The Role of Optimum Connectivity in Image Segmentation: Can the algorithm learn object information during the process?

Alexandre Xavier Falcão

Laboratory of Image Data Science Institute of Computing — University of Campinas

afalcao@ic.unicamp.br







- The role of optimum connectivity in image segmentation.
- Research goals.
- The methodology used for the DGCI'2019 paper [1] based on the Image Foresting Transform (IFT) [6].
- Experiments and results.
- Conclusion and research directions.

• Semantic segmentation involves object detection and object delineation.

- Semantic segmentation involves object detection and object delineation.
- Semantic segmentation models (e.g., deep neural networks [25, 26]) have provided successful object detection and identification.

- Semantic segmentation involves object detection and object delineation.
- Semantic segmentation models (e.g., deep neural networks [25, 26]) have provided successful object detection and identification.
- However, object delineation cannot be solved by simply thresholding some probability map derived from the network.

- Semantic segmentation involves object detection and object delineation.
- Semantic segmentation models (e.g., deep neural networks [25, 26]) have provided successful object detection and identification.
- However, object delineation cannot be solved by simply thresholding some probability map derived from the network.
- Indeed, the need for optimum connectivity in this context has already been recognized [33].

Example of the problem when the scene contains objects of different sizes and shapes [26].



Object categories: dinning table, broccoli, bowl, and person.

Example of the problem when the scene contains objects of different sizes and shapes [26].



Object categories: dinning table, broccoli, bowl, and person.

Example of the problem when the scene contains objects of different sizes and shapes [26].



Object categories: dinning table, broccoli, bowl, and person.

User input for effective object delineation cannot be usually reduced to a single intervention [27].



User input for effective object delineation cannot be usually reduced to a single intervention [27].



The problem is even more critical when objects with similar image properties appear connected to each other.



The problem is even more critical when objects with similar image properties appear connected to each other.



It is paramount that user intervention for correction be simple, fast, and effective by quickly adapting the model or creating a new one for the specific object of interest [7].



The object is an optimum-path forest rooted at its internal markers.

Learn object information from each given image and the users' actions during interactive segmentation with minimum user effort.



Learn object information from each given image and the users' actions during interactive segmentation with minimum user effort.



The object model may be active in its learning process, specific for each given image, and generalized for new images when the number of examples is high enough. One example of a generalized model is a multi-object statistical atlas adaptive for anomalous MR-image segmentation [15].



The shape model is built from normal examples (images and masks), but it can identify anomalous regions in test images.

Research goals: shape models and optimum connectivity

The model estimates the markers and the objects are delineated by optimum connectivity.





MR-image segmentation of the left and right brain hemispheres, and the cerebellum without pons, medulla, and spinal cord.

Research goals: shape models and optimum connectivity

The model estimates the markers and the objects are delineated by optimum connectivity.





MR-image segmentation of the left and right brain hemispheres, and the cerebellum without pons, medulla, and spinal cord.

Research goals: optimum connectivity only

Marker estimation and object delineation can also be mostly based on optimum connectivity in some cases [14].



CT-image segmentation of the left and right lungs, and traquea-and-bronchi in anomalous images.

Research goals: optimum connectivity only

Marker estimation and object delineation can also be mostly based on optimum connectivity in some cases [14].



CT-image segmentation of the left and right lungs, and traquea-and-bronchi in anomalous images.

Research goals: fast segmentation correction

Finally, the segmentation result from any method can be converted into an optimum-path forest rooted at computed markers [21, 22] for fast interactive corrections in a differential way [12, 13].



CT-image segmentation of foot bones.

Methodology adopted for the conference paper

We have used the Image Foresting Transform (IFT) for the design of image operators based on optimum connectivity [5, 6].



The block in red extracts object information prior delineation [1, 10, 11].

Image Foresting Transform (IFT)



 Image elements t may be pixels, superpixels, or objects from some image set, each represented by a feature vector (left).

Image Foresting Transform (IFT)



- Image elements *t* may be pixels, superpixels, or objects from some image set, each represented by a feature vector (left).
- An adjacency relation A transforms the set of image elements into a graph in the image domain (or feature space) (center).

Image Foresting Transform (IFT)



- Image elements t may be pixels, superpixels, or objects from some image set, each represented by a feature vector (left).
- An adjacency relation A transforms the set of image elements into a graph in the image domain (or feature space) (center).
- A connectivity function f assigns to any path π_t with terminus t a cost f(π_t) and the minimization

$$V(t) = \min_{orall \pi_t \in \Pi_t} \{f(\pi_t)\}$$

is solved by propagating paths in a non-decreasing order of costs, leading to an optimum-path forest rooted at the minima of the cost map V (right).

Superpixels (supervoxels in 3D) can be defined by combining parametric and geometric image properties to balance boundary adherence and shape regularity [2].



Each superpixel is one optimum-path tree rooted at a representative seed pixel.

Superpixel segmentation

The method is called Iterative Spanning Forest (ISF) [2].



 Initial seed estimation is important to locate relevant segments, while seed recomputation improves delineation.

Superpixel segmentation

The method is called Iterative Spanning Forest (ISF) [2].



- Initial seed estimation is important to locate relevant segments, while seed recomputation improves delineation.
- Object-based ISF (OISF) incorporates object information in the connectivity function to increase boundary adherence for a specific object of interest [3].

Superpixel segmentation

The method is called Iterative Spanning Forest (ISF) [2].



- Initial seed estimation is important to locate relevant segments, while seed recomputation improves delineation.
- Object-based ISF (OISF) incorporates object information in the connectivity function to increase boundary adherence for a specific object of interest [3].
- Recursive ISF (RISF) applies ISF recursively on superpixel graphs to obtain a hierarchical image segmentation [4].

Clustering of superpixels and object saliency estimation

Superpixels can considerably reduce the processing time for data clustering.



Clusters are optimum-path trees in the feature space rooted at each dome of a probability density function [1, 8, 9].

Clustering is used to select the most suitable markers for object saliency estimation.



The method must avoid pixels from background markers that fall into clusters most populated by pixels from object markers.

Clustering of superpixels and object saliency estimation



- A seed set S is defined by eliminating the undesired background pixels [1, 10].
- The saliency value of each pixel t is $\frac{V_b(t)}{V_o(t)+V_b(t)}$, where $V_o(t)$ and $V_b(t)$ are costs of optimum paths in the feature space from object and background seeds in S, respectively.

• Object delineation can use connectivity functions that optimally cuts the graph based on its arc weights [18, 19].

- Object delineation can use connectivity functions that optimally cuts the graph based on its arc weights [18, 19].
- The arc weights are usually computed prior delineation as a function of image and object properties w/o or w/ shape constraints [10, 20, 34, 35].

- Object delineation can use connectivity functions that optimally cuts the graph based on its arc weights [18, 19].
- The arc weights are usually computed prior delineation as a function of image and object properties w/o or w/ shape constraints [10, 20, 34, 35].
- We show here that object information can also be extracted from the *growing trees* for dynamic arc-weight assignment a method called Dynamic Trees [1, 16].

Example of Dynamic Trees — each region is an optimum-path tree, growing in a non-decreasing order of path costs.



Object delineation

Example of Dynamic Trees — each region is an optimum-path tree, growing in a non-decreasing order of path costs.



- The experiments involve robot [1] and real [16] users selecting seeds on Grab-Cut images.
- The robot selects one seed (disk with radius 1) per iteration at the center of an error component and stops after 15 iterations or under trivial criteria.



- The experiments involve robot [1] and real [16] users selecting seeds on Grab-Cut images.
- The robot selects one seed (disk with radius 1) per iteration at the center of an error component and stops after 15 iterations or under trivial criteria.



- The experiments involve robot [1] and real [16] users selecting seeds on Grab-Cut images.
- The robot selects one seed (disk with radius 1) per iteration at the center of an error component and stops after 15 iterations or under trivial criteria.



- The experiments involve robot [1] and real [16] users selecting seeds on Grab-Cut images.
- The robot selects one seed (disk with radius 1) per iteration at the center of an error component and stops after 15 iterations or under trivial criteria.



• Real users are represented by two benchmarks of scribbles, then this experiment involves a single intervention for object delineation.



Examples of scribbles from Gulshan's [28] and Andrade's [29] datasets, respectively.

Compared object delineation algorithms and arc-weight models include.

- Watershed Cut [30] (WS Cut): the dissimilarity between adjacent nodes based on image properties.
- Min-Cut [32]: the similarity between adjacent nodes based on image properties.
- Dynamic Trees [16]: the dissimilarity between target node and expanding optimum-path tree.
- Closest Dyn. Trees [16]: the dissimilarity between target node and expanding optimum-path forest.

All of them can add pairwise dissimilarity/similarity between adjacent nodes based on object saliency.

Mean Dice coefficient and number of iterations required for convergence with the robot.

Method	Mean Dice (%)	Iterations
Dyn. Trees + Obj. Sal. (w_6)	$\textbf{0.968} \pm \textbf{0.0295}$	11.8
WS Cut + Obj. Sal. (w ₂)	0.961 ± 0.0421	12.6
Dyn. Trees (w ₅)	0.961 ± 0.0491	11.9
DEC	0.942 ± 0.0814	4.00
WS Cut (w_1)	0.933 ± 0.0854	13.8
Min-Cut + Obj. Sal. (w ₄)	0.932 ± 0.0845	13.3
Min-Cut (w ₃)	0.918 ± 0.108	13.6

Results using the robot's markers

Mean Dice coefficient per iteration on the unseen test set with the robot.



Object saliency was not used for arc-weight assignment in the results above.

Mean Dice coefficient and mean execution time (seconds) for a single user intervention using Gulshan's markers [28].

Method	Mean Dice (%)	Time (secs)
Dyn. Trees (w ₅)	$\textbf{84.0} \pm \textbf{1.3}$	0.042 ± 0.016
Closest Dyn. Tree (w ₉)	81.9 ± 1.8	4.633 ± 2.573
Min-Cut (w ₃)	76.2 ± 1.6	0.230 ± 0.167
WS Cut (w_1)	75.6 ± 1.6	$\textbf{0.038} \pm \textbf{0.012}$
$PW_{q=2}$	72.3 ± 1.7	0.966 ± 0.300

Object saliency was not used for arc-weight assignment in the results above.

Mean Dice coefficient and mean execution time (seconds) for a single user intervention using Andrade's markers [29].

Method	Mean Dice (%)	Time (secs)
Closest Dyn. Trees (w ₉)	$\textbf{95.4} \pm \textbf{0.04}$	8.460 ± 3.802
Dyn. Trees (w ₅)	92.1 ± 0.08	0.046 ± 0.018
Min-Cut (w ₃)	90.6 ± 0.08	0.123 ± 0.075
$PW_{q=2}$	89.9 ± 0.08	1.015 ± 0.299
WS Cut (w_1)	89.5 ± 0.09	$\textbf{0.039} \pm \textbf{0.012}$

Object saliency was not used for arc-weight assignment in the results above.

Qualitative results

Results obtained by using the Andrade's scribbles dataset.



Qualitative results

Results obtained by using the Andrade's scribbles dataset.



Original



Min-Cut









Closest Dyn. Tree

- Optimum connectivity is important to separate objects with similar image properties.
- It is possible to create semantic models during interactive image segmentation (prior and during object delineation).
- The object information from that semantic model can considerably improve superpixel and object delineation.

• Which are the semantic models that can learn object information from a single (few) image (s)? [10]

- Which are the semantic models that can learn object information from a single (few) image (s)? [10]
- Can we improve a semantic model from one or multiple users' inputs, w/o or w/ shape constraints [20], along a collaborative segmentation process? [17]

- Which are the semantic models that can learn object information from a single (few) image (s)? [10]
- Can we improve a semantic model from one or multiple users' inputs, w/o or w/ shape constraints [20], along a collaborative segmentation process? [17]
- Can we use a semantic model to suggest relevant markers for its learning process? [23]

- Which are the semantic models that can learn object information from a single (few) image (s)? [10]
- Can we improve a semantic model from one or multiple users' inputs, w/o or w/ shape constraints [20], along a collaborative segmentation process? [17]
- Can we use a semantic model to suggest relevant markers for its learning process? [23]
- How can we exploit superpixel graphs to improve delineation and reduce user effort? [24]

- Which are the semantic models that can learn object information from a single (few) image (s)? [10]
- Can we improve a semantic model from one or multiple users' inputs, w/o or w/ shape constraints [20], along a collaborative segmentation process? [17]
- Can we use a semantic model to suggest relevant markers for its learning process? [23]
- How can we exploit superpixel graphs to improve delineation and reduce user effort? [24]
- Can we further reduce user effort in segmentation correction when resuming segmentation into an optimum-path forest? [21]

FAPESP, CNPq, Université of Paris-Est, Jean Cousty, my co-author in [1], Jordão Bragantini, and some other currently most active collaborators on the IFT: Krzysztof Ciesielski, Paulo Miranda, Fábio Cappabianco, João Paulo Papa, Thiago Spina, Ananda Chowdhury, Silvio Guimarães, Felipe Belém, Felipe Galvão, John Vargas-Muñoz, Samuel Martins, and Azael Sousa. FAPESP, CNPq, Université of Paris-Est, Jean Cousty, my co-author in [1], Jordão Bragantini, and some other currently most active collaborators on the IFT: Krzysztof Ciesielski, Paulo Miranda, Fábio Cappabianco, João Paulo Papa, Thiago Spina, Ananda Chowdhury, Silvio Guimarães, Felipe Belém, Felipe Galvão, John Vargas-Muñoz, Samuel Martins, and Azael Sousa.

Thank you!

- A.X. Falcão and J. Bragantini. The Role of Optimum Connectivity in Image Segmentation: Can the algorithm learn object information during the process?, 21st IAPR International Conference on Discrete Geometry for Computer Imagery, LNCS, 2019, to appear.
- [2] J.E. Vargas-Muñoz, A. S. Chowdhury, E. B. Alexandre, F. L. Galvão, P. A. V. Miranda and A. X. Falcão. An Iterative Spanning Forest Framework for Superpixel Segmentation, *IEEE Transactions on Image Processing*, 10.1109/TIP.2019.2897941, 2019, in press.
- F. de C. Belém, S.J.F. Guimarães and A.X. Falcão. Superpixel Segmentation by Object-based Iterative Spanning Forest. 23rd Iberoamerican Congress on Pattern Recognition, CIARP 2018: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, doi 10.1007/978-3-030-13469-3_39, LNCS 11401, Madrid, Spain, pp. 334-341, 2019.

- [4] F.L. Galvão, A. X. Falcão, and A.S. Chowdhury. RISF: Recursive Iterative Spanning Forest for Superpixel Segmentation. 31st SIBGRAPI: Conference on Graphics, Patterns and Images, 10.1109/SIBGRAPI.2018.00059, Parana, pp. 408-415, 2018.
- [5] K.C. Ciesielski, A.X. Falcão, and P.A.V. Miranda. Path-value functions for which Dijkstra's algorithm returns optimal mapping. *Journal of Mathematical Imaging and Vision*, 10.1007/s10851-018-0793-1, vol. 60, pp. 1025-1036, 2018.
- [6] A. X. Falcão, J. Stolfi and R. de Alencar Lotufo. The image foresting transform: theory, algorithms, and applications. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10.1109/TPAMI.2004.1261076, 26(1), pp. 19-29, 2004.
- T.V. Spina, P.A.V. de Miranda, and A.X. Falcão. Hybrid approaches for interactive image segmentation using the Live Markers paradigm. *IEEE Trans. on Image Processing*, doi: 10.1109/TIP.2014.2367319, 23(12), pp. 5756–5769, 2014.

- [8] F.A.M. Cappabianco, A.X. Falcão, C.L. Yasuda, and J.K. Udupa. Brain Tissue MR-Image Segmentation via Optimum-Path Forest Clustering. *Computer Vision and Image Understanding*. 10.1016/j.cviu.2012.06.002, 116(10), pp. 1047–1059, 2012.
- [9] L.M. Rocha , F.A.M. Cappabianco and A.X. Falcão. Data Clustering as an Optimum-Path Forest Problem with Applications in Image Analysis. *Intl. Journal of Imaging Systems and Technology*, doi.wiley.com/10.1002/ima.20191, 19(2), pp. 50–68, 2009.
- [10] T.V. Spina , P.A.V. de Miranda and A.X. Falcão. Intelligent understanding of user interaction in image segmentation. *Intl. Journal of Pattern Recognition and Artificial Intelligence*. 10.1142/S0218001412650016, 26(2), pp. 1265001-1–1265001-26, 2012.
- [11] P.A.V. Miranda , A.X. Falcão and J.K. Udupa. Synergistic Arc-Weight Estimation for Interactive Image Segmentation

using Graphs. *Computer Vision and Image Understanding*. 10.1016/j.cviu.2009.08.001, vol. 114, issue 1, pp 85-99, 2010.

- [12] M. A. T. Condori, F. A. M. Cappabianco, A. X. Falcão and P. A. V. D. Miranda. Extending the Differential Image Foresting Transform to Root-Based Path-Cost Functions with Application to Superpixel Segmentation, 30th SIBGRAPI Conference on Graphics, Patterns and Images, 10.1109/SIBGRAPI.2017.8, pp. 7-14, 2017.
- [13] A.X. Falcão and F.P.G. Bergo . Interactive Volume Segmentation with Differential Image Foresting Transforms. *IEEE Trans. on Medical Imaging*, 10.1109/TMI.2004.829335, 23(9), pp. 1100–1108, 2004.
- [14] A.M. Sousa, E. Bagatin, G.S. Meirelles, and A.X. Falcão. A computational method to aid the detection and annotation of pleural lesions in CT images of the thorax. SPIE on Medical Imaging: Image Processing, San Diego, CA, Feb 2019, to appear.

- [15] S.B. Martins, T.V. Spina, and A.X. Falcão. A multi-object statistical atlas adaptive for deformable registration errors in anomalous medical image segmentation. *SPIE on Medical Imaging: Image Processing*, 10.1117/12.2254477, Vol. 10133, 8 pages, 2017.
- [16] J. Bragantini, S.B. Martins, C. Castelo-Fernández, and A.X. Falcão. Graph-based Image Segmentation using Dynamic Trees. 23rd Iberoamerican Congress on Pattern Recognition, CIARP 2018: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications, doi 10.1007/978-3-030-13469-3_55, LNCS 11401, Madrid, Spain, pp. 470–478, 2019.
- [17] T.V. Spina, J. Stegmaier, A.X. Falcão, E. Meyerowitz, and A. Cunha. SEGMENT3D: A Web-based Application for Collaborative Segmentation of 3D Images Used in the Shoot Apical Meristem. *IEEE Intl. Symp. on Biomedical Imaging (ISBI)*. 10.1109/ISBI.2018.8363600, pp. 391-395, 2018.

[18] K.C. Ciesielski, J.K. Udupa, A. X. Falcão, and P. A. V. Miranda. Fuzzy Connectedness Image Segmentation in Graph Cut Formulation: A Linear-Time Algorithm and a Comparative Analysis. *Journal of Mathematical Imaging and Vision*, 44(3), pp. 375–398, 2012.

[19] Bejar H.H.C., Mansilla L.A.C., Miranda P.A.V., Efficient Unsupervised Image Segmentation by Optimum Cuts in Graphs. 23rd Iberoamerican Congress on Pattern Recognition, CIARP 2018: Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications. LNCS 11401, doi 10.1007/978-3-030-13469-3_42, pp. 359-367, 2019.

[20] L. M. C. Leon and P. A. V. D. Miranda, Multi-Object Segmentation by Hierarchical Layered Oriented Image Foresting Transform, *30th SIBGRAPI Conference on Graphics, Patterns and Images* (SIBGRAPI), doi: 10.1109/SIBGRAPI.2017.17, pp. 79–86, 2017. [21] A.C.M. Tavares, P.A.V. Miranda, T.V. Spina, and A.X. Falcão. A Supervoxel-based Solution to Resume Segmentation for Interactive Correction by Differential Image-Foresting Transforms. 13th International Symposium on Mathematical Morphology and its Application to Signal and Image Processing, LNCS 10225, 10.1007/978-3-319-57240-6_9, pp. 107–118, 2017.

- [22] P.A.V. Miranda, A.X. Falcão, G. Ruppert and F. Cappabianco. How to Fix any 3D Segmentation Interactively via Image Foresting Transform and its use in MRI Brain Segmentation. 8th IEEE Intl. Symp. on Biomedical Imaging: From Nano to Macro (ISBI), 10.1109/ISBI.2011.5872811, pp. 2031–2035, 2011.
- [23] T.V. Spina, S.B. Martins, and A.X. Falcão. Interactive Medical Image Segmentation by Statistical Seed Models. XXIX SIBGRAPI - Conference on Graphics, Patterns and Images, doi: 10.1109/SIBGRAPI.2016.045, pp. 273–280, 2016.

- [24] P. Rauber, A.X. Falcão, T.V. Spina, and P.J. de Rezende. Interactive Segmentation by Image Foresting Transform on Superpixel Graphs. Proc. of the XXVI SIBGRAPI - Conference on Graphics, Patterns and Images, 10.1109/SIBGRAPI.2013.27, pp. 131–138, 2013.
- [25] He, K., Zhang, X., Ren, S., and Sun, J., Deep Residual Learning for Image Recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016.
- [26] He, K., Gkioxari, G., Dollár, P., and Girshick, R., Mask R-CNN. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2961-2969, 2017.
- [27] Maninis, K. K., Caelles, S., Pont-Tuset, J., and Van Gool, L., Deep Extreme Cut: From Extreme Points to Object Segmentation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 616-625, 2018.

- [28] Gulshan, V., Rother, C., Criminisi, A., Blake, A., and Zisserman, A., Geodesic star convexity for interactive image segmentation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3129-3136, 2010.
- [29] Andrade, F., and Carrera, E. V., Supervised evaluation of seed-based interactive image segmentation algorithms. In 20th IEEE Symposium on Signal Processing, Images and Computer Vision (STSIVA), pp. 1-7, 2015.
- [30] Cousty, J., Bertrand, G., Najman, L., and Couprie, M., Watershed Cuts: Thinnings, Shortest Path Forests, and Topological Watersheds, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 32 (5), pp. 925-939, 2010.
- [31] Couprie, C., Grady, L., Najman, L., and Talbot, Power Watershed: A Unifying Graph-Based Optimization Framework, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 33(7), pp. 1384–1399, 2011.

[32] Boykov, Y. and Kolmogorov, V., An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision, *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 26(9), pp. 1124–1137, 2004.

- [33] N. Xu, B. Price, S. Cohen, J. Yang and T. Huang. Deep Interactive Object Selection. *IEEE Conf. on Computer Vision* and Pattern Recognition (CVPR), 10.1109/CVPR.2016.47, pp. 373-381, 2016.
- [34] L. A. C. Mansilla and P. A. V. Miranda, Oriented Image Foresting Transform Segmentation: Connectivity Constraints with Adjustable Width, 29th SIBGRAPI Conference on Graphics, Patterns and Images, 10.1109/SIBGRAPI.2016.047, pp. 289-296, 2016.
- [35] P. A. V. Miranda and L. A. C. Mansilla. Oriented Image Foresting Transform Segmentation by Seed Competition, *IEEE Transactions on Image Processing*, 10.1109/TIP.2013.2288867, 23(1), pp. 389-398, 2014.