Semi-Supervised Segmentation with Topological Penalization for Quantifying Retinal Degeneration in Histological Sections

Benjamin Dorra, Giacomo Nardi, Antoine Habis, Jean-Philippe Annereau, Jean-Christophe Olivo-Marin, Luis Briseño-Roa, Meas-Yedid Vannary

B. Dorra et al., ”Semi-Supervised Segmentation with Topological Penalization for Quantifying Retinal Degeneration in Histological Sections,” 2024 IEEE International Symposium on Biomedical Imaging (ISBI), Athens, Greece, 2024, pp. 1-5, doi: 10.1109/ISBI56570.2024.10635431. keywords: Atrophy;Image segmentation;In vivo;Correlation;Manuals;Retina;Mice;WSI;histology;retina;segmentation;deep learning;CNN;semi-supervision;topological penalization,

SEMI-SUPERVISED SEGMENTATION WITH TOPOLOGICAL PENALIZATION FOR QUANTIFYING RETINAL DEGENERATION IN HISTOLOGICAL SECTIONS

Benjamin Dorra[⋆†] *Giacomo Nardi*[⋆] *Antoine Habis*[⋆] *Jean-Philippe Annereau*[†]
Jean-Christophe Olivo-Marin[⋆] *Luis Briseño-Roa*[†] *Vannary Meas-Yedid*[⋆]

[⋆] Institut Pasteur, Universite de Paris Cité, CNRS UMR 3691, BioImage Analysis Unit, Paris, France

[†] Medetia Pharmaceuticals, Imagine Institute, 75015 Paris, France

ABSTRACT

Retinal degeneration is a hallmark of various degenerative ocular diseases, and its accurate quantification is crucial for non-clinical pharmacological studies. Here, we describe a method to assess retinal atrophy by measuring the layers' thickness in histological sections stained with hematoxylin and eosin. We use a semi-supervised approach with topological penalization to segment layers followed by thickness measurements along the layers' medial axes. The method is tested on mice for four layers, obtaining a high correlation with manual measurements collected by biologists.

***Index Terms*—** WSI, histology, retina, segmentation, deep learning, CNN, semi-supervision, topological penalization

1. INTRODUCTION

Retinal atrophy is a thinning of the retina that appears in different eye diseases, leading to visual impairment in non-clinical models (mice) and patients. Accurate measurement of retinal atrophy is essential in studying these diseases. It is critical for pharmacological research, in particular drug discovery.

The retina is composed of several layers, each with distinct functions (see Fig. 1): the ganglion cell layer processes visual information; the inner nuclear layer (INL) contains bipolar, amacrine, and horizontal cells that amplify, extract and compress signal; the outer plexiform layer (OPL) contains the synaptic terminals of photoreceptors, horizontal cells, and bipolar cells, connecting those different cells; the outer nuclear layer (ONL) holds the nuclei of photoreceptor cells; and the inner (IS) and outer segments (OS) capture the light [1].

Retinal atrophy is not homogeneous and degradation of different layers can occur at different rates in degenerative ocular diseases. For medical and research purpose it is therefore useful to separately measure thickness for different layers. For instance, the ONL thickness is often used on its own [2]. Measuring thickness for a layer usually requires manually drawing line segments (also called calipers), whose lengths are used as local thickness values. Different methods have been developed to automate this process across several imaging modalities. The main principle is to segment a layer before locally measuring border-to-border distance.

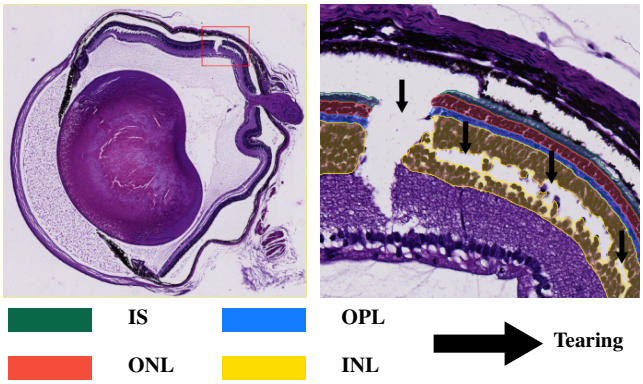

Fig. 1: Retina section (left), layers structure with tearing (right).

In fluorescence microscopy, the authors of [3] use nuclei detection and morphological operations to construct the contours of the ONL, and the thickness is measured as the distance between borders along the medial axis. ThicknessTool [2] uses thresholding and morphological operations to segment ONL and INL and thoroughly validates its measurements against manual values.

In optical coherence tomography (OCT), segmentation and thickness measurement of retina layers has been widely studied: Most state-of-the-art methods use a variant of U-Net [4, 5] to segment every retinal layer, and some include semi-supervision [6] or topological constraints [7, 8].

Histological retina images offer great value for preclinical research due to their high resolution and contrast. A method is developed in [9] to segment and measure retina layer thickness in histological images: ONL and INL layers are segmented via a deep learning model, and thickness is measured on local maxima of the distance function to the layer's boundary. The authors have access to relatively few labeled data (14 annotated WSIs, all healthy), and because they use deep learning models trained in a supervised fashion this causes a

severe performance drop in atrophied retinas. The measurement algorithm introduced in [2] is adapted to images from multiple modalities.

This paper proposes a novel approach for retina layer segmentation and thickness measurement in histological images. Like [9], we have relatively few annotated retina sections (9 healthy, 12 atrophied) for training. However we have access to many more unlabeled WSIs, and we exploit them with a semi-supervised segmentation approach. In parallel, we use topological penalization. We show that adding these components improves agreement with manual measurements in atrophied retina sections of LCA mice models.

2. METHOD

This section introduces the segmentation and thickness measurement methods. The segmentation part details three different components used in the segmentation method. Four segmentation models are tested, one for every component and one using a combined approach. For each model, the same thickness measurement approach is used to locally measure border-to-border distance. Training the segmentation models uses a set of labeled images (denoted by Ω_l), and a set of unlabeled ones (Ω_u).

2.1. Segmentation

We formalize the segmentation problem as a 5-class pixel classification task. Each class represents the retina layers previously introduced: IS, ONL, OPL, INL, and background. We associate to each image $I \in \mathbb{R}^{3\times H\times W}$ a prediction $P = (P_1, .., P_5) \in \mathbb{R}^{5\times H\times W}$, where, for each class $c \in \{1,..., 5\}$, $P_c \in \mathbb{R}^{H\times W}$ is the predicted probabilities for each pixel to belong to the class c. Let us additionally define the operator O that associates to P_c a binary mask:

$$O(P_c)_{i,j} = \mathbb{1}_{\{(i,j)\,:\,c=\text{argmax}_{k\in\{1,\dots,C\}}P_{k,i,j}\}}(i,j) \quad (1)$$

where $\mathbb{1}$ is the indicator function, $i = 1, ..., H$, and $j = 1, ..., W$. In the following, we denote by P (or P_c for class c) the prediction of the models and the related ground truth by G (or G_c for class c).

BASE. The baseline approach uses a U-Net [4] architecture. For every image $I_l \in \Omega_l$ the model is trained to minimize the baseline supervised loss function L_{BASE}:

$$L_{BASE} = L_{Dice}(P, G) + L_{CE}(P, G) \quad (2)$$

where L_{CE} is the standard cross-entropy loss and L_{Dice} is the per-class Dice loss [10] defined as

$$L_{Dice} = \frac{1}{5}\sum_{c=1}^{5} 1 - DSC(P_c, G_c) \quad (3)$$

with DSC the Dice coefficient defined by using the Hadamard product $\odot$:

$$DSC(P, G) = \frac{2\times\sum P\odot G}{\sum P + \sum G}. \quad (4)$$

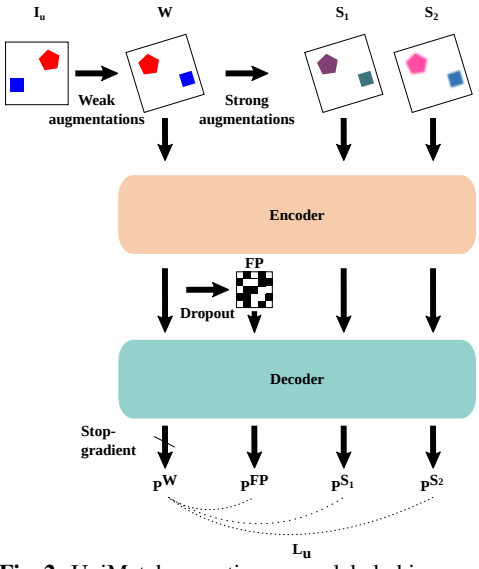

Fig. 2: UniMatch operations on unlabeled images.

SSL. This method is based on semi-supervised learning (SSL): we exploit the set of unlabeled images (Ω_u) via the UniMatch framework [11], illustrated in Fig. 2. The method is trained using pseudo-labels for unlabeled images that are iteratively generated as the prediction of the current network. UniMatch contains different branches. The first branch weakly augments the initial image I_u to generate a new image W based on standard geometric transformation (random rotation and flipping). The second branch generates the images S_1, S_2 based on color modifications of W without modifying its geometry (random blurring, perturbation for saturation, contrast, brightness and hue). Finally, an additional perturbation branch (FP) acts within the latent space by randomly dropping features of the encoded W before decoding (dropout with probability 0.5). UniMatch uses the following loss:

$$L_{UM} = L_u(P^{S_1}, P^W) + L_u(P^{S_2}, P^W) + 2L_u(P^{FP}, P^W) \quad (5)$$

where

$$L_u(P, P^W) = L_{Dice}(P, O(P^W)) + L_{CE}(P, O(P^W)) \quad (6)$$

The SSL method is trained with the following loss:

$$L_{SSL} = \mathbb{1}_{\Omega_l} L_{BASE} + \mathbb{1}_{\Omega_u} L_{UM} \quad (7)$$

TOPO. The last component introduces a topological penalization to learn the spatial organization of the different layers more efficiently. This is achieved by the following constraint: pixels of non-adjacent layers should not be in contact with each other. Here, we penalize adjacency between classes {(IS,OPL), (IS,INL), (ONL,INL), (OPL,background)}. Formally, let NAC be a set of paired non-adjacent classes, and let the mask $M_A = O(P_A)$ be the binary prediction for class A. Then proximity between two classes A and B can be estimated using a morphological dilation with a kernel K, and the topological interactions loss L_{TI} [12] is computed as follows

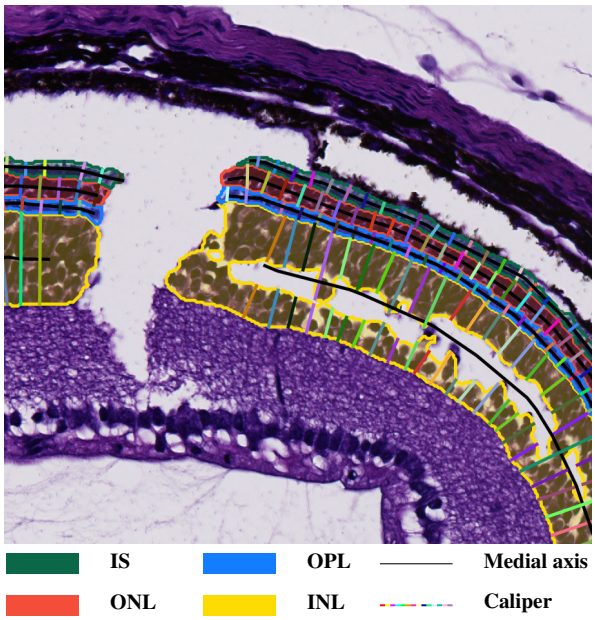

Fig. 3: Segmented layers with medial axes and calipers.

:

$$L_{TI}(P,G) = L_{CE}(P \odot V, G \odot V)$$
$$V = \bigcup_{(A,B)\in \text{NAC}} M_A \odot (M_B * K) \cup M_B \odot (M_A * K) \quad (8)$$

where $*$ is the convolution operation. Here V is the set of pixels violating the non-adjacency rule. This defines the following loss acting on the labeled dataset:

$$L_{TOPO} = \mathbb{1}_{\Omega_l}(L_{BASE} + \lambda_{TI} L_{TI}) \quad (9)$$

SSL-TOPO. Finally, we can consider one last loss combining topological penalization and semi-supervised learning on top of the base approach:

$$L_{SSL-TOPO} = \mathbb{1}_{\Omega_l} L_{TOPO} + \mathbb{1}_{\Omega_u} L_{UM} \quad (10)$$

2.2. Thickness measurement

For all segmentation models, we use the segmentation maps to measure a thickness profile along each layer. However our images present frequent tissue discontinuities in layers, especially the INL. Such discontinuities make it difficult to reliably estimate thickness, and are a problem for atrophied retinas that tend to be especially fragile. These tears or holes in tissue are artifacts from the slicing and fixation process, do not reflect a biological reality and are ignored by biologists in their measurements. To detect the holes, we annotate the tissue discontinuities as background for training. Then, any background region enclosed within a layer contour is considered as a hole. Next, we compute the thickness of each layer along the medial axis, removing the hole width if needed (see Fig. 3). The Straight Medial Axis Calculator plugin of Icy [13, 14] is used to compute the medial axis of each layer. To obtain a single thickness value for the whole layer, parts near the optic nerve and at the boundary of the retina should not be measured: the layers there are thinner than in the rest of the retina and taking them into account skews the measurement. The thickness value for the whole layer is therefore computed as the mean of all values between the 25th and 75th quantiles of the distribution. In addition, removing values outside the quantiles helps filters out outliers in the segmentation and anomalous calipers caused by abrupt changes in the medial axis.

2.3. Data and implementation details

We use 21 labeled retina sections for training (9 healthy, 12 atrophied) and 4 for testing (atrophied). 671 unlabeled slides are used for semi-supervised training. All slides were stained in H&E and scanned with 40x magnification at a resolution of 0.227 μm/pixel. Because of their large size, the training and inference are performed on image patches, managed using TIAToolbox [15] and Sleek-patch [16]. The labels are handled with Cytomine [17] and Rasterio [18].

U-Net training is performed using 100 epochs with stochastic gradient descent for the full dataset. The learning rate follows a cosine decay starting at 10^{-2}. Weight decay is 10^{-4}, momentum 0.9 and batch size is 16. For S_1 and S_2 branches in UniMatch, the CutMix augmentation [19] is used in addition to blurring and color perturbations.

For topological penalization, λ_{TI} is set at 10^{-4}, and the dilation kernel uses 4-connectivity.

Manual thickness measurements for the IS, INL, OPL, and ONL are available on 61 sections of the atrophied retina. None of these sections are included in the training or validation set. For the thickness measurement, we apply a morphological closing of 4 μm, sample every 10 μm along the medial axis, and the calipers have a max length of 400 μm. Operations on contours use OpenCV [20] and Shapely [21].

3. RESULTS AND DISCUSSION

We show in Table 1 the results of layers segmentation for the different models. We use two metrics for model comparison: the Dice score $DSC(O(P_c), G_c)$ (4) measures the overlap between prediction and annotation for class c; the 75% Hausdorff distance (HD75) computes border distances up to the 75th quantile of the distances distribution:

$$\text{HD75}(O(P_c), G_c) = \max\{q_{75}(O(P_c), G_c), q_{75}(G_c, O(P_c))\} \quad (11)$$

where, for two sets of pixels M_1, M_2, $q_{75}(M_1, M_2)$ denote the 75th quantile of the distribution of the distances $d(p, M_2)$ with $p \in M_1$. Using the 75th quantile was necessary as the segmentation from the BASE model presents large spurious connected components. Because of these large outlier regions, both the regular Hausdorff distance and the 95% Hausdorff distance are too sensitive to outliers, and the removal of small connected components is ineffective to accurately compare the contours of segmented layers with the ground truth.

Method	Metric	IS	ONL	OPL	INL	Average
SSL-TOPO	DSC↑	**0.79**	**0.90**	**0.80**	**0.91**	**0.85**
	HD75↓	**2.1**	**1.9**	**2.1**	**3.1**	**2.3**
SSL	DSC↑	**0.79**	0.89	0.79	**0.91**	**0.85**
	HD75↓	2.2	**1.9**	**2.1**	3.2	2.4
TOPO	DSC↑	0.74	0.89	0.74	0.86	0.81
	HD75↓	2.6	2.1	2.4	4.4	2.9
BASE	DSC↑	0.75	0.89	0.70	0.83	0.79
	HD75↓	3.8	1.9	2.9	5.9	3.6

Table 1: Segmentation results based on Dice score (DSC) and 75% Haussdorff distance (HD75) in μm between prediction and manual ground truth. BASE is the regular supervised approach, TOPO adds penalization for predictions that don't respect the structure of the retina, and SSL uses UniMatch to perform semi-supervised learning. SSL-TOPO combines SSL and TOPO.

Method	Metric	IS	ONL	OPL	INL	Average
SSL-TOPO	R↑	**0.94**	**0.98**	**0.88**	**0.83**	**0.91**
	MAE↓	1.3	**1.5**	**0.7**	2.0	**1.4**
SSL	R↑	**0.94**	**0.98**	0.86	0.82	0.90
	MAE↓	**1.2**	1.7	**0.7**	**1.9**	**1.4**
TOPO	R↑	0.92	0.64	0.76	0.77	0.77
	MAE↓	**1.2**	2.1	1.2	4.1	2.2
BASE	R↑	0.40	0.84	0.76	0.77	0.69
	MAE↓	5.1	1.7	1.2	4.1	3.0

Table 2: Correlation (R) and mean absolute error (MAE) in μm between predictions and manual thickness values.

We use the MONAI [22] implementation for both metrics. Using UniMatch for semi-supervision leads to a clear improvement (+7.6% DSC, -33% HD75) and the topological penalization also improves results (+2.5% DSC, -19% HD75). Combining the two leads to our best result.

Article [9] reports an accuracy of 94.4% for ONL and INL, defined as the percentage of pixels correctly segmented. For ONL and INL, we obtain 97.6% (SSL-TOPO) and 97.2% (BASE) accuracy, suggesting an advantage for our method. However direct comparison is difficult, as we do not have access to their models or data. We use accuracy here for the sake of comparison as it is the only metric reported by [9], but it is a problematic metric for segmentation, as it is strongly influenced by the proportion of classes, especially background.

As a final validation of both our segmentation and measurement method, we report Pearson R correlation and mean absolute error (MAE) between our results and manual measurements from a biologist on 61 retina sections.

The results for correlation are presented in Table 2. We notice that including topological interactions improves results (+12% R, -26% MAE). Using UniMatch considerably improves correlation (+30.4%) with manual measurements and reduces MAE (-53%). Combining both slightly improves this result (+1% R). Our best method reaches 0.91 average correlation with manual measurements, with 0.98 for the ONL and 0.83 for the INL. For reference, the authors of [2] report a correlation with manual measurements of 0.88 for the ONL and 0.84 for the INL on fluorescence images.

Finally, similarly to [12], we test the influence of the λ_{TI} parameter on the SSL-TOPO method in Table 3. The results for 10^{-2} show that setting the value too high will degrade performance, with worse results for every layer and an especially high Hausdorff distance for the INL caused by outliers under the 75th quantile.

λ_{TI}	Metric	IS	ONL	OPL	INL	Average
10^{-5}	DSC↑	0.77	0.89	0.79	**0.91**	0.84
	HD75↓	2.2	2.1	**2.1**	3.3	2.4
10^{-4}	DSC↑	**0.79**	**0.90**	**0.80**	**0.91**	**0.85**
	HD75↓	**2.1**	**1.9**	2.1	**3.1**	**2.3**
10^{-3}	DSC↑	**0.79**	**0.90**	0.79	**0.91**	**0.85**
	HD75↓	2.3	**1.9**	**2.0**	3.3	2.4
10^{-2}	DSC ↑	0.76	0.86	0.70	0.82	0.79
	HD75 ↓	2.6	3.0	3.0	75.6	21.0

Table 3: Influence of the topological penalization weight λ_{TI} on segmentation performance, with the SSL-TOPO method.

The semi-supervised approach based on UniMatch considerably improves segmentation and thickness quantification performance compared to the baseline method. The role of UniMatch is mostly to increase the model robustness to irrelevant changes in input images, improving generalization. Moreover, topological penalization encourages the model to learn global structural interactions, resulting in a more biologically consistent segmentation. Our approach results in high correlation and low error in mice with atrophied retinas, even for very thin layers.

4. CONCLUSION

We present a technique for assessing the thickness of retinal layers in H&E-stained histological images. Our approach correlates well with manual measurements, even for very thin layers. It estimates retinal atrophy within mice cohorts, with the goal to aid in non-clinical in vivo pharmacology studies, and drug discovery at large. Our contribution is threefold: we introduce the first method for segmenting up to four retinal layers in histological images, a novel approach for assessing layer thickness that accounts for tissue discontinuities, and a strategy that uses unlabeled images and topological priors to mitigate the prevalent challenge of annotation scarcity. This should enable researchers to adopt our method for studies using different stains or imaging modalities with relatively few labels, and potentially for different types of tissue with a layered structures.

The fragility of atrophied retinas exposes them to tissue tearing during the slicing and fixation process, making thickness measurement difficult. Our method presents a step forward in this direction by properly handling tears included within the closed contour of a layer, as holes in the segmentation. However some of the discontinuities open the contour and cannot be considered as holes. Further research is needed to manage these discontinuities. Our method uses semi-supervision and topological interactions to improve segmentation quality. Despite the improvement compared to the base model, segmentation maps still present outlier regions. To reduce outliers, additional penalization could be used such as contour regularization, and a multi-resolution approach could be adopted to increase available context.

5. COMPLIANCE WITH ETHICAL STANDARDS

All animal procedures were performed with approval from the French Ministry of Research (APAFIS project authorization 45924-2023092513468026) in compliance with guidelines for animal experiments in France and conducted in accordance with the ethical principles.

6. ACKNOWLEDGMENTS

This work was supported by the DIM ELICIT Île-de-France (21006574), Medetia Pharmaceuticals and the Labex IBEID (ANR-10-LABX-62-IBEID). We thank France De Malglaive (Medetia), Iris Barny (Medetia), and Jean-Michel Rozet (Int. Imagine; INSERM) for sharing the images and expertise.

7. REFERENCES

[1] Ching-Hwa Sung and Jen-Zen Chuang, "The cell biology of vision," *Journal of Cell Biology*, vol. 190, no. 6, pp. 953–963, Sept. 2010.

[2] Daniel E. Maidana, Shoji Notomi, Takashi Ueta, et al., "ThicknessTool: Automated ImageJ retinal layer thickness and profile in digital images," *Scientific Reports*, vol. 10, no. 1, pp. 18459, Oct. 2020.

[3] J Byun, MR Verardo, N Vu, et al., "Quantifying structural distortions in retinal tissue before and after injury," in *Workshop on Multiscale Biological Imaging, Data Mining and Informatics*. Citeseer, 2006.

[4] Olaf Ronneberger, Philipp Fischer, and Thomas Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Nassir Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, Eds., Cham, 2015, Lecture Notes in Computer Science, pp. 234–241, Springer International Publishing.

[5] Ignacio A. Viedma, David Alonso-Caneiro, Scott A. Read, and Michael J. Collins, "Deep learning in retinal optical coherence tomography (OCT): A comprehensive survey," *Neurocomputing*, vol. 507, pp. 247–264, Oct. 2022.

[6] Xiaoming Liu, Jun Cao, Tianyu Fu, et al., "Semi-Supervised Automatic Segmentation of Layer and Fluid Region in Retinal Optical Coherence Tomography Images Using Adversarial Learning," *IEEE Access*, vol. 7, pp. 3046–3061, 2019.

[7] Bo Wang, Wei Wei, Shuang Qiu, et al., "Boundary Aware U-Net for Retinal Layers Segmentation in Optical Coherence Tomography Images," *IEEE Journal of Biomedical and Health Informatics*, vol. 25, no. 8, pp. 3029–3040, Aug. 2021.

[8] Timo Kepp, Jan Ehrhardt, Mattias P Heinrich, et al., "Topology-preserving shape-based regression of retinal layers in oct image data using convolutional neural networks," in *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*. IEEE, 2019, pp. 1437–1440.

[9] Maria Cristina De Vera Mudry, Jim Martin, Vanessa Schumacher, and Raghavan Venugopal, "Deep Learning in Toxicologic Pathology: A New Approach to Evaluate Rodent Retinal Atrophy," *Toxicologic Pathology*, vol. 49, no. 4, pp. 851–861, June 2021.

[10] Fausto Milletari, Nassir Navab, and Seyed-Ahmad Ahmadi, "V-net: Fully convolutional neural networks for volumetric medical image segmentation," in *2016 fourth international conference on 3D vision (3DV)*. Ieee, 2016, pp. 565–571.

[11] Lihe Yang, Lei Qi, Litong Feng, et al., "Revisiting Weak-to-Strong Consistency in Semi-Supervised Semantic Segmentation," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2023, pp. 7236–7246.

[12] Saumya Gupta, Xiaoling Hu, James Kaan, et al., "Learning Topological Interactions for Multi-Class Medical Image Segmentation," in *European Conference on Computer Vision*. Springer, 2022, pp. 701–718.

[13] Fabrice de Chaumont, Stéphane Dallongeville, Nicolas Chenouard, et al., "Icy: An open bioimage informatics platform for extended reproducible research," *Nature Methods*, vol. 9, no. 7, pp. 690–696, July 2012.

[14] Daniel Felipe González Obando, Jean-Christophe Olivo-Marin, Laurent Wendling, and Vannary Meas-Yedid, "Vector-based morphological operations on polygons using straight skeletons for digital pathology," in *Discrete Geometry for Computer Imagery: 21st IAPR International Conference, DGCI 2019, Marne-la-Vallée, France, March 26–28, 2019, Proceedings 21*. Springer, 2019, pp. 249–261.

[15] Johnathan Pocock, Simon Graham, Quoc Dang Vu, et al., "TIAToolbox as an end-to-end library for advanced tissue image analytics," *Communications Medicine*, vol. 2, no. 1, pp. 120, sep 2022.

[16] D. Mandache, E. Benoit à La Guillaume, Y. Badachi, et al., "The lifecycle of a neural network in the wild: A multiple instance learning study on cancer detection from breast biopsies imaged with novel technique," in *2022 IEEE International Conference on Image Processing (ICIP)*, 2022, pp. 3601–3605.

[17] Raphaël Marée, Loïc Rollus, Benjamin Stévens, et al., "Collaborative analysis of multi-gigapixel imaging data using cytomine," *Bioinformatics*, vol. 32, no. 9, pp. 1395–1401, 2016.

[18] Sean Gillies et al., "Rasterio: geospatial raster i/o for Python programmers," 2013–, https://github.com/rasterio/rasterio.

[19] Sangdoo Yun, Dongyoon Han, Seong Joon Oh, et al., "Cutmix: Regularization strategy to train strong classifiers with localizable features," in *Proceedings of the IEEE/CVF international conference on computer vision*, 2019, pp. 6023–6032.

[20] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[21] Sean Gillies, Casper van der Wel, Joris Van den Bossche, et al., "Shapely," Oct. 2023, https://github.com/shapely/shapely.

[22] M Jorge Cardoso, Wenqi Li, Richard Brown, et al., "Monai: An open-source framework for deep learning in healthcare," *arXiv preprint arXiv:2211.02701*, 2022.