Title: **Visualizing the Loss Landscape of Neural Networks: Insights into Trainability and Generalization**

*Presented by : Juanita AD & Thywill E. Attipoe*

The paper investigates the relationship between the geometry of neural network loss functions (the "loss landscape") and their trainability and generalization performance. The authors introduce a novel "filter normalization" technique to visualize high-dimensional loss surfaces in a meaningful way. Using this technique and other visualization methods, they analyze the impact of:

\* Network architecture: Specifically, the presence or absence of skip connections (like in ResNet architectures) significantly affects the loss landscape's smoothness and convexity. Skip connections promote "flat" minima which generalize better. Deeper networks without skip connections tend towards chaotic and highly non-convex loss landscapes.

\* Training parameters: Batch size and weight decay influence the sharpness and scale of the loss function's minima, but these effects are often artifacts of scale invariance that are misleading if not properly normalized.

\* Width of the network: Wider networks (more filters per layer) tend to have less chaotic and more convex loss landscapes.

In summary , this research provides compelling evidence that the geometry of the loss landscape significantly affects the trainability and generalization ability of neural networks. The use of skip connections and wider networks appears to be crucial for creating more benign, convex-like loss landscapes, making training easier and resulting in models that generalize better. However, further research is needed to fully understand the complex interplay between loss function geometry, optimization algorithms, and generalization performance. The authors’ findings challenge existing assumptions and offer a new perspective on fundamental questions in deep learning.

<https://docs.google.com/presentation/d/1EpvVC1mv0Q79-fzHX9YybzE0J8O7loI3sirKi-1Lgdg/edit?usp=sharing>-ppt link

<https://www.kaggle.com/datasets/fedesoriano/cifar10-python-in-csv/data>-dataset

<https://papers.nips.cc/paper/7875-visualizing-the-loss-landscape-of-neural-nets->research paper

Reference:

Authors: Hao Li, Zheng Xu, Gavin Taylor, Christoph Studer, Tom Goldstein\* Authors: Hao Li, Zheng Xu, Gavin Taylor, Christoph Studer, Tom Goldstein