Robust Seeded Image Segmentation Using Adaptive Label Propagation and Deep Learning-Based Contour Orientation

Aldimir José Bruzadin[1], Marilaine Colnago[2], Rogério Galante Negri[3], and Wallace Casaca[1(✉)]

[1] IBILCE, São Paulo State University, São José do Rio Preto - SP, Brazil
{aldimir.j.bruzadin,wallace.casaca}@unesp.br
[2] IQ - São Paulo State University, Araraquara - SP, Brazil
marilaine.colnago@unesp.br
[3] ICT, São Paulo State University, São José dos Campos - SP, Brazil
rogerio.negri@unesp.br

Abstract. Deep Learning has become a popular tool for addressing complex tasks in many computer vision applications. Label diffusion methods have also been a very effective technique for getting accurate segmentations of real-world images, as they combine user autonomy, versatility and accurateness through a user-friendly interface. In this paper, we propose a seeded segmentation framework for partitioning real-world images by combining deep contour learning and graph-based label propagation models. More precisely, our approach takes a CNN-type contour detection network to learn graph edge weights, which are used as input to solve a coupled energy minimization problem that diffuses the user-selected annotations to the desired targets. To accurately extract deep features from image contours while generating diffusion maps, we train a deep learning architecture that integrates a hierarchical neural network, a graph-based label propagation model and a loss function, allowing the coupled training mechanism to refine the results until convergence. We attest to the effectiveness and accuracy of the proposed approach by conducting both quantitative and qualitative assessments with existing seeded image segmentation methods.

Keywords: Seeded Segmentation · Contour Learning

1 Introduction

Image Segmentation refers to the process of splitting an image into disjointed partitions [7]. This task is a critical and mandatory step in many Computer

This research has been funded by the São Paulo Research Foundation (FAPESP – grants 2013/07375-0, #2021/03328-3 and #2021/01305-6), the Coordination for the Improvement of Higher Education Personnel (CAPES - Funding Code 001), and the National Council for Scientific and Technological Development (CNPq – grants #316228/2021-4 and #305220/2022-5).

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2023
O. Gervasi et al. (Eds.): ICCSA 2023, LNCS 13957, pp. 19–31, 2023.
https://doi.org/10.1007/978-3-031-36808-0_2

Vision applications, including image classification [24], object detection [13] and pattern recognition [22]. In this context, a very effective approach for clustering digital images is the seeded segmentation application, which consists of manually selecting a small set of user-labaled pixels to initialize the segmentation process [27]. In fact, seeded segmentation methods have proved to be effective and robust in obtaining precise image partitions, as it offers users autonomy, adaptability, and user-friendliness while capturing the marked targets among numerous objects in the image [6].

Some of the most widely recognized seed-based segmentation methods are *Random Walker* (RW) [14] and *Watershed* (WS) [9], as they are formulated in terms of the solid theory of energy minimization on graphs [19,26]. The RW solves a graph-based quadratic optimization problem to proceed with the image label propagation, and it has still served as inspiration for the development of current segmentation approaches [1,11,17], while the WS method stands out as a valuable technique devoted to capturing objects that overlap on the image.

Deep learning (DL) is another consolidated and powerful approach to partition images, which achieves very accurate segmentations for a variety of image-driven domains such as image registration [4] and remote sensing [25]. Indeed, the advance promoted by DL can be attributed to its ability to extract relevant features from large datasets of training images, allowing the trained models to continuously learn and improve over time. Popular deep learning architectures such as Convolutional Neural Networks (CNNs) can be successfully used to extract high-level features from images, hence improving the segmentation task significantly.

In the past few years, there has been growing interest in integrating deep learning techniques with graph-based label propagation strategies to improve the segmentation task in an intuitive and user-friendly way. For example, Kucharski et al. [20] utilize a CNN-based methodology to predict markers, which are used in conjunction with a WS transform to segment corneal endothelium. Hu et al. [18] take a seed-guided segmentation algorithm together with two parallel CNN-type networks to reduce the user intervention. Despite the satisfactory results, their methodology relies on specific sets of clicks rather than brushed pixels. Roth et al. [28] propose a weakly supervised approach that applies the RW model to generate pseudo-masks based on extreme points, so that these pseudo-masks are then used to train a network to produce the final masks. While their method is flexible enough to process medical data, it is not suitable to partition real-world digital images. Can et al. [5] also utilize the RW model as part of their advances, to obtain semi-automated segmentations of prostate and cardiac data by applying a Fully Convolutional Network (FCN). Similarly, Cerrone et al. [8] present a learning architecture that unifies a CNN-type network and the RW model for segmenting Electron Microscopy (EM) data, while Wolf et al. [32] focus on addressing the same segmentation problem (i.e., EM data), but by taking shape priors via a WS-driven framework coupled with a CNN instead of an RW-based one.

As previously discussed, a prevalent trait present in most seeded segmentation methods based on DL is that they are designed to be effective for very particular imaging modalities (e.g., medical data). Also, the requirement for specific data sources for training is another challenge difficult to overcome in practical application scenarios. Lastly, most purely DL-based segmentation approaches do not allow the user to interact and customize the segmentation process according to their needs and preferences, thus setting aside the autonomy and flexibility offered by different interactive segmentation tools.

In this paper, we introduce a new DL-based seeded segmentation framework for partitioning real-world images that is capable of tackling most of the aforementioned issues. The proposed approach combines the precision of deep contour learning with the high adaptability of seeded image segmentation towards delivering a fully learning approach that takes the so-called Convolutional Oriented Boundaries (COB) [21] network, to learn graph edge weights, whose values are inputted to the RW label propagation method to generate diffusion maps. In our approach, we also employ a gradient computation scheme to accelerate the training process, while still ensuring an appropriate balance between segmentation precision and computational efficiency.

2 Learning Label Propagation from Deep Oriented Boundaries and Seeded Maps

In this section, we present the main steps of our learning architecture for segmenting real-world images (see Fig. 1). First, we apply the COB network to recursively learn a set of edge weights from image contour maps. This set is then inputted together with the seed mask into the RW model to propagate the prior labeled data to the image targets, producing diffusion maps. Next, a cost function evaluates the diffusion maps so as to provide updated parameters to the COB network. Such a procedure is repeated until convergence to build the final segmentation model.

2.1 Deep Contour Orientation Mapping

To properly capture different types of contours from the target image, we utilize the *Convolutional Oriented Boundaries* (COB) network [21]. COB is a robust CNN-type network that enables deep learning of multiscale boundary representations of the image. The network builds upon the ResNet [16], but with the fully connected layers removed. Instead, it employs a series of convolutional layers paired that are combined with ReLU activation functions, arranged into five stages, where each stage extracts feature maps of similar sizes so as to obtain specific scales. The fine-scale mappings capture high-level contours, while the coarse ones detect lower-level boundaries, thus resulting in a multi-scale segmentation of the image.

Furthermore, COB enables for the extraction of multiple oriented contours representations of the image. To train our model, we utilize both horizontal- and

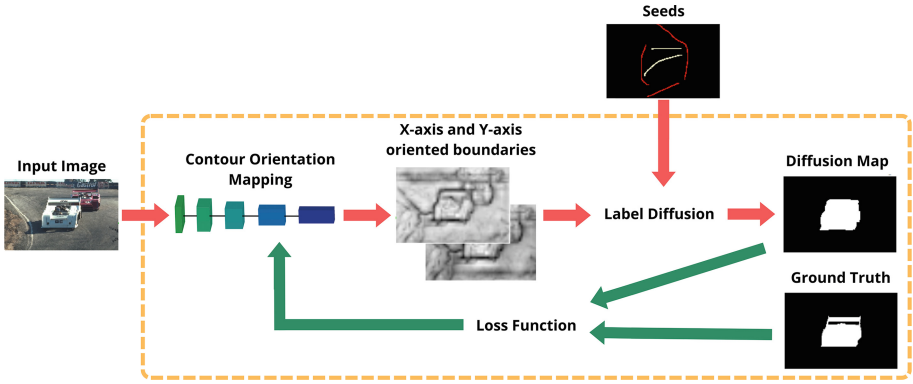

Fig. 1. Top: Initially, the input image is processed by a Convolutional Neural Network, generating deep contour orientation maps, which are taken to guide the label propagation step. The diffusion map is then obtained for the image targets by using a label propagation model. Bottom: The loss function is calculated from the diffusion map, while the backpropagation algorithm is applied to refine the network parameters.

vertical-oriented contour maps. The edge weights learned from the boundary maps are then assigned to the label diffusion graph, as described in Sect. 2.2. For additional information regarding the COB network, we refer to [16,21].

2.2 Weighted Graph Construction

After obtaining the contour orientation maps from the COB network [21], we build a weighted graph $G = (V, E, W_E)$, where the set V consists of image pixels, i.e., each node $k \in V$ corresponds to a single pixel P_k from the image. The set E is defined as the collection of edges connecting adjacent pixels by means of a 4-neighborhood stencil. The weight of each edge w_{kj} changes as the learning process advances, as it is obtained from the contour orientation maps. Specifically, for edges in the x-direction, we designate weight w_{kj} to correspond to the leftmost pixel value between P_k and P_j concerning the horizontal contour map. Similarly, for edges in the y-direction,we designate weight w_{kj} to denote the value of the topmost pixel between P_k and P_j in relation to the vertical contour map.

2.3 Label Propagation Model

Given a node $k \in V$, we can define its local neighborhood set $N(k)$ as the set of vertices j that are connected to k, i.e., $N(k) = j : (k, j) \in E$. The weighted valency of node k, represented by d_k, is calculated as the sum of the weights of all edges that are incident to k, which can be expressed as $d_k = \sum_{j \in N(k)} w_{kj}$.

We denote the *graph Laplacian matrix* by $\mathbf{L} = (L_{kj})$, which is computed as:

$$L_{kj} = \begin{cases} -w_{kj}, \text{ if } & j \in N(k) \\ d_k, \text{ if } & k = j \\ 0, \text{ otherwise} & \end{cases}. \tag{1}$$

Laplacian matrix $\mathbf{L}$ (1) can be decomposed in terms of the diagonal matrix $\mathbf{D}$, $D_{kk} = d_k$, and the weighted adjacency matrix $\mathbf{W}$, $W_{kj} = w_{kj}, (k, j) \in E$, and 0 otherwise: $\mathbf{L} = \mathbf{D} - \mathbf{W}$.

Define the set of seeded nodes by $S \subset V$, thus U $= V \backslash S$ gives the subset of unseeded nodes. We permute the columns and rows of matrix $\mathbf{L}$ in order to separate the labeled from the unseeded nodes:

$$\mathbf{P}_\pi \mathbf{L} \mathbf{P}_\pi^{\mathrm{T}} = \left[\begin{array}{c|c} \mathbf{L}_{\mathrm{S}} & \mathbf{R} \\ \hline \mathbf{R}^{\mathrm{T}} & \mathbf{L}_{\mathrm{U}} \end{array}\right], \tag{2}$$

where $\mathbf{P}_\pi$ is the permutation matrix. Submatrices $\mathbf{L}_{\mathrm{S}}$ and $\mathbf{L}_{\mathrm{U}}$ are symmetric due to the symmetry of matrix $\mathbf{L}$, while $\mathbf{R}$ defines the weighted adjacency matrix that encompasses both seeded and unseeded nodes in its rows and columns, respectively.

By taking Eq. (2), one can formulate a suitable graph-based energy minimization problem, and subsequently solve it. In our learning scheme, we choose the Random Walker (RW) method as the label propagation method.

In order to segment the image into binary segments, we partition the set S into disjoint subsets of the foreground (F) pixels and background (B) ones, i.e., $S = F \cup B$. Thus, the use of the RW-based label diffusion method leads to the minimization problem of the following energy equation [14]:

$$E(\mathbf{x}) = \left[\mathbf{x}_{\mathrm{S}}^{\mathrm{T}} \ \mathbf{x}_{\mathrm{U}}^{\mathrm{T}}\right] \left[\begin{array}{c|c} \mathbf{L}_{\mathrm{S}} & \mathbf{R} \\ \hline \mathbf{R}^{\mathrm{T}} & \mathbf{L}_{\mathrm{U}} \end{array}\right] \left[\begin{array}{c} \mathbf{x}_{\mathrm{S}} \\ \hline \mathbf{x}_{\mathrm{U}} \end{array}\right], \tag{3}$$

where $\mathbf{x} = \left[\begin{array}{c} \mathbf{x}_{\mathrm{S}} \\ \hline \mathbf{x}_{\mathrm{U}} \end{array}\right] \in \mathbb{R}^{|V|}$ is a vector with real entries where each x_k, $k \in U$, determines the probability that a random walker starting at an unlabeled node to reach a foreground node in S. Scalars x_F and x_B in $\mathbf{x}_{\mathrm{S}} = [(x_F)_{1\times|F|} \ (x_B)_{1\times|B|}]^T$ match the foreground and background tags in Eq. (3).

Since Eq. (3) is a quadratic form, its global minimum can be obtained by solving the following system of linear equations. This is achieved by differentiating $E(\mathbf{x})$ w.r.t. $\mathbf{x}_{\mathrm{U}}$ and setting the resulting expression equal to zero:

$$\mathbf{L}_{\mathrm{U}} \mathbf{x}_{\mathrm{U}} = -\mathbf{R}^{\mathrm{T}} \mathbf{x}_{\mathrm{S}}. \tag{4}$$

The existence of a unique solution in $\mathbf{x}_{\mathrm{U}}$ for Equation (4) is guaranteed only if the corresponding affinity graph G is connected and $S \neq \emptyset$ [14]. In our approach, we assume $x_F = 1$ and $x_B = 0$ in Equation (4) to obtain $\mathbf{x}_{\mathrm{U}}$ as a genuine diffusion map.

2.4 Loss Function Minimization

Our methodology is recursively trained by solving Eq. (4) and minimizing the discrepancy between the diffusion map $\mathbf{x}_{\mathrm{U}} = (x_i)$ and the ground-truth map $\mathbf{x}^*_{\mathrm{U}} = (x^*_i) \in \{0, 1\}$. We take as cost function the *Cross-Entropy Error*, which takes into account the *softmax* q_i of $\mathbf{x}_{\mathrm{U}}$:

$$l(\mathbf{x}^*_{\mathrm{U}}, \mathbf{x}_{\mathrm{U}}) = -\frac{1}{|U|} \sum_{k \in U} \Big[x^*_k \log(q_k) + (1 - x^*_k) \log(1 - q_k) \Big]. \tag{5}$$

When the loss value is high, the model's estimate deviates significantly from the ground-truth data. Conversely, a low loss value corresponds to diffusion maps that are closer to the desired segmentation.

2.5 Implementation Aspects, Gradient Simplification and Training

In order to computationally apply the Gradient Descent (GD) method while training our model, we adopt the approach outlined in [8] to obtain a formula for the partial derivatives that can be used to calculate the partial derivatives of the term $l(\mathbf{x}^*_{\mathrm{U}}, \mathbf{x}_{\mathrm{U}})$.

From Eq. (4), $l(\mathbf{x}^*_{\mathrm{U}}, \mathbf{x}_{\mathrm{U}}) = l\big(\mathbf{x}^*_{\mathrm{U}}, -\mathbf{L}_{\mathrm{U}}^{-1}\mathbf{R}^{\mathrm{T}}\mathbf{x}_{\mathrm{S}}\big)$, where $\mathbf{L}_{\mathrm{U}} = \mathbf{L}_{\mathrm{U}}(\mathbf{w}(\Theta))$ and $\mathbf{R} = \mathbf{R}(\mathbf{w}(\Theta))$ are given in terms of the edge weights $\mathbf{w} = (w_1, w_2, \ldots, w_{|W_E|})$, $w_k \in W_E$, and the neural network parameters, Θ. While the chain rule allows us to get the derivative of the cost function w.r.t. Θ, it also introduces the term $\partial \mathbf{x}_{\mathrm{U}} / \partial \mathbf{w} \in \mathbb{R}^{|U| \times |E|}$, which requires specific manipulations to determine an explicit computation formula. By differentiating Equation (4) w.r.t. $\mathbf{w}$ and rearranging the output, the following linear system is generated:

$$\mathbf{L}_{\mathrm{U}} \frac{\partial \mathbf{x}_{\mathrm{U}}}{\partial \mathbf{w}} = -\left(\frac{\partial \mathbf{L}_{\mathrm{U}}}{\partial \mathbf{w}} \mathbf{x}_{\mathrm{U}} + \frac{\partial \mathbf{R}^{\mathrm{T}}}{\partial \mathbf{w}} \mathbf{x}_{\mathrm{S}} \right). \tag{6}$$

The system of equations (6) incurs significant computational cost, as it demands solving $|E|$ linear systems during every GD iteration to obtain $\partial \mathbf{x}_{\mathrm{U}} / \partial \mathbf{w}$. To deal with this issue, instead of solving (6) for all edges, we pick a random subset of representative edges $\tilde{E} \subset E$, $|\tilde{E}| \ll |E|$, and only solve the linear systems that correspond to $\tilde{E}$, keeping the gradients as zero for the remaining edges. In particular, we select the subset $\tilde{E}$ so that the following heuristics are satisfied:

1. Edges (k, l) can not have any seeded node, i.e., $k \notin N(l)$, $\forall l \in S$.
2. Edges in $\tilde{E}$ must be placed as close as feasible to the object area, i.e., within the pixel portion delimited by the background seeds enclosing the object.

By applying such a simplification scheme, the gradient dimension $\partial \mathbf{x}_{\mathrm{U}} / \partial \mathbf{w}$ can be reduced from $|U| \times |E|$ to $|U| \times |\tilde{E}|$, which means that only sparse linear systems of size $|\tilde{E}|$ have to be solved. Figure 2 shows the described gradient simplification scheme.

Lastly, regarding our training design, the COB network integrated with the RW model was trained on the *Geostar Dataset* [15] by taking a total of 100

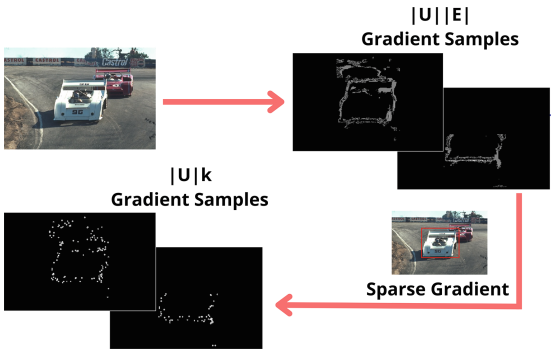

Fig. 2. Gradient simplification scheme.

epochs. This benchmark consists of 152 real-world images, each of which is paired with its corresponding ground-truth and seeded mask.

2.6 Segmentation

After the model is trained to propagate labels on regular images, the segmentation process is then accomplished by implementing a standard binarization technique like *Otsu* or any other suitable thresholding method on diffusion map $\mathbf{x}_{\mathrm{U}}$. Our approach assigns the labels $y_k \in \{x_B, x_F\} = \{0, 1\}$, where $k \in U$, based on the following cutting scheme:

$$y_k = \begin{cases} 1, \text{ if } \quad x_k \geq 1/2 \\ 0, \text{ otherwise} \end{cases} . \tag{7}$$

3 Results and Discussion

In this section, we provide a comprehensive evaluation of the proposed framework by comparing it to a range of existing state-of-the-art seed-based segmentation methods. More specifically, we perform both qualitative and quantitative experiments involving the following seeded segmentation methods: Laplacian Coordinates (LC) [6], Normalized Lazy Random Walker (NLRW) [3], One Cut (ONE) [30], and the classic Random Walker (RW) [14] method.

To assess the segmentation results, we take the popular *Grab-Cut Dataset* [29], which gathers a variety of real-world images with their counterparts of ground-truth segmentations, as well as several digital images taken from the well-known *Berkeley Image Segmentation Benchmark Database* [2]. Concerning the seeded maps used in our experimental assessments, we take the challenge set of *Grab-Cut sparse seeds* comprising small portions of foreground and background pixels of the images. Lastly, we use publicly available versions of the segmentation algorithms and made adjustments to their parameters based on the specific recommendations provided by the authors.

3.1 Quantitative Evaluation

To numerically gauge the quality of the resulting segmentations, we apply four traditional evaluation metrics used in the computer vision literature: *F1-Score* (F1) [10], *Rand Index* (RI) [31], *Variation of Information* (VOI) [23], and *Boundary Displacement Error* (BDE) [12]. To facilitate a comprehension of the metrics quantification, the symbol (↑) indicates a higher value is better, while (↓) implies that a lower score is better.

We start by numerically assessing the segmentation outputs in Table 1. From the tabulated measurements (mean and standard deviation values) in Table 1, one can see that the proposed framework achieves the best average scores for the evaluation metrics F1, RI, and BDE, while maintaining superior or competitive across all assessments in terms of std. Regarding the VOI metric, our approach achieves the second-lowest error rate, with only a slight difference compared to the best performer as determined by the ONE method.

Table 1. F1, RI, VOI and BDE metrics calculated for the segmentation methods. Bold values indicate the best scores.

Metrics	**Methods**				
	Proposed	**LC**	**NLRW**	**ONE**	**RW**
(↑) F1	**0.915** ± 0.080	0.907 ± 0.066	0.897 ± 0.088	0.866 ± 0.197	0.867 ± 0.095
(↑) RI	**0.941** ± 0.051	0.933 ± 0.071	0.920 ± 0.079	0.928 ± 0.095	0.898 ± 0.087
(↓) VOI	0.328 ± 0.210	0.381 ± 0.264	0.393 ± 0.289	**0.312** ± 0.250	0.468 ± 0.304
(↓) BDE	**5.985** ± 4.372	6.677 ± 6.343	6.894 ± 6.328	8.353 ± 13.931	9.117 ± 7.651

We also compare, in Fig. 3, the capability of the segmentation techniques in clustering the category of photographic images.

While the proposed framework remains highly accurate and stable, with very few instances of large variations in scores, the NLRW, ONE, and RW methods encounter challenges in dealing with many images. Although the LC method was competitive in most cases, it was still outperformed by the proposed framework in most of these cases. This is also observed in the VOI plot, which shows that although the ONE method performed better for most images, the proposed framework remained competitive in the majority of the cases.

3.2 Qualitative Evaluation

Aiming at assessing the qualitative performance of the segmentation algorithms, we provide in Fig. 4 the segmentations of a few real-world images. To enhance the visual inspection, we highlight the ground-truth segments (in green) and the outputs generated by each method (in red).

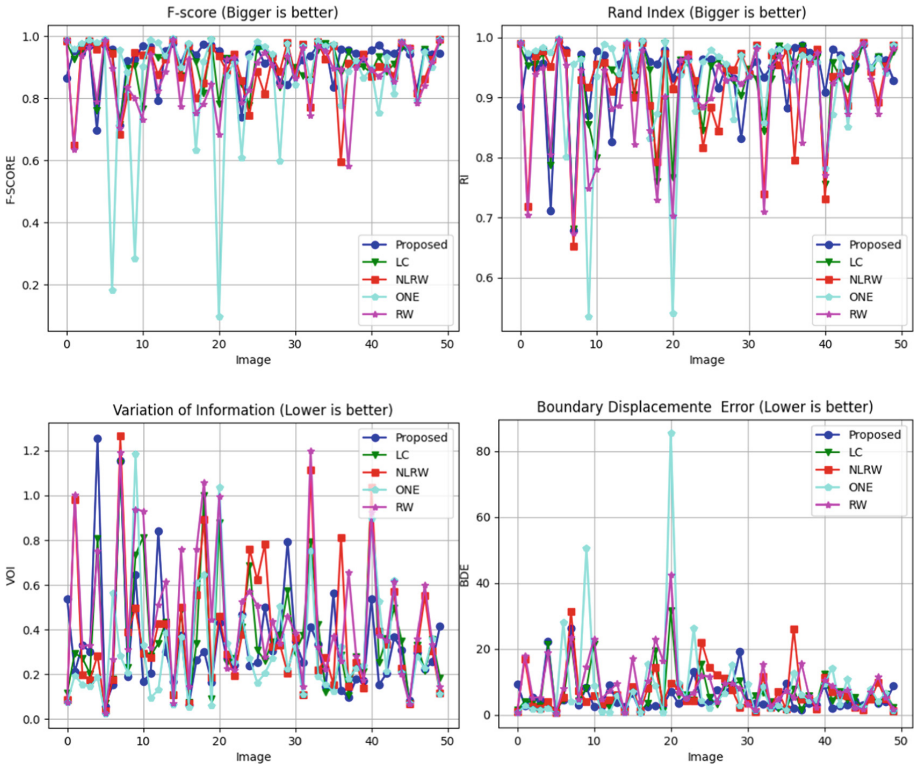

Fig. 3. Evaluation metrics calculated for several digital images. Segmentations produced by the proposed framework (dark blue), LC (green), NLRW (red), ONE (light blue), and RW (lilac). (Color figure online)

By visually examining the segmentation results, one can check that the outputs produced by our approach present smooth boundaries, leading to a better quality in terms of contour adherence and accurate object detection. In contrast, accurately capturing the targets within the images can pose obstacles for other segmentation algorithms.

From a more detailed visual inspection, one can observe that in the first three cases, the NLRW, ONE, and RW methods failed to accurately fit the targets, either by missing the object boundaries or propagating the seeds incorrectly. While our approach and LC method achieved better segmentations, the weak edges present in the first and the last images were not properly segmented by the LC method, as well as the narrow regions in the second image.

Overall, our approach delivers smoother partitions not only in terms of adherent contours but also in identifying small details of object edges.

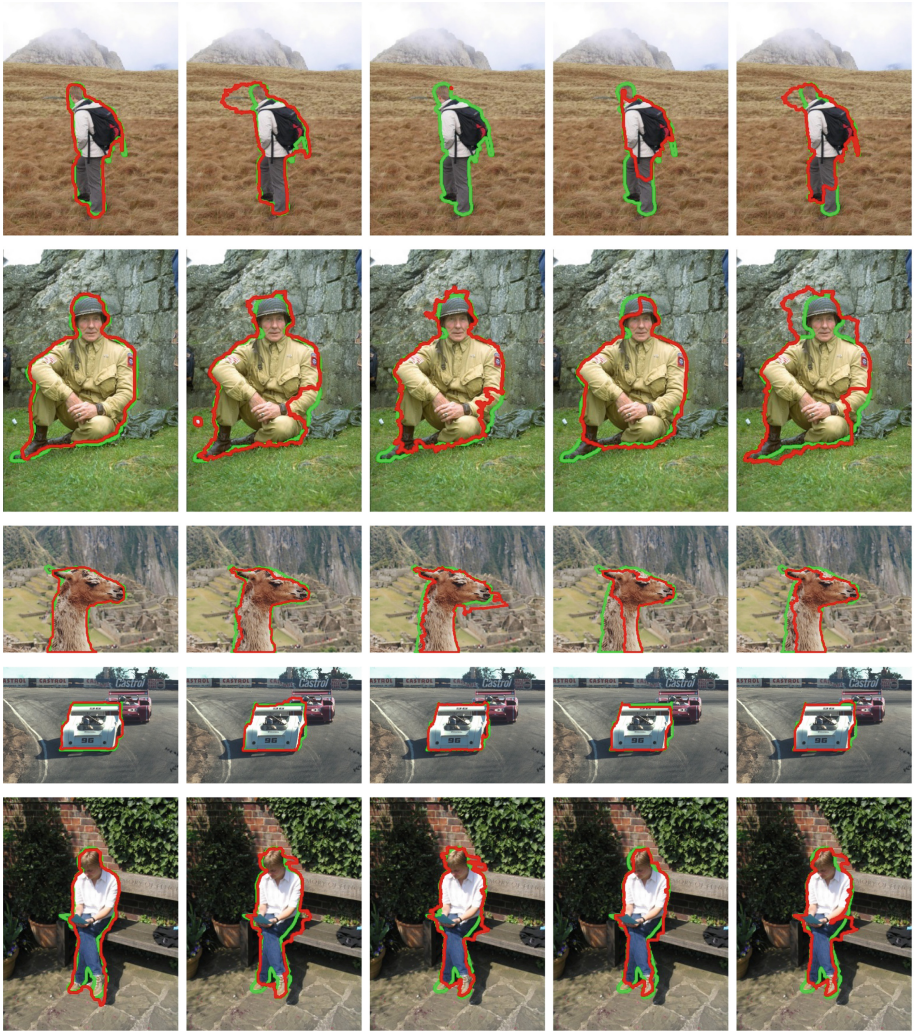

Fig. 4. Qualitative results. From left to right: segmentations produced by the proposed framework, LC, NLRW, ONE and RW.

4 Conclusion

We proposed in this paper a deep learning-based seeded image segmentation framework for real-world images that integrates a deep contour learning network with a label propagation method in a recursive and effective manner. The proposed framework enables the customized learning of the label propagation task for a variety of real-world scenes and photographic images, as shown and discussed in the experiments.

We compared the assertiveness and the fitting capability of our approach against existing segmentation methods from a battery of verification tests and metrics. As shown in the experiments, the trained model can robustly learn the graph edge weights so that the threshold between the foreground and the background segments can be accurately detected, resulting in smooth boundaries and high-quality segmentations concerning contour adherence and object detection.

As future work, we intend to extend our framework to deal with both the segmentation and classification tasks simultaneously. Furthermore, we aim to evaluate the accuracy of other label propagation models and alternative contour learning networks in an effort to improve the ability of the segmentation model to distinguish objects in complex scenarios, including the particular application of remote sensing.

Acknowledgment. The authors would like to thank the São Paulo Research Foundation (FAPESP – grants #2013/07375-0, #2021/03328-3 and #2021/01305-6), the Coordination for the Improvement of Higher Education Personnel (CAPES - Funding Code 001), and the National Council for Scientific and Technological Development (CNPq – grants #316228/2021-4 and #305220/2022-5) for providing resources that greatly contributed to the development of this research.

References

1. Aletti, G., Benfenati, A., Naldi, G.: A semiautomatic multi-label color image segmentation coupling dirichlet problem and colour distances. J. Imaging. **7**(10) (2021)
2. Arbeláez, P., Maire, M., Fowlkes, C., Malik, J.: Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **33**(5), 898–916 (2011)
3. Bampis, C.G., Maragos, P., Bovik, A.C.: Graph-driven diffusion and random walk schemes for image segmentation. IEEE Trans. Image Process. **26**(1), 35–50 (2017)
4. Benvenuto, G.A., Colnago, M., Dias, M.A., Negri, R.G., Silva, E.A., Casaca, W.: A fully unsupervised deep learning framework for non-rigid fundus image registration. Bioengineering **9**(8), 369 (2022)
5. Can, Y.B., Chaitanya, K., Mustafa, B., Koch, L.M., Konukoglu, E., Baumgartner, C.F.: Learning to segment medical images with scribble-supervision alone. In: Stoyanov, D., et al. (eds.) DLMIA/ML-CDS -2018. LNCS, vol. 11045, pp. 236–244. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-00889-5_27
6. Casaca, W., Gois, J.P., Batagelo, H.C., Taubin, G., Nonato, L.G.: Laplacian coordinates: theory and methods for seeded image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **43**(8), 2665–2681 (2021)
7. Casaca, W., Nonato, L.G., Taubin, G.: Laplacian coordinates for seeded image segmentation. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 384–391 (2014)
8. Cerrone, L., Zeilmann, A., Hamprecht, F.A.: End-to-end learned random walker for seeded image segmentation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 12559–12568 (2019)
9. Cousty, J., Bertrand, G., Najman, L., Couprie, M.: Watershed cuts: minimum spanning forests and the drop of water principle. IEEE Trans. Pattern Anal. Mach. Intell. **31**(8), 1362–1374 (2009)

10. Estrada, F.J., Jepson, A.D.: Benchmarking image segmentation algorithms. Int. J. Comput. Vision **85**(2), 167–181 (2009)
11. Fischer, M., Hepp, T., Gatidis, S., Yang, B.: Self-supervised contrastive learning with random walks for medical image segmentation with limited annotations. Computerized Medical Imaging and Graphics, p. 102174 (2023)
12. Freixenet, J., Muñoz, X., Raba, D., Martí, J., Cufí, X.: Yet another survey on image segmentation: region and boundary information integration. In: Heyden, A., Sparr, G., Nielsen, M., Johansen, P. (eds.) ECCV 2002. LNCS, vol. 2352, pp. 408–422. Springer, Heidelberg (2002). https://doi.org/10.1007/3-540-47977-5_27
13. Gao, G., Xu, G., Yu, Y., Xie, J., Yang, J., Yue, D.: MSCFNet: a lightweight network with multi-scale context fusion for real-time semantic segmentation. IEEE Trans. Intell. Transp. Syst. **23**(12), 25489–25499 (2021)
14. Grady, L.: Random walks for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **28**(11), 1768–1783 (2006)
15. Gulshan, V., Rother, C., Criminisi, A., Blake, A., Zisserman, A.: Geodesic star convexity for interactive image segmentation. In: 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 3129–3136 (2010)
16. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778 (2016)
17. Hu, J., Chen, Z., Zhang, R., Yang, M., Zhang, S.: Robust random walk for leaf segmentation. IET Image Proc. **14**(6), 1180–1186 (2020)
18. Hu, Y., Soltoggio, A., Lock, R., Carter, S.: A fully convolutional two-stream fusion network for interactive image segmentation. Neural Netw. **109**, 31–42 (2019)
19. Kim, K.I., Tompkin, J., Pfister, H., Theobalt, C.: Context-guided diffusion for label propagation on graphs. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2776–2784 (2015)
20. Kucharski, A., Fabijańska, A.: CNN-watershed: a watershed transform with predicted markers for corneal endothelium image segmentation. Biomed. Signal Process. Control **68**, 102805 (2021)
21. Maninis, K.K., Pont-Tuset, J., Arbeláez, P., Gool, L.V.: Convolutional oriented boundaries: from image segmentation to high-level tasks. IEEE Trans. Pattern Anal. Mach. Intell. **40**(4), 819–833 (2018)
22. Markovic, M., Malehmir, R., Malehmir, A.: Diffraction pattern recognition using deep semantic segmentation. Near Surface Geophys. **20**(5), 507–518 (2022)
23. Meilă, M.: Comparing clusterings: an axiomatic view. In: Proceedings of the 22nd International Conference on Machine Learning, p. 577–584 (2005)
24. Negri, R.G., da Silva, E.A., Casaca, W.: Inducing contextual classifications with kernel functions into support vector machines. IEEE Geosci. Remote Sens. Lett. **15**(6), 962–966 (2018)
25. Neupane, B., Horanont, T., Aryal, J.: Deep learning-based semantic segmentation of urban features in satellite images: a review and meta-analysis. Rem. Sens. **13**(4), 808 (2021)
26. Tai, X.-C., Bae, E., Chan, T.F., Lysaker, M. (eds.): Energy Minimization Methods in Computer Vision and Pattern Recognition, EMMCVPR 2015. LNCS, vol. 8932. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-14612-6
27. Ramadan, H., Lachqar, C., Tairi, H.: A survey of recent interactive image segmentation methods. Comput. Visual Med. **6**, 355–384 (2020)
28. Roth, H.R., Yang, D., Xu, Z., Wang, X., Xu, D.: Going to extremes: weakly supervised medical image segmentation. Mach. Learn. Knowl. Extract. **3**(2), 507–524 (2021)

29. Rother, C., Kolmogorov, V., Boykov, Y., Blake, A.: Interactive foreground extraction using graph cut. In: Advances in Markov Random Fields for Vision and Image Processing (2011)
30. Tang, M., Gorelick, L., Veksler, O., Boykov, Y.: Grabcut in one cut. In: 2013 IEEE International Conference on Computer Vision, pp. 1769–1776 (2013)
31. Warrens, M.J., van der Hoef, H.: Understanding the rand index. In: Imaizumi, T., Okada, A., Miyamoto, S., Sakaori, F., Yamamoto, Y., Vichi, M. (eds.) Advanced Studies in Classification and Data Science, pp. 301–313 (2020)
32. Wolf, S., Schott, L., Kothe, U., Hamprecht, F.: Learned watershed: End-to-end learning of seeded segmentation. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 2011–2019 (2017)