

GlobScope: Toward a Global View of the Loss Landscape

Mashiat Mustaq

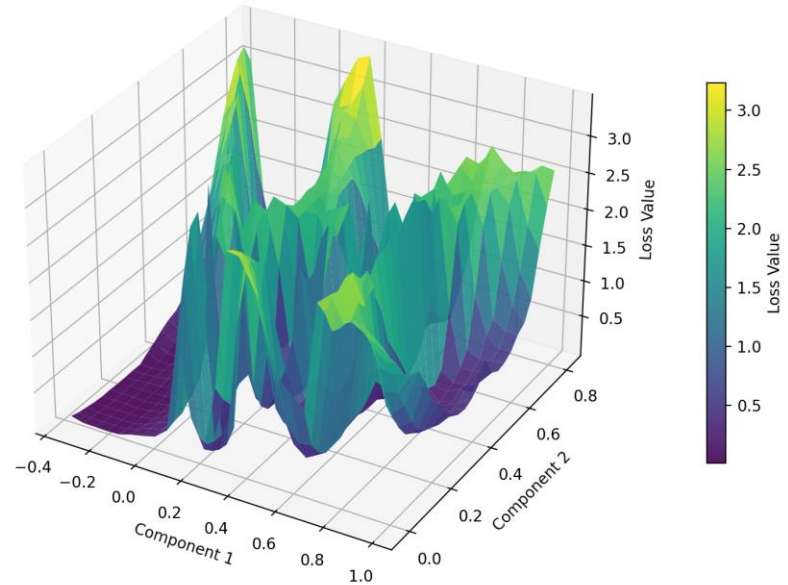
Xavier M. Tricoche

Purdue University

Definition: Loss Landscape

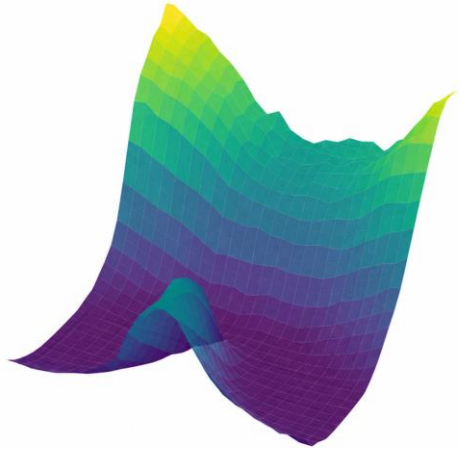
The surface formed by plotting a neural network's loss value against all possible parameter values

- Defined by loss function, $L(\theta)$ over the space of trainable parameters θ
- Extremely high dimensional
- Reduction to lower dimensions for visualization

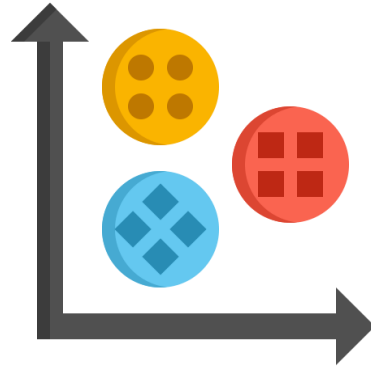


2D Loss Landscape of a trained ResNet56 on CIFAR10

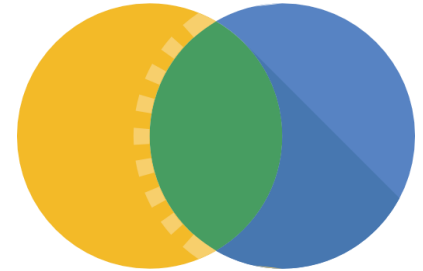
Motivation: Why is it important?



Generalization

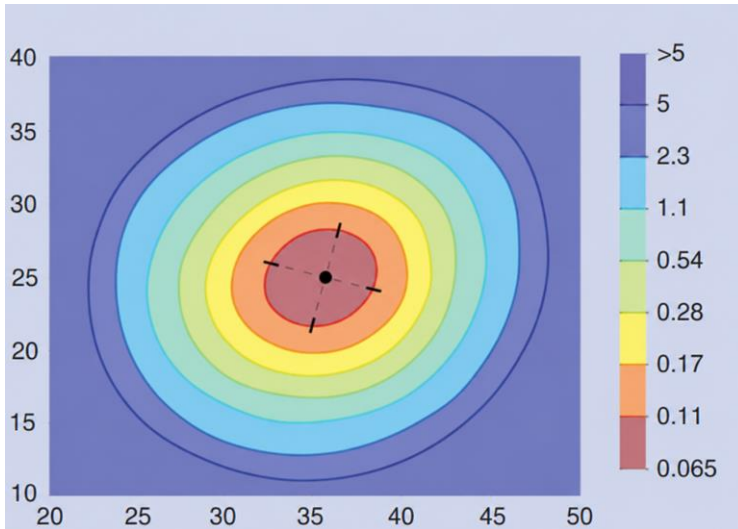


Ensembling



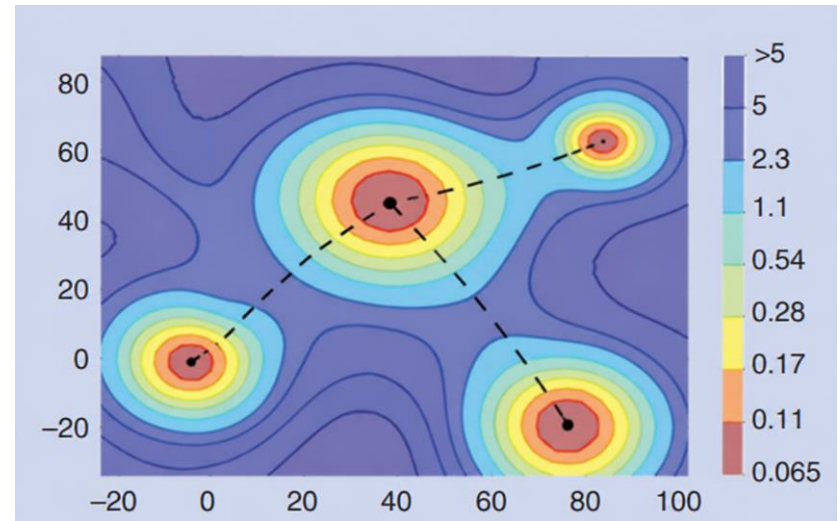
Model Similarity

Definition: Local VS Global Loss Landscape



Local

- Trajectory of a single model
- Single basin



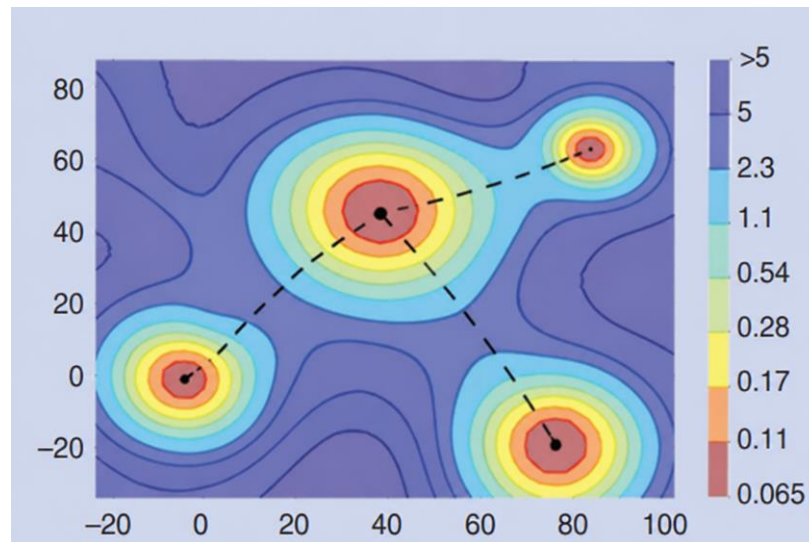
Global

- Trajectories of different models
- Multiple basins

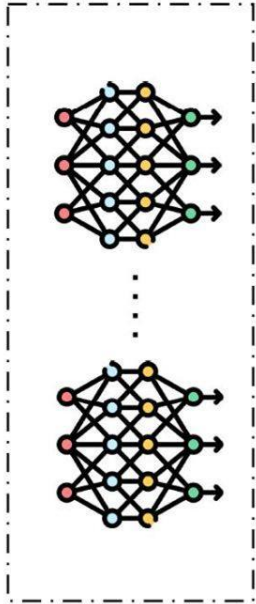
Goal: GlobScope

Construct a **low dimensional global loss landscape** visualization that:

- **Reconstructs** the full model in high dimensions with low error.
- Extends **beyond 2D** visual analysis.
- **Preserves** key high-dimensional **structure** in the low-dimensional view



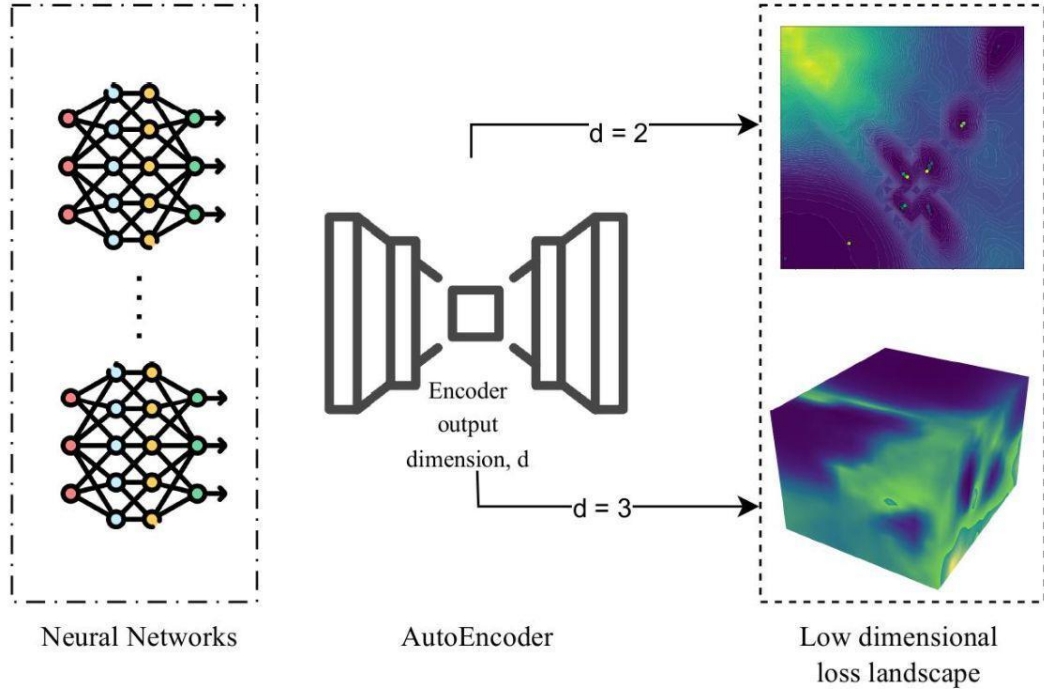
GlobScope: Pipeline



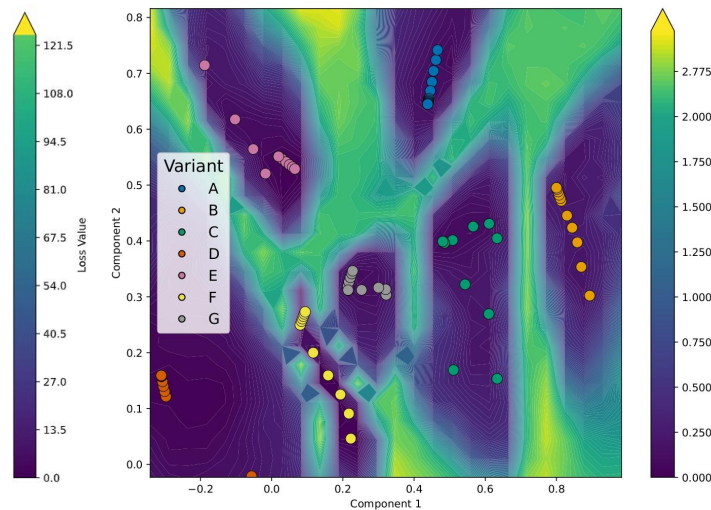
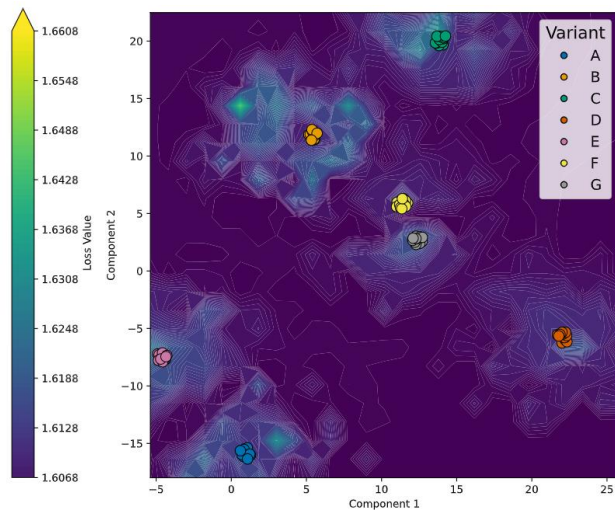
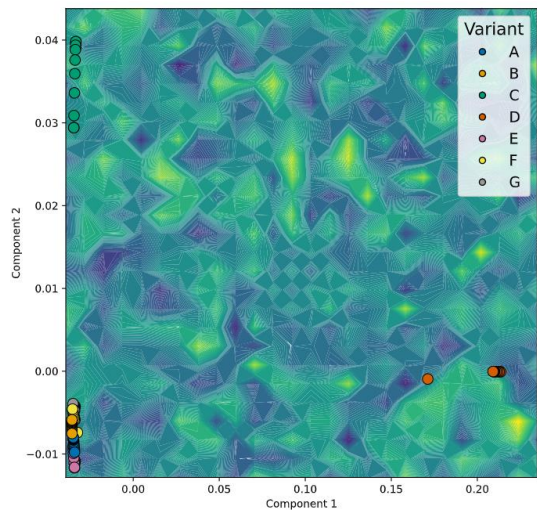
Neural Networks

Input

GlobScope: Pipeline



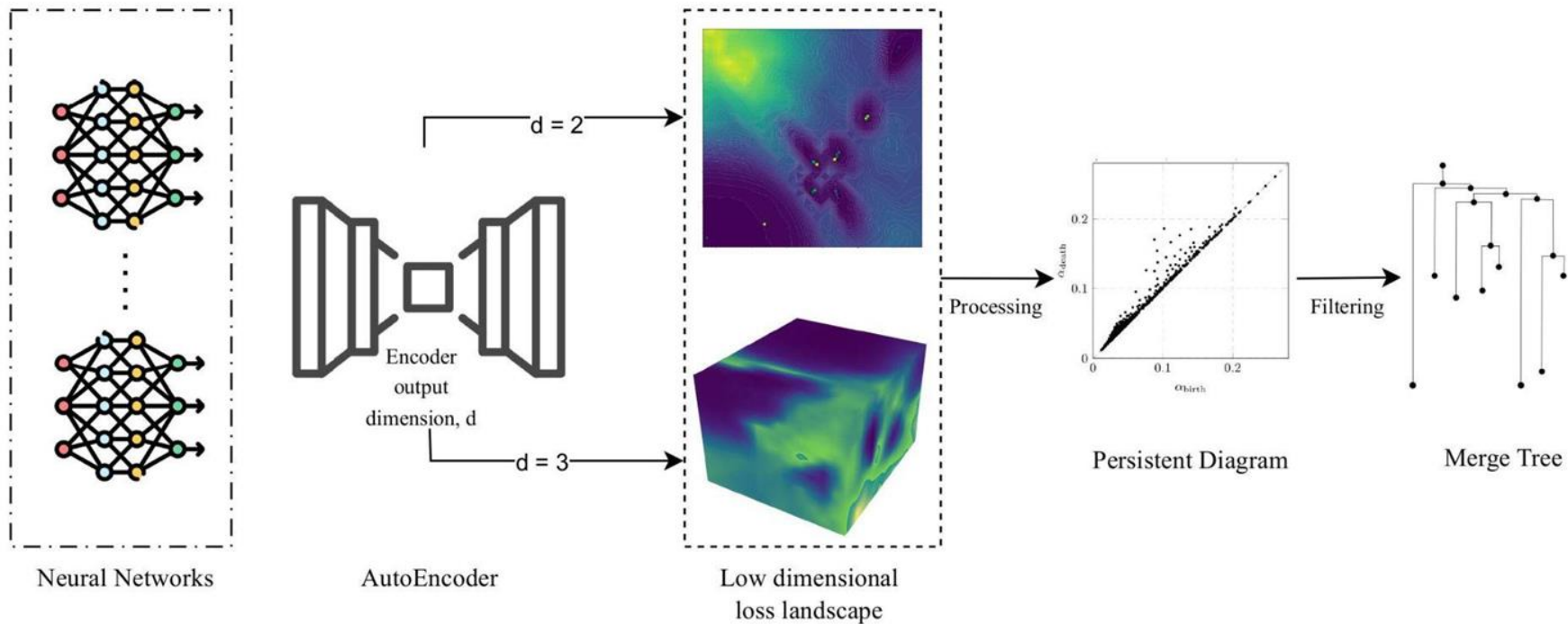
Comparison with Baseline Methods



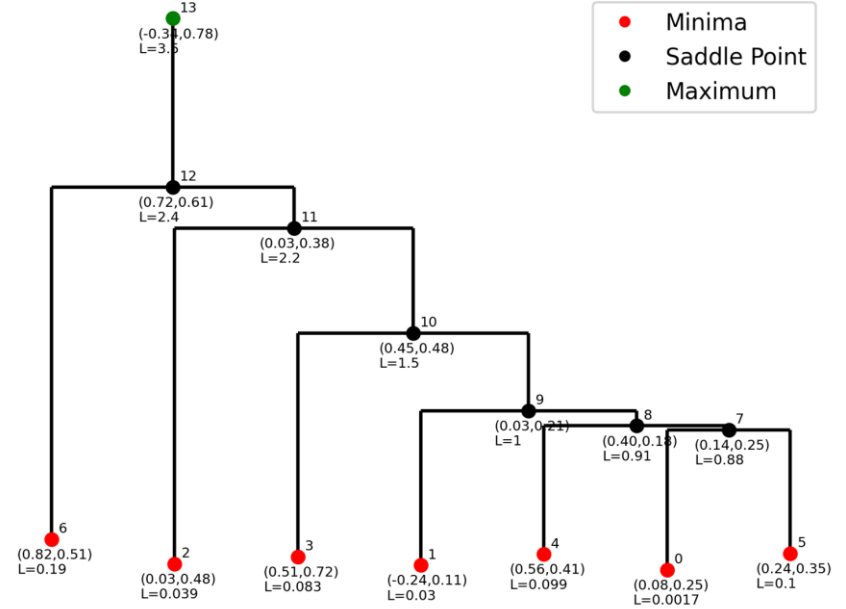
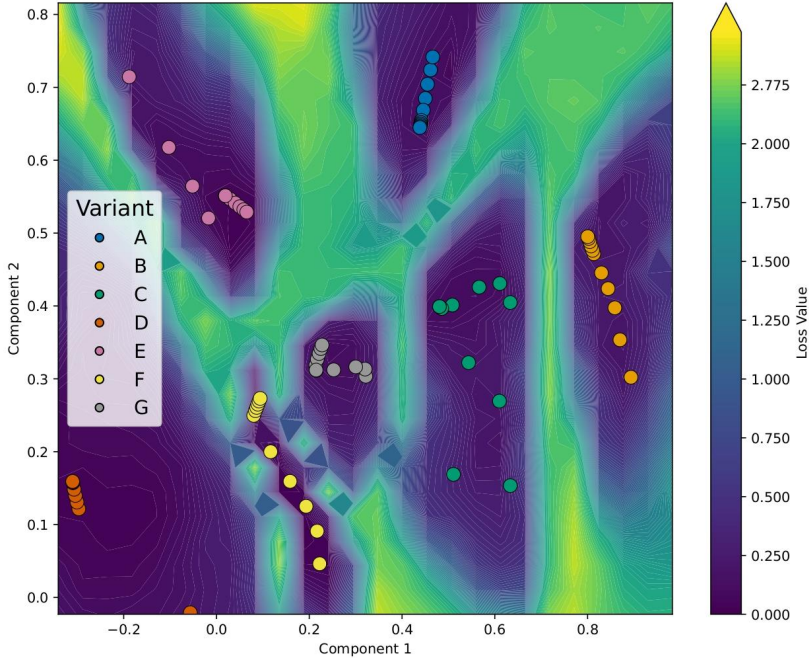
Kernel PCA	UMAP	AutoEncoder
0.00427	0.501484	0.000007

Reconstruction Loss

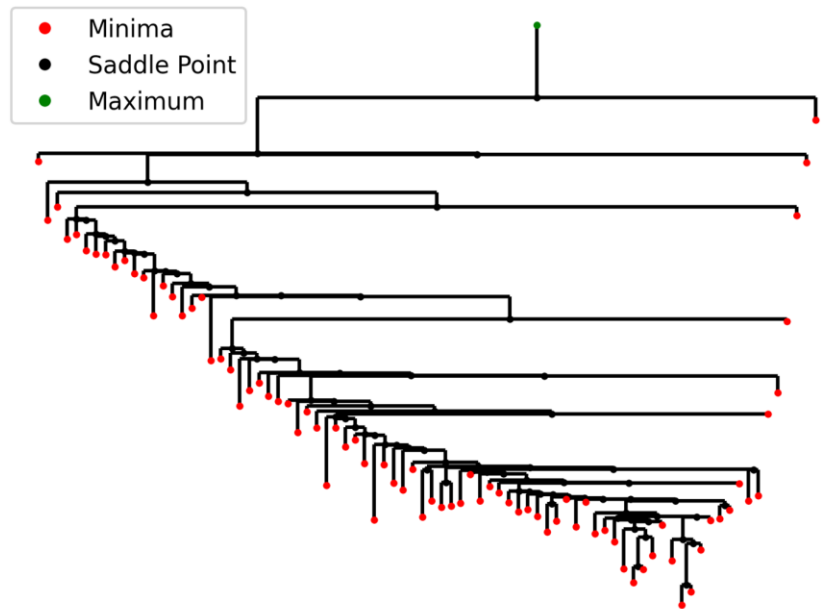
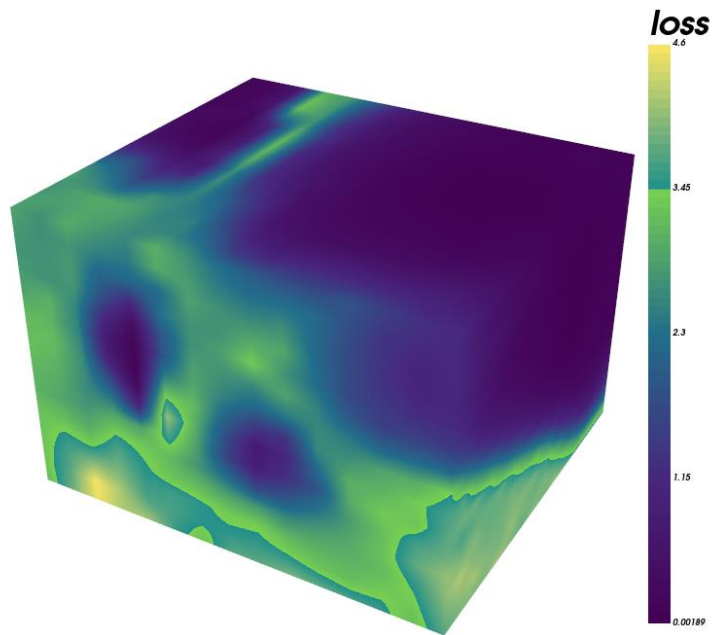
GlobScope: Pipeline



Merge Tree for 2D Loss Landscape



Merge Tree for 3D Loss Landscape

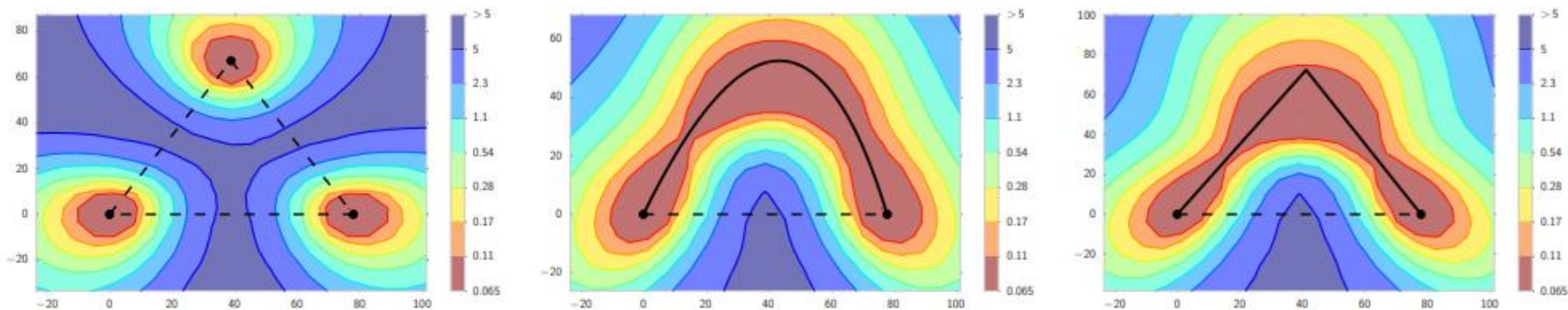


Is GlobScope consistent with the results of existing applications?

Evaluation - Mode Connectivity

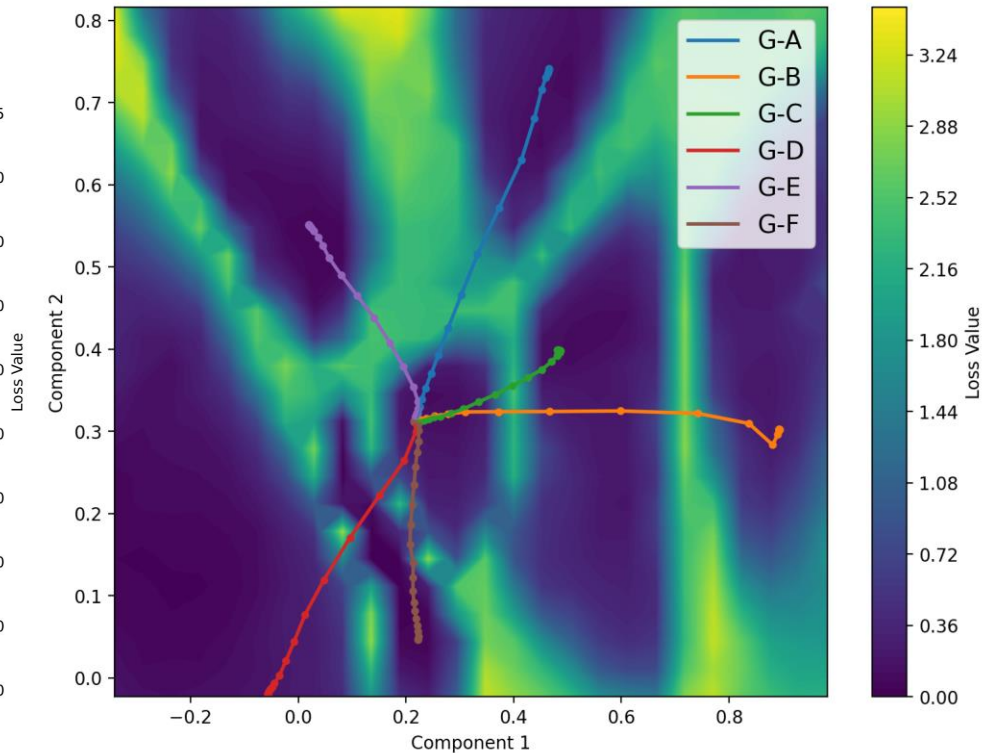
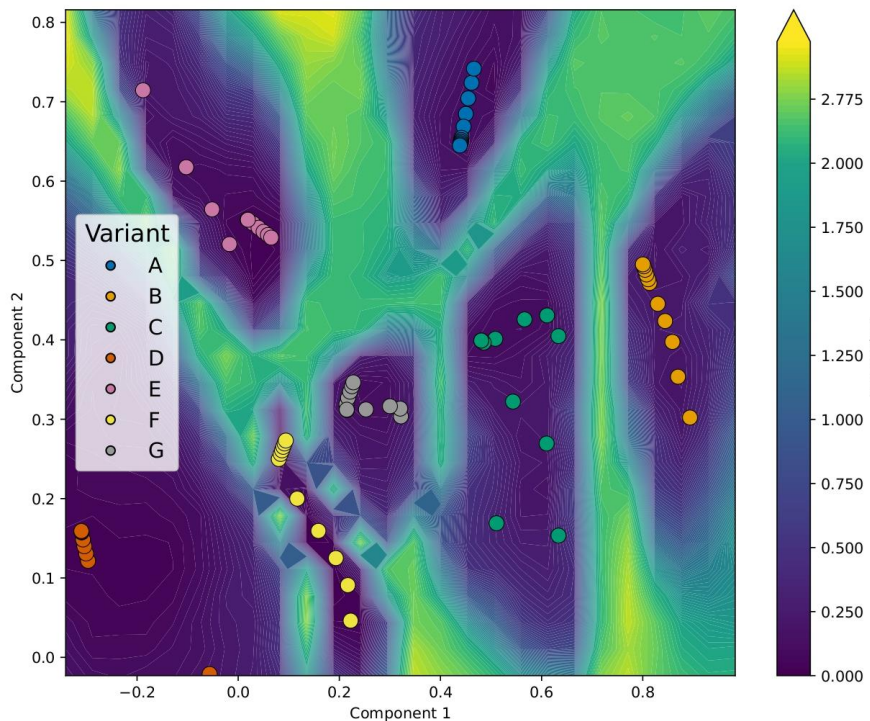
Two optima in a neural network's parameter space are connected by a **low-loss curve** along which both training and test accuracies remain nearly constant, a phenomenon known as **Mode Connectivity**

- Garipov et al. [2018]



Reference: Garipov et al (NEURIPS 2018)

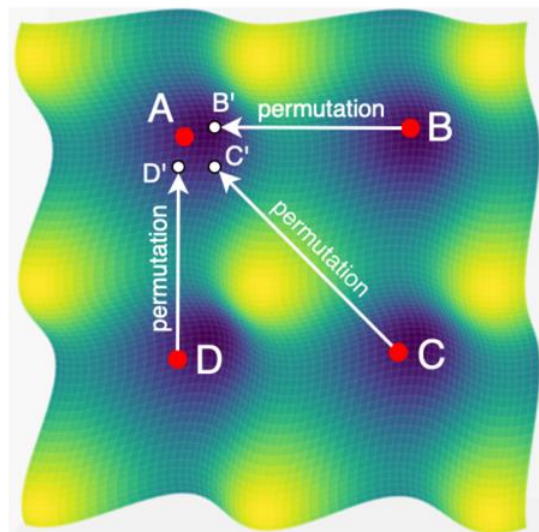
Evaluation - Mode Connectivity Path Visualization



Evaluation - Permutation Symmetry

Definition: Permuting the order of neurons in a neural network, along with their corresponding incoming and outgoing weights, does not change the function.

- If permutation symmetry is taken into account, differently trained neural network models may lie in the same basin
- Weight alignment algorithms aim to identify the neuron permutations that reveal this equivalence.

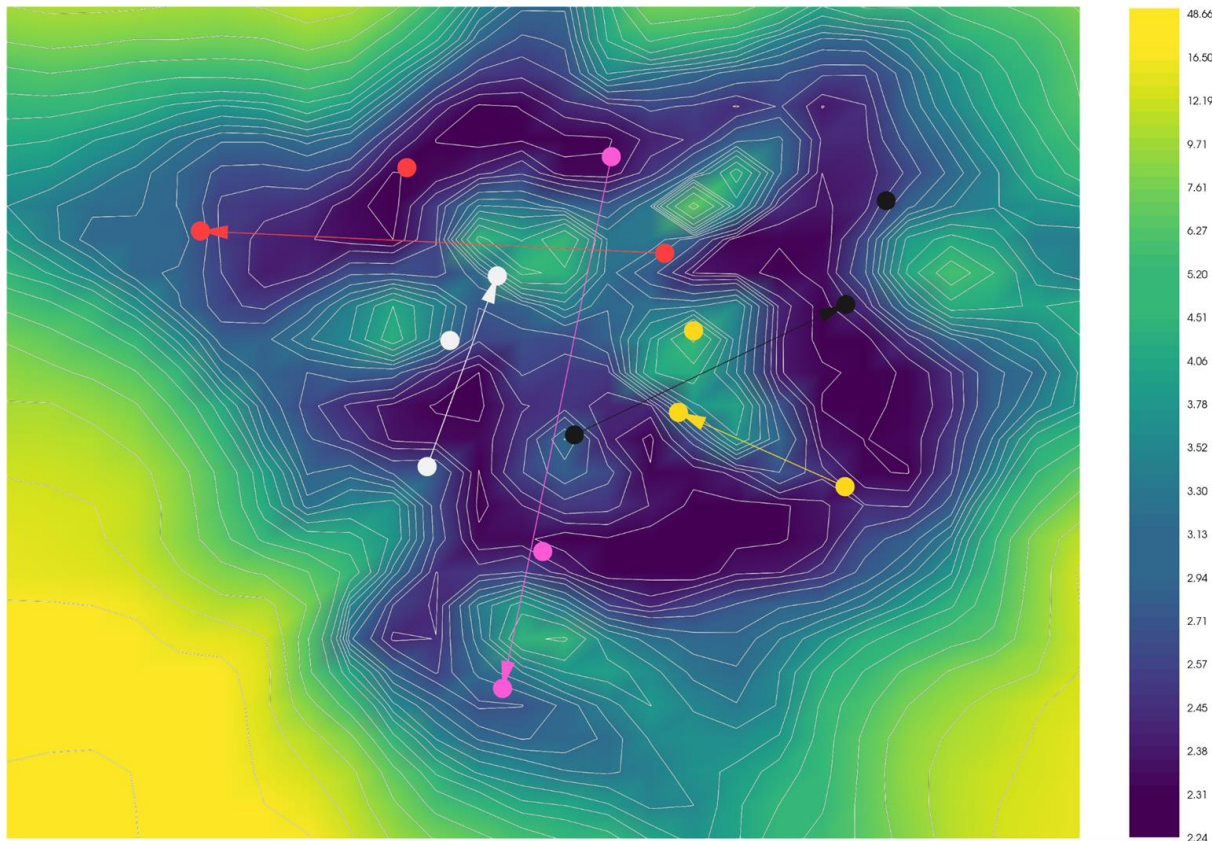


Reference: Entezari et al (ICLR 2022)

Evaluation - Permutation Symmetry Visualization

Weight Alignment
Algorithm: DEEP-
ALIGN (ICML 2024)

MLP trained on MNIST



Computational Requirements

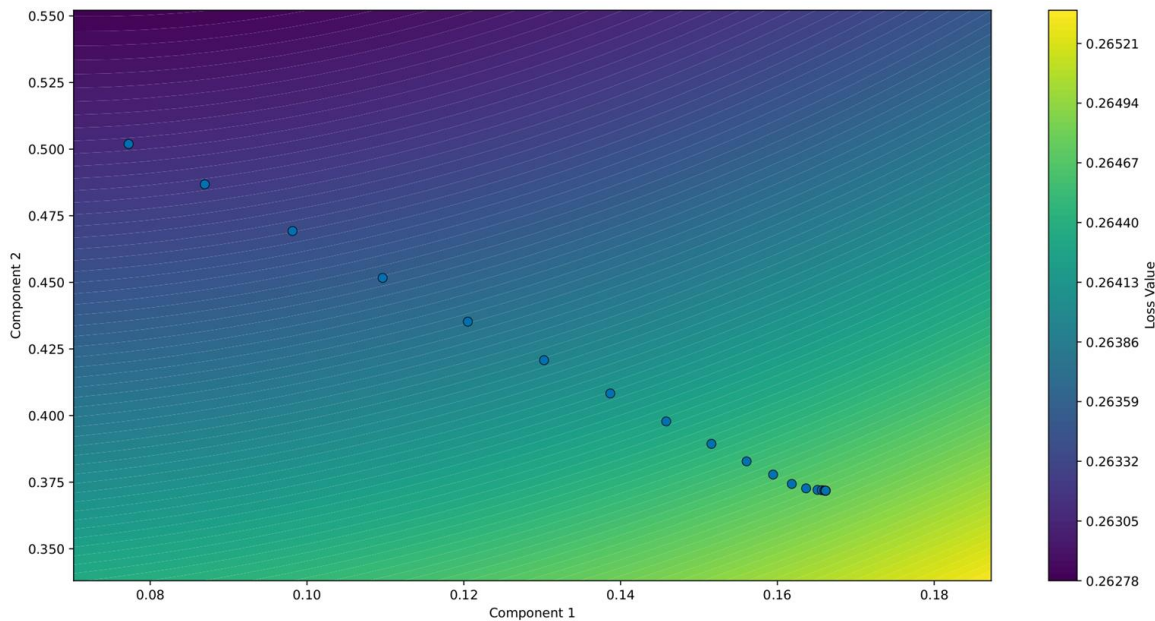
Model	# Params	Peak Mem.
MLP-MNIST	1.35×10^5	~ 0.5 GB
ResNet56-CIFAR10	8.56×10^5	~ 3.3 GB
CNN-CIFAR10	4.73×10^6	~ 18.0 GB
CNN-CIFAR100	4.91×10^6	~ 18.7 GB

GlobScope: Contributions

- Visualization of global loss landscape with low error.
- Visualization of the relation of multiple minima or basins.
- Enabling analysis of loss landscape beyond two dimensions.
- Consistent with established methods.

Future Work

- Visualizing trajectory of larger models
- Applications of loss landscape geometry to develop new algorithms



Ongoing work: Trajectory of a Vision Transformer (10^7 params)