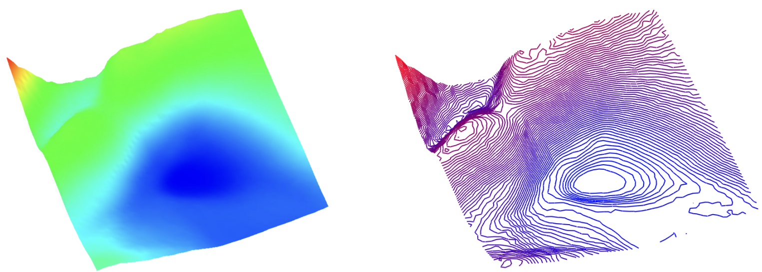
CS 4/553 Term Project: Visualizing the Loss Landscape of Neural Nets

Charles Ison*
Oregon State University

Matthew Morgan*
Oregon State University



(a) (b)

Figure 1: The loss landscapes of two different models trained on the same dataset for the same number of epochs: (a) ResNet50, and (b) VGG11.

Abstract

TODO: Designing rotational symmetries on surfaces is a necessary task for a wide variety of graphics applications, such as surface parameterization and remeshing, painterly rendering and pen-and-ink sketching, and texture synthesis. In these applications, the *topology* of a rotational symmetry field such as *singularities* and *separatrices* can have a direct impact on the quality of the results. In this paper, we present a design system that provides control over the topology of rotational symmetry fields on surfaces.

As the foundation of our system, we provide comprehensive analysis for rotational symmetry fields on surfaces and present efficient algorithms to identify singularities and separatrices. We also describe design operations that allow a rotational symmetry field to be created and modified in an intuitive fashion by using the idea of basis fields and relaxation. In particular, we provide control over the topology of a rotational symmetry field by allowing the user to remove singularities from the field or to move them to more desirable locations.

At the core of our analysis and design implementations is the observation that N-way rotational symmetries can be described by symmetric N-th order tensors, which allows an efficient vector-based representation that not only supports coherent definitions of arithmetic operations on rotational symmetries but also enables many analysis and design operations for vector fields to be adapted to rotational symmetry fields.

To demonstrate the effectiveness of our approach, we apply our design system to pen-and-ink sketching and geometry remeshing.

*e-mail: {isonc|morgamat}@eecs.oregonstate.edu

Keywords: loss landscape, scalar field topology, neural networks, gradient descent

1 Introduction

Neural networks are machine learning models constructed from interconnected nodes. Backpropagation through gradient descent is one commonly used algorithm for training neural networks. By visualizing neural network loss functions as scalar fields, can we better understand the impact of various network architectures on the model's trainability.

In this project, we implement the paper "Visualizing the Loss Landscape of Neural Nets" (`https://arxiv.org/pdf/1712.09913.pdf`) for Option 2 (implementing a published research paper). By mapping the loss of a neural network across two dimensions, we create scalar field visualizations of various neural network architectures that the paper authors call "Loss Landscapes." In these visualizations, the loss function's results serve as the scalar value and a two dimensional reduction of the model's weights serves the directional elements of the field. The specific loss functions used in the original paper are cross entropy (`https://pytorch.org/docs/stable/generated/torch.nn.CrossEntropyLoss.html`) and mean squared error (`https://pytorch.org/docs/stable/generated/torch.nn.MSELoss.html`). The dimensionality reduction is done using two different approaches: principal component analysis and random vector selection. For the later approach, two random vectors within the input space are created by sampling a random Gaussian distribution. These vectors are used as the two dimensions of the scalar field input. The resulting loss landscapes generated can then provide insights into how certain architectural decisions can improve or worsen a network's trainability.

2 Previous Work

There have been numerous studies on the ability to optimize neural loss functions in order to improve training time and model performance. Less work has been conducted on visualizing the losses though to gain intuition about how certain architectural decisions impact the loss field's convexity. One of the original works in the space was published in 1997 by [Hochreiter and Schmidhuber 1997]. The authors defined the "flatness" of a loss landscape as the size of the connected region around the minimum where the training loss remains low which is visualized by this work. In the time since then, technology rapidly advanced and along with it the complexity of neural networks. The paper we are implementing, ”Visualizing the Loss Landscape of Neural Nets”, is one of the first to visualize the loss landscapes of modern architectures using both normalized random direction iteration and principal component analysis as dimensionality reduction techniques. Since the paper was released in 2018, numerous other works have cited it. However, many of the citations are not from works attempting to advance the visualizations, but rather to learn from the visualizations and develop more generalizable and accurate models. The visualizations have come in handy for researchers studying reinforcement learning models such as [Bekci and Gümüş 2020] and [Plaat 2022].

There are some recent works which produce new visualizations of loss functions. One such by [Huang et al. 2020], ”Understanding Generalization Through Visualizations”, use visualization methods to give intuitions about why certain architectures generalize better than others. They first visualize loss as a scalar field with height and color representing the output. But they also use a colored dot plot and a ”Swissroll decision boundary” to show the difference in models that perform well versus models that perform poorly at generalizing.

An interesting extension of this work by [Linse et al. 2022] visualizes large neural networks in virtual reality. The approach allows for high interactivity, but requires a virtual reality headset and powerful computer to render.

3 Background

TODO: Magnam nulla non possimus est error aliquam non iste. Hic praesentium perspiciatis numquam. Necessitatibus aut mollitia porro quod. Libero voluptatum illo exercitationem incidunt voluptate. Vero ut sapiente quia nisi possimus non. Consequuntur molestiae voluptas eum illo ut quo qui. Et quod dolore maxime. Quo omnis ut a iure ut facere. Asperiores et explicabo aliquid omnis ratione laborum libero. Aperiam fuga facilis soluta non suscipit. Saepe velit ut iusto dolorum beatae quae expedita. Facere quia aut architecto velit quibusdam iste. Explicabo a provident debitis. Quia dolore numquam qui laboriosam consectetur necessitatibus. Quisquam officiis placeat vitae necessitatibus aut. Iure libero quae nostrum expedita quis voluptates harum. Delectus natus amet nihil voluptatem corporis maxime. Necessitatibus natus aut tempore aut et. Eos facilis aut id commodi quis id maiores. Labore distinctio nobis deserunt perferendis perferendis perferendis laborum temporibus. Dignissimos nostrum laudantium enim unde numquam temporibus ipsa. Qui ut officia enim est deserunt.

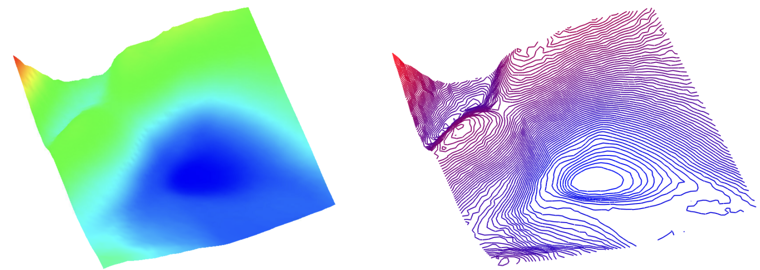
4 Results

4.1 Results 1

TODO: Results Magnam nulla non possimus est error aliquam non iste. Hic praesentium perspiciatis numquam. Necessitatibus aut mollitia porro quod. Libero voluptatum illo exercitationem incidunt voluptate. Vero ut sapiente quia nisi possimus non. Consequuntur molestiae voluptas eum illo ut quo qui. Et quod dolore maxime. Quo omnis ut a iure ut facere. Asperiores et explicabo aliquid omnis ratione laborum libero. Aperiam fuga facilis soluta non suscipit.

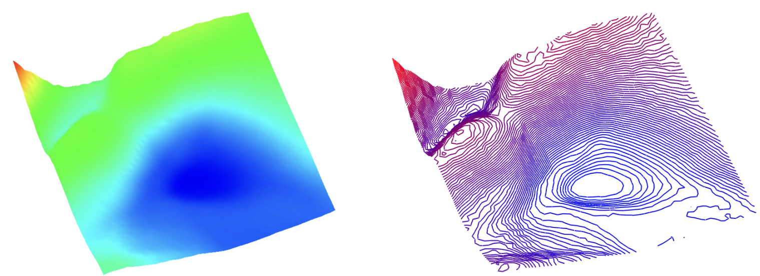
4.2 Results 2

TODO: Results Magnam nulla non possimus est error aliquam non iste. Hic praesentium perspiciatis numquam. Necessitatibus aut mollitia porro quod. Libero voluptatum illo exercitationem incidunt voluptate. Vero ut sapiente quia nisi possimus non. Consequuntur molestiae voluptas eum illo ut quo qui. Et quod dolore maxime. Quo omnis ut a iure ut facere. Asperiores et explicabo aliquid omnis ratione laborum libero. Aperiam fuga facilis soluta non suscipit.



(a) (b)

Figure 2: The loss landscapes of two different models trained on the same dataset for the same number of epochs: (a) ResNet50, and (b) VGG11.



(a) (b)

Figure 3: VGG11 random directions visualized using two techniques: color + height and contours + height.

5 Evaluation

TODO: Magnam nulla non possimus est error aliquam non iste. Hic praesentium perspiciatis numquam. Necessitatibus aut mollitia porro quod. Libero voluptatum illo exercitationem incidunt voluptate. Vero ut sapiente quia nisi possimus non. Consequuntur molestiae voluptas eum illo ut quo qui. Et quod dolore maxime. Quo omnis ut a iure ut facere. Asperiores et explicabo aliquid omnis ratione laborum libero. Aperiam fuga facilis soluta non suscipit. Saepe velit ut iusto dolorum beatae quae expedita. Facere quia aut architecto velit quibusdam iste. Explicabo a provident debitis. Quia dolore numquam qui laboriosam consectetur necessitatibus. Quisquam officiis placeat vitae necessitatibus aut. Iure libero quae nostrum expedita quis voluptates harum. Delectus natus amet nihil voluptatem corporis maxime. Necessitatibus natus aut tempore aut et. Eos facilis aut id commodi quis id maiores. Labore distinctio nobis deserunt perferendis perferendis perferendis laborum temporibus. Dignissimos nostrum laudantium enim unde numquam temporibus ipsa. Qui ut officia enim est deserunt.

6 Division of Tasks

Charles took responsibility for creating and training the ResNet and VGG models on CIFAR-10 dataset. He then implemented the dimensionality reduction techniques of principal component analysis and normalized random direction iteration. Next, by iteratively manipulate the models weights during training and testing, Charles constructed the loss landscape data as tensors from the PyTorch models. Matthew took responsibility for writing the code to export the loss landscape data from the PyTorch models to PLY files which could be visualized in OpenGL. Once in OpenGL, Matthew applied the techniques of mapping the landscape to color, visualizing the loss as height, and drawing contour lines across the scalar field. Additionally, Matthew took data from the model's training iterations and plotted the gradient descent of the model as a curve on the scalar field. Last, he compared and contrasted the different techniques to find the optimal combinations and discussed how they might be interpreted. Charles also contributed to the OpenGL visualizations by extracting and showing the critical points, as well as evaluating their correctness.

7 Conclusions

TODO: Magnam nulla non possimus est error aliquam non iste. Hic praesentium perspiciatis numquam. Necessitatibus aut mollitia porro quod. Libero voluptatum illo exercitationem incidunt voluptate. Vero ut sapiente quia nisi possimus non. Consequuntur molestiae voluptas eum illo ut quo qui. Et quod dolore maxime. Quo omnis ut a iure ut facere. Asperiores et explicabo aliquid omnis ratione laborum libero. Aperiam fuga facilis soluta non suscipit. Saepe velit ut iusto dolorum beatae quae expedita. Facere quia aut architecto velit quibusdam iste. Explicabo a provident debitis. Quia dolore numquam qui laboriosam consectetur necessitatibus. Quisquam officiis placeat vitae necessitatibus aut. Iure libero quae nostrum expedita quis voluptates harum. Delectus natus amet nihil voluptatem corporis maxime. Necessitatibus natus aut tempore aut et. Eos facilis aut id commodi quis id maiores. Labore distinctio nobis deserunt perferendis perferendis perferendis laborum temporibus. Dignissimos nostrum laudantium enim unde numquam temporibus ipsa. Qui ut officia enim est deserunt.

References

BEKCI, R. Y., AND GÜMÜŞ, M., 2020. Visualizing the loss landscape of actor critic methods with applications in inventory optimization.

HOCHREITER, S., AND SCHMIDHUBER, J. 1997. Flat minima. *Neural computation 9*, 1, 1–42.

HUANG, W. R., EMAM, Z., GOLDBLUM, M., FOWL, L., TERRY, J. K., HUANG, F., AND GOLDSTEIN, T. 2020. Understanding generalization through visualizations. In *Proceedings on "I Can't Believe It's Not Better!" at NeurIPS Workshops*, PMLR, J. Zosa Forde, F. Ruiz, M. F. Pradier, and A. Schein, Eds., vol. 137 of *Proceedings of Machine Learning Research*, 87–97.

LI, H., XU, Z., TAYLOR, G., STUDER, C., AND GOLDSTEIN, T. 2018. Visualizing the loss landscape of neural nets. In *Advances in Neural Information Processing Systems*, Curran Associates, Inc., S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, Eds., vol. 31.

LINSE, C., ALSHAZLY, H., AND MARTINETZ, T. 2022. A walk in the black-box: 3d visualization of large neural networks in virtual reality. *Neural Computing and Applications*, 1–16.

PLAAT, A. 2022. Deep reinforcement learning. *arXiv preprint arXiv:2201.02135*.