

Visualizing radiological data bias through persistence images

Yashbir Singh, Colleen Farrelly, Quincy A. Hathaway and Gunnar Carlsson

ABSTRACT

Persistence images, derived from topological data analysis, emerge as a powerful tool for visualizing and mitigating biases in radiological data interpretation and AI model development. This technique transforms complex topological features into stable, interpretable representations, offering unique insights into medical imaging data structure. By providing intuitive visualizations, persistence images enable the identification of subtle structural differences and potential biases in data acquisition, interpretation, and AI model training. Persistence images can also facilitate stratified sampling, matching statistics, and noise filtration, enhancing the accuracy and equity of radiological analysis. Despite challenges in computational complexity and workflow integration, persistence images show promise in developing more accurate, equitable, and trustworthy AI systems in radiology, potentially improving patient outcomes and personalized healthcare delivery.

INTRODUCTION

In the ever-evolving landscape of medical imaging and artificial intelligence (AI), the challenge of identifying and mitigating biases in radiological data interpretation remains paramount. As we strive for more accurate and equitable healthcare outcomes, an innovative approach from topological data analysis (TDA) has emerged: persistence images. This technique offers a promising avenue for visualizing complex radiological data and uncovering hidden biases that may influence diagnosis and treatment decisions.

Understanding persistence images

Persistence images are a stable vector representation of topological features derived from persistence diagrams, which are core tools in TDA [1]. While persistence diagrams capture the birth and death of topological features across different scales, persistence images transform this information into a format more amenable to machine learning techniques [2]. The key advantage of persistence images lies in their stability: small perturbations in the input data result in only small changes in the persistence image [3]. This property makes them particularly valuable in medical imaging, where variations in image acquisition or patient positioning should not dramatically alter the underlying topological features being analyzed.

Interpretable visualizations of complex radiological data

One of the most significant benefits of persistence images is their ability to provide interpretable

visualizations of complex radiological data. By encoding topological features as intensity values in a 2D image, persistence images offer a more intuitive representation of the data's structure compared to raw persistence diagrams or traditional statistical summaries [4]. In the context of radiology, these visualizations can help highlight subtle structural differences in images that might be indicative of specific pathologies or biases in data interpretation. For instance, persistence images could reveal consistent differences in the topological features of lung CT scans between different demographic groups, potentially uncovering biases in how these images are acquired [5].

Identifying and addressing biases

The use of persistence images in radiological analysis opens up new possibilities for identifying and addressing biases in both data interpretation and AI model training:

Matching statistics and stratified sampling

Persistence images can be used to compute matching statistics between different subsets of radiological data. This approach allows for a more nuanced comparison of data distributions, helping to identify potential biases in sampling or data collection [6]. By visualizing these matching statistics through persistence images, researchers and clinicians can more easily identify underrepresented groups or oversampled categories in their datasets. This insight can then inform stratified sampling strategies to create more balanced and representative training sets for AI models, mitigating class bias and improving overall

model performance; further, this can better inform clinicians and researchers about the limitations of a model's generalizability [7].

Mitigating implicit bias in interpretation

Persistence images offer a powerful tool for breaking down radiological data by various demographic factors. By generating and comparing persistence images for different groups (e.g., age, gender, ethnicity), we can visually identify systematic differences in image features that might be indicative of bias in human interpretation [8]. For example, in patients with perceived low risk of breast cancer (e.g., young, no hormone replacement therapy, negative family history, no germline mutation) [9], persistence images can offer an unbiased view of breast morphology and the likelihood of an asymmetry being malignant. Supporting clinicians in cases where implicit biases (such as perceived low risk patients) may occur will ultimately make interpretation more equitable [10].

Filtration of noise

One of the inherent strengths of persistence images is their ability to filter out noise while preserving meaningful topological features. This property is particularly valuable in radiological imaging, where image artifacts and noise can significantly impact interpretation [11]. By focusing on persistent topological features and representing them in a stable format, persistence images can help radiologists and AI models distinguish between genuine anatomical structures and noise or artifacts. This filtration effect can lead to more reliable and less biased interpretations of complex imaging data [12].

Applications in AI model training and evaluation

The vector representation provided by persistence images makes them particularly well-suited for integration into machine learning pipelines. This opens up several exciting possibilities for improving AI model training and evaluation in radiology:

1. Feature engineering: Persistence images can serve as topological feature vectors, complementing traditional image features in AI model training. This can lead to more robust models that are sensitive to both local and global structural characteristics of medical images [13].
2. Model interpretability: By visualizing the persistence images of the input data alongside model predictions, we can gain insights into which topological features are most influential in the model's decision-making process. This enhanced interpretability can help identify potential biases in the model's behavior [14].
3. Quality assurance: Persistence images can be used as a quality assurance tool for both input data and model outputs. By comparing the persistence images of model predictions with those of ground truth data, we can quickly identify systematic errors or biases in the model's performance across different patient subgroups [15].

Challenges and future directions

While persistence images offer significant promise in visualizing and addressing biases in radiological data, several challenges remain:

1. Computational complexity: Generating persistence images for large-scale radiological datasets can be computationally intensive. Developing more efficient algorithms and leveraging high-performance computing resources will be crucial for widespread adoption [16].
2. Integration with existing workflows: Incorporating persistence image analysis into established radiological workflows will require careful planning and validation. User-friendly tools and interfaces will be essential to make this technology accessible to clinicians who may not have expertise in TDA [17].
3. Standardization: As with any new analytical technique, establishing standards for generating, interpreting, and comparing persistence images across different institutions and imaging modalities will be crucial for ensuring reproducibility and comparability of results [18].

CONCLUSIONS

Persistence images represent a powerful new tool in our ongoing efforts to visualize, understand, and mitigate biases in radiological data interpretation and AI model development. By providing stable, interpretable visualizations of complex topological features, they offer unique insights into the structure of medical imaging data that can complement existing analytical approaches. As we continue to refine and validate this technique, persistence images have the potential to play a crucial role in developing more accurate, equitable, and trustworthy AI systems in radiology. By helping us visualize and address hidden biases, they can contribute to improved patient outcomes and more personalized healthcare delivery. The journey towards truly unbiased radiological analysis is ongoing, but with innovative approaches like persistence images, we are illuminating the path forward, one topological feature at a time.

AUTHOR CONTRIBUTIONS

Yashbir Singh: Drafting of the manuscript Colleen and Quincy A Hathaway: Critical review.

CONFLICTS OF INTEREST

Authors have no conflicts of interest to declare.

FUNDING

No funding was used for this paper.

Yashbir Singh: Department of Radiology, Mayo Clinic, Rochester, MN 55905, USA

Correspondence to: Yashbir Singh,
email: singh.yashbir@mayo.edu

Keywords: persistence images; data interpretation; radiology; topology

Received: October 28, 2024

Published: November 12, 2024

REFERENCES

1. Edelsbrunner H, et al. Computational topology: an introduction. American Mathematical Society. 2010.
2. Adams H, et al. Journal of Machine Learning Research. 2017; 18:1–35. <https://doi.org/10.48550/arXiv.1507.06217>.
3. Singh Y, et al. Insights Imaging. 2023; 14:58. <https://doi.org/10.1186/s13244-023-01413-w>. [PubMed]
4. Chazal F, et al. An introduction to Topological Data Analysis: fundamental and practical aspects for data scientists. arXiv. 2017; arXiv:1710.04019.
5. Qaiser T, et al. Med Image Anal. 2019; 55:1–14. <https://doi.org/10.1016/j.media.2019.03.014>. [PubMed]
6. Gao M, et al. Learning topology from synthetic data for unsupervised tumor segmentation in brain MRI. arXiv. 2019; arXiv:1910.05351.
7. Larrazabal AJ, et al. Proc Natl Acad Sci U S A. 2020; 117:12592–94. <https://doi.org/10.1073/pnas.1919012117>. [PubMed].
8. Crawford K. The trouble with bias. Conference on Neural Information Processing Systems, invited speaker. 2017.
9. Joy JE, et al. The National Academies Collection: Reports funded by National Institutes of Health. 2005. <https://doi.org/10.17226/11016>. [PubMed]
10. Obermeyer Z, et al. Science. 2019; 366:447–53. <https://doi.org/10.1126/science.aax2342>. [PubMed]
11. Cohen-Steiner D, et al. Discrete & Computational Geometry. 2007; 37:103–20. <https://doi.org/10.1007/s00454-006-1276-5>.
12. Chazal F, et al. Geometriae Dedicata. 2014; 173:193–214. <https://doi.org/10.1007/s10711-013-9937-z>.
13. Hofer C, et al. Deep learning with topological signatures. In Advances in Neural Information Processing Systems. 2017; 1634–44.
14. Reininghaus J, et al. A stable multi-scale kernel for topological machine learning. In Proceedings of the IEEE conference on computer vision and pattern recognition. 2015; 4741–48.
15. Levenson RM, et al. Lab Invest. 2024; 104:102060. <https://doi.org/10.1016/j.labinv.2024.102060>. [PubMed].
16. Bauer U. Journal of Applied and Computational Topology. 2021; 5:391–423. <https://doi.org/10.1007/s41468-021-00071-5>.
17. Lum PY, et al. Sci Rep. 2013; 3:1236. <https://doi.org/10.1038/srep01236>. [PubMed]
18. Wasserman L. Annual Review of Statistics and Its Application. 2018; 5:501–32. <https://doi.org/10.1146/annurev-statistics-031017-100045>.

Copyright: © 2024 Singh et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.