adapted in deep segmentation networks and systematically improve the results of the networks. Trained with the new loss function, our framework outperforms state-of-the-art methods in a matter of Variation of Information and Adapted Rand Index, with very limited parameter sensitivity.

References

[1] David Acuna, Amlan Kar, and Sanja Fidler. Devil is in the edges: Learning semantic boundaries from noisy annotations. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2019.

[2] I. Arganda-Carreras, H.S. Seung, A. Vishwanathan, and D. Berger. 3D segmentation of neurites in EM images challenge. In *International Symposium on Biomedical Imaging*, 2013.

[3] Ignacio Arganda-Carreras, Srinivas C Turaga, Daniel R Berger, Dan Cireşan, Alessandro Giusti, Luca M Gambardella, Jürgen Schmidhuber, Dmitry Laptev, Sarvesh Dwivedi, Joachim M Buhmann, et al. Crowdsourcing the creation of image segmentation algorithms for connectomics. *Frontiers in Neuroanatomy*, 9:142, 2015.

[4] M. Bai and R. Urtasun. Deep watershed transform for instance segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 5221–5229, 2017.

[5] Alberto Bailoni, Constantin Pape, Steffen Wolf, Anna Kreshuk, and Fred A Hamprecht. Proposal-free volumetric instance segmentation from latent single-instance masks. In *Pattern Recognition: DAGM GCPR 2020*, volume 12544 of *LNCS*, pages 331–344. Springer, 2021.

[6] Maxim Berman, Amal Rannen Triki, and Matthew B. Blaschko. The Lovász-Softmax loss: A tractable surrogate for the optimization of the intersection-over-union measure in neural networks. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 4413–4421. IEEE, 2018.

[7] Gedas Bertasius, Jianbo Shi, and Lorenzo Torresani. Semantic segmentation with boundary neural fields. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 3602–3610, 2016.

[8] Gedas Bertasius, Lorenzo Torresani, Stella X. Yu, and Jianbo Shi. Convolutional random walk networks for semantic image segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 858–866, 2017.

[9] Serge Beucher and Fernand Meyer. The morphological approach to segmentation: The watershed transformation. In *Mathematical Morphology in Image Processing*, pages 433–481. CRC Press, 2018.

[10] Shubhankar Borse, Ying Wang, Yizhe Zhang, and Fatih Porikli. InverseForm: A loss function for structured boundary-aware segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 5901–5911, 2021.

[11] Yuri Y. Boykov and Marie-Pierre Jolly. Interactive graph cuts for optimal boundary & region segmentation of objects in nd images. In *Proceedings of the International Conference on Computer Vision*, volume 1, pages 105–112, 2001.

[12] Lorenzo Cerrone, Alexander Zeilmann, and Fred A Hamprecht. End-to-end learned random walker for seeded image segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 12559–12568, 2019.

[13] Bowen Cheng, Ross Girshick, Piotr Dollar, Alexander C. Berg, and Alexander Kirillov. Boundary IoU: Improving object-centric image segmentation evaluation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 15334–15342, 2021.

[14] Krzysztof Chris Ciesielski, Robin Strand, Filip Malmberg, and Punam K Saha. Efficient algorithm for finding the exact minimum barrier distance. *Computer Vision and Image Understanding*, 123:53–64, 2014.

[15] CREMI. Miccai challenge on circuit reconstruction from electron microscopy images. *https://cremi.org*, 2017.

[16] Henghui Ding, Xudong Jiang, Ai Qun Liu, Nadia Magnenat Thalmann, and Gang Wang. Boundary-aware feature propagation for scene segmentation. In *Proceedings of the International Conference on Computer Vision*, pages 6819–6829, 2019.

[17] Jan Funke, Fabian Tschopp, William Grisaitis, Arlo Sheridan, Chandan Singh, Stephan Saalfeld, and Srinivas C Turaga. Large scale image segmentation with structured loss based deep learning for connectome reconstruction. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(7):1669–1680, 2018.

[18] Leo Grady. Random walks for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(11):1768–1783, 2006.

[19] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross Girshick. Mask R-CNN. In *Proceedings of the International Conference on Computer Vision*, pages 2961–2969, 2017.

[20] Christoph D. Hofer, Roland Kwitt, Marc Niethammer, and Andreas Uhl. Deep learning with topological signatures. In Isabelle Guyon, Ulrike von Luxburg, Samy Bengio, Hanna M. Wallach, Rob Fergus, S. V. N. Vishwanathan, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems 2017, December 4-9, 2017, Long Beach, CA, USA*, pages 1634–1644, 2017.

[21] Xiaoling Hu, Fuxin Li, Dimitris Samaras, and Chao Chen. Topology-preserving deep image segmentation. In Hanna M. Wallach, Hugo Larochelle, Alina Beygelzimer, Florence d'Alché-Buc, Emily B. Fox, and Roman Garnett, editors, *Advances in Neural Information Processing Systems 32: Annual Conference on Neural Information Processing Systems 2019, NeurIPS 2019, December 8-14, 2019, Vancouver, BC, Canada*, pages 5658–5669, 2019.

[22] Xiaoling Hu, Yusu Wang, Fuxin Li, Dimitris Samaras, and Chao Chen. Topology-aware segmentation using discrete morse theory. In *9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021*. OpenReview.net, 2021.

[23] Xiaoming Huang and Yujin Zhang. Water flow driven salient object detection at 180 fps. *Pattern Recognition*, 76:95–107, 2018.

[24] Shruti Jadon. A survey of loss functions for semantic segmentation. In *Proceedings of the IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology*, pages 1–7, 2020.

[25] Michael Kampffmeyer, Nanqing Dong, Xiaodan Liang, Yujia Zhang, and Eric P. Xing. Connnet: A long-range relation-aware pixel-connectivity network for salient segmentation. *IEEE Transactions on Image Processing*, 28(5):2518–2529, 2018.

[26] Davood Karimi and Septimiu E. Salcudean. Reducing the Hausdorff distance in medical image segmentation with convolutional neural networks. *IEEE Transactions on Medical Imaging*, 39(2):499–513, 2019.

[27] Tsung-Wei Ke, Jyh-Jing Hwang, Ziwei Liu, and Stella X. Yu. Adaptive affinity fields for semantic segmentation. In *European Conference on Computer Vision*, pages 587–602, 2018.

[28] Hoel Kervadec, Jihene Bouchtiba, Christian Desrosiers, Eric Granger, Jose Dolz, and Ismail Ben Ayed. Boundary loss for highly unbalanced segmentation. In *International conference on medical imaging with deep learning*, pages 285–296. PMLR, 2019.

[29] Hoel Kervadec, Jihene Bouchtiba, Christian Desrosiers, Eric Granger, Jose Dolz, and Ismail Ben Ayed. Boundary loss for highly unbalanced segmentation. In *Proceedings of the International Conference on Medical Imaging with Deep Learning*, pages 285–296, 2019.

[30] Diederik P. Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In Yoshua Bengio and Yann LeCun, editors, *3rd International Conference on Learning Representations, ICLR 2015, San Diego, CA, USA, May 7-9, 2015, Conference Track Proceedings*, 2015.

[31] Iasonas Kokkinos. Pushing the boundaries of boundary detection using deep learning. In *Proceedings of the International Conference on Learning Representations*, 2016.

[32] Tao Lei, Xiaohong Jia, Tongliang Liu, Shigang Liu, Hongying Meng, and Asoke K Nandi. Adaptive morphological reconstruction for seeded image segmentation. *IEEE Transactions on Image Processing*, 28(11):5510–5523, 2019.

[33] Liangzhi Li, Manisha Verma, Yuta Nakashima, Hajime Nagahara, and Ryo Kawasaki. Iternet: Retinal image segmentation utilizing structural redundancy in vessel networks. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, pages 3656–3665, 2020.

[34] Xiangtai Li, Xia Li, Li Zhang, Guangliang Cheng, Jianping Shi, Zhouchen Lin, Shaohua Tan, and Yunhai Tong. Improving semantic segmentation via decoupled body and edge supervision. In Andrea Vedaldi, Horst Bischof, Thomas Brox, and Jan-Michael Frahm, editors, *Computer Vision - ECCV 2020 - 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part XVII*, volume 12362 of *Lecture Notes in Computer Science*, pages 435–452. Springer, 2020. doi: 10.1007/978-3-030-58520-4_26.

[35] Yun Liu, Ming-Ming Cheng, Xiaowei Hu, Kai Wang, and Xiang Bai. Richer convolutional features for edge detection. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2017.

[36] Jonathan Long, Evan Shelhamer, and Trevor Darrell. Fully convolutional networks for semantic segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 3431–3440, 2015.

[37] Ange Lou, Shuyue Guan, and Murray H. Loew. DC-UNet: Rethinking the U-Net architecture with dual channel efficient CNN for medical image segmentation. In *Medical Imaging 2021: Image Processing*, volume 11596, page 115962T, 2021.

[38] Michael Maire, Stella X. Yu, and Pietro Perona. Object detection and segmentation from joint embedding of parts and pixels. In *Proceedings of the International Conference on Computer Vision*, pages 2142–2149, 2011.

[39] Shervin Minaee, Yuri Y. Boykov, Fatih Porikli, Antonio J. Plaza, Nasser Kehtarnavaz, and Demetri Terzopoulos. Image segmentation using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021. Early Access.

[40] Agata Mosinska, Pablo Marquez-Neila, Mateusz Koziński, and Pascal Fua. Beyond the pixel-wise loss for topology-aware delineation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 3136–3145, 2018.

[41] Minh Ôn Vũ Ngoc, Nicolas Boutry, Jonathan Fabrizio, and Thierry Géraud. A minimum barrier distance for multivariate images with applications. *Computer Vision and Image Understanding*, 197:102993, 2020.

[42] Hyeonwoo Noh, Seunghoon Hong, and Bohyung Han. Learning deconvolution network for semantic segmentation. In *Proceedings of the International Conference on Computer Vision*, pages 1520–1528, 2015.

[43] Sebastian Nowozin and Christoph H. Lampert. Global connectivity potentials for random field models. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 818–825. IEEE, 2009.

[44] Martin Ralf Oswald, Jan Stühmer, and Daniel Cremers. Generalized connectivity constraints for spatio-temporal 3d reconstruction. In *European Conference on Computer Vision*, pages 32–46. Springer, 2014.

[45] Yanwei Pang, Yazhao Li, Jianbing Shen, and Ling Shao. Towards bridging semantic gap to improve semantic segmentation. In *Proceedings of the International Conference on Computer Vision*, pages 4230–4239, 2019.

[46] Lorenzo Porzi, Samuel Rota Bulo, and Peter Kontschieder. Improving panoptic segmentation at all scales. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 7302–7311, 2021.

[47] Brian L. Price, Bryan Morse, and Scott Cohen. Geodesic graph cut for interactive image segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 3161–3168. IEEE, 2010.

[48] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*, volume 9351 of *LNCS*, pages 234–241. Springer, 2015.

[49] Carsten Rother, Vladimir Kolmogorov, and Andrew Blake. "GrabCut" interactive foreground extraction using iterated graph cuts. *ACM Transactions on Graphics*, 23(3): 309–314, 2004.

[50] Ali Kemal Sinop and Leo Grady. A seeded image segmentation framework unifying graph cuts and random walker which yields a new algorithm. In *Proceedings of the International Conference on Computer Vision*, pages 1–8. IEEE, 2007.

[51] Robin Strand, Krzysztof Chris Ciesielski, Filip Malmberg, and Punam K Saha. The minimum barrier distance. *Computer Vision and Image Understanding*, 117(4):429–437, 2013.

[52] Carole H. Sudre, Wenqi Li, Tom Vercauteren, Sebastien Ourselin, and M. Jorge Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep learning in medical image analysis and multimodal learning for clinical decision support (Proc. of DLMIA 2017, ML-CDS 2017)*, volume 10553 of *LNCS*, pages 240–248. Springer, 2017.

[53] Towaki Takikawa, David Acuna, Varun Jampani, and Sanja Fidler. Gated-scnn: Gated shape cnns for semantic segmentation. In *Proceedings of the International Conference on Computer Vision*, pages 5229–5238, 2019.

[54] Chufeng Tang, Hang Chen, Xiao Li, Jianmin Li, Zhaoxiang Zhang, and Xiaolin Hu. Look closer to segment better: Boundary patch refinement for instance segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 13926–13935, 2021.

[55] Wei-Chih Tu, Shengfeng He, Qingxiong Yang, and Shao-Yi Chien. Real-time salient object detection with a minimum spanning tree. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 2334–2342, 2016.

[56] Srinivas C. Turaga, Kevin L. Briggman, Moritz Helmstaedter, Winfried Denk, and H. Sebastian Seung. Maximin affinity learning of image segmentation. In Yoshua Bengio, Dale Schuurmans, John D. Lafferty, Christopher K. I. Williams, and Aron Culotta, editors, *Advances in Neural Information Processing Systems 22: 23rd Annual Conference on Neural Information Processing Systems 2009. Proceedings of a meeting held 7-10 December 2009, Vancouver, British Columbia, Canada*, pages 1865–1873. Curran Associates, Inc., 2009.

[57] Chi Wang, Yunke Zhang, Miaomiao Cui, Jinlin Liu, Peiran Ren, Yin Yang, Xuansong Xie, XianSheng Hua, Hujun Bao, and Weiwei Xu. Active boundary loss for semantic segmentation. *CoRR*, abs/2102.02696, 2021.

[58] Jan D Wegner, Javier A Montoya-Zegarra, and Konrad Schindler. A higher-order crf model for road network extraction. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 1698–1705, 2013.

[59] Marco Wiering and Martijn Van Otterlo. *Reinforcement Learning*, volume 12 of *Adaptation, Learning, and Optimization*. Springer, 2012.

[60] S. Wolf, Y. Li, C. Pape, A. Bailoni, A. Kreshuk, and F. A. Hamprecht. The semantic mutex watershed for efficient bottom-up semantic instance segmentation. In *European Conference on Computer Vision*, pages 208–224, 2020.

[61] Steffen Wolf, Lukas Schott, Ullrich Kothe, and Fred Hamprecht. Learned watershed: End-to-end learning of seeded segmentation. In *Proceedings of the International Conference on Computer Vision*, Oct 2017.

[62] Steffen Wolf, Alberto Bailoni, Constantin Pape, Nasim Rahaman, Anna Kreshuk, Ullrich Köthe, and Fred A. Hamprecht. The mutex watershed and its objective: Efficient, parameter-free graph partitioning. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020. Early Access.

[63] Saining Xie and Zhuowen Tu. Holistically-nested edge detection. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 1395–1403, 2015.

[64] Ziyun Yang, Somayyeh Soltanian-Zadeh, and Sina Farsiu. Biconnet: An edge-preserved connectivity-based approach for salient object detection. *CoRR*, abs/2103.00334, 2021.

[65] Zhiding Yu, Weiyang Liu, Yang Zou, Chen Feng, Srikumar Ramalingam, B. V. K. Vijaya Kumar, and Jan Kautz. Simultaneous edge alignment and learning. In *European Conference on Computer Vision*, 2018.

[66] Yuhui Yuan, Jingyi Xie, Xilin Chen, and Jingdong Wang. Segfix: Model-agnostic boundary refinement for segmentation. In *European Conference on Computer Vision*, pages 489–506. Springer, 2020.

[67] Yun Zeng, Dimitris Samaras, Wei Chen, and Qunsheng Peng. Topology cuts: A novel min-cut/max-flow algorithm for topology preserving segmentation in N-D images. *Computer Vision and Image Understanding*, 112(1):81–90, 2008.

[68] Jianming Zhang, Stan Sclaroff, Zhe Lin, Xiaohui Shen, Brian Price, and Radomir Mech. Minimum barrier salient object detection at 80 fps. In *Proceedings of the International Conference on Computer Vision*, pages 1404–1412, 2015.

[69] Juyong Zhang, Jianmin Zheng, and Jianfei Cai. A diffusion approach to seeded image segmentation. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 2125–2132. IEEE, 2010.

[70] Yifan Zhao, Jia Li, Yu Zhang, and Yonghong Tian. Multi-class part parsing with joint boundary-semantic awareness. In *Proceedings of the International Conference on Computer Vision*, 2019.

[71] Mingmin Zhen, Jinglu Wang, Lei Zhou, Shiwei Li, Tianwei Shen, Jiaxiang Shang, Tian Fang, and Long Quan. Joint semantic segmentation and boundary detection using iterative pyramid contexts. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pages 13666–13675, 2020.

[72] Peng Zhou, Brian Price, Scott Cohen, Gregg Wilensky, and Larry S. Davis. Deepstrip: High-resolution boundary refinement. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, 2020.

[73] Zongwei Zhou, Md Mahfuzur Rahman Siddiquee, Nima Tajbakhsh, and Jianming Liang. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pages 3–11. Springer, 2018.