means the model becomes more stable, as perturbations result in smaller changes in loss. In Figure B.6, we consider how the loss landscape changes across different learning rates. When looking at the loss landscapes for three different random seeds, after zooming in, we observe consistent variation in the depth and shape of the loss landscape as the learning rate is varied. This variation is also reflected in the test accuracy scores shown in the heat map. Interestingly, we observe deeper basins when the learning rate is too small or too big, indicating that the trained models are less stable compared to those with shallower basins.

5. Conclusion and Future Work

In this paper, we introduced a new topological landscape profile representation of neural network loss landscapes. To demonstrate the many different ways this new representation of loss landscapes can be used, we explored several different machine learning examples, including image segmentation (e.g., UNet-CRF) and scientific machine learning (e.g., PINNs). Along the way, we provided new insights into how loss landscapes vary across distinct hyperparameter spaces, finding that the topology of the loss landscape is simpler for better-performing models and that this topology is more variable near transitions from low to high model performance. Moreover, by using a merge tree to extract the most important features from a computed loss landscape, we are able to construct a new representation encoding these features. By separating this new representation from the original space in which the loss landscape was sampled, our approach opens up the door to visualizing higher-dimensional loss landscapes.

While we only explore up to four dimensions here, our approach can be extended to any number of dimensions. The limiting factor is sampling, which requires exponentially many more resources as the number of dimensions increases. However, future advances towards more efficient sampling could be combined with our current approach to reveal the higher-dimensional structure of loss functions. Complementary advances in sampling more global loss landscapes (combining multiple independently trained models) could also benefit from our new representations. In that case, we would expect to see more distinct basins in our topological landscape profiles.

6. Acknowledgments

This work was supported by the U.S. Department of Energy, Office of Science, Advanced Scientific Computing Research (ASCR) program under Contract Number DE-AC02-05CH11231 to Lawrence Berkeley National Laboratory and Award Number DE-SC0023328 to Arizona State University ("Visualizing High-dimensional Functions in Scientific Machine Learning"). This research used resources at the National Energy Research Scientific Computing Center (NERSC), a U.S. Department of Energy, Office of Science, User Facility under NERSC Award Number ASCR-ERCAP0026937.

References

Matthew Avaylon, Robbie Sadre, Zhe Bai, and Talita Perciano. Adaptable deep learning and probabilistic graphical model system for semantic segmentation. *Advances in Artificial Intelligence and Machine Learning*, 02:288–302, 2022.

Talha Bin Masood, Joseph Budin, Martin Falk, Guillaume Favelier, Christoph Garth, Charles Gueunet, Pierre Guillou, Lutz Hofmann, Petar Hristov, Adhitya Kamakshidasan, Christopher Kappe, Pavol Klacansky, Patrick Laurin, Joshua Levine, Jonas Lukasczyk, Daisuke Sakurai, Maxime Soler, Peter Steneteg, Julien Tierny, Will Usher, Jules Vidal, and Michal Wozniak. An overview of the Topology ToolKit. In *Topological Methods in Data Analysis and Visualization VI*, pages 327–342. Springer International Publishing, 2021.

Hamish Carr, Jack Snoeyink, and Ulrike Axen. Computing contour trees in all dimensions. *Computational Geometry*, 24(2):75–94, 2003. ISSN 0925-7721. Special Issue on the Fourth CGC Workshop on Computational Geometry.

Junbum Cha, Sanghyuk Chun, Kyungjae Lee, Han-Cheol Cho, Seunghyun Park, Yunsung Lee, and Sungrae Park. Swad: Domain generalization by seeking flat minima. *Advances in Neural Information Processing Systems*, 34:22405–22418, 2021.

Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. BERT: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*, 2018.

Josip Djolonga, Jessica Yung, Michael Tschannen, Rob Romijnders, Lucas Beyer, Alexander Kolesnikov, Joan Puigcerver, Matthias Minderer, Alexander D'Amour, Dan Moldovan, Gelly Sylvain, Neil Houlsby, Xiaohua Zhai, and Mario Lucic. On robustness and transferability of convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 16458–16468, 2021.

Wei Dong, Charikar Moses, and Kai Li. Efficient k-nearest neighbor graph construction for generic similarity measures. In *Proceedings of the 20th international conference on World wide web*, pages 577–586, 2011.

Ian J Goodfellow, Oriol Vinyals, and Andrew M Saxe. Qualitatively characterizing neural network optimization problems. *arXiv preprint arXiv:1412.6544*, 2014.

Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, pages 770–778, 2016.

Christian Heine, Heike Leitte, Mario Hlawitschka, Federico Iuricich, Leila De Floriani, Gerik Scheuermann, Hans Hagen, and Christoph Garth. A Survey of Topology-based Methods in Visualization. *Computer Graphics Forum*, 35(3):643–667, 2016.

Daniel Jiwoong Im, Michael Tao, and Kristin Branson. An empirical analysis of the optimization of deep network loss surfaces. *arXiv preprint arXiv:1612.04010*, 2016.

Jerzy W Jaromczyk and Godfried T Toussaint. Relative neighborhood graphs and their relatives. *Proceedings of the IEEE*, 80(9):1502–1517, 1992.

Arham Khan, Todd Nief, Nathaniel Hudson, Mansi Sakarvadia, Daniel Grzenda, Aswathy Ajith, Jordan Pettyjohn, Kyle Chard, and Ian Foster. Sok: On finding common ground in loss landscapes using deep model merging techniques. *arXiv preprint arXiv:2410.12927*, 2024.

Aditi Krishnapriyan, Amir Gholami, Shandian Zhe, Robert Kirby, and Michael W Mahoney. Characterizing possible failure modes in physics-informed neural networks. *Advances in Neural Information Processing Systems*, 34:26548–26560, 2021.

Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, 60(6):84–90, 2017.

Alexey Kurakin, Ian Goodfellow, and Samy Bengio. Adversarial machine learning at scale. *arXiv preprint arXiv:1611.01236*, 2016.

Hao Li, Zheng Xu, Gavin Taylor, Christoph Studer, and Tom Goldstein. Visualizing the loss landscape of neural nets. *Advances in neural information processing systems*, 2018.

Yinhan Liu, Myle Ott, Naman Goyal, Jingfei Du, Mandar Joshi, Danqi Chen, Omer Levy, Mike Lewis, Luke Zettlemoyer, and Veselin Stoyanov. RoBERTa: A Robustly Optimized BERT Pretraining Approach. *arXiv preprint arXiv:1907.11692*, 2019.

Charles H Martin and Michael W Mahoney. Implicit Self-Regularization in Deep Neural Networks: Evidence from Random Matrix Theory and Implications for Learning. *The Journal of Machine Learning Research*, 22(1):7479–7551, 2021.

Charles H Martin, Tongsu Peng, and Michael W Mahoney. Predicting trends in the quality of state-of-the-art neural networks without access to training or testing data. *Nature Communications*, 12(1):4122, 2021.

Patrick Oesterling, Christian Heine, Gunther H. Weber, and Gerik Scheuermann. Visualizing nd point clouds as topological landscape profiles to guide local data analysis. *IEEE Transactions on Visualization and Computer Graphics*, 19(3):514–526, 2013.

Omkar M Parkhi, Andrea Vedaldi, Andrew Zisserman, and CV Jawahar. Cats and dogs. In *2012 IEEE conference on Computer Vision and Pattern Recognition*, pages 3498–3505. IEEE, 2012.

Pratik Rathore, Weimu Lei, Zachary Frangella, Lu Lu, and Madeleine Udell. Challenges in training pinns: A loss landscape perspective. *arXiv preprint arXiv:2402.01868*, 2024.

Mansi Sakarvadia, Aswathy Ajith, Arham Khan, Nathaniel Hudson, Caleb Geniesse, Kyle Chard, Yaoqing Yang, Ian Foster, and Michael W. Mahoney. Mitigating memorization in language models. *arXiv preprint arXiv:2410.02159*, 2024.

Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 2017.