Research Statement

Xiaoling Hu (xihu3@mgh.harvard.edu)
https://huxiaoling.github.io/

My research interest is **Biomedical AI**, which lies in the intersection of medical imaging, computer vision, and machine learning, and my dream is to build AI systems that can efficiently assist in disease diagnosis and treatment. Impressive progress has been witnessed in data-driven approaches with a large amount of training data available. These techniques have been widely adapted to medical image analysis. The brute-force way of leveraging samples *without considering data structure / topology / geometry* property limits model applicability to data with complex structures, which is common in biomedical scenarios. On the other hand, *the obtained data are usually not ideal, especially in healthcare scenarios*. The reliance on perfect training sets has become one of the key bottlenecks in creating systems that can attain satisfactory performances in practice.

To address all these limitations, **I strive to create novel algorithms with both concrete theoretical foundations and strong empirical results for imaging data under different contexts, especially biomedical scenarios**. In particular, my research mainly focuses on three themes: **1)** Topology-Driven Deep Image Analysis, **2)** Uncertainty Estimation and Its Applications, and **3)** Learning with Imperfect Data. In the following, I will highlight my recent research on the above themes, and discuss ongoing/future research directions.

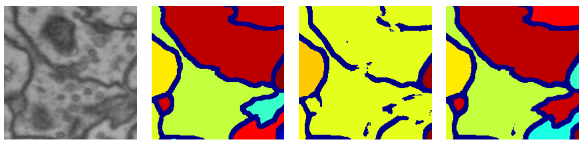
Research Progress

1 Topology-Driven Deep Image Analysis

Despite the strong predictive power of deep learning methods, they are only learning pixel-wise representations, thus creating significant barriers in scalable annotation and downstream analysis. The first direction of my research is to explore beyond pixel-wise representations: *How can we learn topological representations to understand the topology / structures for image tasks*?

I have focused on designing novel topology-aware deep image segmentation algorithms for 2D [1, 2, 3, 4, 5, 6] and 3D [7, 8] images as well as images with multiple foreground classes [9]. Also, I have proposed to use topological priors to recover the triggers in trojan detection [10].

Learn to segment with correct topology/geometry. Image segmentation, i.e., assigning labels to all pixels of an input image, is crucial in many computer vision tasks. State-of-the-art deep segmentation methods learn high-quality feature representations through an end-to-end trained deep network and achieve satisfactory per-pixel accuracy. However, these segmentation algorithms are still prone to make errors on fine-scale structures, such as small object instances, instances with multiple connected components, and thin connections. These fine-scale structures may be crucial in analyzing the *functionality* of the objects. For example, accurate extraction of thin parts such as ropes and handles is crucial in planning robot actions, e.g., dragging or grasping. In biomedical images, correct delineation of thin objects such as neuron membranes and vessels is crucial in providing accurate morphological and structural quantification of the underlying system. A broken connection or a missing component may only induce marginal per-pixel error, but can cause catastrophic functional mistakes. See Fig. 1 for an example.



(a) Ori (b) GT (c) Baseline (d) Ours

Figure 1: Illustration of the importance of topological correctness in a neuron image segmentation task. **(a)** an input neuron image. **(b)** ground truth segmentation. **(c)** result of a baseline method without topological guarantee. **(d)** My method.

I have created algorithms to learn to segment with correct topology [1, 2, 3, 4, 5, 6, 7, 8, 9]. In this setting, the goal is to preserve the topological/geometrical properties of the segmentation masks, besides the per-pixel accuracy, which is important for the downstream analysis. TopoLoss [1] is a *pioneer end-to-end deep segmentation network with guaranteed topological correctness*. In particular, based on *persistent homology* [11, 12, 13], I propose a *differentiable topological loss* that enforces the segmentation results to have the same topology as the ground truth, i.e., having the same Betti number (number of connected components and handles). A neural network trained with such loss will achieve high topological fidelity without sacrificing per-pixel accuracy. I have shown that when the topological loss is decreased to zero, the segmentation is guaranteed to be topologically correct, i.e., have identical topology as the ground truth.

2 Uncertainty Estimation and Its Applications

Another exciting topic that I have been working on is uncertainty estimation. Despite the power of deep networks, their overconfidence is a very common issue. Accurate confidence estimation is important in practice. In autonomous driving and computer-aided diagnosis, analyzing low-confidence samples/regions can help identify subpopulations of events or patients that deserve extra consideration. *A good model should both know what it knows and what it does not know*.

I have created novel algorithms for uncertainty estimation for both curvilinear structure data [5, 6] and unlabeled data [14, 15]. The structure-wise uncertainty estimation will be able to facilitate semi-automatic interactive annotation/proofreading.

Structure-aware uncertainty estimation. Instead of learning pixel-wise feature representations, I propose to directly model and reason about the structures. More specifically, I propose *novel deep learning-based method that directly learns the topological/structural representation of images* [5, 6]. To move from pixel space to structure space, I apply the classic discrete Morse theory [16, 17] to decompose an image into a Morse complex, consisting of structural elements like branches, patches, etc. These structural elements are hypothetical structures one can infer from the input image.

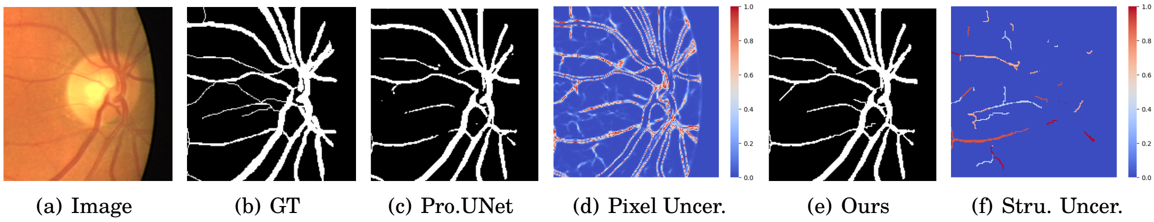


Figure 2: Illustration of structural segmentation and structure-level uncertainty. Compared with Probabilistic-UNet [18] (Fig. 2(c)-(d)), the proposed method is able to generate structure-preserving segmentation maps (Fig. 2(e)), and structure-level uncertainty (Fig. 2(f)).

For further reasoning with structures, in [5], I propose to learn a probabilistic model over the structure space. The challenge is that the space consists of exponentially many branches and is thus of very high dimension. To reduce the learning burden, I introduce the theory of persistent homology for structure pruning. Each branch has its own persistence measuring its relative saliency. By continuously thresholding the complete Morse complex in terms of persistence, we obtain a sequence of Morse complexes parameterized by the persistence threshold, ϵ. This parametric probabilistic model over structure space allows me to make direct structural predictions via sampling and to estimate the empirical structure-level uncertainty via sampling. The benefit is two-fold: First, direct prediction of structures will ensure the model outputs always have structural integrity, even at the inference stage. Second, the samples from the probabilistic model are all feasible structural hypotheses based on the input image, with certain variations at uncertain locations. The uncertainty maps provide hints for downstream human-in-the-loop proofreading. See Fig. 2 as an illustration.

I further propose Probabilistic DMT (Prob. DMT) [6], which represents each structure as sample skeletons drawn from an underlying generative process guided by the likelihood map (the skeleton resulting from the original DMT being one of the samples), to model intra-structural uncertainty and inter-structural uncertainty simultaneously.

Confidence estimation using unlabeled data. Instead of purely relying on labeled data, which is not well suited for a semi-supervised setting, I propose the first confidence estimation method specifically designed for the semi-supervised setting [14]. The first challenge is to leverage the vast majority of unlabeled data for confidence learning. For data without labels, my idea is to use the consistency of the predictions throughout the training process. An initial investigation suggests that the consistency of predictions tends to be correlated with sample confidence on both labeled and unlabeled data.

Having established training consistency as an approximation of confidence, the next challenge is that consistency can only be evaluated on data available during training. To this end, I propose to re-calibrate the model's prediction by aligning it with consistency. In particular, I propose a novel *Consistency Ranking Loss* to regulate the model's output after the softmax layer so it has a similar ranking of model confidence output as the ranking of the consistency. After the re-calibration, we expect the model's output to correctly account for its confidence on test samples. I both theoretically and empirically validate the effectiveness of the proposed Consistency Ranking Loss. I further apply the proposed Consistency Ranking Loss to crowd counting tasks and achieve SOTA performances on multiple datasets [15].

3 Learning with Imperfect Data

Training powerful deep neural networks usually requires a large amount of data. While, as mentioned above, *the obtained data are usually not ideal, especially in healthcare scenarios*. Learning with imperfect data is of great importance in biomedical domains.

I have proposed novel algorithms to better estimate confidence for unlabeled data [14], as well as solving practical problems under imperfect data scenarios, including crowd counting [15], lesion segmentation [19] and brain tumor segmentation [20]. More importantly, I feel honored to have had the opportunity to collaborate with radiologists and ophthalmologists to work on challenging practical problems. I am particularly excited about the recent work [20], which learns robust representations for missing-modality data by using the meta-learning framework. The learned representations can be used for downstream tasks, such as classification and segmentation.

Weakly supervised learning for lesion segmentation on OCT images. Big data can transform the studies in biomedical research to generate greater scientific insights, if expert labeling is available to facilitate supervised learning. However, data annotation can be labor-intensive and cost-prohibitive if pixel-level precision is required. In the recent development of weakly supervised semantic segmentation (WSSS) with image-level labeling, the focus is mainly on natural images, but such an algorithm cannot produce comparable accuracy for medical images. Challenges include substantial variation in lesion scales and occurrences. Specifically, I propose a novel anomaly attention mechanism (AAM) for segmenting common pathologic biomarkers on retinal optical coherence tomography (OCT) given image-level labels only [19]. AAM adopts the GANomaly model and self-attention approach to highlight noisy lesions as anomalies and leverages the global contextual information in the training process.

Brain tumor segmentation with missing modalities. In the medical vision domain, different imaging modalities provide complementary information. However, in practice, not all modalities are available during inference. Previous approaches, e.g., knowledge distillation or image synthesis, often assume the availability of full modalities for all patients during training; this is unrealistic and impractical owing to the variability in data collection across sites. I propose a novel approach to learn enhanced modality-agnostic representations by employing a novel meta-learning strategy in training, even when only a fraction of full modality patients are available [20]. Meta-learning enhances partial modality representations to full modality representations by meta-training on partial modality data and meta-testing on limited full modality samples. Additionally, I co-supervise this feature enrichment by introducing an auxiliary adversarial learning branch. More specifically, a missing modality detector is used as a discriminator to mimic the full modality setting.

Ongoing and Future Work

To summarize, my research lies in the intersection of medical imaging, computer vision, and machine learning. I mainly focus on developing algorithms that explore the properties of complex data and learn with imperfect data. In addition to these themes, I am interested in all other challenges that need to be addressed in general medical AI contexts. Specifically, I am interested in dealing with the following key topics:

- ***Efficiency of Using Data***: How to efficiently make use of the data by taking the data structure property into consideration? By exploring the structural/topological/geometric priors, we might be able to guide the efficient training of deep neural networks. However, until very recently, researchers have focused only on pure data-driven methods. Incorporating structural/geometric/topological information into deep neural networks will be especially critical for biomedical data. I have begun to make progress in this direction [1, 2, 3, 4, 5, 6, 7, 8, 9, 10].

- ***Uncertainty Estimation and Its Applications***: How confident is the model of its predictions? While most existing algorithms claim to achieve good performances, it's important to study how reliable the models are in their predictions, which is also important in biomedical contexts. *A model should both know what it knows and what it does not know*. My recent work on structure-wise uncertainty estimation for curvilinear structure data [5, 6] and confidence estimation using unlabeled data [14] take a step towards this direction. It is natural to ask an interesting research question by extending the recent work [14]: *How are we able to make use of the uncertainty estimation in unsupervised contexts*? I have tried to apply the uncertainty estimation approach to crowding counting tasks and achieve good performances [15]. I believe the proposed reliable uncertainty estimation method will benefit even more number of unsupervised scenarios.

- ***Deep Learning Based Quantification Analysis***: How do the accurate structures affect the downstream analysis, such as the quantification analysis? Based on my current work, I am ready for the next step: quantity how the accurate topology-geometry aware segmentation results affect the downstream analysis. By extracting topology/geometry-informed features, we can do a number of interesting analyses, such as diagnosing retinal diseases, predicting aortic aneurysm eruption risk, and so on. Fig. 3 is an illustration pipeline for the downstream analysis. This is an ongoing direction and will be explored further.

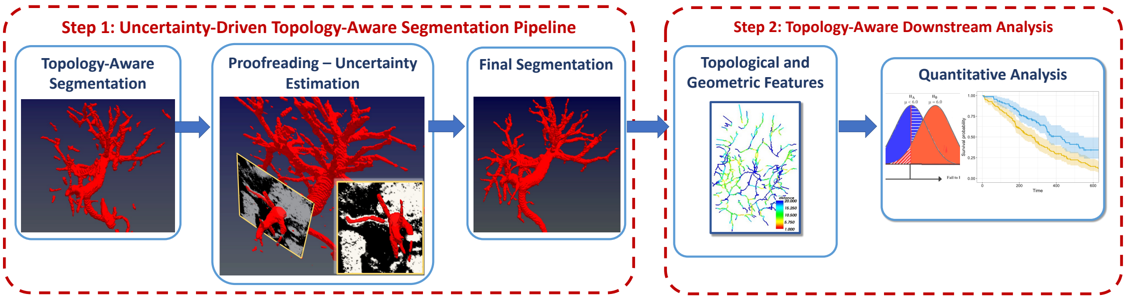


Figure 3: Workflow for uncertainty-driven topology-aware segmentation, and the downstream topology/geometry aware analysis.

- ***Learning with Imperfect Data***: Are we able to achieve satisfactory performances even under scenarios with imperfect data? The superiority of current data-driven methods heavily relies on a large amount of labeled training data. While in practice especially in biomedical contexts, the gathering of labeled data can be cost-prohibitive and time-consuming, which usually requires domain knowledge. The dependency on diverse, high-quality training datasets substantially limits model applicability to complex scenes where data from are imperfect including *missing modality*, and/or *human labeling limited*. Furthermore, the uncertainty estimation I have been working on might provide hints for scenarios with imperfect data. For example, *are we able to focus on the most unconfident samples/regions to facilitate the training with imperfect data?* I have just stepped into this interesting direction [14, 15, 19, 20].

Over the next few years, I will continue to build intelligent AI systems that can assist in diagnosis and disease treatment. I am excited to ask the right (meaningful and impactful) research questions and create innovative and effective solutions to those questions from both theoretical and empirical aspects. Also, I enjoy working towards these challenges with collaborators in medical imaging, computer vision, machine learning, and other related fields including computational geometry, radiology, ophthalmology, digital pathology, cell science, and so on.

References

[1] **Xiaoling Hu**, F. Li, D. Samaras, and C. Chen, "Topology-preserving deep image segmentation," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2019.

[2] **Xiaoling Hu**, Y. Wang, F. Li, D. Samaras, and C. Chen, "Topology-aware segmentation using discrete Morse theory," in *International Conference on Learning Representations (ICLR)*, 2021.

[3] **Xiaoling Hu**, "Image segmentation with homotopy warping," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2022.

[4] **Xiaoling Hu**, X. Chen, Y. Liu, E. Z. Chen, T. Chen, and S. Sun, "Deep statistic shape model for myocardium segmentation," *arXiv preprint arXiv:2207.10607*, 2022.

[5] **Xiaoling Hu**, D. Samaras, and C. Chen, "Learning probabilistic topological representation using discrete Morse theory," in *International Conference on Learning Representations (ICLR)*, 2023.

[6] S. Gupta, Y. Zhang, **Xiaoling Hu**, P. Prasanna, and C. Chen, "Topology-aware uncertainty for image segmentation," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2023.

[7] J. Yang*, **Xiaoling Hu***, C. Chen, and C. Tsai, "3d topology-preserving segmentation with compound multi-slice representation," in *International Symposium on Biomedical Imaging (ISBI)*, 2021.

[8] J. Yang*, **Xiaoling Hu***, C. Chen, and C. Tsai, "A topological-attention convlstm network and its application to em images," in *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2021.

[9] S. Gupta*, **Xiaoling Hu***, J. Kaan, M. Jin, M. Mpoy, K. Chung, G. Singh, M. Saltz, T. Kurc, J. Saltz, *et al.*, "Learning topological interactions for multi-class medical image segmentation," in *European Conference on Computer Vision (ECCV)*, 2022.

[10] **Xiaoling Hu**, X. Lin, M. Cogswell, Y. Yao, S. Jha, and C. Chen, "Trigger hunting with a topological prior for trojan detection," in *International Conference on Learning Representations (ICLR)*, 2022.

[11] H. Edelsbrunner and J. L. Harer, *Computational topology: an introduction*. American Mathematical Society, 2022.

[12] H. Edelsbrunner, D. Letscher, and A. Zomorodian, "Topological persistence and simplification," in *Proceedings 41st Annual Symposium on Foundations of Computer Science*, 2000.

[13] A. Zomorodian and G. Carlsson, "Computing persistent homology," *Discrete & Computational Geometry*, 2005.

[14] C. Li, **Xiaoling Hu**, and C. Chen, "Confidence estimation using unlabeled data," in *International Conference on Learning Representations (ICLR)*, 2023.

[15] C. Li, **Xiaoling Hu**, S. Abousamra, and C. Chen, "Calibrating uncertainty for semi-supervised crowd counting," in *International Conference on Computer Vision (ICCV)*, 2023.

[16] J. Milnor, *Morse theory.(AM-51)*. Princeton university press, 1963.

[17] R. Forman, "A user's guide to discrete morse theory," *Sém. Lothar. Combin*, 2002.

[18] S. Kohl, B. Romera-Paredes, C. Meyer, J. De Fauw, J. R. Ledsam, K. Maier-Hein, S. Eslami, D. Jimenez Rezende, and O. Ronneberger, "A probabilistic u-net for segmentation of ambiguous images," in *Advances in Neural Information Processing Systems (NeurIPS)*, 2018.

[19] J. Yang, G. Merve Demirci, **Xiaoling Hu**, N. Mehta, and C.-L. Tsai, "Weakly supervised learning for lesion segmentation on oct images with anomaly attention mechanism," *Major revision for Medical Image Analysis*, 2023.

[20] A. Konwer, **Xiaoling Hu**, J. Bae, X. Xu, C. Chen, and P. Prasanna, "Enhancing modality-agnostic representations via meta-learning for brain tumor segmentation," in *International Conference on Computer Vision (ICCV)*, 2023.