

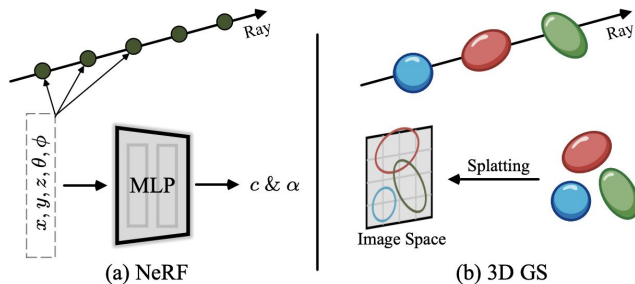
Autoencoders for 2D Gaussian Splats

Rok Mokotar, Federico Harjes Ruiloba

Background

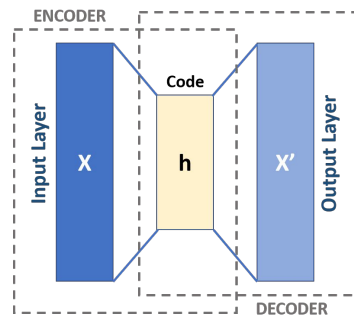
Gaussian Splats

- ML-based scene representation and novel view synthesis -> **NeRF**: high-quality, continuous representation
- Slow optimization due to the cost of training a NN
- **3D Gaussian Splats**: much faster and similar quality, discrete representation



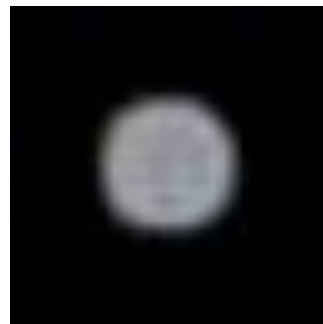
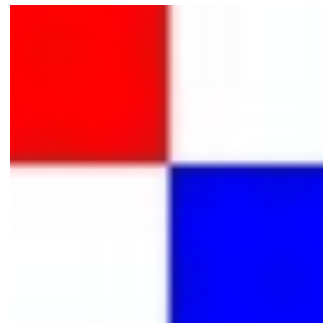
Autoencoders

- Encoder learns a compact representation of the input in the *latent space*
- Decoder then reconstructs the input from the latent space
- Useful for feature extraction and dimensionality reduction, among others



Methodology: Training a Splat Dataset

- **Dataset:** CIFAR-10
- **Splat generation:** gsplat
- **Hyperparameter Tuning:** Optuna
- **Added features:**
 - Selective learning of splat parameters
 - Support for different rasterization techniques
 - Bilateral guided radiance support
- **Initialization strategies:**
 - Random 3D initialization
 - Grid-based initialization
 - KNN-based initialization



Methodology: Autoencoding Gaussian Splats

Model Architectures

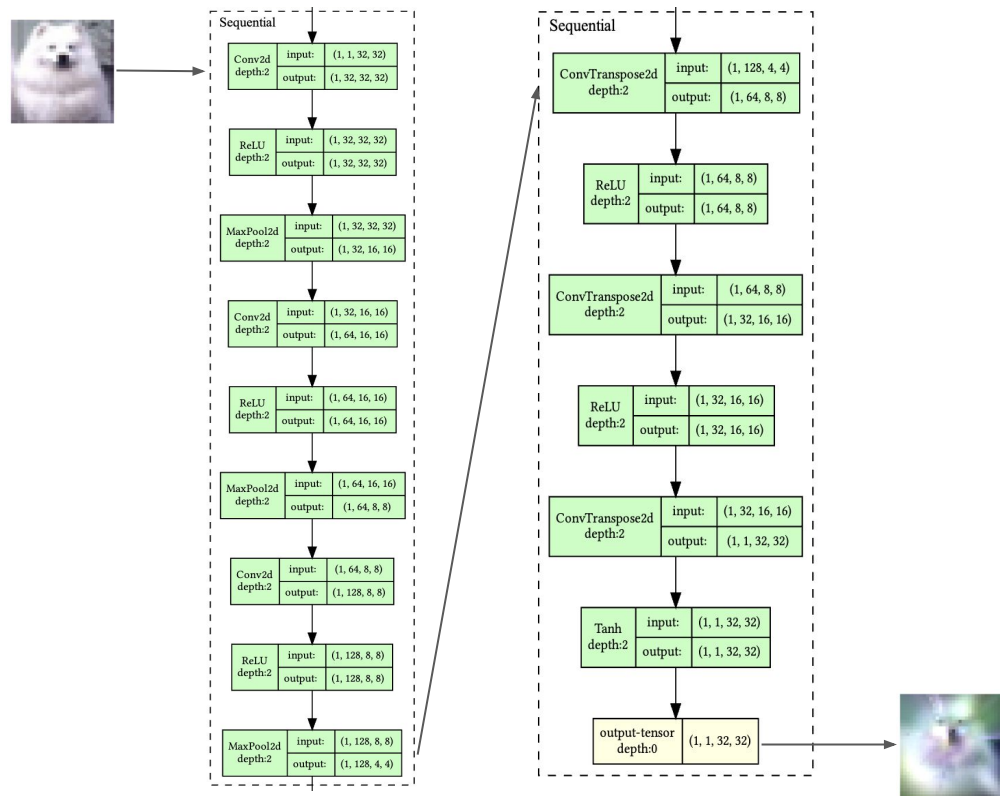
- Simple (Deep) AE
- Convolutional AE
- ResNet-18 AE
- (Abstract AE)

Splat Representations

- Vector-based encoding
- Full-image encoding
- Single-channel encoding
- Independent parameter models

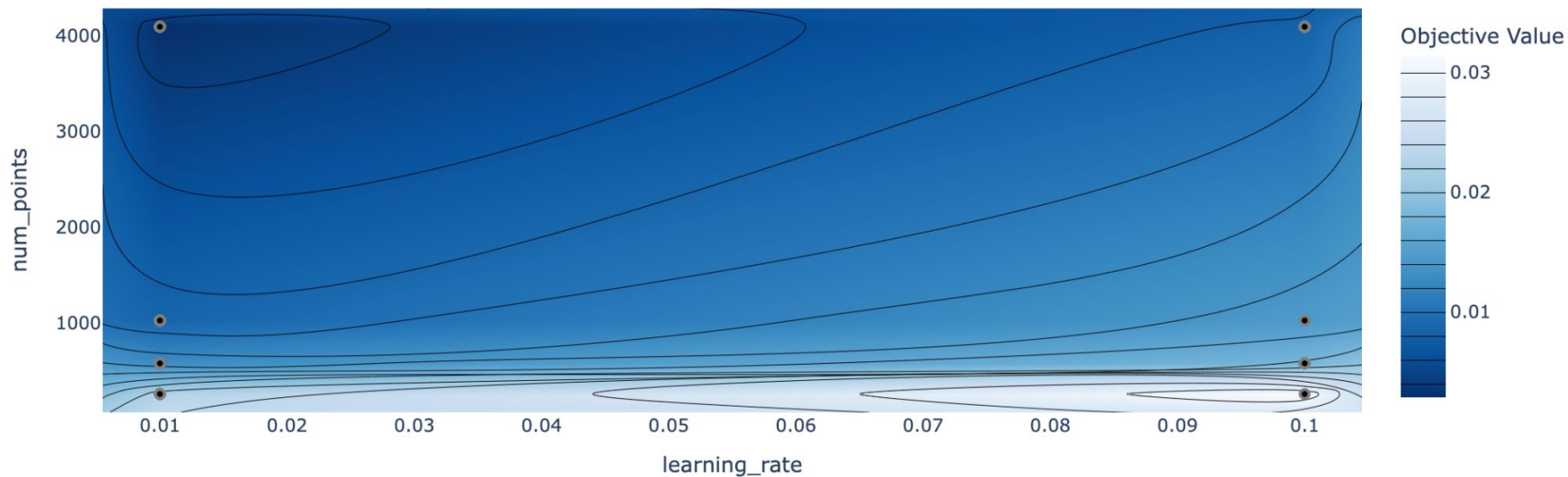
Hyperparameter Tuning (Optuna)

Slurm: Distributed Training over 10 machines with either 2060 RTX or CPU

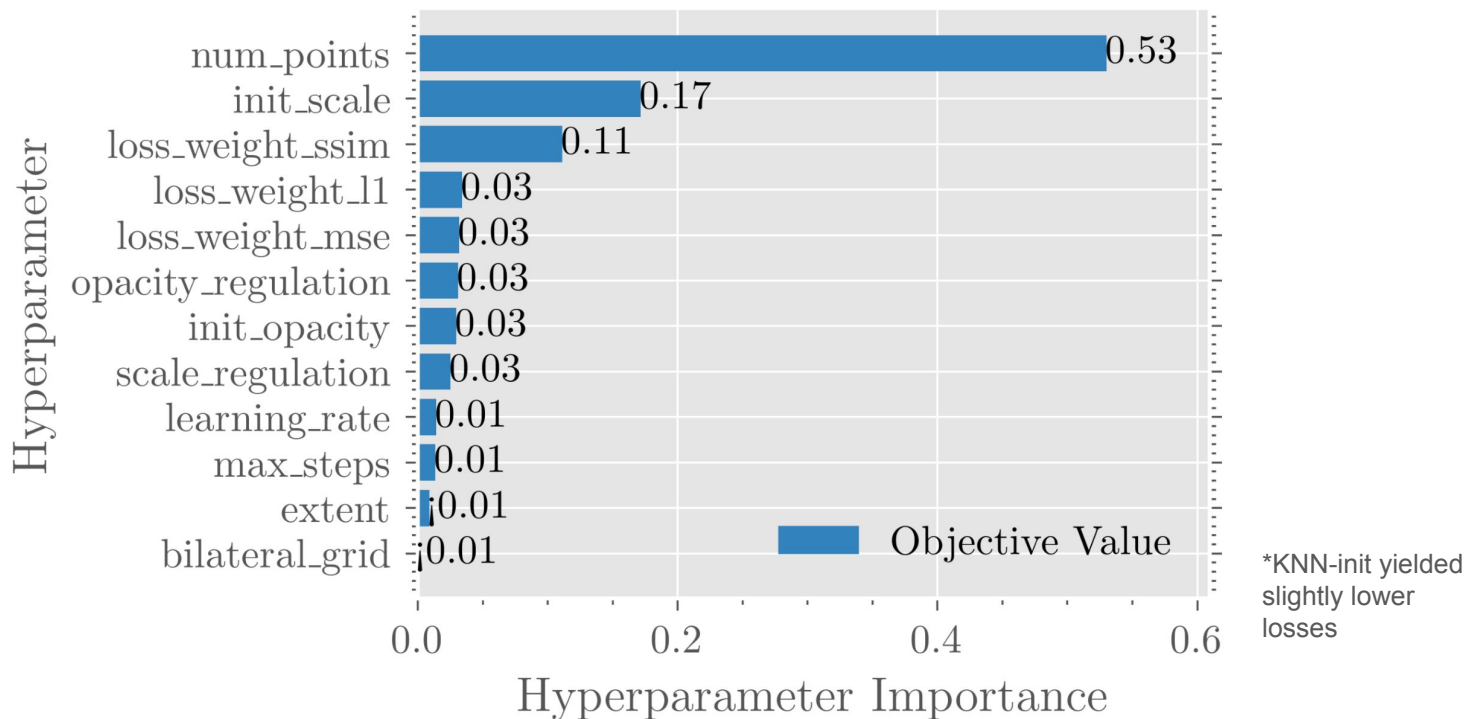


Results: Gaussian Splats - Hyperparameter Tuning

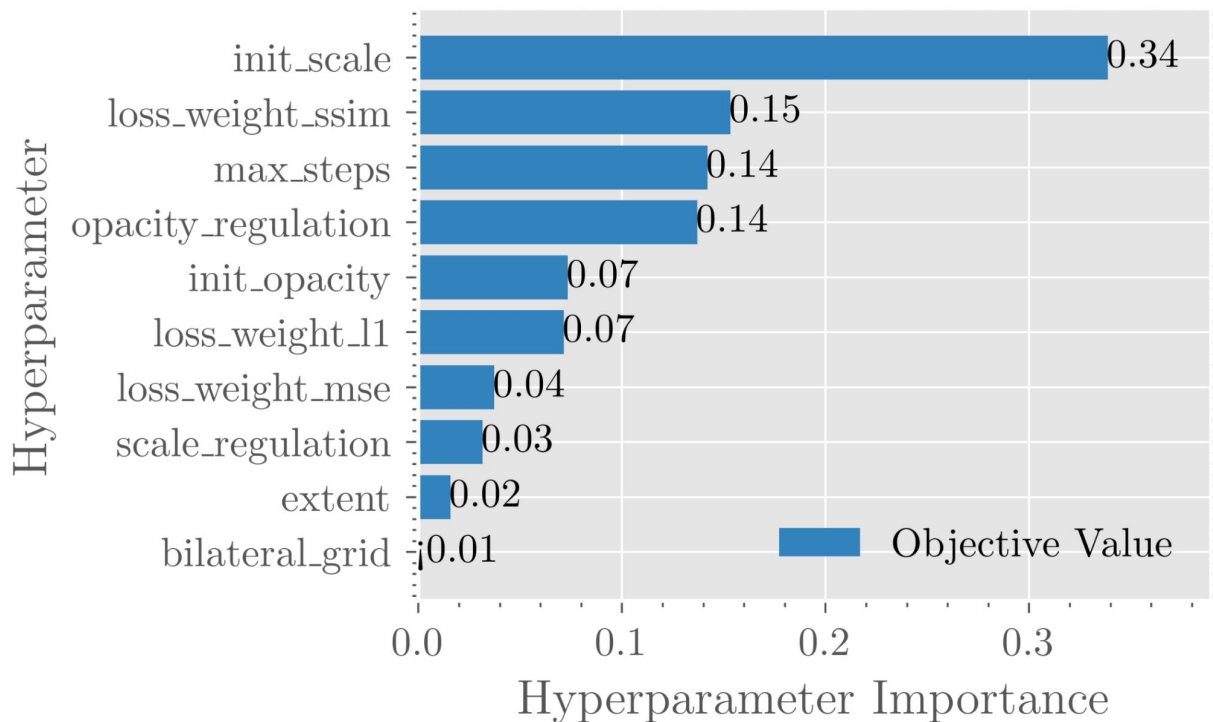
Contour Plot



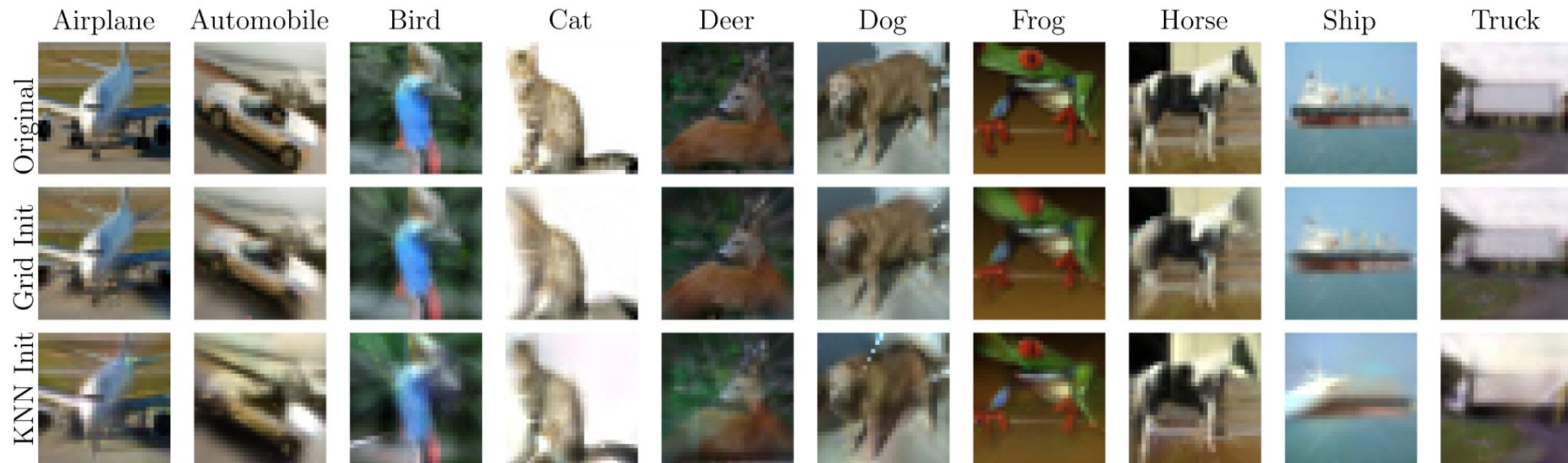
Results: Gaussian Splats - Hyperparameter Tuning



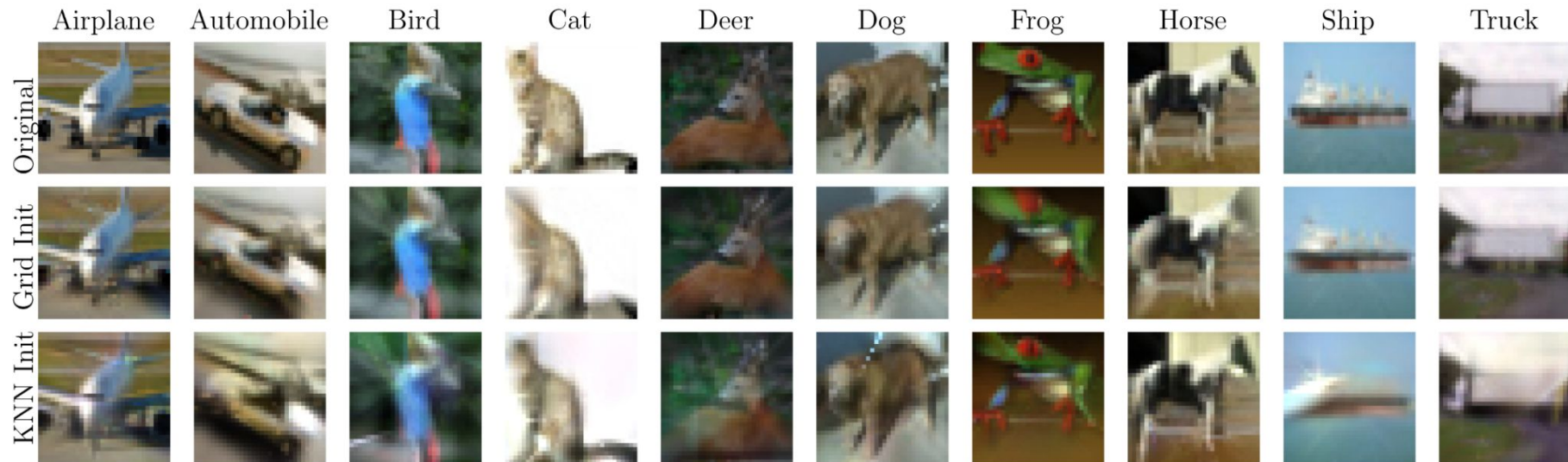
Results: Gaussian Splats - Hyperparameter Tuning



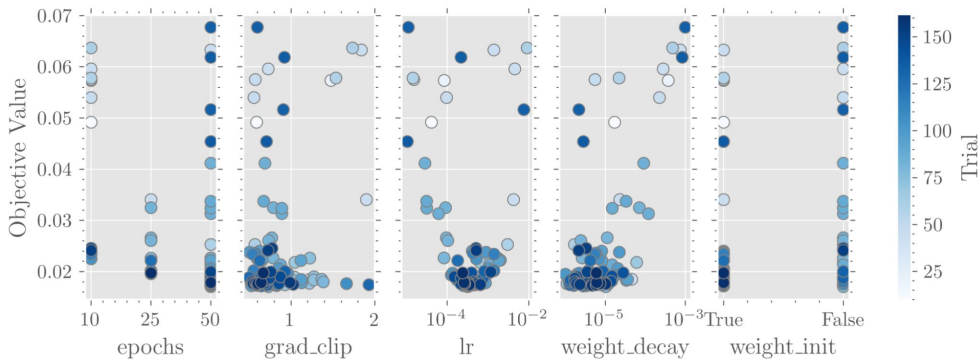
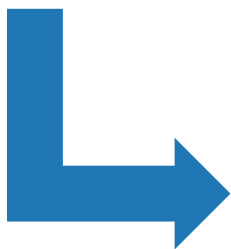
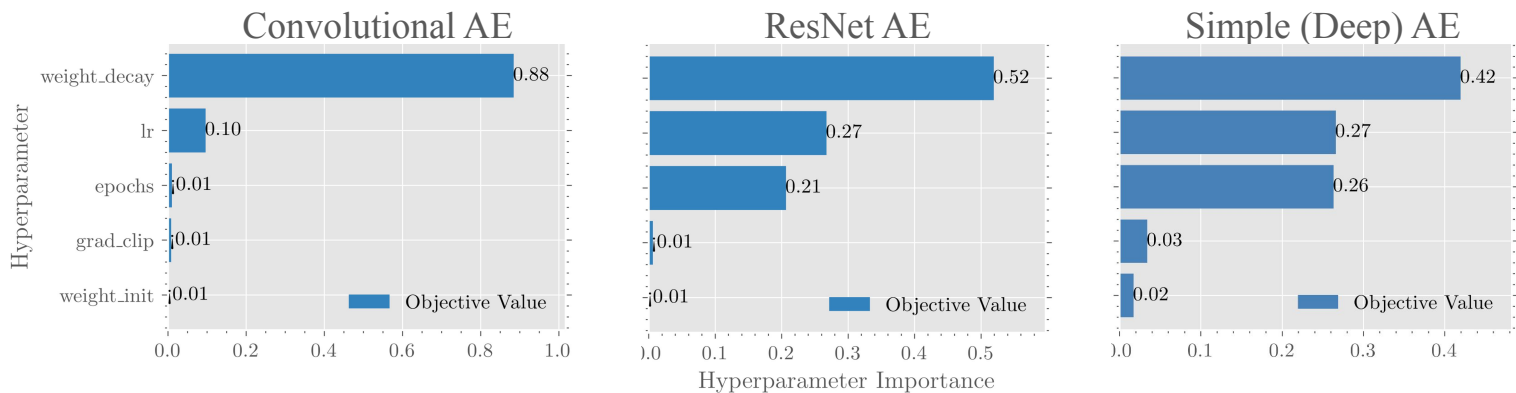
Results: Gaussian Splats - Full Dataset



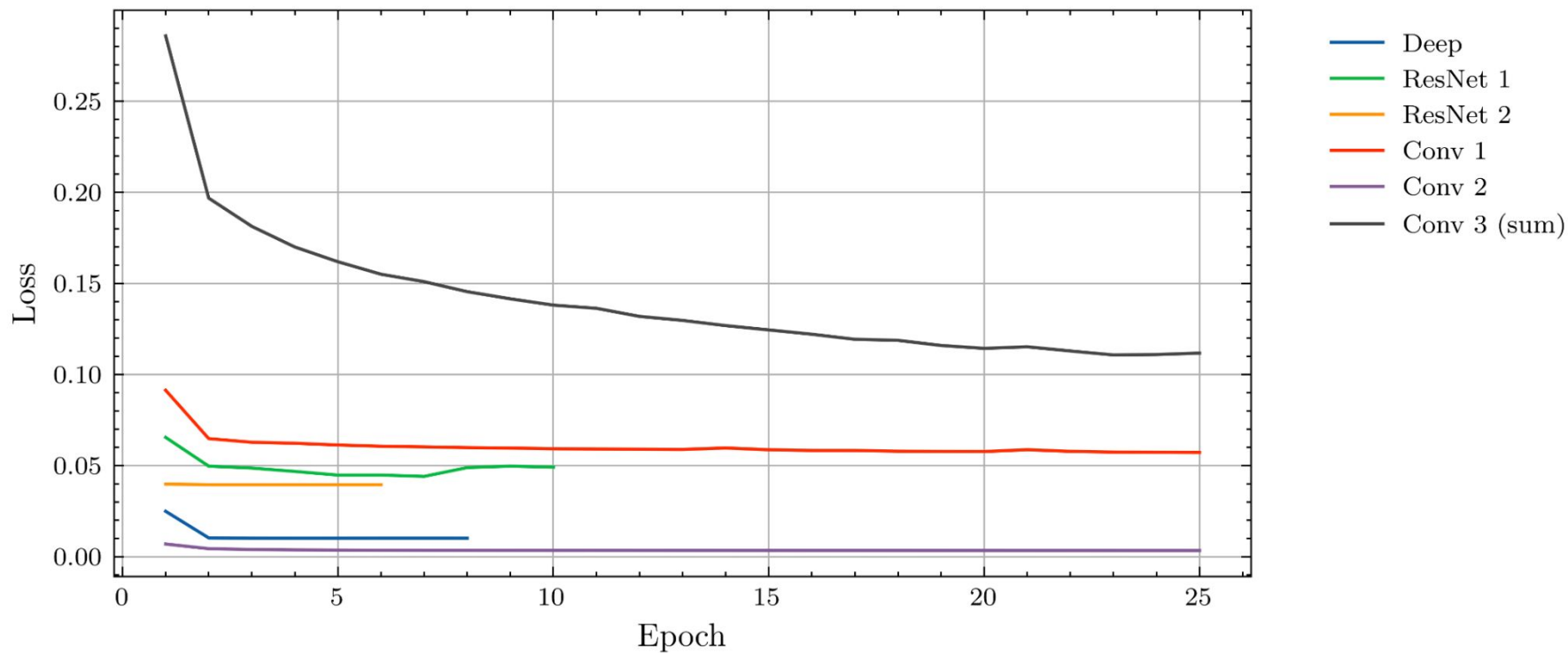
Results: Gaussian Splats - Full Dataset



Results: Autoencoding GS - Hyperparameter Tuning



Results: Autoencoding GS - Loss Evolution



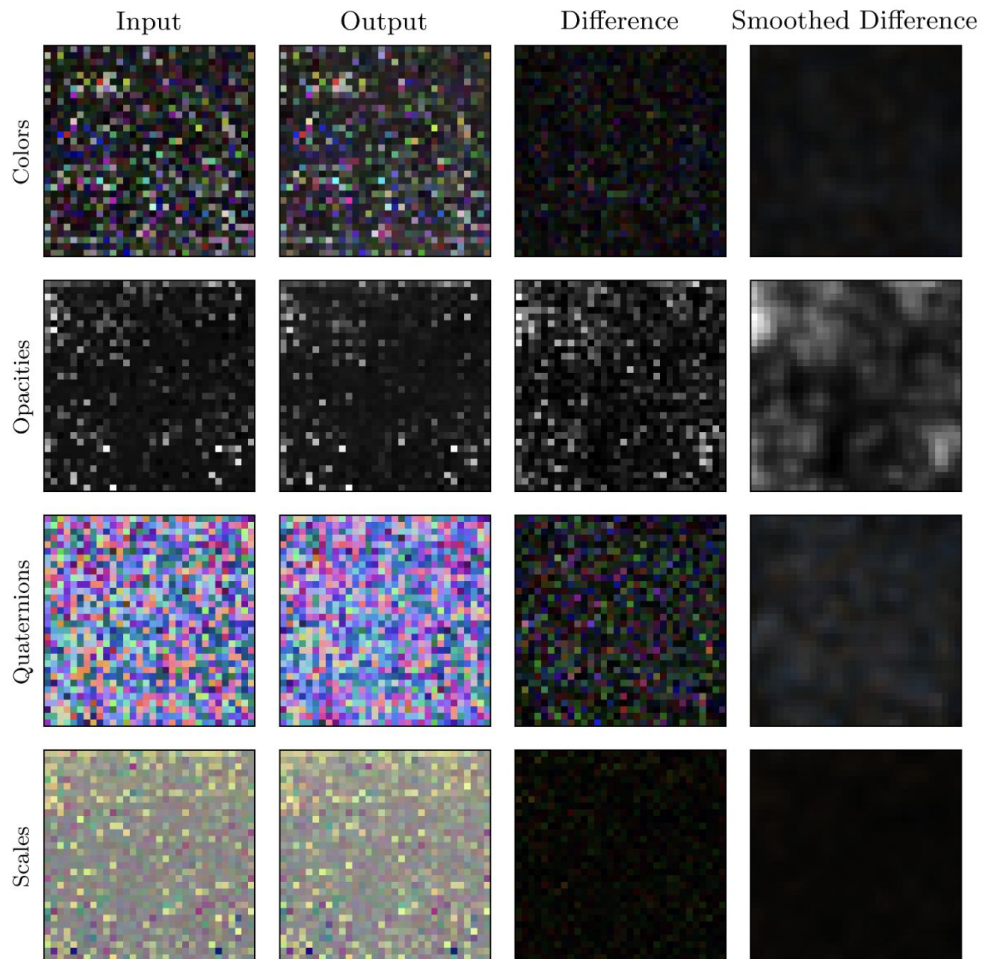
Results: Autoencoding GS - Reconstructed Images



Results: Autoencoding GS - Latent Space Dimension

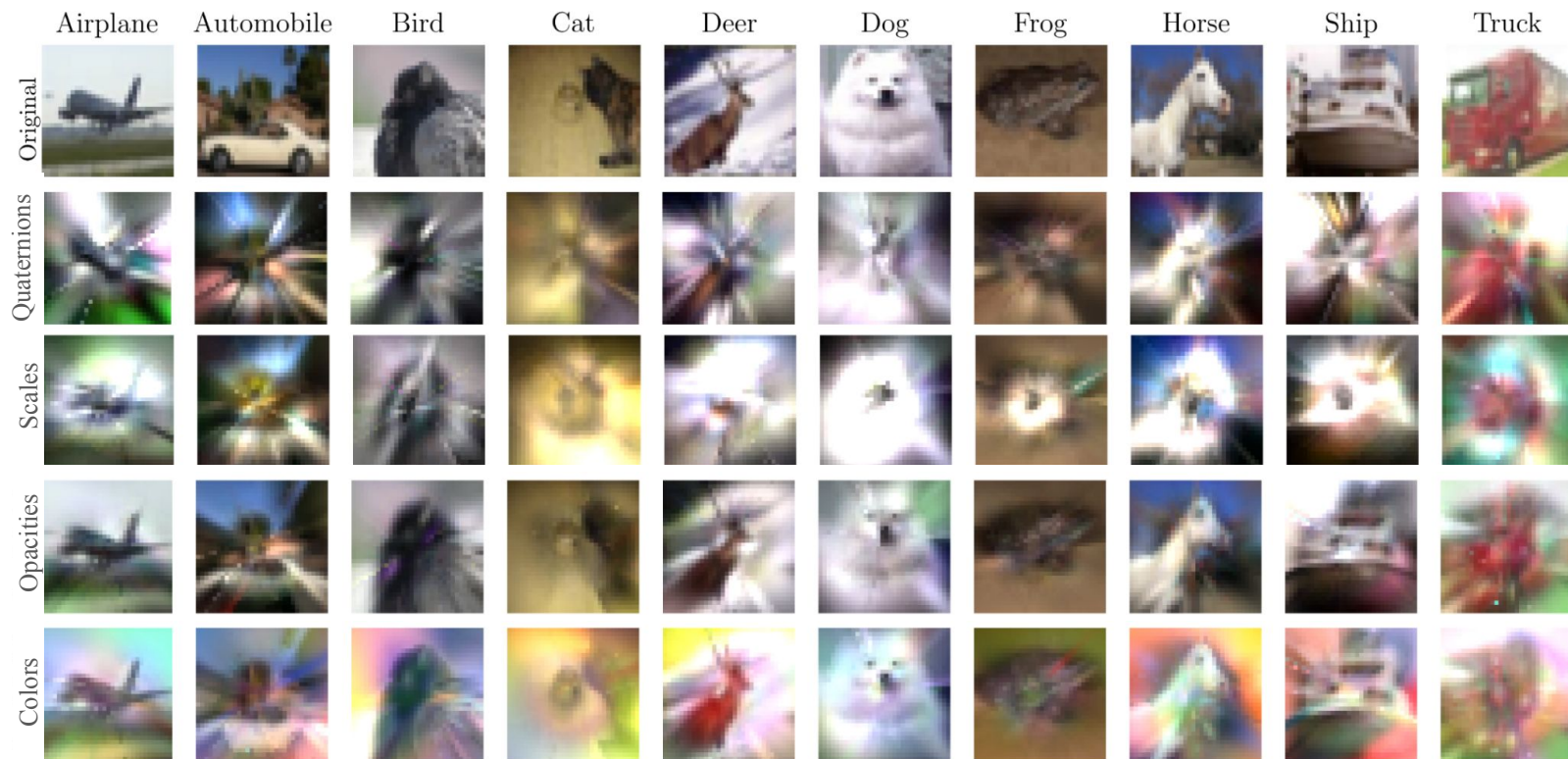


Individual Parameters



* $\sigma = 1$ for the
smoothed difference
** quaternions were
converted into a
rotation matrix

Results: Autoencoding GS - Individual Parameters



Discussion: Interpreting the Results - Insights and Shortcomings

- **Gaussian Splatting:** highly accurate reconstruction of the CIFAR-10 dataset
- **Autoencoding:** small losses + similar visualizations for all parameters
 - ⚡ deficient output quality across all AE variants
- **Possible causes:**
 - Small variations in splat representation might cause large variation in reconstructed image
 - Issues with reconstruction
 - Latent space might be too small (size = 16 for the best model)
- Unlikely that the model is incapable of

Discussion: Future Work

- **Improved loss function:** include image reconstruction fidelity
- Further exploration of the impact of the latent space size
- Implementation of more complex architectures, such as HiP
- Generative modelling on the latent space
- Gaussian splat generation
- Exploration of the latent space

Methodology: Training Process

Model Architectures

- Simple (Deep) AE
- Convolutional AE
- ResNet-18 AE
- (Abstract AE)

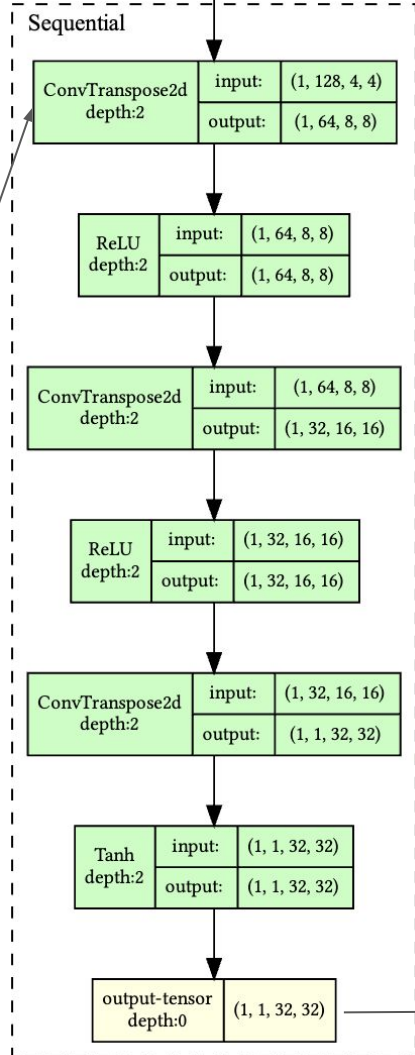
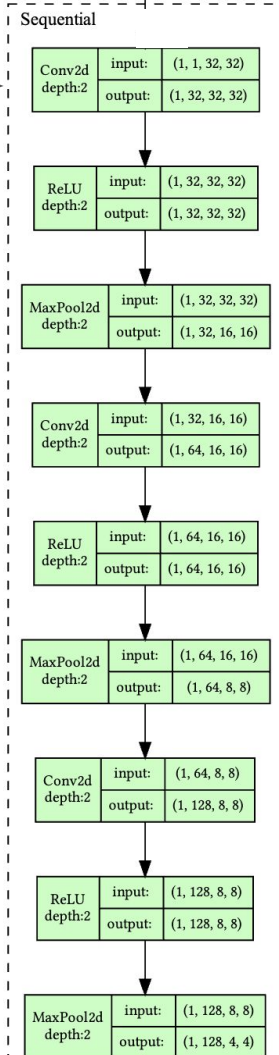
Splat Representations

- Vector-based encoding
- Full-image encoding
- Single-channel encoding
- Independent parameter models

Hyperparameter Tuning (Optuna)



Conv Model



ResNet Model

