CONTEXTLOSS: CONTEXT INFORMATION FOR TOPOLOGY-PRESERVING SEGMENTATION

Benedict Schacht[1,2], *Imke Greving*[2], *Simone Frintrop*[1], *Berit Zeller-Plumhoff*[2,3], *Christian Wilms*[1]

[1]University of Hamburg, [2]Helmholtz-Zentrum Hereon, [3]University of Rostock

ABSTRACT

In image segmentation, preserving the topology of segmented structures like vessels, membranes, or roads is crucial. For instance, topological errors on road networks can significantly impact navigation. Recently proposed solutions are loss functions based on critical pixel masks that consider the whole skeleton of the segmented structures in the critical pixel mask. We propose the novel loss function ContextLoss (CLoss) that improves topological correctness by considering topological errors with their whole context in the critical pixel mask. The additional context improves the network focus on the topological errors. Further, we propose two intuitive metrics to verify improved connectivity due to a closing of missed connections. We benchmark our proposed CLoss on three public datasets (2D & 3D) and our own 3D nano-imaging dataset of bone cement lines. Training with our proposed CLoss increases performance on topology-aware metrics and repairs up to 44 % more missed connections than other state-of-the-art methods. We make the code publicly available[1] [2].

***Index Terms*—** Segmentation, Topology-Preserving, Elongated Structures, Loss Function, Bone Cement Line

1. INTRODUCTION

Loss functions are a core component of neural networks, which determine what the network is optimized for. In image segmentation, this translates directly to how well the predictions of the network can resemble the structural properties of the ground truth. The two most commonly used loss functions in image segmentation are the pixel-wise loss functions Dice loss (L_{Dice}) and Cross-Entropy loss (L_{CE}). However, these losses optimize for a maximum overlap of predictions and ground truth by weighting all pixel contributions equally [1, 2]. This is a problem if the topology is of particular interest because the topology is not explicitly considered. This can promote topological errors, eg. in downstream applications related to the segmentation of road systems, blood vessels, or bone features, see Fig. 1.

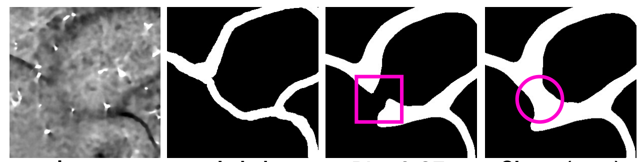

Fig. 1: **Motivation.** Image segmentation in bone research, where our method improves topological correctness over training with Dice and Cross-Entropy loss (Dice & CE).

Related work [3–8] implements topology sensitivity by applying a pixel-wise loss function (L_{pixel}) like L_{Dice} or L_{CE} on a critical pixel mask of topologically important pixels and combining it with an additional L_{pixel} that is applied on all pixels. Applying L_{pixel} on the critical pixel mask ensures higher network attention on the critical pixels than on the normal pixels. Related work defines the critical pixel mask by including the complete morphological skeletons of ground truth and predictions [3, 4, 8], or a slightly dilated skeleton [5]. [4, 5, 8] extend [3] mainly by different L_{pixel} or removing small structures during training.

While the morphological skeleton is an excellent measure for topological correctness, the complete skeleton typically includes many already correctly predicted pixels. However, the network must be trained to learn a better representation of the topological errors [6, 7]. [6] and [7] identify critical pixel masks at the topological errors without skeletonization at the cost of high runtime. However, their identified critical pixels are either singular pixels [6] or a line of 1-pixel width [7] in the center of the topological errors. Considering only a few singular pixels in the critical pixel mask is a problem because the network can obtain better results if more pixels are included in the mask [9]. We argue that [6] and [7] are missing important context information by including only few critical pixels.

In this paper, we propose the novel loss function CLoss for topology-preserving segmentation. CLoss features an extensive critical pixel mask of the topological errors and particularly considers more context pixels of the topological errors. The additional context pixels are acquired by com-

[1]https://gitlab.com/Benedict_S/ContextLoss
[2]Supplement: https://dx.doi.org/10.60864/jkny-a610

bining skeletonization and the distance transform. Utilizing skeletonization in the critical pixel selection makes CLoss naturally fast and compatible with 3D data. Our loss can be used on 2D and 3D data with any arbitrary segmentation network to improve topological correctness in a plug-and-play manner. Additionally, we combine our proposed CLoss with a dedicated training strategy to exploit topological post-processing. For validation, we isolate the contribution of our proposed critical pixel mask from the impact of L_{pixel}. We benchmark on four datasets, including our own novel 3D bone Cement Line Dataset (CLD). CLD contains boundary layers of bone tissue which challenge segmentation algorithms with low contrast, diffuse borders, and image artifacts. Further, we propose two intuitive metrics to verify improved connectivity due to a closing of missed connections, which complement existing metrics and are more robust to artifacts from evaluation. In this paper, our contributions can be summarized as follows:

- We propose the novel topology-preserving loss function CLoss which utilizes a critical pixel mask to consider the whole context of topological errors (Sec. 2.1).
- We propose a new 3D bone Cement Line Dataset (CLD) for topology-preserving segmentation (Sec. 3).
- We propose two metrics which are sensitive to missed connections (Sec. 2.2).

2. METHODOLOGY

2.1. ContextLoss

We propose the novel loss function CLoss to promote topology-preserving segmentation with arbitrary segmentation networks. CLoss extracts a critical pixel mask of topologically important locations, which is then considered in the total cost function. In addition, we introduce a pretraining and topological fine-tuning strategy. The results of both stages are combined in a post-processing to generate the final results, see Fig. 2.

Critical pixel mask. The acquisition of the critical pixel mask is illustrated in Fig. 3. The critical pixel mask contains the complete context of all topological errors. We determine the skeleton at the topological critical locations, which is then extended to include all context pixels, using the distance transform. Missed connections are processed as displayed, which we describe below. False positive connections are processed similarly, as mentioned in the caption of Fig. 3.

We first obtain the skeleton of the label. Then, we obtain the skeleton part at the topological error and the correctly predicted skeleton part by splitting the label skeleton with the prediction mask. Then, we apply the distance transform to both splitting results, which calculates the distance of all pixels to the respective split skeleton part. Based on the re-

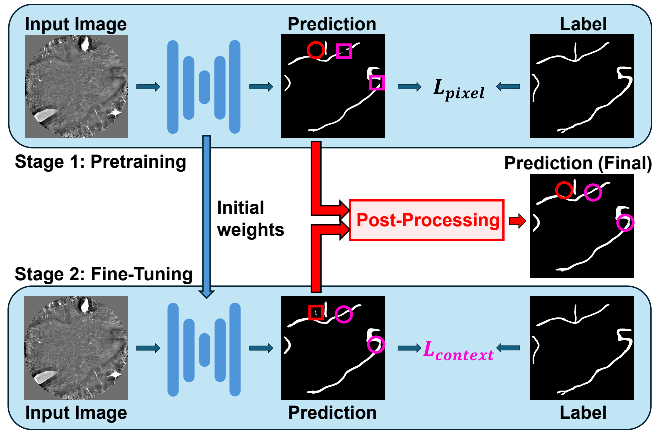

Fig. 2: **Our proposed method.** An arbitrary segmentation network is pretrained with a standard pixel-wise loss (L_{pixel}) and then fine-tuned with our topology-sensitive CLoss ($L_{context}$). We combine the inference results of the pretraining for topological post-processing with the inference results from the fine-tuning. Our method improves topological correctness (circles) of topological errors (squares).

sults of the distance transforms, the context extraction keeps only the pixels closest to the skeleton split of the topological error (see bottom right in Fig. 3). The context pixel selection is further cropped to label size to focus on the context pixels of the topological error. The resulting critical pixel mask (M) contains all context pixels of missed and false positive connections. See our pseudo-code in the supplement for details (Sec. A.1).

Cost function. Our proposed CLoss is based on a pixel-wise loss function (L_{pixel}), which is adapted for topology-sensitivity. For L_{pixel}, we choose the loss combination of L_{Dice} and L_{CE}

$$L_{pixel} = (1 - \alpha)L_{Dice} + \alpha L_{CE}, \quad (1)$$

with $\alpha \in [0, 1]$ as a weight. For this work, we use $\alpha = 0.5$ for a more robust performance compared to individual L_{Dice} and L_{CE} [1, 10]. We combine L_{pixel} with the critical pixel mask M to our proposed CLoss

$$L_{context} = (1 - \gamma)L_{pixel} + \gamma L_{pixel} \odot M, \quad (2)$$

with the Hadamard product $\odot$. $\gamma \in [0, 1]$ is a weight to adjust the impact of the topology-sensitive term that contains M.

Training with CLoss. First, a segmentation network is trained with regular L_{pixel}. Subsequently, CLoss is used for fine-tuning to improve topological correctness, see Fig. 2. Empirically, we found a fine-tuning length with 5% of the pretraining length to be most efficient. We include ablations in the supplement (Sec. A.10).

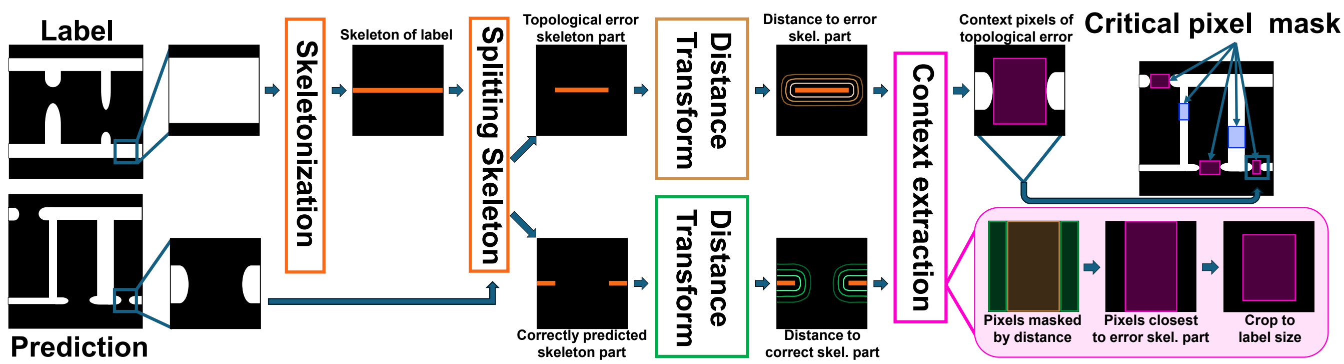

Fig. 3: **Workflow for extracting the critical pixel mask.** CLoss considers all pixels of missed connections (pink mask) and false positive connections (blue mask) as critical pixels, which captures the complete context of the topological errors. We display the critical pixel extraction for the pink mask. Extracting the blue mask is analogue to the pink mask extraction, just with interchanged label and prediction notation, and cropping to prediction size instead of label size during context extraction.

Topological Post-Processing. We apply topological post-processing to the fine-tuning result, which only keeps structures already present in the pretraining, see Fig. 2. Intuitively, this ensures topological fine-tuning without adding new foreground structures. We include a formula and more details in the supplement (Sec. A.3)

2.2. Proposed Metrics

Empirically, missed connections (gaps) in the predictions are far more frequent than false positive connections in our used datasets, see Fig. 5. Hence, we propose additional metrics to better evaluate gap closing, adding desirable topological information [11, 12]. Our proposed metrics complement existing metrics, which usually evaluate for gaps and false positive connections simultaneously. However, our metrics are more robust against evaluation artifacts, see Fig. 4. We include a qualitative example in the supplement (Sec. A.2).

e_0-Gt. Recent work [3–7, 13, 14] reports the betti matching error $\mathrm{e}_0 = |\beta_0^{X_{bin}} - \beta_0^{Y}|$, with the number of connected components of the binary predictions $\beta_0^{X_{bin}}$ and ground truth β_0^{Y}. e_0 is typically evaluated on patches, so opposite differences in β_0 at different locations don't balance each other out. However, patch evaluation can lead to cropping artifacts with thick predictions, see Fig. 4. We propose

$$\mathrm{e}_0\text{-Gt} = |\beta_0^{X_{bin} \odot Y} - \beta_0^{Y}|, \tag{3}$$

which is similar to e_0, but the binary predictions are additionally masked with the ground truth. Therefore, e_0-Gt is not affected by artifacts from thickened predictions. A lower score indicates more closed gaps.

AGS - Accuracy on Ground Truth Skeleton. Recent work [3, 5, 8, 13] reports the clDice metric, which evaluates the intersection of skeletonized prediction (S_X) and

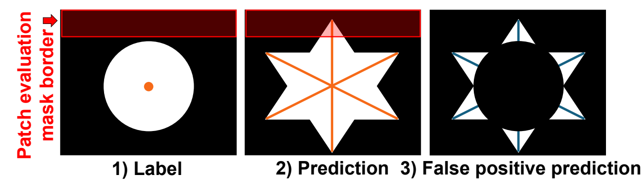

Fig. 4: **Evaluation artifacts.** 1) has the same topology (e_0) as 2). However, the higher thickness of 2), compared to 1), leads to a cropping artifact in the patch-based evaluation (red area), which falsely suggests an unfavorable e_0. Our e_0-Gt is robust to this artifact from patch evaluation. Moreover, the complex contour of 2) evokes many additional skeleton pixels (orange). Paired with the higher thickness of 2), this leads to multiple false positive skeleton pixels (blue), resulting in a bad clDice score. Our AGS is robust to this contour artifact.

skeletonized ground truth (S_Y). clDice complements e_0, as it is sensitive to the partial reparation of gaps and false positive connections, which β_0-based metrics are not. However, S_X is overly sensitive to skeletonization artifacts related to a complex contour or surface structure, see Fig. 4. We propose

$$\mathrm{AGS} = \frac{sum(X_{bin} \odot S_Y)}{sum(S_Y)}, \tag{4}$$

which measures the accuracy of the binary predictions (X_{bin}) on S_Y. AGS is robust to artifacts in S_X, since AGS doesn't utilize S_X. A higher score indicates a better closing of gaps.

3. CEMENT LINE DATASET (CLD)

Cement lines form and separate structural units in the bone, and are thought to play an essential role in crack propagation, energy absorption, and bone mineralization [15]. Their topology is a key property for their function [15, 16], therefore automated topology-preserving segmentation creates a

Table 1: **Quantitative results.** Our dataset, proposed method and proposed metrics are indicated in italic.

Dataset	Method	Weight γ	Dice↑	clDice↑ [3]	e↓	e_1↓	e_0↓	e_0-Gt↓	*AGS*↑
Roads	nnU-Net [10]		79.69	89.34	1.181	0.895	0.286	0.702	86.46
	Dice & CE		79.77	89.37	1.156	0.949	0.207	0.699	86.41
	clDice [3]	0.5	79.50	89.11	1.126	0.897	0.230	0.710	87.07
	Compound clDice	0.3	79.73	89.34	1.108	0.909	0.199	0.673	87.05
	CLoss	*0.08*	**79.82**	**89.47**	1.065	0.880	0.185	0.656	87.70
		0.1	79.57	89.21	**0.990**	0.810	**0.180**	0.617	87.88
		0.2	79.12	89.13	0.994	**0.788**	0.205	**0.494**	**89.48**
HRF-Retina	nnU-Net [10]		**82.33**	81.88	0.528	0.281	0.247	2.742	77.75
	Dice & CE		82.23	81.69	0.554	0.294	0.260	2.780	77.38
	clDice [3]	0.5	82.15	83.09	0.426	0.256	0.170	2.475	82.23
	Compound clDice	0.5	**82.33**	82.98	0.405	0.250	**0.155**	2.429	81.25
	CLoss	*0.08*	82.22	83.20	0.447	0.259	0.187	2.347	82.28
		0.1	82.20	83.27	0.449	0.262	0.188	2.469	82.61
		0.2	80.57	**85.06**	**0.397**	**0.225**	0.172	**1.378**	**90.39**
Vessap	nnU-Net [10]		92.89	**95.83**	28.95	1.413	27.54	14.28	96.03
	Dice & CE		**93.13**	95.81	26.67	1.280	25.39	13.15	96.26
	clDice [3]	0.4	92.70	94.86	29.00	1.240	27.76	9.44	97.74
	Compound clDice	0.1	93.11	95.80	26.10	**1.200**	24.90	11.68	97.00
	CLoss	*0.08*	92.91	95.01	26.09	1.227	24.86	9.97	97.85
		0.1	92.28	94.57	27.02	1.300	25.72	9.32	98.03
		0.2	91.69	93.13	**25.07**	1.827	**23.24**	**6.22**	**99.04**
CLD	nnU-Net [10]		69.67	83.97	3.243	1.085	2.158	2.493	79.19
	Dice & CE		70.27	85.23	3.422	1.075	2.347	2.447	82.02
	clDice [3]	0.5	70.88	86.22	3.374	1.042	2.333	1.825	87.11
	Compound clDice	0.5	70.61	85.45	3.165	1.065	2.099	2.150	82.66
	CLoss	*0.08*	**70.98**	86.22	3.158	1.063	2.095	2.061	84.47
		0.1	70.75	86.26	3.109	1.069	2.040	2.040	84.44
		0.2	70.44	**86.83**	**3.020**	1.047	**1.973**	1.899	85.98
	CLoss (Dice)	*0.08*	69.43	86.40	3.275	**0.969**	2.306	**1.525**	**90.51**

substantial benefit in this research field to improve the understanding of cement lines. 3D nano-imaging[3] of cement lines can specifically improve our understanding of cement lines on the nano-scale. Unfortunately, nano-imaging quality is degraded by noise and various imaging artifacts [18]. Moreover, cement lines have low contrast with respect to their surroundings and diffuse borders. Our CLD consists of 17 3D images of shape $1024^2 \times 600$, featuring a voxel size of $(45.6\,\text{nm})^3$. More details on the annotation process and image acquisition are included in the supplement (Sec. A.4). The characteristics of CLD are distinctly different from current 2D topology-benchmark datasets like Roads [19] and HRF-Retina [20], where the foreground class usually has adequate contrast and clear borders. CLD also features a layer/membrane-like foreground structure that is not included in the 3D topology-benchmark dataset Vessap [21], which has a tubular structure.

[3] Synchrotron radiation-based full-field transmission X-ray nanotomography operated in Zernike phase contrast microscopy mode [17].

4. EXPERIMENTS

Datasets. Following other work on topology-preserving segmentation [3–8, 13], we use the two public 2D datasets, Massachusetts Roads (Roads) [19] and HRF-Retina [20]. Additionally, we evaluate on two 3D datasets, the public Vessap [21] dataset, and our CLD (Sec. 3). More details on the training splits are included in the supplement (Sec. A.5).

Compared Methods. We use the nnU-Net [10, 22], which provides an optimized standard U-Net [23] for all experiments. We compare CLoss against Dice & Cross-Entropy loss (Dice & CE), clDice [3], and compound clDice. We introduce compound clDice, which corresponds to clDice but has the pixel-wise loss (L_{pixel}) from CLoss. We don't compare against [4, 5, 8], which mainly differ from clDice in L_{pixel}, and [6], which has already been outperformed by clDice. More details on the omitted methods and method optimizations are included in the supplement (Sec. A.6).

Evaluation Metrics. We evaluate all methods with the pixel-wise metric Dice score and the topology-aware metrics clDice [3], Betti Number errors (e, e_0, and e_1), e_0-Gt (ours) and AGS (ours). More details on the evaluation are included in the supplement (Sec. A.7).

Implementation Details. The same skeletonization method [3] (50 iterations) is used for all corresponding methods. We apply our proposed topological post-processing (Sec. 2.1) to all fine-tuning results to highlight the impact of our proposed loss function. Ablations for our proposed topological post-processing are provided in the supplement (Sec. A.10). We perform five-fold cross-validations for all methods. Further details are included in the supplement (Sec. A.8).

4.1. Results and Discussion

The quantitative results of our experiments are displayed in Tab. 1. Our proposed CLoss achieves the best overall topology performance on all datasets, indicated by the combined betti error e. This is supported by the superior gap closing of CLoss for all datasets, indicated by e_0-Gt and AGS. For HRF-Retina, CLoss repairs 44 % more missed connections than clDice (e_0-Gt). CLoss achieves the best performance for the clDice metric in all datasets except for Vessap, which can be attributed to skeletonization artifacts connected to the complex surface structure of Vessap in combination with thickened predictions from CLoss, see Fig. 4. CLoss consistently outperforms compound clDice on the topology metrics, which only differs from CLoss in the critical pixel mask, and to clDice only in L_{pixel}. Hence, we conclude a superior critical pixel mask of our CLoss compared to clDice in the scope of our experiments. Losses with L_{Dice} for L_{pixel} have a slight advantage on CLD, so we added CLoss (Dice) for comparison with clDice, which has the same L_{pixel} as clDice. We observe for all datasets, that superior gap closing (better e_0-Gt and AGS) is linked with a stronger weighting of CLoss, which is intuitively expected. Better overall topological performance (e) and gap closing (e_0-Gt and AGS) is also linked with a decreased Dice score for all datasets.

Our quantitative results are supported by our qualitative results in Fig. 5, where we compare CLoss ($\gamma = 0.2$) with Dice & CE. The predictions with CLoss have significantly higher topological correctness than those of Dice & CE, highlighted by the green arrows. For HRF-Retina, CLoss repairs especially many missed connections and doesn't introduce visible false positive connections. All predictions of CLoss have a slightly higher thickness than the ground truth and Dice & CE, see Fig. 5. This observed higher thickness confirms that the decreasing Dice score for higher weights of CLoss in Tab. 1 originates mainly from an increased thickness of the predictions and is compensated by increasing topological correctness. CLoss runtime with soft skeletonization [3] is comparable to clDice, which is orders of magnitude faster than [6, 7, 13], as compared in [13]. We include a more detailed discussion, runtimes and ablation results in the supplement (Sec. A.9, Sec. A.10).

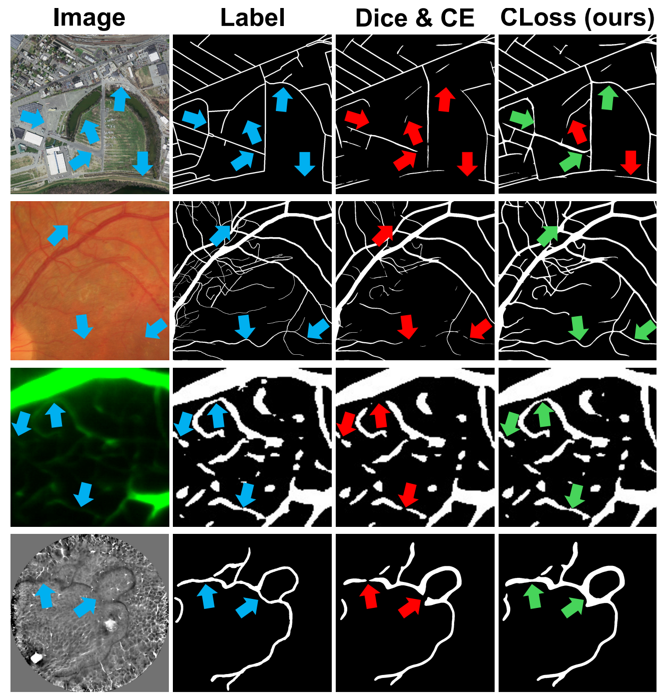

Fig. 5: **Qualitative results.** The sampled patches are from Roads, HRF-Retina, Vessap, and CLD (up to down). Arrows indicate high (green) and low (red) topological correctness with respect to the input image/label (blue).

5. CONCLUSION

In this work, we propose the novel topology-preserving loss function CLoss. CLoss is based on a critical pixel mask, which considers the whole context of topological errors. We implement CLoss with a dedicated training strategy that allows for topological post-processing. CLoss can be used on 2D and 3D datasets with any arbitrary segmentation network. Further, we benchmark on our own 3D dataset CLD, which features low contrast, diffuse borders, and image artifacts. Additionally, we propose two intuitive metrics to verify improved connectivity due to a closing of gaps. We demonstrate the superiority of our context-based critical pixel mask over the critical pixel mask of the entire skeletons on all of our four experiment datasets, where CLoss repairs up to 44 % more missed connections than other state-of-the-art methods.

Acknowledgement. We acknowledge the support by DASHH (Data Science in Hamburg - HELMHOLTZ Graduate School for the Structure of Matter) with the Grant-No. HIDSS-0002.

6. REFERENCES

[1] J. Ma, J. Chen, M. Ng, R. Huang, Y. Li, C. Li, X. Yang, and A. L. Martel, "Loss odyssey in medical image segmentation," *MedIA*, vol. 71, 2021.

[2] B. Liu, J. Dolz, A. Galdran, R. Kobbi, and I. Ben Ayed, "Do we really need dice? the hidden region-size biases of segmentation losses," *MedIA*, vol. 91, 2024.

[3] S. Shit, J. C. Paetzold, A. Sekuboyina, I. Ezhov, A. Unger, A. Zhylka, J. W. Pluim, U. Bauer, and B. H. Menze, "clDice - a Novel Topology-Preserving Loss Function for Tubular Structure Segmentation," in *CVPR*, 2021.

[4] M. J. Menten, J. C. Paetzold, V. A. Zimmer, S. Shit, I. Ezhov, R. Holland, M. Probst, J. A. Schnabel, and D. Rueckert, "A skeletonization algorithm for gradient-based optimization," in *ICCV*, 2023.

[5] Y. Kirchhoff, M. R. Rokuss, S. Roy, B. Kovacs, C. Ulrich, T. Wald, M. Zenk, P. Vollmuth, J. Kleesiek, F. Isensee, and K. Maier-Hein, "Skeleton recall loss for connectivity conserving and resource efficient segmentation of thin tubular structures," in *ECCV*, 2024.

[6] X. Hu, F. Li, D. Samaras, and C. Chen, "Topology-preserving deep image segmentation," *NeurIPS*, vol. 32, 2019.

[7] X. Hu, "Structure-aware image segmentation with homotopy warping," in *NeurIPS*, 2022, vol. 35.

[8] C. Acebes, A. H. Moustafa, O. Camara, and A. Galdran, "The centerline-cross entropy loss for vessel-like structure segmentation: Better topology consistency without sacrificing accuracy," in *MICCAI*, 2024.

[9] K. Gotkowski, C. Lüth, P. F. Jäger, S. Ziegler, L. Krämer, S. Denner, S. Xiao, N. Disch, K. H. Maier-Hein, and F. Isensee, "Embarrassingly simple scribble supervision for 3d medical segmentation," *arXiv preprint arXiv:2403.12834*, 2024.

[10] F. Isensee, P. F. Jaeger, S. A. A. Kohl, J. Petersen, and K. H. Maier-Hein, "nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation," *Nature Methods*, vol. 18, no. 2, 2021.

[11] L. Maier-Hein and et al., "Metrics reloaded: recommendations for image analysis validation," *Nature Methods*, vol. 21, no. 2, 2024.

[12] Alexander H. Berger, Laurin Lux, Alexander Weers, Martin Menten, Daniel Rueckert, and Johannes C. Paetzold, "Pitfalls of topology-aware image segmentation," *arXiv preprint arXiv:2412.14619*, 2024.

[13] L. Lux, A. H. Berger, A. Weers, N. Stucki, D. Rueckert, U. Bauer, and J. C. Paetzold, "Topograph: An efficient graph-based framework for strictly topology preserving image segmentation," in *ICLR*, 2025.

[14] L. Li, Q. Ma, C. Ouyang, Z. Li, Q. Meng, W. Zhang, M. Qiao, V. Kyriakopoulou, J. V. Hajnal, D. Rueckert, and B. Kainz, "Robust segmentation via topology violation detection and feature synthesis," in *MICCAI*, 2023.

[15] B. A. Lamarche, J. S. Thomsen, C. M. Andreasen, T. L. Andersen, and W. B. Lievers, "Trabecular bone structural units and their cement lines change with age, bone volume fraction, structure, and strength in female human vertebrae," *JBMR Plus*, vol. 9, no. 1, 2024.

[16] T. A. Grünewald, A. Johannes, N. K. Wittig, J. Palle, A. Rack, M. Burghammer, and H. Birkedal, "Bone mineral properties and 3D orientation of human lamellar bone around cement lines and the Haversian system," *IUCrJ*, vol. 10, no. 2, 2023.

[17] E. Longo, L. Sancey, S. Flenner, A. Kubec, A. Bonnin, C. David, M. Müller, and I. Greving, "X-ray Zernike phase contrast tomography: 3D ROI visualization of mm-sized mice organ tissues down to sub-cellular components," *Biomed. Opt. Express*, vol. 11, no. 10, 2020.

[18] S. Sefa, J. Espiritu, H. Ćwieka, I. Greving, S. Flenner, O. Will, S. Beuer, D.C F. Wieland, R. Willumeit-Römer, and B. Zeller-Plumhoff, "Multiscale morphological analysis of bone microarchitecture around mg-10gd implants," *Bioactive Materials*, vol. 30, 2023.

[19] Volodymyr Mnih, *Machine Learning for Aerial Image Labeling*, Ph.D. thesis, University of Toronto, 2013.

[20] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust Vessel Segmentation in Fundus Images," *International Journal of Biomedical Imaging*, vol. 2013, no. 1, 1 2013.

[21] M. I. Todorov, J. C. Paetzold, O. Schoppe, G. Tetteh, V. Efremov, K. Völgyi, M. Düring, M. Dichgans, M. Piraud, B. Menze, and A. Ertürk, "Automated analysis of whole brain vasculature using machine learning," *bioRxiv*, 2019.

[22] F. Isensee, T. Wald, C. Ulrich, M. Baumgartner, S. Roy, K. Maier-Hein, and P. F. Jäger, " nnU-Net Revisited: A Call for Rigorous Validation in 3D Medical Image Segmentation ," in *MICCAI*, 2024.

[23] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *MICCAI*, 2015.