Towards an Agnostic Superpixel Segmentation Framework

Felipe Belém[*,1,2,3], Benjamin Perret[2], Silvio Jamil F. Guimarães[1], Jean Cousty[2], and Alexandre X. Falcão[3]

[1] IMScience, Pontifical Catholic University of Minas Gerais – Minas Gerais, Brazil 30535–901
Email: {felipebelem,sjamil}@pucminas.br
[2] LIDS, University of Campinas – São Paulo, Brazil 13083–970
Email: {felipe.belem,afalcao}@ic.unicamp.br
[3] LIGM, Université Gustave-Eiffel, CNRS – Marne-la-Valée, France F-77454
Email: {felipe.belem,benjamin.perret,jean.cousty}@esiee.fr

***Abstract*—Superpixel segmentation partitions multiple objects into disjoint parts so that their delineation can be accurately achieved by their grouping, and it has been used as an intermediary step for solving multiple problems. However, state-of-the-art algorithms face a significant challenge of effective and efficient segmentation irrespective of the problem's domain (object and background characteristics, and user's desires). In this work, we address such challenge by proposing several contributions. One of such is a novel superpixel segmentation framework, named Superpixels through Iterative CLEarcutting (SICLE), which generalizes two other contributions of this work. In SICLE, three independent steps are defined: (i) seed oversampling; (ii) superpixel generation using the Image Foresting Transform (IFT) framework; and (iii) seed removal. From (i), where a significantly high amount of seeds is selected, steps (ii) and (iii) are performed for generating superpixels from a refined seed set until achieving the desired number of superpixels. SICLE overcomes domain shifts primarily through steps (ii) and (iii), where the user may provide an objective function for optimization. Experimental results show that SICLE variants surpass several state-of-the-art algorithms concerning speed and accuracy for distinct domains while generating a series of segmentations in a single execution. Still, in SICLE, the contours from a preceding scale might not be present in the subsequent one leading to hierarchical violations. Thus, we studied eight possible cases when analyzing pairwise subsequent segmentations, and we conceived three measures for estimating the hierarchiness of a multiscale segmentation: (i) nestedness; (ii) inflation ratio; and (iii) refinement error. From our results, it is possible to verify if a multiscale is a hierarchy and, when it is not the case, to analyze and state the nature and extent of the hierarchical violations that prevent it from being hierarchical.**

I. Introduction

In superpixel segmentation, the effective delineation of the object is achieved by generating a collection of connected *picture elements* (*i.e.*, *pixels*) by optimizing some objective function (*e.g.*, color similarity). Thus, with meticulous superpixel selection, the object is accurately built. This and other properties favor their usage in many applications [1]–[3]. When a single segmentation is insufficient, one may generate several superpixel segmentations (*i.e.*, *multiscale* segmentation) with different resolutions (*i.e.*, *scales*) in which higher quantity of superpixels (resp. lower) produces a finer (resp. coarser) resolution, providing detailed (resp. global) information of the object. If every superpixel is contained in only one region of its subsequent scale, then it is a *hierarchy*.

Each problem has its own *domain* that can be broadly defined by three main axes: (i) data (*e.g.*, data availability and object characteristics); (ii) hardware (*e.g.*, limitations); and (iii) user interaction (*e.g.*, implementation and adaptation). From (i), we focus on two aspects: *bias* and *morphology*. Bias is all prior knowledge of the data and its objects (*e.g.*, color, location), and morphology involves the set of morphological requirements (*e.g.*, shape, displacement), both crucial for solution. Note that both may differ significantly between domains and goals. In this work, we do not consider interactive methods.

However, most methods in literature are highly optimized for specific domains (*i.e.*, natural datasets) due to lack of diversity in their evaluation. Thus, shifting the domain may not sustain performance. One can perceive such behavior in classical approaches [4], [5], where only two morphologies may be achieved (*i.e.*, compact or irregular), or even in content-sensitive approaches [6], where superpixel resolution in high-feature variability is favored, tending to fail for homogeneous and simple objects. Deep-learning algorithms [7] may easily incorporate new biases through fine-tuning, but, aside from the need for data and specialized hardware (*i.e.*, GPUs), promoting different morphologies can be challenging. Finally, object-based strategies [8], [9] allow prioritizing the delineation of the objects of interest, with the use of prior object information (*e.g.*, *probable location*), at the expense of disregarding non-desired ones. However, they tend to be highly dependent on the quality of the provided information [10].

On tasks where a multiscale segmentation is needed, obtaining such a structure can be difficult, if not combinatorial, even for hierarchical methods, like [11]. While non-hierarchical approaches may "correct" incorrect delineations [12] in subsequent iterations, hierarchical ones provide a set of optimizations the former does not offer due to violations [13]. Yet, to the best of our knowledge, there is no measure to determine the resemblance of a multiscale to a hierarchy (*i.e.*,

* Ph.D. thesis

its *hierarchiness*). Most works require an object ground-truth for assessment [14], [15], or weight contours differently based on a ground-truth [16]. With such hierarchiness measure, one may improve performance by adopting local optimizations or resorting to "relaxed" properties.

In this thesis, we study certain conditions for a superpixel segmentation to be effective and efficient for any domain. Also, since a naïve approach may lead to a combinatorial problem, we study strategies for computing hierarchiness efficiently from a given multiscale segmentation. More specifically, we focused on four main research questions:

Q1. Is it possible to build a superpixel segmentation that assimilates any bias?

Q2. Is it possible to build a superpixel segmentation that promotes any bias?

Q3. Is it possible to build a superpixel segmentation that achive both latter with efficacy and efficiency for any object?

Q4. Is it possible to conceive a measure in which estimates the resemblance of a multiscale segmentation to a hierarchical one?

As contributions, we proposed a framework named *Superpixels through Iterative CLEarcutting* (SICLE), whose pipeline is defined by three steps: (i) seed oversampling; (ii) superpixel generation; and (iii) seed recomputation. SICLE generalizes two superpixel segmentation methods, both also contributions of this thesis: the *Dynamic and Iterative Spanning Forest* (DISF) [12], and the *Object-based DISF* (ODISF) [10]. In (i), an initial set of points (*i.e.*, *seeds*) significantly greater than the desired number of superpixels are sampled. From such seed set, superpixels are generated in step (ii) using the *Image Foresting Transform* (IFT) [17] framework. In (iii), a portion of *irrelevant* seeds (*i.e.*, does not assist in the object delineation) are removed based on some (application-driven) criterion, and the improved set is provided for the subsequent superpixel segmentation. Steps (ii) and (iii) are performed until reaching the exact number of superpixels desired by the user. Note that such segmentation pipeline provides a multiscale segmentation on-the-fly and, from such, we also studied several events when analyzing subsequent pairwise scales, leading to eight distinct region-altering cases and three novel hierarchiness measures: (a) *inflation ratio*; (b) *refinement error*; and (c) *nestedness*. Experimental results show that SICLE variants surpass state-of-the-art methods in efficiency and efficacy, and that our hierarchiness measures assist in analyzing the nature and extent of hierarchical errors (if there are any) in a given multiscale segmentation, such as those provided by SICLE.

This paper is organized as follows. In Section II, we present the SICLE framework and the hierarchiness measures. Section III presents the main results achieved by our work. Finally, in Section IV, our list of contributions is detailed alongside the conclusion of this paper, and possible future work.

II. SICLE and Hierarchiness

Figure 1 presents the pipeline of our framework named *Superpixels through Iterative CLEarcutting* (SICLE), which is a generalization of two previous contributions: the *Dynamic and Iterative Spanning Forest* (DISF) [12]; and the *Object-based DISF* (ODISF) [10]. As the first step, SICLE selects a significantly higher number of initial points $N_0 \in \mathbb{N}$ (*i.e.seeds*) than the desired number $N_f \ll N_0$ of superpixels, so that important object parts are not missed by the sampling strategy. Due to such quantity, aiming for higher precision in such initial selection is not of utmost concern and, consequently, one may opt for faster approaches, as in a grid or a random distribution, rather than slower, yet more precise ones [19]. In fact, based on this premise, the crucial step in SICLE is to accurately remove superpixels for reaching the final quantity N_f while sustaining the effective performance achieved in higher quantities.

From such seed set, superpixels are generated using the *Image Foresting Transform* (IFT) [17] framework, a generalization of the Dijkstra's algorithm for conceiving connectivity-based image operators. Thus, from an *image graph* $G = (V, E)$ in which V and E denote the *vertex* and the *edge* set, respectively, the IFT generates superpixels by minimizing the cost from *path concatenation* from a seed to the remaining vertices. Such minimization considers some *connectivity function* $\boldsymbol{f}_*$, such as the popular ones exemplified by Equation 1:

$$\begin{aligned} &\boldsymbol{f}_{\text{max}}(\rho_v \odot \{v,u\}) = \max\{\boldsymbol{f}_{\text{max}}(\rho_v), \boldsymbol{w}_*(v,u)\} \quad \text{and} \\ &\boldsymbol{f}_{\text{sum}}(\rho_v \odot \{v,u\}) = \boldsymbol{f}_{\text{sum}}(\rho_v) + \boldsymbol{w}_*(v,u) \end{aligned} \tag{1}$$

where ρ_v is an *optimum-path* (from some seed) that has a vertex v as *terminus*, and $\odot$ is the *concatenation* operator. With a proper selection of *edge-cost function* $\boldsymbol{w}_*$, one may generate different superpixel morphologies. For instance, with those presented by Equation 2 and proposed in this work:

$$\begin{aligned} \boldsymbol{w}^{\alpha}_{\text{sum}}(p,q) &= ((\iota + \alpha\|\boldsymbol{O}(\boldsymbol{R}(p)) - \boldsymbol{O}(q)\|_1)\|\boldsymbol{F}(q) - \boldsymbol{F}(\boldsymbol{R}(p))\|_2)^{\beta} \\ &\quad + \|p-q\|_2 \quad \text{and} \\ \boldsymbol{w}^{\alpha}_{\text{root}}(p,q) &= (\|\boldsymbol{F}(\boldsymbol{R}(p)) - \boldsymbol{F}(q)\|_2)^{1+\alpha\|\boldsymbol{O}(\boldsymbol{R}(p))-\boldsymbol{O}(q)\|_1} \end{aligned} \tag{2}$$

one can obtain irregular superpixels with $\boldsymbol{f}_{\text{max}}$ and $\boldsymbol{w}^{\alpha}_{\text{root}}$, or generate compact superpixels $\boldsymbol{f}_{\text{sum}}$ and $\boldsymbol{w}^{\alpha}_{\text{sum}}$, by controlling the superpixels' adherence and irregularity factors (*i.e.*, β and ι, respectively). Moreover, through the *confidence factor* α, one can weight the influence of the prior object information $\boldsymbol{O}$ (*e.g.*, *saliency*) over the image's *features* $\boldsymbol{F}$ whose dissimilarity is calculated with respect to those from the seed mapped by $\boldsymbol{R}$. Thus, paths that cross probable object borders will be significantly penalized, leading to improved delineation when color dissimilarity is insufficient.

In order to achieve the desired number N_f of superpixels specified by the user, several superpixels must be removed. If one conceives an ideal seed selection strategy, SICLE would require only two iterations for segmenting any image. However, accurately determining the set of seeds whose superpixels do not hamper object delineation (*i.e.irrelevant*) is challenging in many situations. Therefore, the strategy in SICLE is to "perturb" seeds in each iteration by removing only a portion of the most irrelevant seeds (based on some

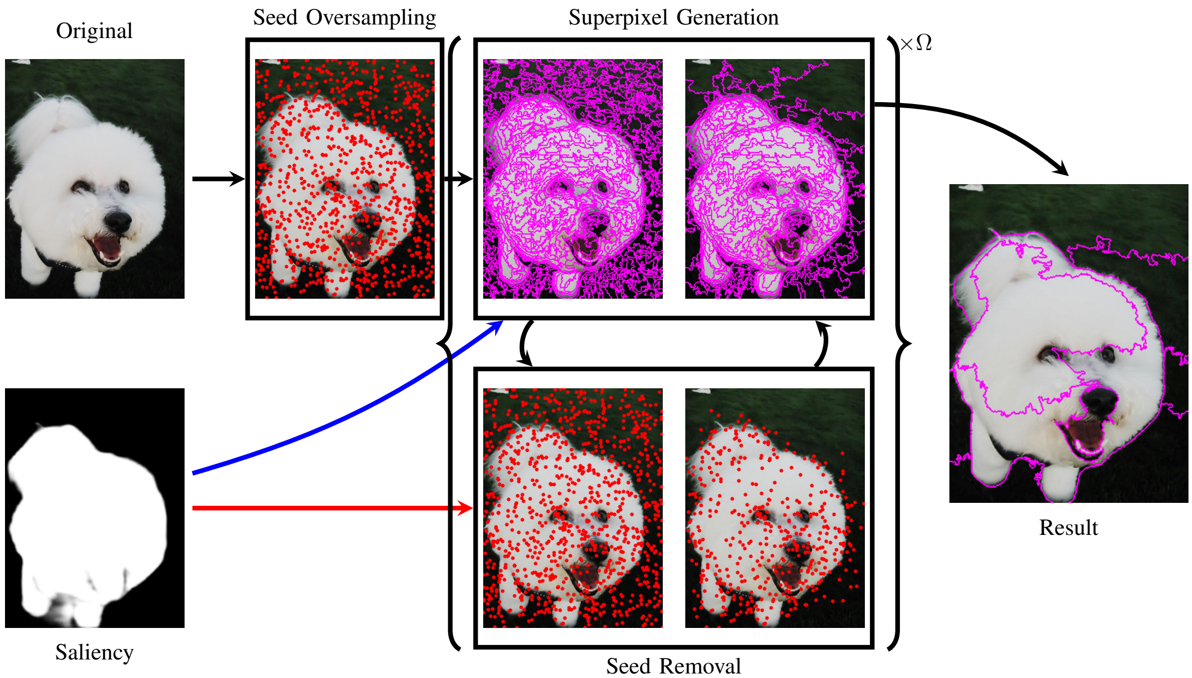


Fig. 1. Pipeline for superpixel generation using SICLE (black, red, and blue arrows) and considering the saliency generated by [18]. DISF's pipeline is indicated by black arrows, whereas ODISF's by black and red arrows. A quantity of 1000 seeds was sampled for generating a segmentation with 10 superpixels.

mathematical criterion), and promoting a novel competition for the subsequent IFT execution. Intuitively speaking, a seed is relevant for delineation if it is consistently marked as such. Such quantity of relevant seeds to be maintained is defined by a *seed curve* $\boldsymbol{M}(i) = \max\{(N_0)^{1-\omega \cdot i}, N_f\}$, where ω is a *seed decay factor*, often determined by $\omega = 1/(\Omega - 1)$ given a maximum number of iterations Ω. We emphasize that SICLE may achieve the final segmentation with significantly fewer iterations than established by Ω.

For determining relevance, SICLE uses some *seed relevance function* $\boldsymbol{V}_*$ by recurring to each seed's superpixel features, with respect to the application domain. As an example, based on the IFT, a seed resulting in a large superpixel tends to be better positioned for conquering (than its pairs). However, since background tends to be homogeneous and significantly greater than the object, the latter strategy may prejudice heterogeneous or small objects. In such case, a combination of size and contrast may also be considered for selecting superpixels near crucial regions for delineation. For instance, while favoring superpixels near smooth borders may prevent superpixel "leaking", selecting those near high-gradient borders may guarantee accurate object delineation while promoting shape constraints. Finally, some applications require an even distribution of superpixels (*e.g.*, feature extraction) and, for that, one may opt for a criterion that combines the size and positioning of superpixels among their immediate superpixel neighbors.

Yet, as one may note, the previous relevance computation is subjected to the user's desire to segment an object: for instance, any previous relevance estimation may be discarded if it is certain to be a background superpixel. Therefore, after computing the relevances, we apply a *seed penalization function* $\boldsymbol{B}_*$ that proportionately penalizes seeds that do not meet some other criterion with respect to the object information $\boldsymbol{O}$ provided. As an example, one may prioritize those superpixels that meet the criterion $\boldsymbol{V}_*$ but within (or nearby) the objects of interest. On the other hand, for maximizing delineation, one strategy is to prioritize those that only fall near the objects' borders. It is possible to opt for a combination of one of the latter, while strictly enforcing dispersed superpixels in the background. Ultimately, such a high amplitude for possible configurations favors distinct and unique morphologies.

Finally, this pipeline produces a multiscale on-the-fly by storing intermediate segmentations while the irrelevant seeds are removed. However, such a series of segmentations does not guarantee the nesting behavior between regions from a finer segmentation to a coarser one (*i.e.*higher and lower quantity of superpixels). On the other hand, performing a naïve approach of analyzing every superpixel for every possible pair of segmentations is a combinatorial task. Therefore, we order the series of segmentations in a non-increasing order of size (*i.e.*, superpixel quantity) and, due to the transitivity property in nesting, we analyze pairwise subsequent segmentations (for optimization). For that, we first propose the concepts of *nucleus* η and *cover* $\widehat{\eta}$, as defined by Equation 3:

$$\begin{aligned} \eta(X, Y_j) &= \{X_i \in X : X_i \subseteq Y_j\} \quad \text{and} \\ \widehat{\eta}(X, Y_j) &= \{X_i \in X : X_i \cap Y_j \neq \emptyset\} \end{aligned} \tag{3}$$

where, for $r \geq c > 0$, $X = \{X_1, \ldots, X_r\}$ and $Y =$

$\{Y_1, \ldots, Y_c\}$ denotes the finer and coarser segmentations, respectively. Based on the latter, we mapped eight possible region-altering cases when analyzing the transition from X to Y:

- *Full Deflation:* $\widehat{\eta}(X, Y_j) = \{X_i\}$, $\eta(X, Y_j) = \emptyset$ and $\eta(Y, X_i) = \{Y_j\}$;
- *Deflation and Split:* $\widehat{\eta}(X, Y_j) = \{X_i\}$, $\eta(X, Y_j) = \emptyset$ and $\cup\eta(Y, X_i) \neq X_i$;
- *Full Split:* $\widehat{\eta}(X, Y_j) = \{X_i\}$, $\eta(X, Y_j) = \emptyset$ and $\cup\eta(Y, X_i) = X_i$;
- *Stability:* $\widehat{\eta}(X, Y_j) = \eta(X, Y_j) = \{X_i\}$;
- *Instability:* $|\widehat{\eta}(X, Y_j)| > 1$ and $\eta(X, Y_j) = \emptyset$;
- *Full Inflation:* $|\widehat{\eta}(X, Y_j)| > 1$ and $|\eta(X, Y_j)| = 1$;
- *Merge and Inflation:* $|\widehat{\eta}(X, Y_j)| > 1$, $|\eta(X, Y_j)| > 1$ and $\cup\eta(X, Y_j) \neq Y_j$; and
- *Full Merge:* $|\eta(X, Y_j)| > 1$ and $\cup\eta(X, Y_j) = Y_j$;

Therefore, in a finer-to-coarser multiscale segmentation, only Full Merge and Stability cases are allowed in a hierarchy; otherwise, it is non-hierarchical.

For estimating the degree of nesting behavior between superpixels (*i.e.*, higher is better) of X, Y, we propose the *Nestedness* ***NE*** measure $\boldsymbol{NE}(X, Y) = \sum_{X_i \in X} \sum_{Y_j \in Y} \mathbf{1}(X_i \subseteq Y_j)/|\Upsilon|$, where $|\Upsilon| = |\cup X| = |\cup Y|$ is the number of pixels in the image, and $\mathbf{1}$ is an *indicator function* which outputs 1 if the expression is true, or 0 otherwise. On the other hand, for calculating the degree from inflating superpixels (without any merging), we propose the *Inflation Ratio* $\boldsymbol{IR}(X, Y) = \sum_{X_i \in X} \sum_{Y_j \in Y} \mathbf{1}(X_i \cap Y_j \neq \emptyset \ \wedge \ X_i \not\subseteq Y_j)/|\Upsilon|$, where $\wedge$ is the logical operator AND. In this case, the goal is to achieve minimum values of IR, being zero when it is a hierarchy. Finally, inspired by [20], we propose the *Refinement Error* measure, which computes the "minimum effort" to alter superpixel borders so that total nesting behavior occurs in between X and Y, being ideal when it is minimum (*i.e.*, zero). For that, we select the minimum effort between including regions X_i in the nucleus of Y_j, or removing them totally from the cover of Y_j. More formally, such measure is defined as $\boldsymbol{RE}(X_i, Y_j) = \sum_{X_i \in X} \sum_{Y_j \in Y} \min\{|X_i \cap Y_j|, |X_i \setminus Y_j|\}/(|X_i||\Upsilon|)$

III. Experimental Results

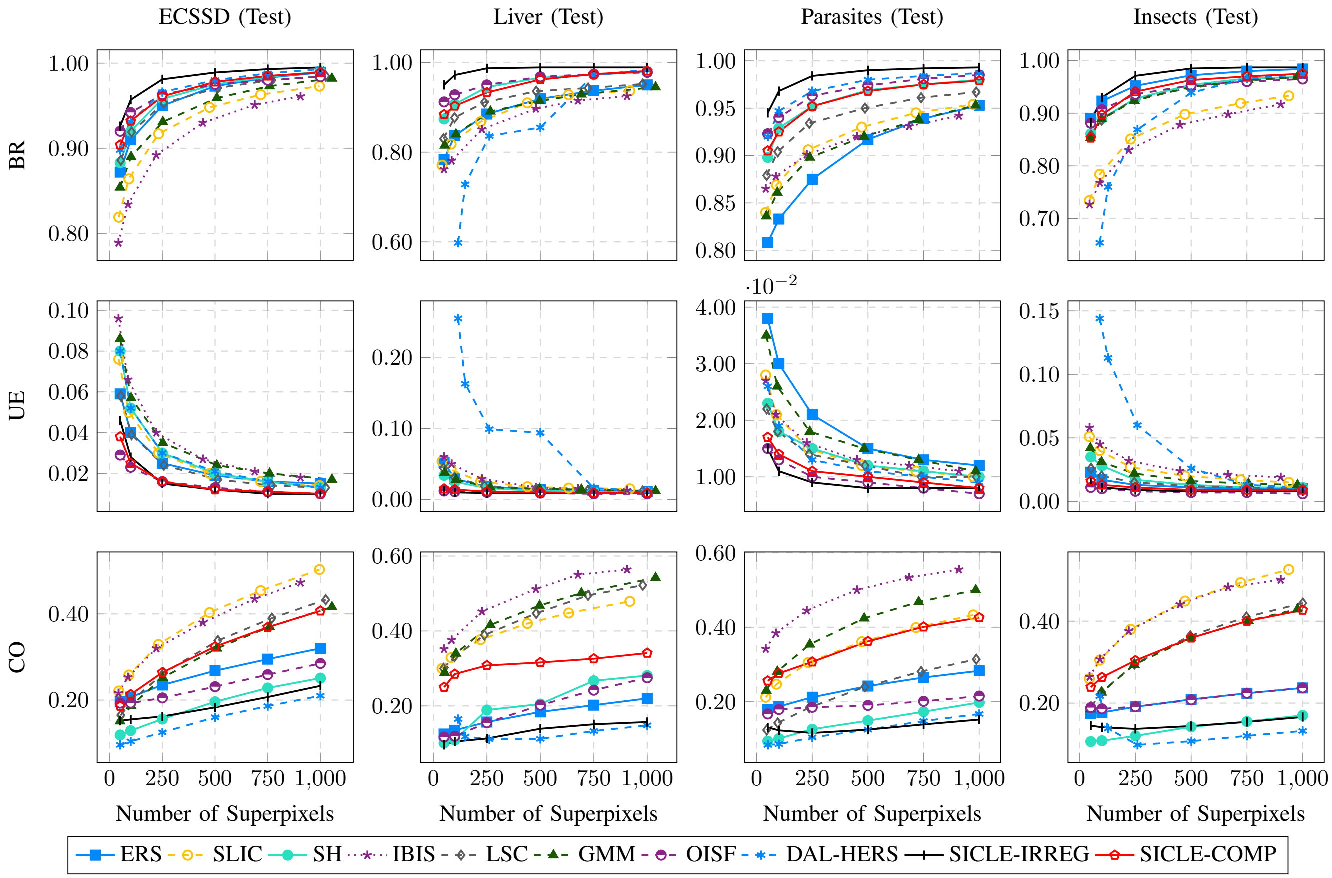
In this section, we provide a brief overview of the results (*c.f.* the thesis for a thorough discussion and analysis). For that, we chose several datasets from clearly distinct domains: natural (ECSSD [21] and Insects [22]) and medical (Liver [23] and Parasites [8]). While ECSSD offers a general-purpose dataset for saliency estimation (*i.e.*, objects and backgrounds with significant differences), Insects impose a challenge for delineation due to the object's thin parts. Similarly, for Liver, the monochromatic CT slices of the liver (and its smooth borders) also offer great difficulty in segmentation. Although the borders of the parasite egg in Parasites are also a hard task to perform, the presence of impurities adds significant difficulty to it. For all datasets, we use the *U^2-Net* [18] saliency estimator due to its accuracy.

Based on their novelty, efficiency or efficacy, we chose several state-of-the-art methods as baselines: (i) ERS [5]; (ii) SLIC [4]; (iii) SH [11]; (iv) IBIS [24]; (v) LSC [25]; (vi) GMM [26]; (vii) OISF [8], [9]; and (viii) DAL-HERS [7]. For all the aforementioned methods, we used their default parameter setting, and we evaluated considering $N_f \in \{50, 100, 250, 500, 750, 1000\}$. Such an interval was selected since $N_f < 50$ leads to an object segmentation setup [27], and that performance often converges before achieving $N_f = 1000$ superpixels in images with similar dimensions [28]. Our best compact and irregular variants are named SICLE-COMP and SICLE-IRREG, respectively.

We also used classical metrics for evaluating the superpixel segmentation. First, we evaluated the computed time in a 64-bit Intel(R) Core(TM) i7-4790S PC with a CPU Speed of 3.20GHz. For delineation, we recur to the *Boundary Recall* (BR) [28] measure, informally defined as the ratio of superpixel borders overlapping object boundaries (*i.e.*, higher is better), and the the *Under-segmentation Error* (UE) [20], which measures the error from multiple object overlap by the superpixels (*i.e.*, lower is better). Finally, we used the popular *Compactness* (CO) [28] measure, which evaluates whether the superpixels present a compact shape, being maximum when it is a circle (*i.e.*, higher is better).

As one may note from Figure 2, SICLE variants present top performance in all measures in all datasets. For SICLE-IRREG, it surpasses all baselines for both BR and UE, irrespective of the domain. Similarly, and for all datasets, SICLE-COMP achieves equivalent performance to several state-of-the-art methods for generating irregular superpixels in terms of BR and UE, while sustaining a good compactness performance. We emphasize that, as seen by the results from OISF (which uses the same saliency map as SICLE), our method manages to properly incorporate the saliency information, while not being influenced by its incorrect estimations. Moreover, we highlight the event of domain internalization from the results of DAL-HERS in both ECSSD and Insects dataset (both natural images), and ERS in Insects and Parasites. For the former, a slight shift in the domain had severe impacts on performance, while the latter did not sustain its performance for a significant domain shift. Based on the time performance, SICLE variants are the fastest amongst the fastest in ECSSD, while being able to generate a multiscale segmentation with minimum (possibly, zero) increase in computed time. In the end, although the optimized SICLE variants managed to surpass all baselines in distinct domains, being an effective off-the-shelf option, we emphasize that the best SICLE variant for each task can be achieved if one knows how to model its objectives (*i.e.*, biases and morphology) mathematically when generating superpixels or determining the relevance of a seed.

In terms of hierarchiness, we evaluated SICLE variants without saliency information for a better analysis. Since SH is a hierarchical method, we used it as a use case for an ideally hierarchical multiscale. Conversely, we used a random-seed Voronoy partition method for generating a worst-case scenario for a multiscale segmentation. Our results have shown that, clearly, irregular non-hierarchical superpixel segmentation (*e.g.*, SICLE IRREG) favors inflation while prejudicing



N_f	SLIC	SH	IBIS	LSC	SICLE-IRREG	SICLE-COMP
50	**0.576±0.029**	0.813±0.028	0.710±0.027	0.702±0.036	**0.501±0.055**	**0.514±0.046**
250	**0.581±0.027**	0.816±0.028	0.700±0.028	0.713±0.036	**0.503±0.040**	**0.511±0.043**
500	**0.583±0.028**	0.813±0.028	0.694±0.026	0.714±0.039	**0.507±0.039**	**0.517±0.042**
1000	**0.583±0.028**	0.812±0.027	0.694±0.028	0.718±0.036	**0.420±0.090**	**0.482±0.060**

Fig. 2. Quantitative results with the U^2-Net [18] saliency estimator. The table presents the speed performance, in seconds, of the fastest methods in the ECSSD dataset. The values depicted in bold, blue, and red indicate, respectively, the first, second, and third fastest performances for each N_f.

nesting behavior in between regions. On the other hand, SICLE-COMP, by enforcing compactness, reduces inflation and promotes nesting. However, it was noted that, even with such discrepant behaviors, both achieved similar (and reduced) refinement error. Thus, as we argue, both SICLE-COMP and SICLE-IRREG would demand slight modifications to make their multiscale segmentation a hierarchical one.

IV. Conclusion and Future Work

In this work, we propose *Superpixels through Iterative CLEarcutting* (SICLE), a multiscale superpixel segmentation framework that provides theoretical foundations for domain agnosticism. It selects a significantly greater number of seeds and, iteratively, generates superpixels, subsequently removing those deemed irrelevant. This approach allows for effective and efficient performance, surpassing several state-of-the-art methods in delineation, leaking prevention, and compactness. Since SICLE produces a multiscale segmentation on-the-fly, we also studied several events when performing a pairwise analysis and conceived three *hierarchiness* measures: *nestedness*; *inflation ratio*; and *refinement error*. From these contributions, we are able to provide substantial insight into the hierarchical violations in the multiscale, if there are any.

Revisiting the research questions, we reached the following conclusions. From Q1, we argue that assimilating any bias is a viable option (rather than all biases at once), and that SICLE provides the tools for such through the seed relevance computation (which is application-based). From Q2, we again argue that the promotion of morphology is rather a specific than an all-encompassing task. As a property of the IFT, not only is SICLE able to produce the two most common superpixel shapes (*i.e.*, compact and irregular), but it may enforce several others if one provides the set of connectivity and edge-cost functions. Therefore, the need for adaptation is limited to offering new functions, rather than to substantially alter the generation algorithm. With our measures, we were able to assess Q4 by evaluating a given multiscale and investigating the occurrences that impede it from being hierarchical.

Finally, instead of accurately segmenting all possible objects, we argue that SICLE manages to accurately segment the object of interest (as raised by Q3), irrespective of its characteristics. Such property lies on the perception that the representation of the probable object location stays unchanged irrespective of its own characteristics, meaning that changing the object of interest (thus, possibly changing bias) does not lead to significant adaptations in the algorithm.

We were able to publish this work in several renowed conferences and journals:

[12] **F.C. Belém, S.J.F. Guimarães and A.X. Falcão**, "Superpixel Segmentation Using Dynamic and Iterative Spanning Forest". IEEE Signal Processing Letters, 2020.

[29] C. Jerônimo, F.C. Belém, S.A. Carneiro, Z.K.G. Patrocínio Jr., L. Najman, A.X. Falcão, S.J.F. Guimarães, "Graph-Based Supervoxel Computation from Iterative Spanning Forest". Discrete Geometry and Mathematical Morphology (DGMM), 2021.

[10] **F.C. Belém, B.Perret, J.Cousty, S.J.F. Guimarães and A.X. Falcão**, "Towards a Simple and Efficient Object-based Superpixel Delineation Framework". 34th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2021.

[30] I.B. Barcelos, F.C. Belém, P.A.V. Miranda, A.X. Falcão, Z.K.G. Patrocínio Jr., and S.J.F. Guimarães, "Towards Interactive Image Segmentation by Dynamic and Iterative Spanning Forest", Discrete Geometry and Mathematical Morphology (DGMM), 2021.

[31] I.B. Barcelos, F.C. Belém, L.M. João, A.X. Falcão and S.J.F. Guimarães, "Improving color homogeneity measure in superpixel segmentation assessment". 35th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2022.

[32] **F.C. Belém, I.B. Barcelos, L. M. João, B. Perret, J. Cousty, S.J.F. Guimarães and A.X. Falcão**, "Fast and Effective Superpixel Segmentation Using Accurate Saliency Estimation". Discrete Geometry and Mathematical Morphology (DGMM), 2022.

[33] **F.C. Belém, I.B. Barcelos, L. M. João, B. Perret, J. Cousty, S.J.F. Guimarães and A.X. Falcão**, "Novel Arc-Cost Functions and Seed Relevance Estimations for Compact and Accurate Superpixels". Journal of Mathematical Imaging and Vision, 2023.

[34] I.B. Barcelos, F.C. Belém, L.M. João, Z.K.G. Patrocínio Jr., A.X. Falcão and S.J.F. Guimarães, "A Comprehensive Review and New Taxonomy on Superpixel Segmentation". ACM Computer Surveys, 2024;

[35] D. Vieira, I.B. Barcelos, F.C. Belém, Z.K.G. Patrocínio Jr., A.X. Falcão and S.J.F. Guimarães, "Streaming Graph-based Supervoxel Computation Based on Dynamic Iterative Spanning Forest". Progress In Pattern Recognition, Image Analysis, Computer Vision and Applications (CIARP), 2024.

[36] **F.C. Belém, F. Faria, A.X. Falcão, Z.K.G. Patrocínio Jr., S.J.F. Guimarães, B. Perret, and J. Cousty**, "Measuring Hierarchiness of Image Segmentations". 37th Brazilian Conference on Graphics, Patterns and Images (SIBGRAPI), 2024.

[37] **F.C. Belém, B.Perret, J.Cousty, S.J.F. Guimarães and A.X. Falcão**, "Efficient Multiscale Object-based Superpixel Segmentation Framework". Journal of the Brazilian Computer Society, 2025.

[38] L. Lacerda, F.C. Belém, A.X. Falcão, Z.K.G. Patrocínio Jr., and S.J.F. Guimarães, "Seed-based Superpixel Re-Segmentation for Improving Object Delineation". Iberoamerican Conference on Pattern Recognition (CIARP), 2025.

Those depicted in bold are directly associated with this work, whereas the remaining ones are related to it (*e.g.*, applications).

For future work, we intend to further evaluate SICLE for multidimensional applications and conceive new connectivity functions that ensure nesting behavior between subsequent segmentations. Also, we intend to provide "relaxed" and novel hierarchiness measures to be used in optimization and learning tasks.

Acknowledgements

This study was financed in part by Coordenação de Aperfeiçoamento de Ensino Superior (Capes) – Finance Code 001 –, by the Brazilian National Council for Scientific and Technological Development (CNPq), grants 407242/2021-0, 131000/2018-7, 306573/2022-9, 442950/2023-3, PQ 303808/2018-7, 310075/2019-0, and 304711/2023-3. This study was also financed, in part, by the São Paulo Research Foundation (FAPESP), Brasil. Process Number 2023/14427-8, 2014/12236-1, and 2013/07375-0 –, and by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) – Grant APQ-01079-23, Grant APQ-05058-23, PPM-00006-18, and PCE-00417-24.

References

[1] C. Castelo-Fernandez and A. Falcão, "Learning visual dictionaries from class-specific superpixel segmentation," in *Computer Analysis of Images and Patterns*, 2019, pp. 171–182.

[2] S. Dhore and D. Abin, "Chest x-ray segmentation using watershed and super pixel segmentation technique," in *4th International Conference on Communication information and Computing Technology (ICCICT)*, 2021, pp. 1–4.

[3] Y. Yu, Y. Makihara, and Y. Yagi, "Pedestrian segmentation based on a spatio-temporally consistent graph-cut with optimal transport," *Transactions on Computer Vision and Applications*, vol. 11, no. 1, p. 10, 2019.

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

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

[6] J. Wu, C. Liu, and B. Li, "Texture-aware and structure-preserving superpixel segmentation," *Computers and Graphics*, vol. 94, pp. 152–163, 2021.

[7] H. Peng, A. Aviles-Rivero, and C. Schonlieb, "Hers superpixels: Deep affinity learning for hierarchical entropy rate segmentation," in *IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 2022, pp. 72–81.

[8] F. Belém, S. Guimarães, and A. Falcão, "Superpixel segmentation by object-based iterative spanning forest," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, vol. 11401, 2019, pp. 334–341.

[9] ——, "Superpixel generation by the iterative spanning forest using object information," in *33rd Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2020, pp. 22–28, workshop of Thesis and Dissertations.

[10] F. Belém, J. Cousty, B. Perret, S. Guimarães, and A. Falcão, "Towards a simple and efficient object-based superpixel delineation framework," in *34th Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2021, pp. 346–353.

[11] X. Wei, Q. Yang, Y. Gong, N. Ahuja, and M. Yang, "Superpixel hierarchy," *IEEE Transactions on Image Processing*, vol. 27, no. 10, pp. 4838–4849, 2018.

[12] F. Belém, S. Guimarães, and A. Falcão, "Superpixel segmentation using dynamic and iterative spanning forest," *Signal Processing Letters*, vol. 27, pp. 1440–1444, 2020.

[13] S. Guimarães, Y. Kenmochi, J. Cousty, Z. Patrocínio, and L. Najman, "Hierarchizing graph-based image segmentation algorithms relying on region dissimilarity: the case of the felzenszwalb-huttenlocher method," *Mathematical Morphology - Theory and Applications*, vol. 2, no. 1, pp. 55–75, 2017.

[14] Z. Wu, L. He, Z. Hu, Y. Zhang, and G. Wu, "Hierarchical segmentation evaluation of region-based image hierarchy," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 8, pp. 2718–2727, 2019.

[15] S. Todorovic and M. Nechyba, "Dynamic trees for unsupervised segmentation and matching of image regions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 11, pp. 1762–1777, 2005.

[16] J. Pont-Tuset and F. Marqués, "Supervised evaluation of image segmentation and object proposal techniques," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, pp. 1–1, 2015.

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

[18] X. Qin, Z. Zhang, C. Huang, M. Dehghan, O. Zaiane, and M. Jagersand, "U2-net: Going deeper with nested u-structure for salient object detection," *Pattern Recognition*, vol. 106, p. 107404, 2020.

[19] 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.

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

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

[22] L. Mansilla and P. Miranda, "Oriented image foresting transform segmentation: Connectivity constraints with adjustable width," in *29th Conference on Graphics, Patterns and Images (SIBGRAPI)*, 2016, pp. 289–296.

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

[24] S. Bobbia, R. Macwan, Y. Benezeth, K. Nakamura, R. Gomez, and J. Dubois, "Iterative boundaries implicit identification for superpixels segmentation: A real-time approach," *IEEE Access*, vol. 9, pp. 77 250–77 263, 2021.

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

[26] Z. Ban, J. Liu, and L. Cao, "Superpixel segmentation using gaussian mixture model," *IEEE Transactions on Image Processing*, vol. 27, no. 8, pp. 4105–4117, 2018.

[27] 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.

[28] 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.

[29] C. Jerônimo, F. Belém, S. Carneiro, Z. Patrocínio, L. Najman, A. Falcão, and S. Guimarães, "Graph-based supervoxel computation from iterative spanning forest," in *Discrete Geometry and Mathematical Morphology*, 2021, pp. 404–415.

[30] I. Barcelos, F. Belém, P. Miranda, A. Falcão, Z. Patrocínio, and S. Guimarães, "Towards interactive image segmentation by dynamic and iterative spanning forest," in *Discrete Geometry and Mathematical Morphology*, 2021, pp. 351–364.

[31] I. Barcelos, F. Belém, L. João, A. Falcão, and S. Guimarães, "Improving color homogeneity measure in superpixel segmentation assessment," in *35th Conference on Graphics, Patterns and Images (SIBGRAPI)*, vol. 1, 2022, pp. 79–84.

[32] F. Belém, I. Borlido, L. João, B. Perret, J. Cousty, S. Guimarães, and A. Falcão, "Fast and effective superpixel segmentation using accurate saliency estimation," in *Discrete Geometry and Mathematical Morphology*, 2022, pp. 261–273.

[33] ——, "Novel arc-cost functions and seed relevance estimations for compact and accurate superpixels," *Journal of Mathematical Imaging and Vision*, vol. 65, no. 5, pp. 770–786, 2023.

[34] I. Barcelos, F. Belém, L. João, Z. Patrocínio, A. Falcão, and S. Guimarães, "A comprehensive review and new taxonomy on superpixel segmentation," *ACM Computing Surveys*, vol. 56, no. 8, pp. 1–39, 2024.

[35] D. Vieira, I. Borlido, , F. Belém, Z. Patrocínio, A. Falcão, and S. Guimarães, "Streaming graph-based supervoxel computation based on dynamic iterative spanning forest," in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, 2024, pp. 90–104.

[36] F. Belém, F. Faria, A. Falcão, Z. Patrocínio, S. Guimarães, B. Perret, and J. Cousty, "Measuring hierarchiness of image segmentations," pp. 1–6, 2024, to appear.

[37] F. Belém, B. Perret, J. Cousty, S. Guimarães, and A. Falcão, "Efficient multiscale object-based superpixel framework," pp. 1–16, 2024, submitted to the Journal of the Brazilian Computer Society.

[38] L. Lacerda, F. Belém, A. Falcão, Z. Patrocínio, and S. Guimarães, "Seed-based superpixel re-segmentation for improving object delineation," pp. 1–13, 2024, submitted to the Iberoamerican Conference on Pattern Recognition.