

TopoAI: Topological Approaches for Improved Structural Correctness

Alexander Spiridonov, Alexander Veicht, András Strausz, Richard Danis

Department of Computer Science, ETH Zurich, Switzerland

{aspiridonov, veichta, strausza, rdanis}@ethz.ch

Abstract—Road network extraction from satellite images is a critical task for various downstream applications such as route planning and autonomous driving. While state-of-the-art methods based on deep learning have shown remarkable performance in terms of accuracy or IoU, even minor prediction errors can lead to disconnected roads, impacting certain use cases such as route planning. In this study, we investigate this issue across multiple models, revealing the presence of incorrectly disconnected sections in predicted masks. To address this challenge, we propose three different approaches based on regularization and re-weighting. First, we introduce a novel weighting technique that leverages the inherent properties of the roads. Drawing inspiration from the medical domain, we explore both implicit and explicit measures for topological correctness, namely Soft-clDice and TopologyLayer. Through extensive evaluation, we demonstrate modest yet significant improvements in topological correctness. Our best results are achieved using soft-clDice, yielding an enhancement in topological correctness, measured by clDice from 0.814 to 0.844.

I. INTRODUCTION

The manual extraction of roads from aerial satellite images is a tedious and time-consuming process. As data volumes continue to grow, this approach has become impractical, necessitating computer-aided methods as the preferred alternative.

Initially, filtering-based approaches [1] were used to produce segmentation maps; however, with the rise of Deep Learning, the trend has shifted towards Convolutional Neural Networks [2], [3] and recently towards transformer-based architectures, such as Vision Transformers [4].

Although current state-of-the-art vision models perform well in road segmentation tasks, there are still instances where the results may not be entirely satisfactory for certain applications. For example, in route planning, roads with multiple discontinuities can lead to erroneous routes. Hence, ensuring the topological correctness of predictions, which involves preserving the road connectivity and avoiding holes within positive regions, is crucial.

Accurate and topologically correct road segmentation is challenging because of sparse masks and thin, occasionally partially occluded, streets. The diversity in landscapes, lighting conditions, and satellite image quality further complicates this task. Models optimized solely for prediction performance may struggle to produce masks that are practically useful. In Figure 1, we demonstrate how occlusions and label ambiguities can result in false predictions and unconnected roads.

In this work, we aim to explicitly capture the characteristics of roads and road networks by considering optimization

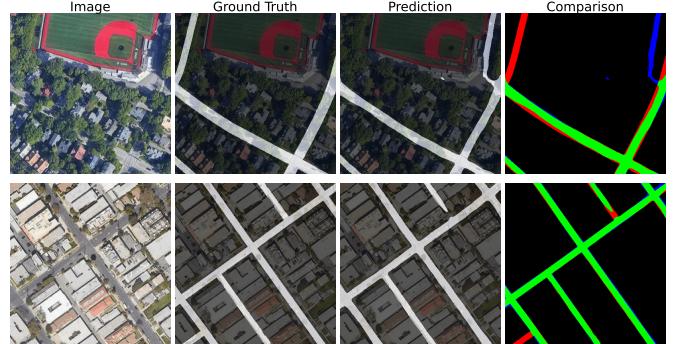


Fig. 1: We see that while the prediction is mostly correct (green), false positives (blue) arise because of label ambiguities. False negatives (red) appear mostly in the occluded regions of the image. The lower example highlights the phenomenon in which a minor mistake in the prediction leads to a break in the road network.

with task-specific weighting and regularization. As our first method, we propose to use a weighting scheme that emphasizes edges and corners of roads by assigning a higher weight to their vicinity. In particular, we use the Canny edge detection method [5] to detect the boundaries of road masks, and take inspiration from GapLoss [6] to extract corners. After the regions of interest are located, we apply Gaussian filters to upweight their neighborhood in a decaying fashion. Furthermore, we propose incorporating topology-aware loss metrics applied in the medical domain to learn structurally correct road network masks. TopologyLayer [7] leverages concepts from the theory of persistent homology to compute the basic topological features and enforce them as regularizers. Although this directly measures the topological structure, it comes with considerable computational costs of $\mathcal{O}(d^3)$ for a d -dimensional image. Therefore, we compare Soft-clDice [8], which acts as a proxy measure by comparing the skeletons of the predicted and ground-truth masks.

We conduct a comprehensive evaluation of the proposed losses on a diverse set of models, aiming to encompass as many varieties as possible to highlight their impacts. All in all, we

- Propose a novel weighting of edges and corners for topological correctness
- Evaluate the performance of clDice and TopologyLayer taken from the medical domain
- Improve structural correctness of predictions for all considered models

II. BACKGROUND

In this section, we introduce three different models used for evaluating the proposed loss metrics. To gain a comprehensive understanding of the proposed objectives, we employ a traditional architecture, UNet++ [9]; a pre-trained image model called UPerNet [10]; and SPIN [11], specifically designed for road segmentation. To achieve topological correctness in the segmentation results, we propose a novel re-weighting technique and explore objectives from the medical domain, evaluating them for road segmentation.

A. Models

We examine three concepts in our investigation of different models: a standard benchmark, a model tailored for road segmentation, and a pretrained model on ImageNet [12] and ADE20k [13] fine-tuned for road segmentation.

UNet [14] and its enhanced counterpart, UNet++ [9], have emerged as strong baselines for various image-segmentation tasks. UNet++, a nested UNet design, effectively addresses the semantic gap between the encoder and decoder by incorporating skip connections with the convolutional layers. These connections not only bridge the gap, but also improve the gradient flow, enabling efficient information propagation throughout the network.

One prominent approach tailored to road segmentation is Spatial and Interaction Space Graph Reasoning (SPIN) [11], which extracts roads from aerial images. The SPIN model incorporates an additional road-orientation prediction task, thereby enhancing its ability to accurately discern road layouts. Furthermore, it generates two graphs from the original feature space: one for spatial space reasoning, capturing connectivity in the spatial space, and another for interaction space reasoning, which seeks to improve the segregation of road delineation by considering the semantics of node clusters. Through graph reasoning at multiple scales of the feature map, SPIN aims to incorporate long-range contextual information, thereby enabling robust road segmentations. To investigate the impact on pretrained models, we employ UPerNet [10] combined with a ConvNeXt-tiny backbone, which combines a Feature Pyramid Network (FPN) and a Pyramid Pooling Module (PPM). To evaluate the effectiveness of fine-tuning these models, multiple experiments were conducted, including training only the decoder, using a frozen backbone, and fine-tuning the entire model after loading the pre-trained weights. Training the decoder solely with a frozen backbone yielded subpar results. However, fine-tuning the entire model on the target dataset resulted in an improved segmentation accuracy.

B. Topological correctness through weighting

Our first approach towards achieving topological correctness leverages the inherent characteristics of roads, such as their distinguishable boundaries and interconnected nature. To enforce this, we propose to extract edges and endpoints of roads from segmentation masks and re-weigh the loss map with an affine combination of them.

We extract edges from the ground-truth masks using the Canny edge detector [5] and apply Gaussian smoothing to articulate their neighborhood. For the endpoints, we extract corners following GapLoss [6] and perform this online on the predicted masks. The idea behind this is to emphasize connectedness and not only up-weight the ends of the ground-truth roads but also stress the regions that are currently misidentified as endpoints. To extract corners, GapLoss [6] initially extracts the skeleton of the predicted mask. Subsequently, a corner is defined as a pixel with only one positive pixel in its closest proximity. Finally, we deviate from GapLoss [6] and up-weight the neighborhood pixels of corners using a Gaussian filter instead of an equi-weight filter to better model the decreasing importance of pixels as we move away from a corner.

Overall, we propose the following novel weighting scheme:

$$\mathcal{W} = \mathbb{1} \cdot (1 - (\alpha + \beta)) + \alpha \hat{\mathcal{W}}_{corner} + \beta \hat{\mathcal{W}}_{edge} \quad (1)$$

$$\mathcal{L} = \text{mean}(\mathcal{W} \odot \mathcal{L}_{CE}) \quad (2)$$

where $\alpha, \beta \leq 1$, $\alpha + \beta \leq 1$, \mathcal{L}_{CE} denotes the element-wise cross-entropy map and $\hat{\mathcal{W}}_{corner}$ and $\hat{\mathcal{W}}_{edge}$ are the normalized corner and edge weight maps respectively.

C. Topological correctness through objectives

We consider two different objectives, which we add to the classical binary cross-entropy loss function as regularizers. Soft-clDice uses intersections between masks and their soft skeletons to compute a topology-preserving metric and loss. The TopologyLayer, on the other hand, directly leverages persistent homology. While the use of persistent homology is more explicit, it comes at a computational cost relative to soft-skeletonization in soft-clDice.

1) *Soft-clDice*: Soft-clDice loss [8] facilitates topologically correct segmentation of tubular structures. These are often found in the medical domain, such as in images of vessels and neurons. However, the tubular-like appearance of roads in satellite imagery motivates the use of soft-clDice for road segmentation. Initially, Shit et al. [8] proposed a novel topology-preserving similarity metric, called clDice. The metric implicitly captures topological correctness by computing the intersections of the predicted and label masks V_P and V_L with their skeletons S_P and S_L .

$$T_{\text{prec}}(S_P, V_L) = \frac{S_P \cap V_L}{|S_P|} \quad T_{\text{sens}}(S_P, V_L) = \frac{S_L \cap V_P}{|S_L|} \quad (3)$$

$$\text{clDice}(V_P, V_L) = 2 \times \frac{T_{\text{prec}}(S_P, V_L) \times T_{\text{sens}}(S_P, V_L)}{T_{\text{prec}}(S_P, V_L) + T_{\text{sens}}(S_P, V_L)} \quad (4)$$

ClDice is provably topology-preserving [8], with a value of one corresponding to an ideally preserved topology. Moreover, Shit et al. [8] introduces a fully differentiable soft skeletonization procedure. Accordingly, the soft-clDice loss is essentially the clDice metric calculated with soft skeletonization. Shit et al. [8] calculate the convex combination between the often-used dice loss [15] and soft-clDice loss.

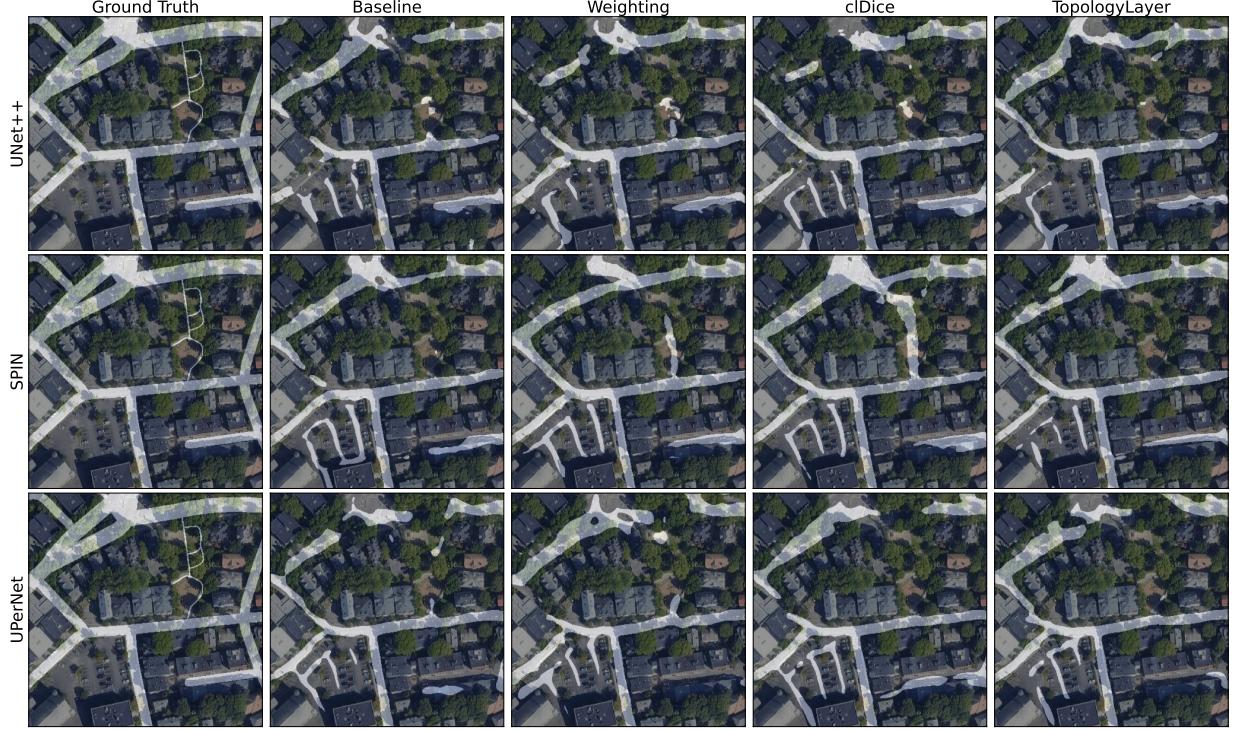


Fig. 2: Visual comparison of the predictions using the generally best performing settings (see I).

2) *TopologyLayer*: With our third approach, we explicitly include the topological landscape for the predicted segmentations as a regularizer in the training objective. Given a 2D mask with each pixel assigned a value in $[0, 1]$, we compute the number of connected components, β_0 , and the number of its holes, β_1 , using the TopologyLayer package [16]. These are the *Betti* numbers of dimensions 0 and 1. By setting a threshold t such that pixels above t are mapped to one and else to zero, we obtain a segmentation mask different from the original. Continuously lowering t from one to zero yields a series of such masks, in which topological features first appear and then gradually vanish. We measure the persistence of each structure indexed by j by the difference between their "birth" $b(j)$ and their "death" $d(j)$. Following [7] we fix $\hat{\beta}_k$ for each dimension and sort the topological features based on their persistence. We can then incorporate prior information about the Betti numbers for each dimension k as follows,

$$\mathcal{L}_k = \sum_{j < \hat{\beta}_k} (1 - (b(j) - d(j))^2) + \sum_{j > \hat{\beta}_k} (b(j) - d(j))^2.$$

As births and deaths can only take on the values present in the predicted mask, which in turn are the outputs of our model, this loss is differentiable. Similar approaches were proposed by [17] and [7]. In contrast to [7], we add this loss to the objective for all labeled instances. Thus, we aim to maximize the persistence of the first $\hat{\beta}_k$ structures and minimize the rest.

III. METHODOLOGY

In the following section, we describe the experimental details, including the dataset and the training parameters.

A. Data

The dataset consists of 144 images accompanied by their respective ground-truth masks. To partition the data, we perform a random split, resulting in 115, 14, and 15 images for training, validation, and testing, respectively. To address the limited number of training images and prevent overfitting, we relied heavily on augmentation, employing rotations, distortions, resizing, color shifts, and jitter techniques. Furthermore, for our leaderboard submissions, we incorporated additional images from the EPFL dataset [18] and the RoadTracer dataset [19] into the training set.

B. Training details

Throughout the experiments for UNet++ and SPIN, we train all models for 300 epochs using a batch size of 8, cross-entropy loss function, and Adam [20] optimizer, with a learning rate of 0.001 and a weight decay of 1e-4. For UPerNet, we applied the AdamW [21] optimizer with a slightly lower learning rate of 3e-4 and additional losses, namely focal [22], mean squared error, and MIoU losses. To prevent overfitting and facilitate learning, the learning rate is dynamically reduced by reducing the learning rate upon plateau with a patience of 40 epochs. We report the results for all configurations in the Appendix. The training duration varied for different models: UNet requires approximately 40 min, SPIN required approximately 3h, and UPerNet took approximately 1h and 15 min to complete the training process.

Model	Loss	Acc	IoU	clDice	Score
UNet++	Baseline	0.926	0.620	0.797	0.906
	Weighting	0.928	0.629	0.804	0.908
	soft-clDice	0.923	0.607	0.784	0.907
	TopoLayer	0.927	0.625	0.805	0.907
SPIN	Baseline	0.936	0.668	0.814	0.915
	Weighting	0.932	0.653	0.805	0.916
	soft-clDice	0.939	0.682	0.844	0.916
	TopoLayer	0.934	0.660	0.818	0.914
UPerNet	Baseline	0.932	0.645	0.820	0.908
	Weighting	0.932	0.652	0.812	0.909
	soft-clDice	0.932	0.653	0.821	0.912
	TopoLayer	0.934	0.662	0.818	0.910

TABLE I: Results using general best settings for the different proposed methods. Specifically, for weighting we use $\alpha, \beta = 0.3$, for TopologyLayer we use $\lambda = 0.01, \hat{\beta}_0 = 3, \hat{\beta}_1 = 5$ and for soft-clDice we set $\lambda = 0.25$

IV. RESULTS

In this section, we analyze our results and compare them with the performance of the proposed method. In Figure 2 we provide a visual overview for the effects of the different methods. Furthermore, we discuss our best scores on the public leaderboard that were achieved by submitting an ensemble of models.

A. Loss Evaluation

1) *Weighting*: Experiments with edge and corner weights showed minor topological improvements. When only corner weights were used, there was a small increase in clDice for SPIN and UPerNet. Edge weights improved UNet++ clDice by 0.003 and SPIN clDice from 0.814 to 0.825. Combining both weights showed a similar performance, with the largest increase in clDice achieved for SPIN.

2) *soft-clDice*: Both SPIN and UPerNet exhibit improved clDice for all $\lambda > 0$. The accuracy and IoU improved marginally. For UNet++, soft-clDice did not yield any improvements in the accuracy, IoU, or clDice.

3) *TopoLayer*: To reduce the computational cost, the predictions were resized to 100×100 for TopologyLayer loss. Lower values of $\hat{\beta}_0$ and $\hat{\beta}_1$ produce the best results. UNet++ showed a slight improvement across all metrics with weights of $\lambda = 0.01$ and $\hat{\beta}_0 = 3, \hat{\beta}_1 = 5$. The performance was worse for other configurations. For SPIN, some runs yielded better clDice scores, but the accuracy and IoU improved for only one configuration. A downward trend in the clDice scores was observed for higher weights. UPerNet did not show an improvement in clDice; accuracy and IoU were only marginally better than those of the baseline.

B. Comparison

Across all the different objectives, soft-clDice performs the best for SPIN as can be seen in Table I. The improvement is largest for the clDice metric, which indicates that the soft-clDice loss indeed helps to enforce topology correctness. We

find that SPIN with soft-clDice loss achieves the highest clDice metric of 0.8438. TopoLayer has the best clDice score for UNet++. For the other models this loss leads to a marginally better clDice score in comparison to the baseline. The weighting based approach has the smallest effect on topological correctness and performance. Overall the three introduced approaches have little impact on performance, even though some improvements in topological correctness are noticeable in terms of clDice score.

C. Leaderboard Submissions

Both SPIN and UPerNet outperform the UNet++ in all metrics by quite some margin with SPIN edginf slightly higher on the public leaderboard. Our most successful model on the leaderboard is an ensemble of SPIN and UPerNet, with an impressive public score of 0.933. We train these models with additional data of the EPFL [18] and RoadTracer datasets [19] to enhance performance. The scores of the models can be seen in Table II.

Model	Acc	IoU	Public Score
UNet++	0.938	0.673	0.920
SPIN	0.945	0.701	0.928
UPerNet	0.946	0.710	0.927
Top 2	-	-	0.933
Top 3	-	-	0.932

TABLE II: Results of the best models and their ensembles.

V. SUMMARY AND OUTLOOK

In this study, we introduce a novel edge weighting technique and adapt methods commonly used in medical image analysis for road segmentation. By conducting this suervey we assess their ability to enforce topological correctness, which we believe to be a crucial property for generating useful segmentations for downstream applications. While overall prediction performance as measured by accuracy and IoU could not be improved by the proposed methods, we find slight improvements in clDice score for most configurations.

While our approaches did not yield significant improvement in the analyzed metrics. We believe further research into topological constraints and losses to be beneficial. Connectedness plays an import role in many segmentation tasks, such as delineating organs in medical datasets. Therefore future work could evaluate the proposed methods on datasets from different domains.

REFERENCES

- [1] P. Arbeláez, M. Maire, C. Fowlkes, and J. Malik, “Contour detection and hierarchical image segmentation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 33, no. 5, pp. 898–916, 2011.
- [2] L. Ichim and D. Popescu, “Road detection and segmentation from aerial images using a cnn based system,” in *2018 41st International Conference on Telecommunications and Signal Processing (TSP)*, pp. 1–5, 2018.
- [3] P. Kaiser, J. D. Wegner, A. Lucchi, M. Jaggi, T. Hofmann, and K. Schindler, “Learning aerial image segmentation from online maps,” *CoRR*, vol. abs/1707.06879, 2017.

- [4] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, “An image is worth 16x16 words: Transformers for image recognition at scale,” *CoRR*, vol. abs/2010.11929, 2020.
- [5] J. Canny, “A computational approach to edge detection,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-8, no. 6, pp. 679–698, 1986.
- [6] W. Yuan and W. Xu, “Gaploss: A loss function for semantic segmentation of roads in remote sensing images,” *Remote Sensing*, vol. 14, p. 2422, May 2022.
- [7] J. R. Clough, I. Öksüz, N. Byrne, V. A. Zimmer, J. A. Schnabel, and A. P. King, “A topological loss function for deep-learning based image segmentation using persistent homology,” *CoRR*, vol. abs/1910.01877, 2019.
- [8] S. Shit, J. C. Paetzold, A. Sekuboyina, I. Ezhov, A. Unger, A. Zhylka, J. P. W. Pluim, U. Bauer, and B. H. Menze, “cldice - a novel topology-preserving loss function for tubular structure segmentation,” in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 16560–16569, June 2021.
- [9] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, and J. Liang, “Unet++: A nested u-net architecture for medical image segmentation,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4*, pp. 3–11, Springer, 2018.
- [10] T. Xiao, Y. Liu, B. Zhou, Y. Jiang, and J. Sun, “Unified perceptual parsing for scene understanding,” in *Proceedings of the European conference on computer vision (ECCV)*, pp. 418–434, 2018.
- [11] W. G. C. Bandara, J. M. J. Valanarasu, and V. M. Patel, “Spin road mapper: Extracting roads from aerial images via spatial and interaction space graph reasoning for autonomous driving,” in *2022 International Conference on Robotics and Automation (ICRA)*, pp. 343–350, IEEE, 2022.
- [12] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “Imagenet classification with deep convolutional neural networks,” *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.
- [13] B. Zhou, H. Zhao, X. Puig, S. Fidler, A. Barriuso, and A. Torralba, “Scene parsing through ade20k dataset,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 633–641, 2017.
- [14] O. Ronneberger, P. Fischer, and T. Brox, “U-net: Convolutional networks for biomedical image segmentation,” in *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, October 5–9, 2015, Proceedings, Part III 18*, pp. 234–241, Springer, 2015.
- [15] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, and M. J. Cardoso, “Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations,” in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, pp. 240–248, Springer International Publishing, 2017.
- [16] R. B. Gabrielsson, B. J. Nelson, A. Dwaraknath, P. Skraba, L. J. Guibas, and G. E. Carlsson, “A topology layer for machine learning,” *CoRR*, vol. abs/1905.12200, 2019.
- [17] X. Hu, F. Li, D. Samaras, and C. Chen, “Topology-preserving deep image segmentation,” *CoRR*, vol. abs/1906.05404, 2019.
- [18] “<https://www.aicrowd.com/challenges/epfl-ml-road-segmentation/>”
- [19] F. Bastani, S. He, S. Abbar, M. Alizadeh, H. Balakrishnan, S. Chawla, S. Madden, and D. DeWitt, “Roadtracer: Automatic extraction of road networks from aerial images,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 4720–4728, 2018.
- [20] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” *arXiv preprint arXiv:1412.6980*, 2014.
- [21] I. Loshchilov and F. Hutter, “Decoupled weight decay regularization,” *arXiv preprint arXiv:1711.05101*, 2017.
- [22] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, “Focal loss for dense object detection,” in *Proceedings of the IEEE international conference on computer vision*, pp. 2980–2988, 2017.

APPENDIX A
DETAILED RESULTS FOR WEIGHTING

$\alpha == \beta$	UNet++			SPIN			UPerNet		
	Acc	IoU	clDice	Acc	IoU	clDice	Acc	IoU	clDice
0	0.926	0.620	0.797	0.936	0.668	0.814	0.932	0.645	0.82
0.1	0.926	0.619	0.797	0.933	0.652	0.811	0.930	0.647	0.806
0.2	0.926	0.621	0.799	0.935	0.663	0.821	0.933	0.657	0.813
0.3	0.928	0.629	0.804	0.932	0.653	0.805	0.932	0.652	0.812
0.4	0.924	0.611	0.783	0.933	0.655	0.815	0.934	0.655	0.82
0.45	0.927	0.627	0.807	0.934	0.657	0.805	0.929	0.637	0.801

TABLE III: Results using both edge and corner weights simultaneously, with equal strength.

α	UNet++			SPIN			UPerNet		
	Acc	IoU	clDice	Acc	IoU	clDice	Acc	IoU	clDice
0	0.926	0.620	0.797	0.936	0.668	0.814	0.932	0.645	0.82
0.2	0.925	0.617	0.792	0.933	0.659	0.822	0.929	0.629	0.808
0.4	0.926	0.619	0.792	0.934	0.663	0.822	0.934	0.661	0.822
0.6	0.927	0.623	0.791	0.934	0.659	0.821	0.926	0.625	0.792
0.8	0.925	0.612	0.791	0.931	0.649	0.809	0.927	0.629	0.801

TABLE IV: Results using corner weights with different strength.

β	UNet++			SPIN			UPerNet		
	Acc	IoU	clDice	Acc	IoU	clDice	Acc	IoU	clDice
0	0.926	0.620	0.797	0.936	0.668	0.814	0.932	0.645	0.82
0.2	0.919	0.577	0.768	0.935	0.664	0.822	0.935	0.669	0.821
0.4	0.921	0.597	0.784	0.935	0.666	0.825	0.934	0.659	0.823
0.6	0.927	0.631	0.800	0.932	0.652	0.813	0.932	0.655	0.811
0.8	0.918	0.600	0.774	0.932	0.654	0.811	0.934	0.658	0.815

TABLE V: Results using edge weights with different strength.

APPENDIX B
DETAILED RESULTS FOR TOPOLOGYLAYER LOSS

λ	$\hat{\beta}_0$	$\hat{\beta}_1$	UNet++			SPIN			UPerNet		
			Acc	IoU	clDice	Acc	IoU	clDice	Acc	IoU	clDice
0	-	-	0.926	0.620	0.797	0.936	0.668	0.814	0.932	0.645	0.82
0.01	3	3	0.922	0.606	0.785	0.935	0.663	0.825	0.932	0.653	0.808
	3	5	0.927	0.625	0.805	0.934	0.660	0.818	0.934	0.662	0.818
	3	7	0.922	0.604	0.779	0.935	0.668	0.828	0.934	0.663	0.818
0.02	3	3	0.922	0.598	0.789	0.937	0.678	0.836	0.929	0.644	0.789
	3	5	0.922	0.587	0.775	0.928	0.629	0.796	0.933	0.650	0.803
	3	7	0.925	0.616	0.791	0.932	0.654	0.813	0.931	0.650	0.784
0.05	3	3	0.920	0.590	0.744	0.929	0.649	0.712	0.931	0.651	0.788
	3	5	0.922	0.599	0.767	0.924	0.612	0.704	0.934	0.660	0.796
	3	7	0.911	0.551	0.729	0.920	0.595	0.628	0.928	0.638	0.707

TABLE VI: Results using the TopologyLayer loss.

APPENDIX C
DETAILED RESULTS FOR SOFT-CLDICE

λ	UNet++			SPIN			UPerNet		
	Acc	IoU	clDice	Acc	IoU	clDice	Acc	IoU	clDice
0	0.926	0.620	0.797	0.936	0.668	0.814	0.934	0.645	0.82
0.25	0.923	0.607	0.784	0.939	0.682	0.844	0.932	0.653	0.821
0.5	0.925	0.617	0.791	0.934	0.658	0.818	0.936	0.666	0.822
0.75	0.920	0.588	0.774	0.934	0.666	0.818	0.935	0.667	0.835
1	0.924	0.603	0.790	0.934	0.664	0.823	0.933	0.651	0.822

TABLE VII: Results using soft-clDice.



Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Declaration of originality

The signed declaration of originality is a component of every semester paper, Bachelor's thesis, Master's thesis and any other degree paper undertaken during the course of studies, including the respective electronic versions.

Lecturers may also require a declaration of originality for other written papers compiled for their courses.

I hereby confirm that I am the sole author of the written work here enclosed and that I have compiled it in my own words. Parts excepted are corrections of form and content by the supervisor.

Title of work (in block letters):

TopoAI: Topological Approaches for Improved Structural Correctness

Authored by (in block letters):

For papers written by groups the names of all authors are required.

Name(s):	First name(s):
Spiridonov	Alexander
Veicht	Alexander
Strausz	András
Danis	Richard

With my signature I confirm that

- I have committed none of the forms of plagiarism described in the '[Citation etiquette](#)' information sheet.
- I have documented all methods, data and processes truthfully.
- I have not manipulated any data.
- I have mentioned all persons who were significant facilitators of the work.

I am aware that the work may be screened electronically for plagiarism.

Place, date

Zurich, 31/07/2023

Signature(s)

For papers written by groups the names of all authors are required. Their signatures collectively guarantee the entire content of the written paper.