Superpixel Generation by the Iterative Spanning Forest Using Object Information

Felipe C. Belém* and Alexandre X. Falcão
University of Campinas
São Paulo, 13083-852, Brazil
Email: {felipe.belem, afalcao}@ic.unicamp.br

Silvio Jamil F. Guimarães
Pontifical Catholic University of Minas Gerais
Minas Gerais, 31980-110, Brazil
Email: sjamil@pucminas.br

***Abstract*—Superpixel segmentation methods aim to partition the image into homogeneous connected regions of pixels (*i.e.*, superpixels) such that the union of its comprising superpixels precisely defines the objects of interest. However, the homogeneity criterion is often based solely on color, which, in certain conditions, might be insufficient for inferring the extension of the objects (*e.g.*, low gradient regions). In this dissertation, we address such issue by incorporating prior object information — represented as monochromatic object saliency maps — into a state-of-the-art method, the *Iterative Spanning Forest* (ISF) framework, resulting in a novel framework named *Object-based ISF* (OISF). For a given saliency map, OISF-based methods are capable of increasing the superpixel resolution within the objects of interest, whilst permitting a higher adherence to the map's borders, when color is insufficient for delineation. We compared our work with state-of-the-art methods, considering two classic superpixel segmentation metrics, in three datasets. Experimental results show that our approach presents effective object delineation with a significantly lower number of superpixels than the baselines, especially in terms of preventing superpixel leaking.**

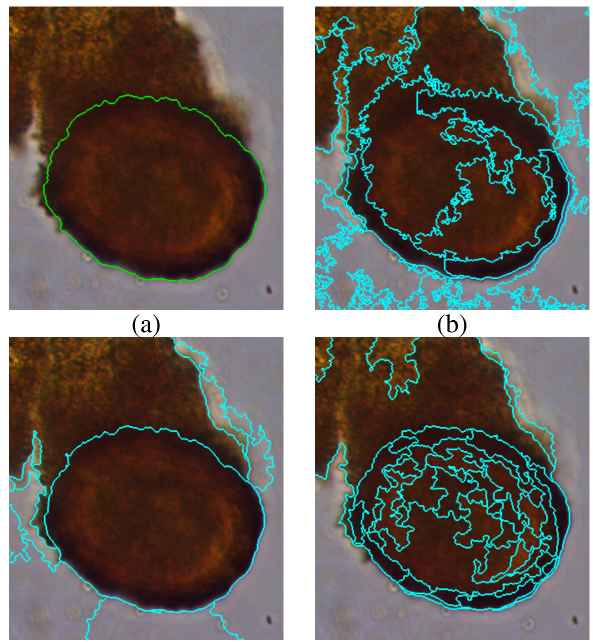
I. Introduction

Due to the absence of fatigue, machines are useful for solving tasks involving a large amount of data, such as automatic object delineation. Although recent works have investigated machine learning strategies [2], [3], they often rely on a substantial quantity of annotated images for training and adaptation; therefore, it is of utmost importance to develop methods which can achieve good results in such scarcity of certified data.

A major class of algorithms partitions the image in numerous disjoint groups of connected pixels (*i.e.*, superpixels) driven by a particular concept of homogeneity (*e.g.*, color similarity). The objective of superpixel segmentation methods is to represent any object in the scene precisely by the union of its superpixels (*i.e.*, semantic segmentation), requiring fewer as possible [4]. One can see that, not only the task of delineation is simplified, but superpixels carry more semantic information (of the object) than pixels. Due to such properties, many works use superpixels as primitives for their respective solutions [2], [3], [5], [6].

However, due to the lack of prior object information (*e.g.*, location), the homogeneity criterion — often defined by heuristics — may not be sufficient in certain circumstances.

*M.Sc. dissertation



(a) (b) (c) (d)

Fig. 1. (a) Original image with the ideal delineation in green. Superpixel segmentation by (b) LSC [1]; and our approach considering object (c) under-sampling; and (d) over-sampling. The number of desired superpixels was 50, and the images were zoomed for visualization purposes.

By grouping similar pixels, the frontiers of the objects of interest could be safely inferred when their contrast with the background is significant. But, as the color transitions tend to become less perceivable, so are the indications of object borders — leading to superpixel leaking and errors in delineation. Although one may argue that a simple increase in the number of superpixels could overcome such issue, this solution may result in a higher workload for subsequent methods within an application. One may see that the prior information regarding the objects' locations — and, thus, their extension — could assist in the delineation task, avoiding the aforementioned consequences. Given an indication of the objects' limits, superpixel leaking could be prevented by constraining the superpixels' growth and, therefore, enforcing the intersection between the objects' frontiers and superpixels' borders. The majority of state-of-the-art methods do not permit the inclusion of prior object information; moreover, the necessary modifications for overcoming such incapability might not be intuitive.

In this dissertation, we propose a generalization of the *Iterative Spanning Forest* (ISF) [7] framework, which allows the inclusion of prior object information, represented as object saliency maps — *i.e.*, monochromatic images whose pixel value indicates the proportional likelihood of belonging to an object of interest. Named *Object-based ISF* (OISF) [8], our framework maintains the properties of ISF, and straightforwardly incorporates the saliency information in each step of the pipeline. OISF-based methods can alter the superpixel resolution within the objects of interest, whilst permitting a better delineation of the object (Figure 1); being a suitable tool for distinct applications, such as feature extraction and semantic segmentation. Note that such improvement can be achieved without the need to increase the number of superpixels. Moreover, we also propose two object-based seed sampling strategies: the *Object Geodesic Grid Sampling* (OGRID) [8], and the *Object Saliency Map by Ordered eXtraction* (OSMOX) [9]. As stated in [9], the simple consideration of one of both leads to a significant improvement for object delineation.

Our contributions can be summarized as follows: (i) a new paradigm for superpixel segmentation; (ii) a novel object-based superpixel segmentation framework; (iii) two object-based superpixel methods; and (iv) a study of the impacts of object-based seed sampling for superpixel segmentation.

This paper is organized as follows. In Section II, we analyze the state-of-the-art in superpixel segmentation and, in Section 2, our approach is presented. The performance of OISF-based approaches is shown in Section IV, alongside a brief discussion of their limitations. Finally, in Section V we conclude and draw possible future work that can be done from ours.

II. Related Works

In this section, we briefly discuss about state-of-the-art proposals for superpixel segmentation, starting by those which do not take into account any prior object information (Subsection II-A), to those which use a particular definition of the latter (Subsections II-B and II-C). Specifically, we discuss the proposals in which are major representatives in their category, with respect to efficiency, popularity, and uniqueness. One can refer to notable surveys [10]–[12] for more information about superpixel segmentation.

The term "superpixel" was first coined in [13], and has been widely used in many applications [11]. Many concepts are used interchangeably between different paradigms [11], [14]; however, in this work, we consider the possibility of generating an arbitrary high number of superpixels (*i.e.*, 500 or more) as primordial for distinction. Aside from the previous two, many authors [1], [4], [15] elect other desirable properties for superpixel segmentation: (i) pixels must be assigned to a unique superpixel; (ii) superpixels must be represented as connected regions of pixels, with effective boundary adherence; and (iii) superpixels should be generated efficiently. Furthermore, as stated in [4], the aforementioned properties should be achieved with a minimum quantity of superpixels. Other properties are commonly associated with aesthetics [10], [11], which are often not relevant.

A. Classic Methods

Many state-of-the-art methods present efficient object delineation performance; however, they are driven by evidence (*e.g.*, color variation), which might be misleading for low-gradient borders. Also, the latter definition not only includes irrelevant (*i.e.*, non-object) borders but assigns the same relevance to the objects'. Finally, even if in possession of prior object information, such methods do not consider it during the superpixel generation, and the necessary modifications for overcoming this drawback might not be intuitive nor possible.

Due to the popularity of the *Simple Linear Iterative Clustering* (SLIC) [15], clustering-based methods constitute a significant group in superpixel segmentation. Although they are usually faster than other approaches, their strategy often leads to connectivity violation [1], [15], or strict rules of seed displacement [16], [17]. Aside from SLIC, the *Linear Spectral Clustering* (LSC) [1] is another example, with effective delineation at the expense of being slower than the primer. Still, LSC requires a post-processing step for ensuring connected superpixels.

Another group of methods models the problem of object delineation as the task of edge insertion/removal in an undirected weighted graph. The *Entropy Rate Superpixels* (ERS) [4] is a popular method with good object delineation; however, it is computationally expensive even if greedy strategies are considered.

Methods based in connectivity, by definition, overcomes both drawbacks of clustering-based ones, whilst being faster than most edge-based algorithms. As an example, the *Simple Non-Iterative Clustering* (SNIC) [16] is an extension of SLIC which generates connected superpixels in a single iteration; however, its performance its only slightly better than SLIC. The *Image Foresting Transform* (IFT) [18] is a framework which computes an optimum-path forest in an image graph; and IFT-based superpixel methods, such as the *Iterative Spanning Forest* (ISF) [7] framework, often present top delineation performance in many datasets and applications [7]–[9], [19]–[21].

B. Deep Learning Methods

Recent advances in deep learning have lead to proposals in the context of superpixel segmentation. Often, such works argument that the inclusion of deep features (*i.e.*, of the object) improves the segmentation results [22]; but this statement requires more evidence, since their performance is not superior to hand-crafted solutions. Moreover, in contrast to the latter, the number of annotated examples for training is crucial for deep-learning approaches, and they are not easily extended to distinct domains without requiring more examples for adaptation.

For all we know, the *Superpixel Sampling Network* (SSN) [23] is the first superpixel generation network. It uses a modification of SLIC in order to be differentiable, but inherits the previous deep learning drawbacks and those of SLIC as well. Although the authors in [22] and [24] classify their works as superpixel methods, we categorize them as feature engineering proposals. As one may see, each core superpixel method remains unaltered, whilst parts of the input are drastically altered — specially through deep learning — for the improvement of the segmentation method.

C. Saliency-based Methods

To the best of our knowledge, aside from our proposal, the work of [25] is the only method that considers saliency values for generating superpixels. It performs a bottom-up merging approach from an initial segmentation with a higher number of superpixels than desired. Its drawbacks are listed as follows: (i) it considers a particular saliency definition; (ii) the errors from the initial segmentation are propagated to the final one; and (iii) it lacks user-control over the superpixel displacement and shape.

III. Object-based Superpixel Segmentation

In this section, we present the *Object-based ISF* (OISF) [8] framework, which is a generalization of the *Iterative Spanning Forest* (ISF) [7] framework. The ISF is an efficient three-staged superpixel segmentation framework in which each component can be defined independently, being the major reason for recent publications shortly after its own [19]–[21]. In OISF, the independence between steps is maintained, with the benefit of incorporating prior object information (*i.e.*, saliency map) within each one (Figure 2); for a suitable set of parameters, OISF obtains segmentations equivalent to ISF's. Therefore, we detail: (i) the initial seed sampling (Subsection III-A); (ii) the superpixel generation (Subsection III-B); and (iii) the seed recomputation (Subsection III-C); by presenting object-based strategies.

A. Seed Sampling

For a given estimation of the object location (*e.g.*, object saliency maps) and a total number of k seeds, one could alter the displacement of the latter through a percentage ρ of object seeds. Thus, higher values of ρ promote a higher superpixel resolution within the objects of interest, due to oversampling — and it is analogous for lower values of ρ (Figure 3). As an example, feature extraction solutions may benefit from object oversampling since a higher superpixel resolution captures most of the objects' nuances. In contrast, a lower resolution may simplify semantic segmentation tasks by reducing the number of superpixels to be evaluated. The following methods describe the necessary steps for obtaining $k_o = \rho k$ object seeds; however, for $k_b = (1-\rho)k$ background ones, the exact same procedure is applied over the complement of the saliency map.

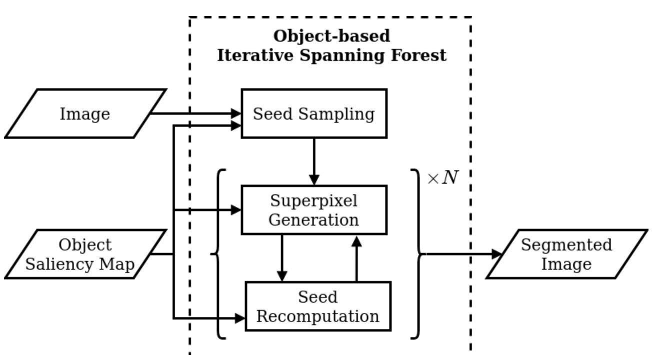


Fig. 2. Pipeline of our proposed object-based superpixel segmentation framework.

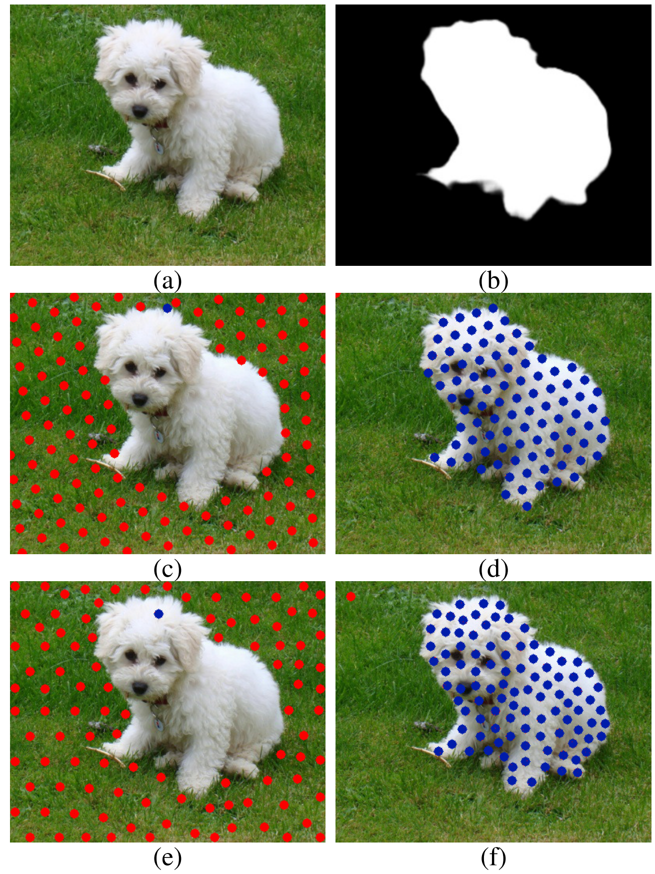


Fig. 3. Object-based seed sampling. (a) Original image. (b) Object saliency map [26]. (c-d) OGRID; and (d-e) OSMOX; considering $k = 100$ and $\rho \in \{0.01, 0.99\}$, respectively. Although the differences are subtle, only OSMOX guarantees the desired number of k seeds.

In [8], we present the *Object Geodesic Grid Sampling* (OGRID) method in which samples seeds equidistantly within the probable objects in the map. The set of probable objects $\mathcal{C}$ is obtained by thresholding for a given minimum certainty value t. Since wider regions require a higher number of seeds, while noises (*i.e.*, small components) should not contain any, the number of internal seeds k_i is proportional to the size of its respective component C_i. Moreover, since each component has a different size and shape, computing the necessary minimum seed distance d_i in C_i is challenging; therefore, d_i is optimized, such that it best approximates k_i. One can see that, due to the latter, OGRID might not guarantee the desired total number of seeds.

For each component, its first *seed candidate* is selected arbitrarily, and it is inserted in a priority queue. The following steps are repeated until the queue is empty, or k_i seeds were sampled. The candidate with the highest priority is removed from the queue, selected as seed, and all of its adjacents — distanced by d_i — are inserted in the queue as seed candidates. Note that the latter events might insert candidates that do not respect the equidistance rule for some seed sampled; therefore, removing them from the queue is mandatory. In contrast, when a seed candidate satisfies the equidistance rule for more than one seed, its priority must be increased accordingly. Finally, as one may see, the object seed set consists of the union of all components' internal seeds.

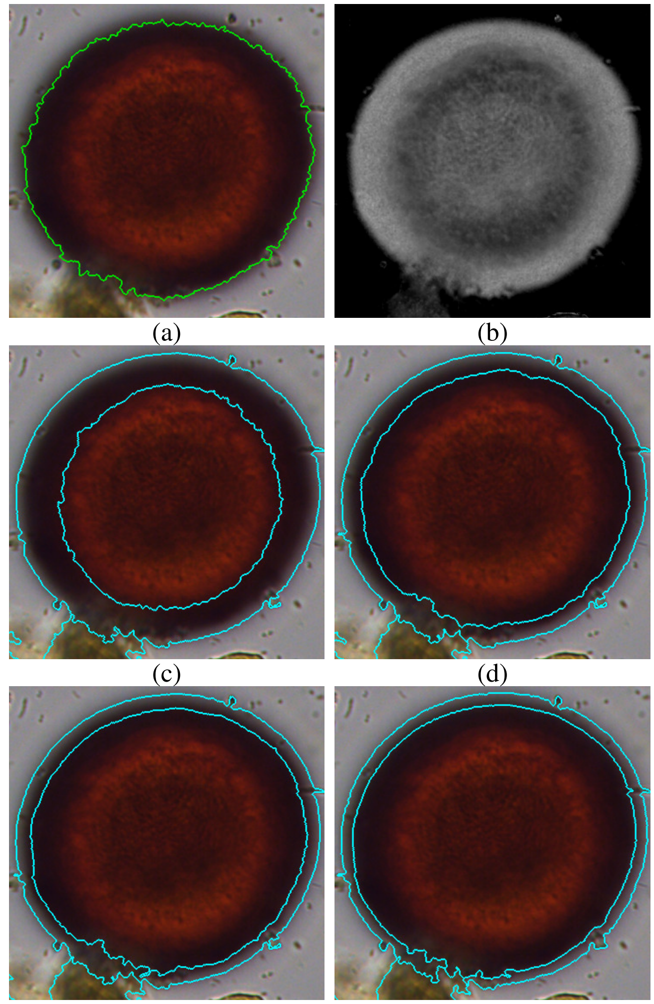
Aside not assuring the desired number of seeds, the selection of the threshold t in OGRID is not intuitive in most cases. Therefore, we propose in [9] the *Object Saliency Map sampling by Ordered eXtraction* (OSMOX) algorithm, which elegantly overcomes both by relaxing the equidistance rule. In OSMOX every pixel is a possible seed candidate: its priority to be selected as one is determined by the summed saliency value within a precalculated neighborhood. All pixels are inserted in a priority queue, and the next steps are repeated until the queue is empty, or k_o object seeds are obtained. The pixel with the highest priority is removed from the queue and selected as seed. In order to promote a fair distribution, all of the seed's adjacents — defined by a disk of radius d — have their priority values decreased by a Gaussian function, and their positions in the queue are updated accordingly.

B. Superpixel Generation

Depending on the quality of the object saliency map, the user should have the ability to control its influence during the generation of superpixels. For instance, for an ideal map (*i.e.*, the ground-truth), one can increase such influence for obtaining superpixels with higher adherence to the primer's borders. Oppositely, for a map with imprecise object borders, its influence must be minimum for avoiding deterioration of the delineation performance — *i.e.*, consider color variation more relevant than saliency.

In OISF, the superpixels are generated through seed competition using the *Image Foresting Transform* (IFT) [18] framework. In IFT, the segmentation task is modeled as the minimization of *path-costs* in the graph, for the generation of an *optimum-path forest* rooted in a seed set. Therefore, as one may see, a superpixel is interpreted as an *optimum-path tree*. Such solution is efficiently obtained through an adaptation of the Dijkstra's algorithm for more general path-cost functions — also known as *connectivity* functions. Due to its performance, the IFT has been widely used in different contexts and purposes [28]–[30].

In the work of [31], the authors propose a new connectivity function for the correction of presegmented binary images. Based on a previous function [32], it imposes a penalization to the path-cost whenever the path crosses a boundary of its originating object. The user can establish the degree of



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

Fig. 4. Object-based superpixel segmentation. (a) Original image, with its ideal delineation in green. (b) Object saliency map [27]. (c-f) OISF with OSMOX sampling considering $k = 50$, $\rho = 0.1$, and $\gamma \in \{0.5, 1, 2, 5\}$, respectively. All images were zoomed for visualization purposes.

such penalization through an importance factor γ. Since the previous work deals with binary images and saliency maps might not be such, we adapted the connectivity function [8] in order to penalize the saliency transitions along the path. Thus, higher importance values impose a harsher penalization, causing an exponential growth of the path-costs and, consequently, reducing the chances of superpixel leaking. Moreover, it indirectly leads to a higher superpixel adherence to the borders of the saliency map (Figure 4).

C. Seed Recomputation

After the generation of superpixels, most state-of-the-art methods compute the seed set for the subsequent iteration, based on the local characteristics of each superpixel [1], [7], [15], [17], [19]. Selecting the superpixel's medoid considers the texture information by selecting the pixel most similar to its superpixel mean feature vector. The objective is to promote homogeneity — *e.g.*, less color dissimilarity amongst its pairs — by selecting the pixel that reinforces its superpixel's characteristics. It is straightforward to notice that, by extending the feature vector of the pixel for including its saliency value, the aforementioned strategy remains unaltered and becomes object-based. Although one may argue over the possibility of increasing the influence of the saliency value over the remaining ones, we understand that the real values of the image (*i.e.*, color) must be more relevant than inference ones (*e.g.*, saliency).

IV. Experimental Results

In this section, we discuss the setup for our experiments (Subsection IV-A) and discuss the quantitative and qualitative results of our approach, compared to the baselines (Subsections IV-B and IV-C, respectively).

A. Experimental Setup

We selected three datasets from distinct domains and two object saliency estimators. For natural images, we considered the *ECSSD* [33] (1000 images) and the *DUT-OMRON* [34] (5168 images) datasets, which offers a challenge by presenting distinct objects and backgrounds. Due to the latter, we chose the *Pyramid Feature Attention* (PFA) [26] since its performance excelled many other approaches, including in both datasets considered. For non-natural images, we selected the *Parasitos* [8] (72 images) dataset, in which the parasite egg (*i.e.*, object of interest) is extremely similar to the impurities that are often attached to it — posing a major difficulty. Since the PFA estimator requires a significant number of examples to be trained, we considered an OPF-based solution (SUP) [27], which trains a pixel-level IFT-based classifier [35] from user-drawn scribbles, in a single image. Given such classifier, one can estimate the saliency of the remaining images in the dataset.

Due to the lack of object-based superpixel segmentation algorithms, we selected the following state-of-the-art methods in object delineation as baselines: (i) SLIC [15]; SNIC [16]; (iii) ERS [4]; (iv) ERGC [36]; (v) LSC [1]; and two ISF instances [7]: (vi) ISF-GRID-ROOT; and (vii) ISF-MIX-MEAN. For simplicity, we maintained the ISF nomenclature for the instances of our approach — *i.e.*, OISF-OGRID, and OISF-OSMOX. For a fair comparison, we optimized the OISF parameters in a small training set (*i.e.*, three images of each dataset), and, for the baseline, the default recommended values were set. For OISF-OSMOX, $\rho = 0.9$ and $\gamma = 2$; whereas for OISF-OGRID, $t = 0.5$, $\rho = 0.9$, and $\gamma = 3$. The performance of each method was evaluated in terms of *Boundary Recall* (BR) [15], and of *Under-segmentation Error* (UE) [37]. In this work, we aim that superpixels accurately delineate the object boundaries (*i.e.*, higher BR), while avoiding the presence of leakings (*i.e.*, lower UE). Finally, we evaluated all methods considering a desired number of superpixels NS varying from 20 to 1000.

B. Quantitative Analysis

The results obtained for each method, in every dataset, are presented in Figure 5. In terms of BR and considering $NS \leq 200$, our approach presents competitive performance with the top methods in both ECSSD and DUT-OMRON. Moreover, for any NS, OISF-based methods still surpass its ISF-based counterparts with a significant margin — an improvement of 5% on average. In Parasitos, our approach managed to effectively delineate (*i.e.*, $BR \geq 90\%$) the parasite egg by requiring a minimum quantity of superpixels (*i.e.*, $NS \approx 20$). Note that, for the baselines, such is achieved at significantly higher values of NS (*i.e.*, $NS \approx 400$).

The charts illustrate the findings of [9]: the use of object-based strategies leads to lower UE values. For ECSSD and DUT-OMRON, our method achieved the best performance for $NS \leq 200$, for a considerable margin — an improvement of 3% on average. For higher values of NS, their performance is still on pair with its ISF-based counterparts. However, for Parasitos, OISF-based methods surpass all baselines for any value of NS, achieving $UE \approx 1\%$ for $NS \approx 200$ — while, for the remaining methods, $NS \approx 1000$ —, reinforcing the premise that our approach manages to delineate the object of interest effectively.

C. Qualitative Analysis

As one can see in Figure 6, in regions which the color transition is smooth — *e.g.*, regions near the head and hand of the child — the best methods for the considered dataset presents major leakings and delineation errors. However, for a given object saliency map, our approach managed not only to avoid both aforementioned errors without requiring a higher number of superpixels but also to increase the superpixel resolution within the object of interest (as desired). This delineation improvement is mainly due to the indirect object delineation obtained by the map: even for a fair one, the contrast between object and background may be a good indication of the primer's extension, thus assisting in a better approximation to the real object borders while reducing the severity of leakings.

D. Limitations

As shown in Figure 5, OISF-based approaches present a rapid convergence in their performance as the number of superpixels increases. We infer that such behavior is caused by one main reason: the saliency map's unchangeable nature. For a given map, OISF-based methods incorporate both correct and incorrect estimations for generating superpixels at all steps of the pipeline. For example, in Figure 6, OISF-based approaches managed to prevent severe leakings, but the superpixel borders could not approximate the real ones at the child's hand. One can see that such errors minimally impact the seed recomputation step since it amortizes by computing the mean feature vector. Although object-based seed sampling strategies might sample seeds in incorrectly estimated regions, a high percentage of object seeds may overcome the latter.

In contrast to the aforementioned steps, the superpixel generation procedure degrades significantly from errors. For a small number of superpixels, the low competition amongst seeds favors the incorporation of significantly dissimilar pixels,

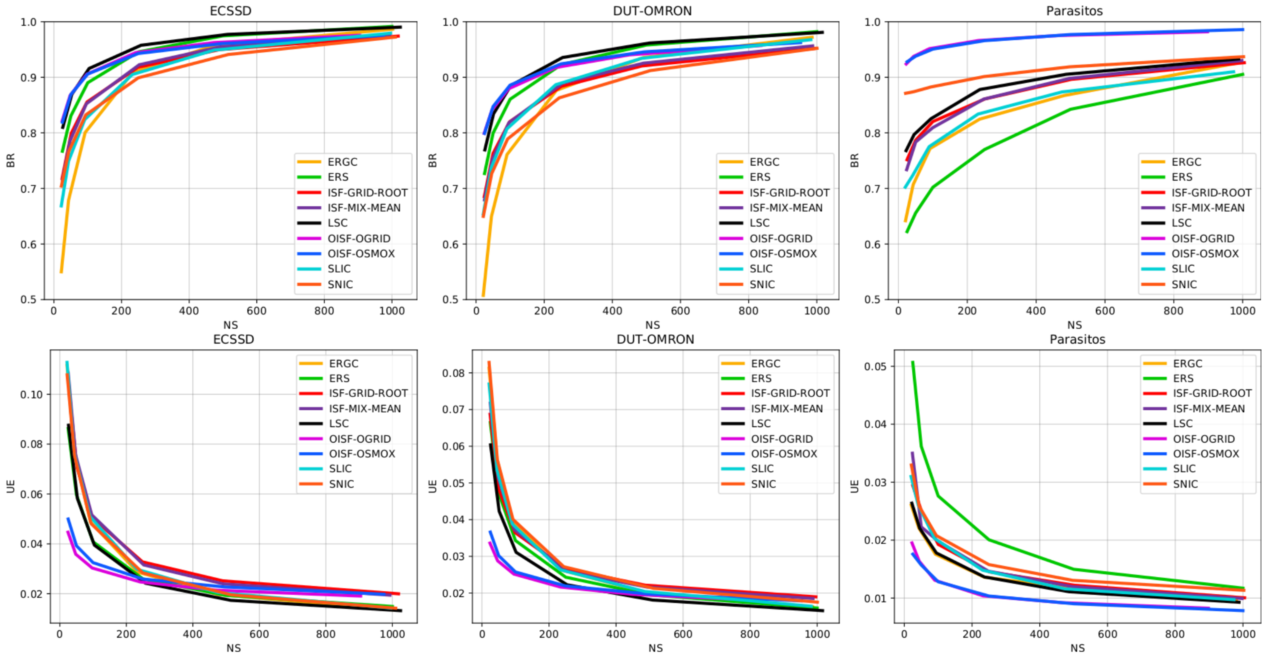


Fig. 5. Results obtained in each dataset for BR and UE. For the baseline, the default parameter configuration was set.

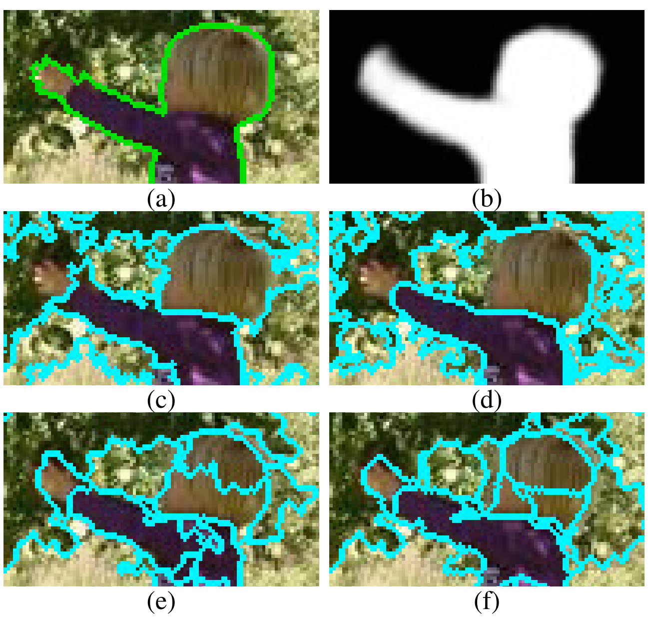


Fig. 6. (a) Original image with the ideal delineation in green. (b) PFA's saliency map. Segmentation results obtained by: (c) ERS; (d) LSC; (e) OISF-OGRID; and (f) OISF-OSMOX. The number of desired superpixels was 50, and the images were zoomed for visualization purposes.

whereas, for higher quantities, the extreme competition makes subtle variations crucial for incorporating similar ones. As one can see, for the same saliency importance factor, it may assist or degrade the delineation performance, for a different number of superpixels. More explicitly: in the absence of seed competition, the γ penalization may prevent the occurrence of superpixel leaking, while for a high number of superpixels, it may intensify the influence of saliency inaccuracies in the path-cost computation. Note that such behavior is repeated throughout the iterations of OISF.

V. Conclusion

In this dissertation, we propose the *Object-based Iterative Spanning Forest* (OISF) framework, which is a generalization of the *Iterative Spanning Forest* (ISF). Our approach considers prior object information represented by monochromatic object saliency maps at each step of its pipeline, allowing the user to control not only the superpixel resolution within the objects of interest but also the superpixel adherence to the saliency map's borders. Both characteristics can be of assistance for different applications of distinct domains, while not being comprised in a particular definition of an object of interest.

Experimental results show that the inclusion of such information leads to a major improvement — especially in preventing superpixel leaking — over its non-object-based counterparts, for a significantly small number of superpixels. For future endeavors, we intend to extend other ISF-based algorithms and evaluate our proposal in medical applications. Moreover, we also desire to study an adaptive solution for establishing the saliency importance factor based on the seed competition environment. This study resulted in the publication of two international conference papers [8], [9].

Acknowledgment

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. The authors thank CNPq (303808/2018-7, 131000/2018-7, 310075/2019-0), FAPESP (2014/12236-1) and FAPEMIG (06-2018) for the financial support.

REFERENCES

[1] Z. Li and J. Chen, "Superpixel segmentation using linear spectral clustering," in *28th Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015, pp. 1356–1363.

[2] Y. Li and L. Shen, "Skin lesion analysis towards melanoma detection using deep learning network," *Sensors*, vol. 18, no. 2, p. 556, 2018.

[3] J. Zhou, J. Ruan, C. Wu, G. Ye, Z. Zhu, J. Yue, and Y. Zhang, "Superpixel segmentation of breast cancer pathology images based on features extracted from the autoencoder," in *11th International Conference on Communication Software and Networks (ICCSN)*, 2019, pp. 366–370.

[4] M. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *24th Conference on Computer Vision and Pattern Recognition (CVPR)*, 2011, pp. 2097–2104.

[5] S. Zhang, H. Wang, W. Huang, and Z. You, "Plant diseased leaf segmentation and recognition by fusion of superpixel, K-means and PHOG," *Optik*, vol. 157, pp. 866–872, 2018.

[6] W. Wu, A. Chen, L. Zhao, and J. Corso, "Brain tumor detection and segmentation in a CRF (conditional random fields) framework with pixel-pairwise affinity and superpixel-level features," *International Journal of Computer Assisted Radiology and Surgery*, vol. 9, no. 2, pp. 241–253, 2014.

[7] J. Vargas-Muñoz, A. Chowdhury, E. Alexandre, F. Galvão, P. Miranda, and A. Falcão, "An iterative spanning forest framework for superpixel segmentation," *Transactions on Image Processing*, vol. 28, no. 7, pp. 3477–3489, 2019.

[8] F. Belém, S. Guimarães, and A. Falcão, "Superpixel segmentation by object-based iterative spanning forest," in *23rd Iberoamerican Congress on Pattern Recognition*, 2018, pp. 334–341.

[9] F. Belém, L. Melo, S. Guimarães, and A. Falcão, "The importance of object-based seed sampling for superpixel segmentation," in *32nd Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2019, pp. 108–115.

[10] A. Schick, M. Fischer, and R. Stiefelhagen, "An evaluation of the compactness of superpixels," *Pattern Recognition Letters*, vol. 43, pp. 71–80, 2014.

[11] D. Stutz, A. Hermans, and B. Leibe, "Superpixels: An evaluation of the state-of-the-art," *Computer Vision and Image Understanding*, vol. 166, pp. 1–27, 2018.

[12] L. Wan, X. Xu, Q. Zhao, and W. Feng, "Spherical superpixels: Benchmark and evaluation," in *14th Asian Conference on Computer Vision (ACCV)*, 2018, pp. 703–717.

[13] X. Ren and J. Malik, "Learning a classification model for segmentation," in *9th International Conference on Computer Vision (ICCV)*, vol. 1, 2003, pp. 10–17.

[14] Z. Ren and G. Shakhnarovich, "Image segmentation by cascaded region agglomeration," in *26th Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 2011–2018.

[15] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 11, pp. 2274–2282, 2012.

[16] R. Achanta and S. Süsstrunk, "Superpixels and polygons using simple non-iterative clustering," in *30th Conference on Computer Vision and Pattern Recognition (CVPR)*, 2017, pp. 4651–4660.

[17] Y. Liu, M. Yu, B. Li, and Y. He, "Intrinsic manifold SLIC: A simple and efficient method for computing content-sensitive superpixels," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 3, pp. 653–666, 2018.

[18] A. Falcão, J. Stolfi, and R. Lotufo, "The image foresting transform: Theory, algorithms, and applications," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 19–29, 2004.

[19] F. Galvão, A. Falcão, and A. Chowdhury, "RISF: recursive iterative spanning forest for superpixel segmentation," in *31st Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2018, pp. 408–415.

[20] C. Castelo-Fernandez and A. Falcão, "Learning visual dictionaries from class-specific superpixel segmentation," in *18th Conference on Computer Analysis of Images and Patterns (CAIP)*, 2019, pp. 171–182.

[21] S. Martins, G. Ruppert, F. Reis, C. Yasuda, and A. Falcão, "A supervoxel-based approach for unsupervised abnormal asymmetry detection in MR images of the brain," in *16th International Symposium on Biomedical Imaging (ISBI)*, 2019, pp. 882–885.

[22] M. Awaisu, L. Li, J. Peng, and J. Zhang, "Fast superpixel segmentation with deep features," in *36th Computer Graphics International Conference (CGI)*, 2019, pp. 410–416.

[23] V. Jampani, D. Sun, M. Liu, M. Yang, and J. Kautz, "Superpixel sampling networks," in *18th European Conference on Computer Vision (ECCV)*, 2018, pp. 352–368.

[24] W. Tu, M. Liu, V. Jampani, D. Sun, S. Chien, M. Yang, and J. Kautz, "Learning superpixels with segmentation-aware affinity loss," in *31st Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 568–576.

[25] L. Xu, L. Zeng, and Z. Wang, "Saliency-based superpixels," *Signal, Image and Video Processing*, vol. 8, no. 1, pp. 181–190, 2014.

[26] T. Zhao and X. Wu, "Pyramid feature attention network for saliency detection," in *32nd Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019, pp. 3080–3089.

[27] T. Spina, P. Miranda, and A. Falcão, "Intelligent understanding of user interaction in image segmentation," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 26, no. 02, p. 1265001, 2012.

[28] A. Falcão, B. Cunha, and R. Lotufo, "Design of connected operators using the image foresting transform," in *Medical Imaging (SPIE-MI)*, vol. 4322, 2001, pp. 468–479.

[29] A. Falcão, L. da Costa, and B. Cunha, "Multiscale skeletons by image foresting transform and its application to neuromorphometry," *Pattern Recognition*, vol. 35, no. 7, pp. 1571–1582, 2002.

[30] A. Sousa, S. Martins, A. Falcão, F. Reis, E. Bagatin, and K. Irion, "Altis: A fast and automatic lung and trachea ct-image segmentation method," *Medical Physics*, vol. 46, no. 11, pp. 4970–4982, 2019.

[31] A. Tavares, P. Miranda, T. Spina, and A. Falcão, "A supervoxel-based solution to resume segmentation for interactive correction by differential image foresting transforms," in *13th International Symposium on Mathematical Morphology (ISMM)*, 2017, pp. 107–118.

[32] L. Mansilla, P. Miranda, and F. Cappabianco, "Image segmentation by image foresting transform with non-smooth connectivity functions," in *26th Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2013, pp. 147–154.

[33] J. Shi, Q. Yan, L. Xu, and J. Jia, "Hierarchical image saliency detection on extended cssd," *Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 4, pp. 717–729, 2015.

[34] C. Yang, L. Zhang, H. Lu, X. Ruan, and M. Yang, "Saliency detection via graph-based manifold ranking," in *26th Conference on Computer Vision and Pattern Recognition (CVPR)*, 2013, pp. 3166–3173.

[35] J. Papa, A. Falcão, and C. Suzuki, "Supervised pattern classification based on optimum-path forest," *International Journal of Imaging Systems and Technology*, vol. 19, no. 2, pp. 120–131, 2009.

[36] P. Buyssens, I. Gardin, S. Ruan, and A. Elmoataz, "Eikonal-based region growing for efficient clustering," *Image and Vision Computing*, vol. 32, no. 12, pp. 1045–1054, 2014.

[37] P. Neubert and P. Protzel, "Superpixel benchmark and comparison," in *Forum Bildverarbeitung*, vol. 6, 2012, pp. 1–12.