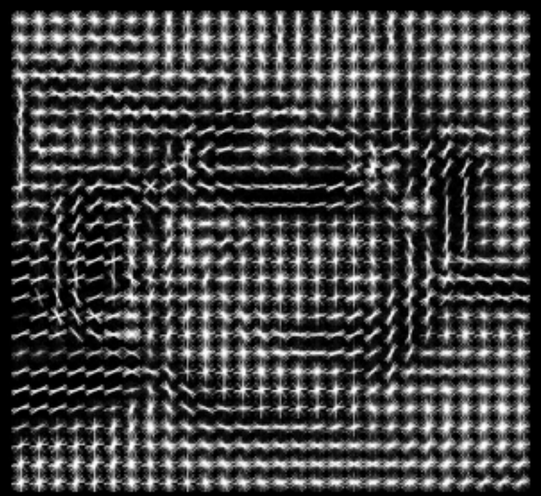
Learning a Predictable and Generative Representation for Objects

Rohit Girdhar, David Fouhey,
Mikel Rodriguez and Abhinav Gupta

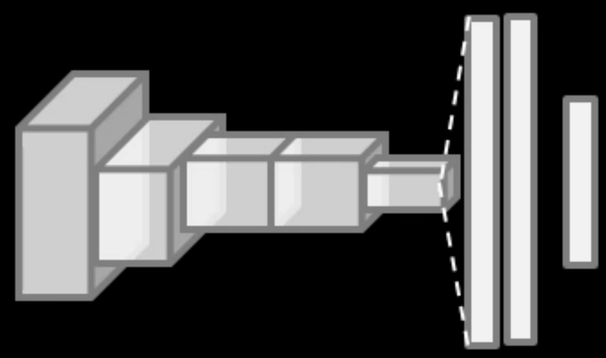
What is the Right Representation for Objects?

SIFT

HOG



CNN



- Good discriminative representation for object classification and retrieval
- However, does not model any 3D properties

What is the Right Representation for Objects?

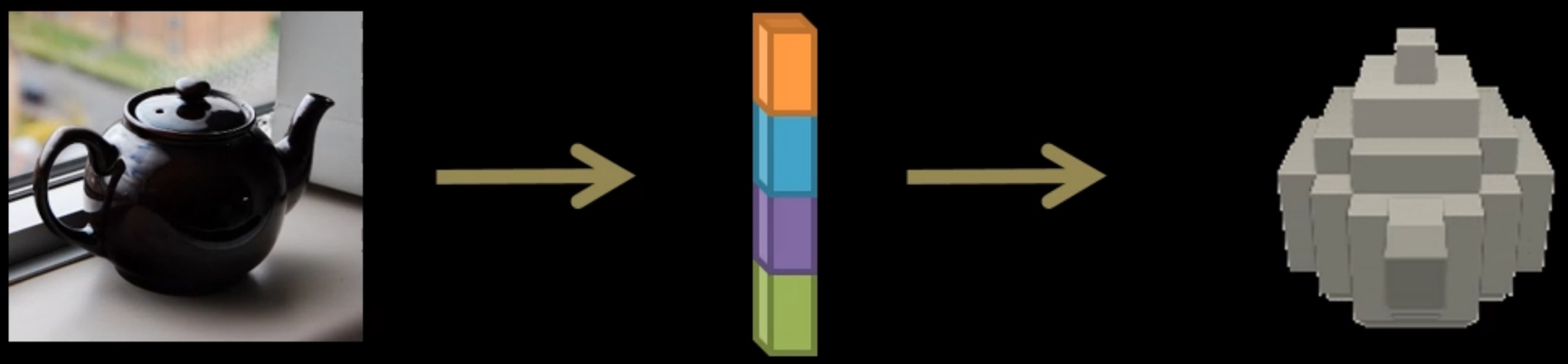
Generative Model
(such as DBN)

- Good 3D representation of objects
- But hard to estimate from a single 2D image

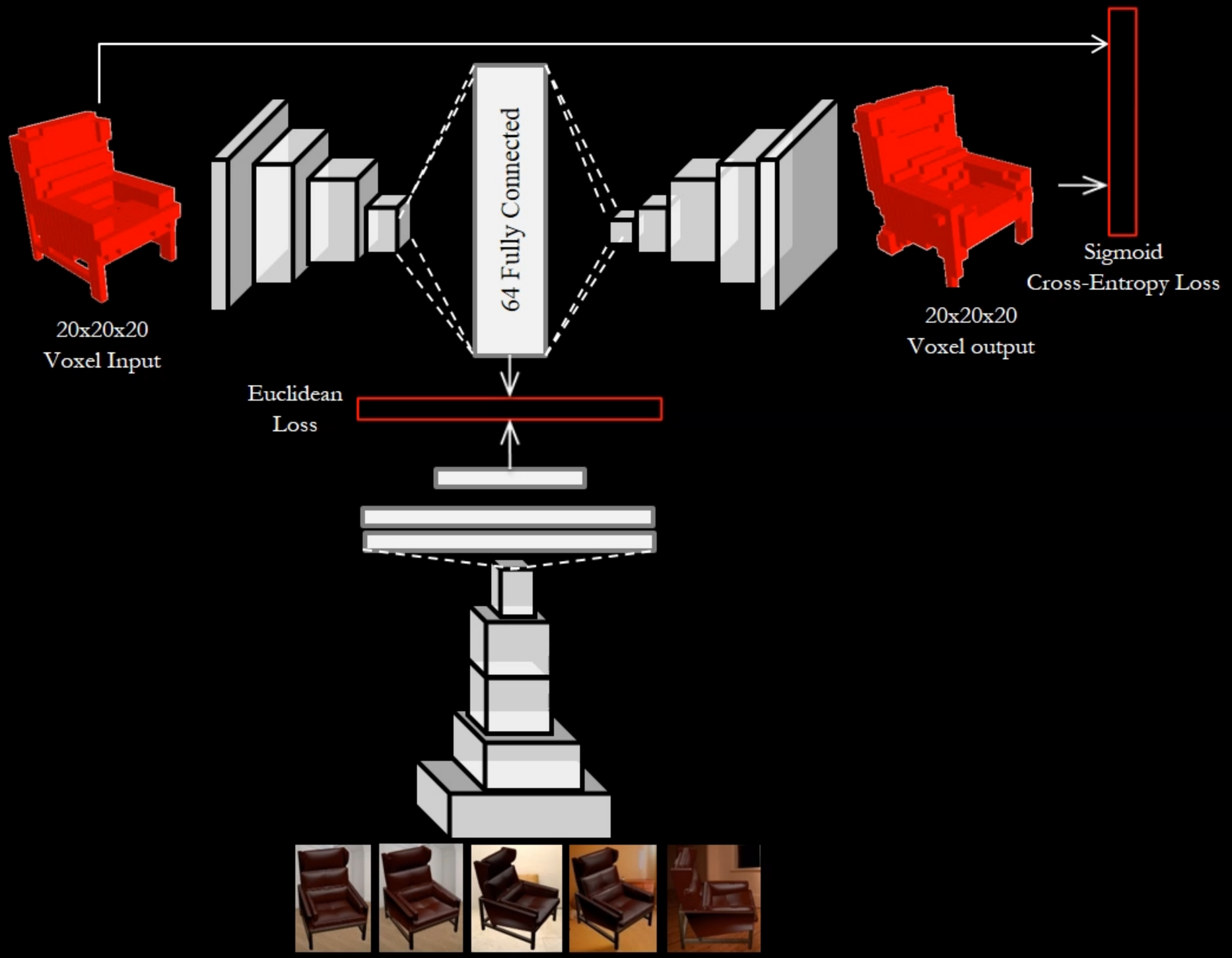
What is the Right Representation for Objects?

We believe object representation should satisfy both properties:

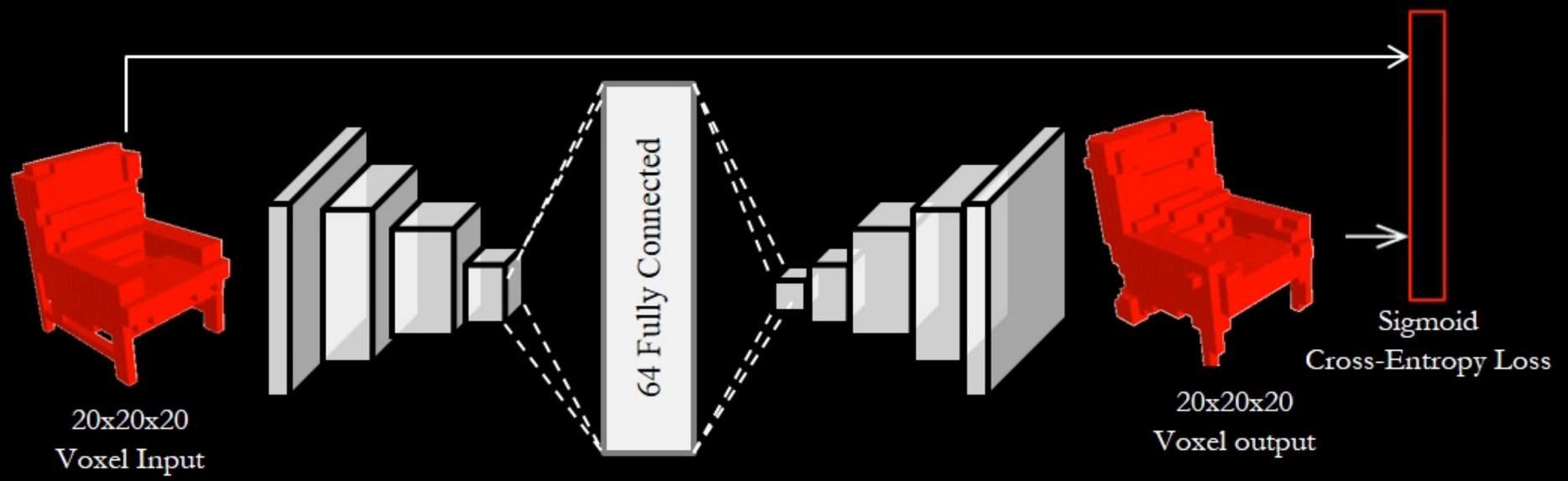
- **Generative in 3D**
- **Predictable from 2D**



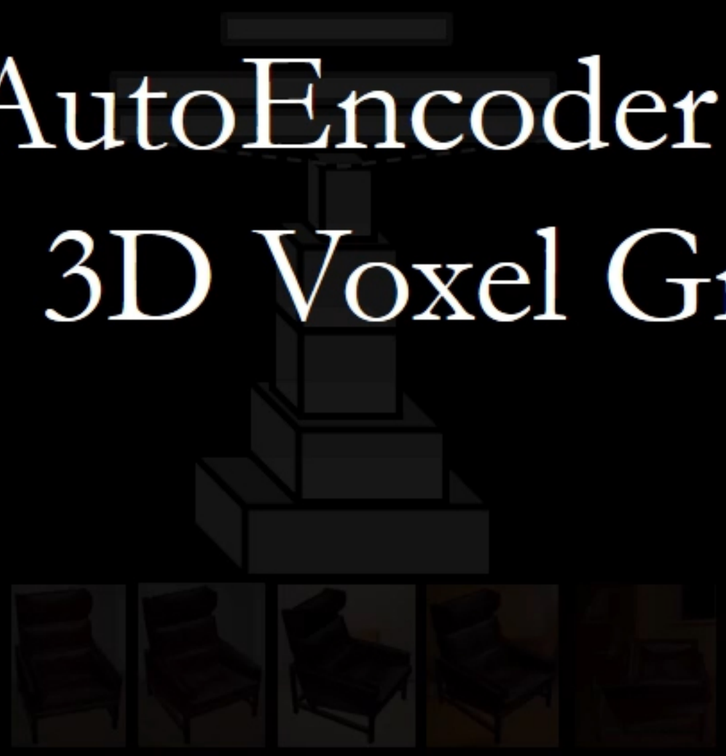
The T-L Embedding Architecture



The T-L Embedding Architecture

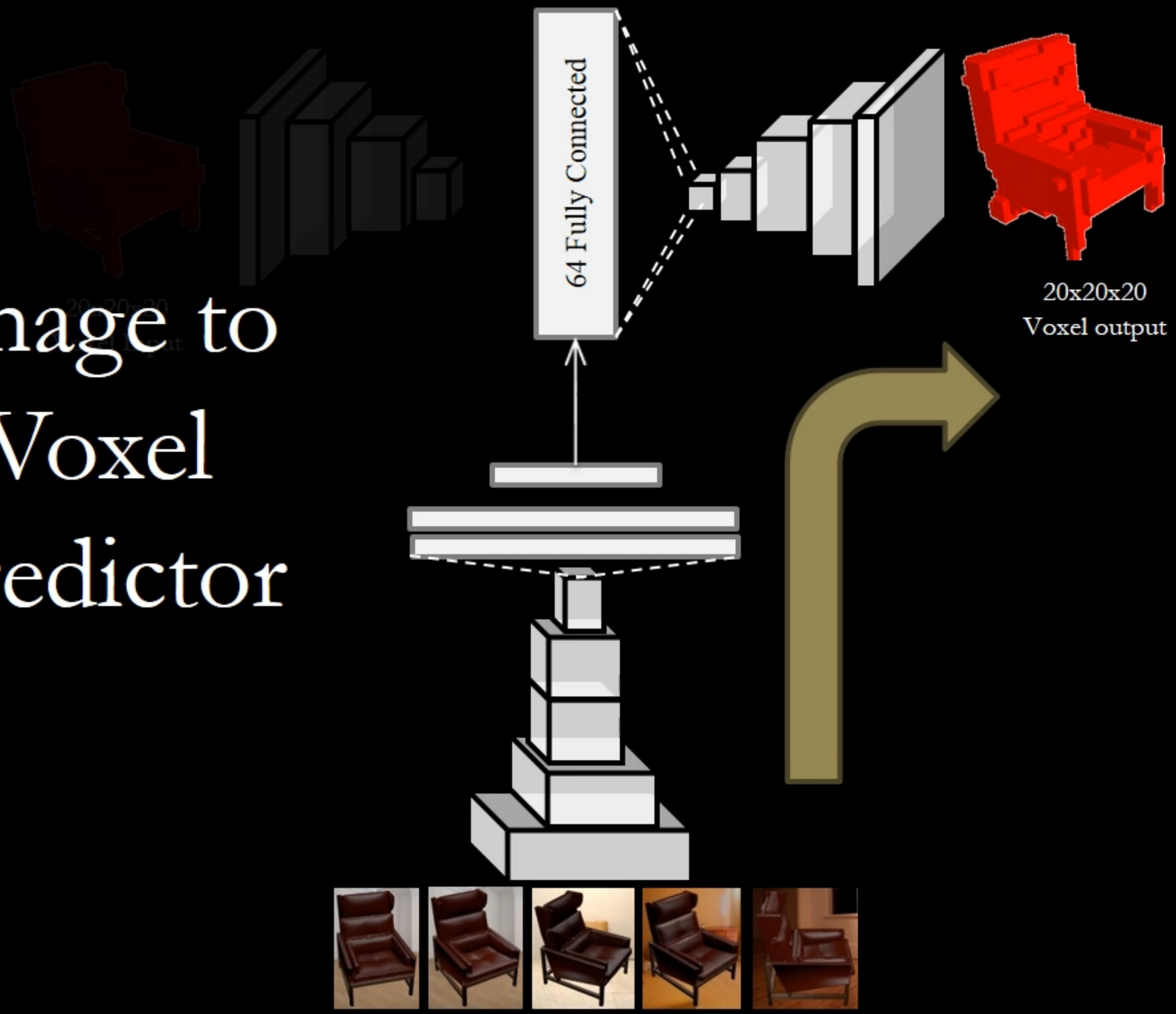


AutoEncoder over 3D Voxel Grids

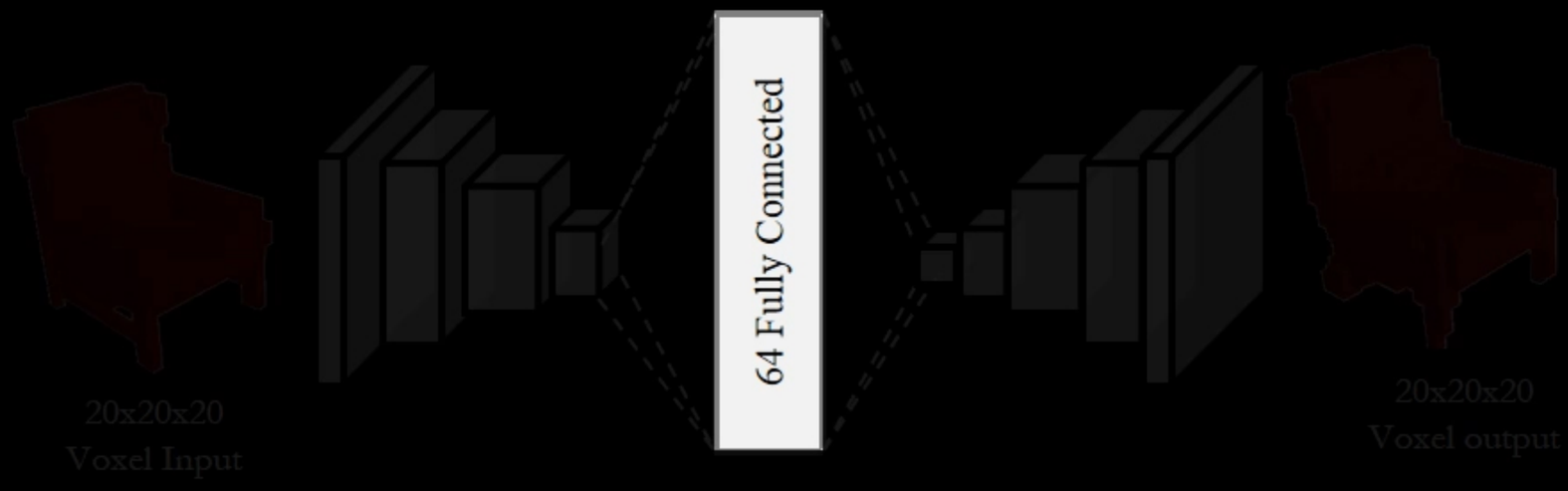


The T-L Embedding Architecture

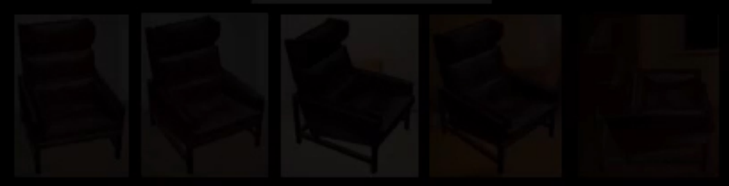
Image to Voxel Predictor



The T-L Embedding Architecture



3D Voxels and Images connected through a 64D latent representation

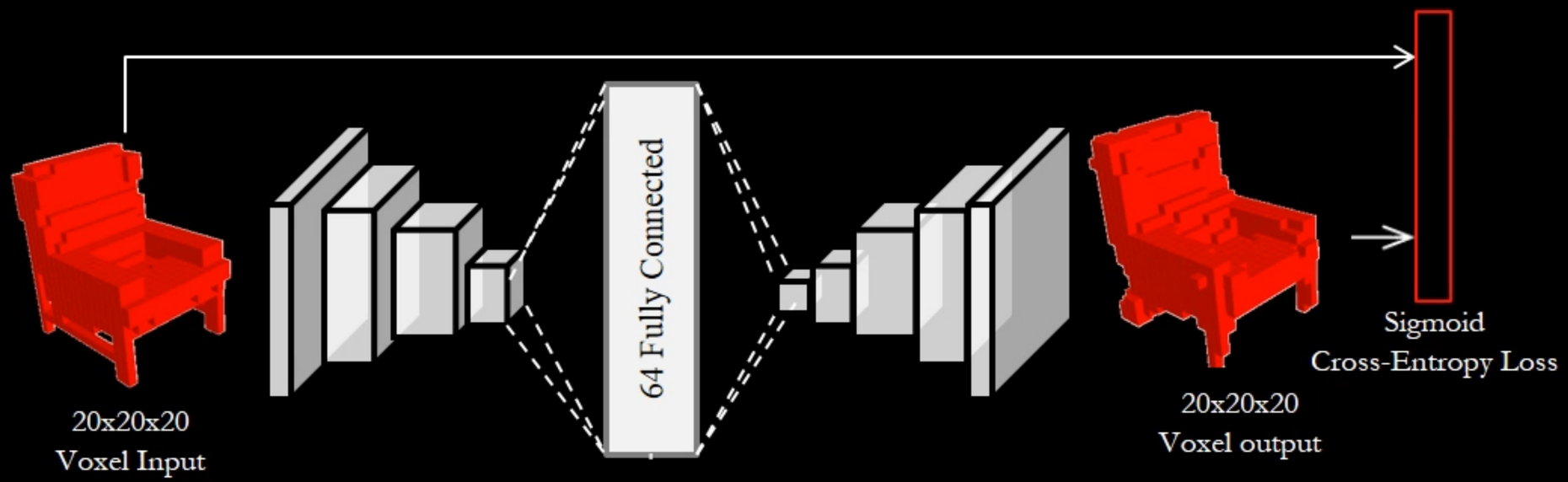


Training Data

- 5 Categories from ShapeNet (no category labels used in learning)
- Rendered onto random indoor backgrounds

Data from Stanford ShapeNet Dataset

Training the TL-Embedding



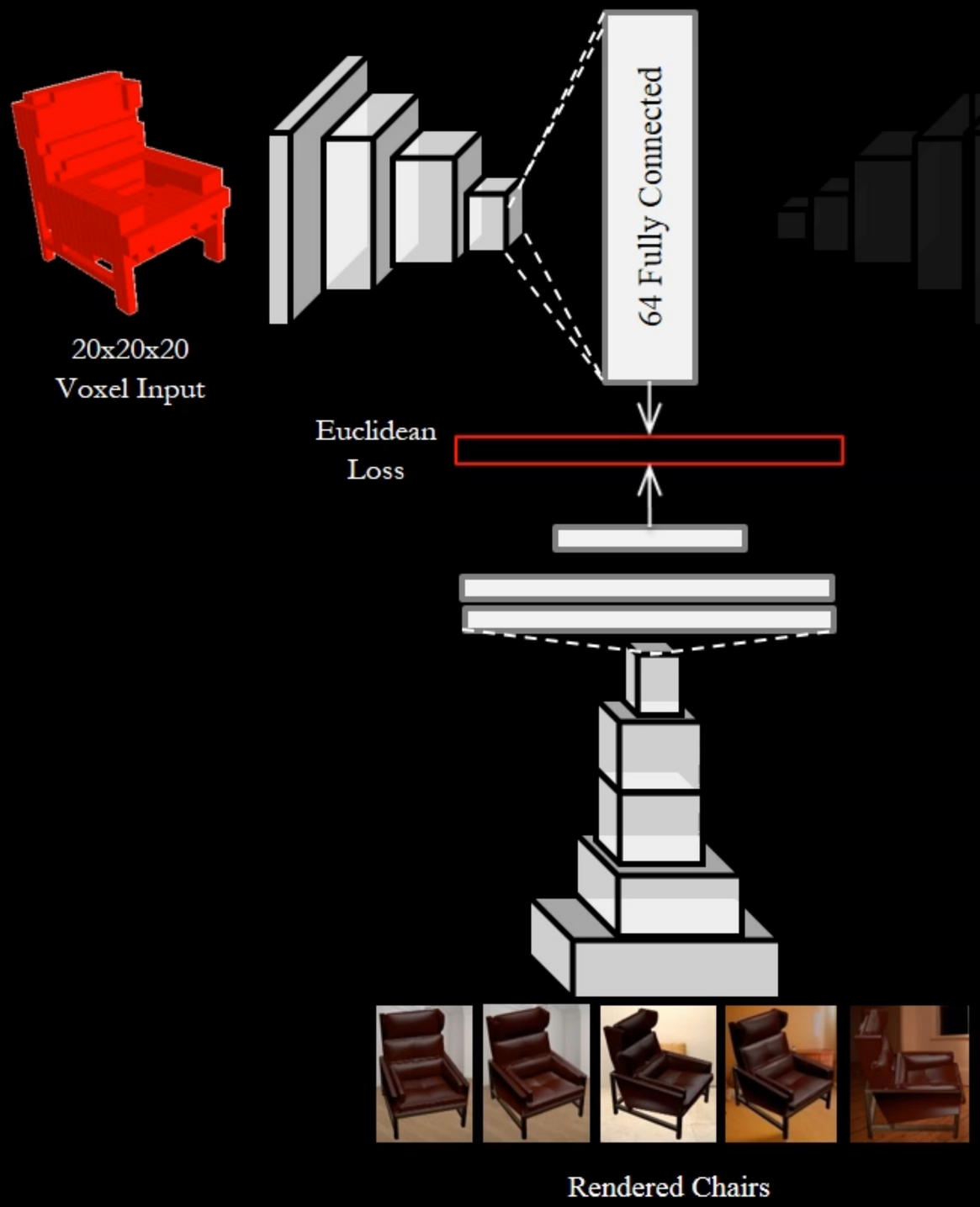
Euclidean
Loss

Train
(T-Network)

Rendered Chairs

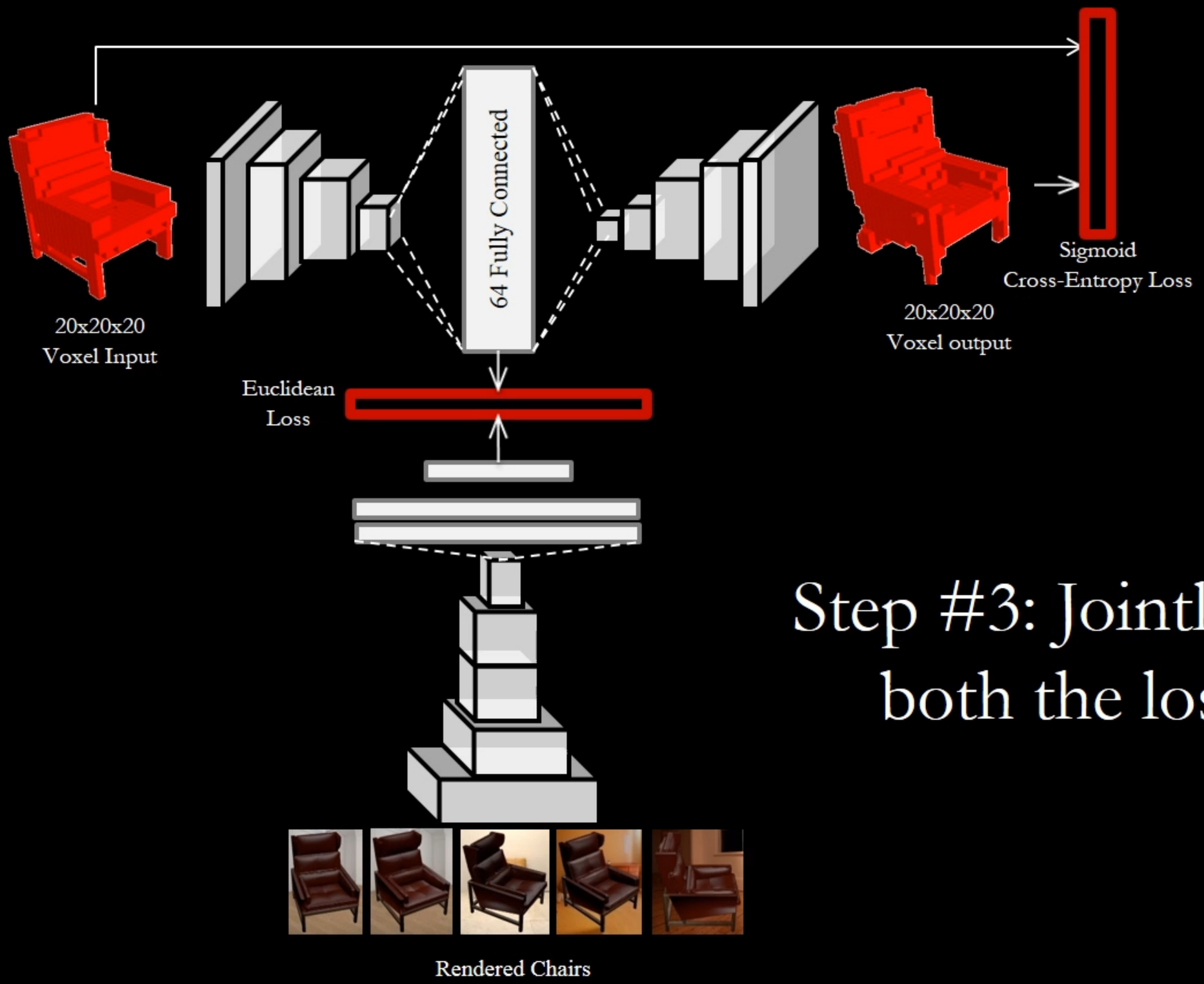
Step #1: Train the **Auto-Encoder**

Training the TL-Embedding



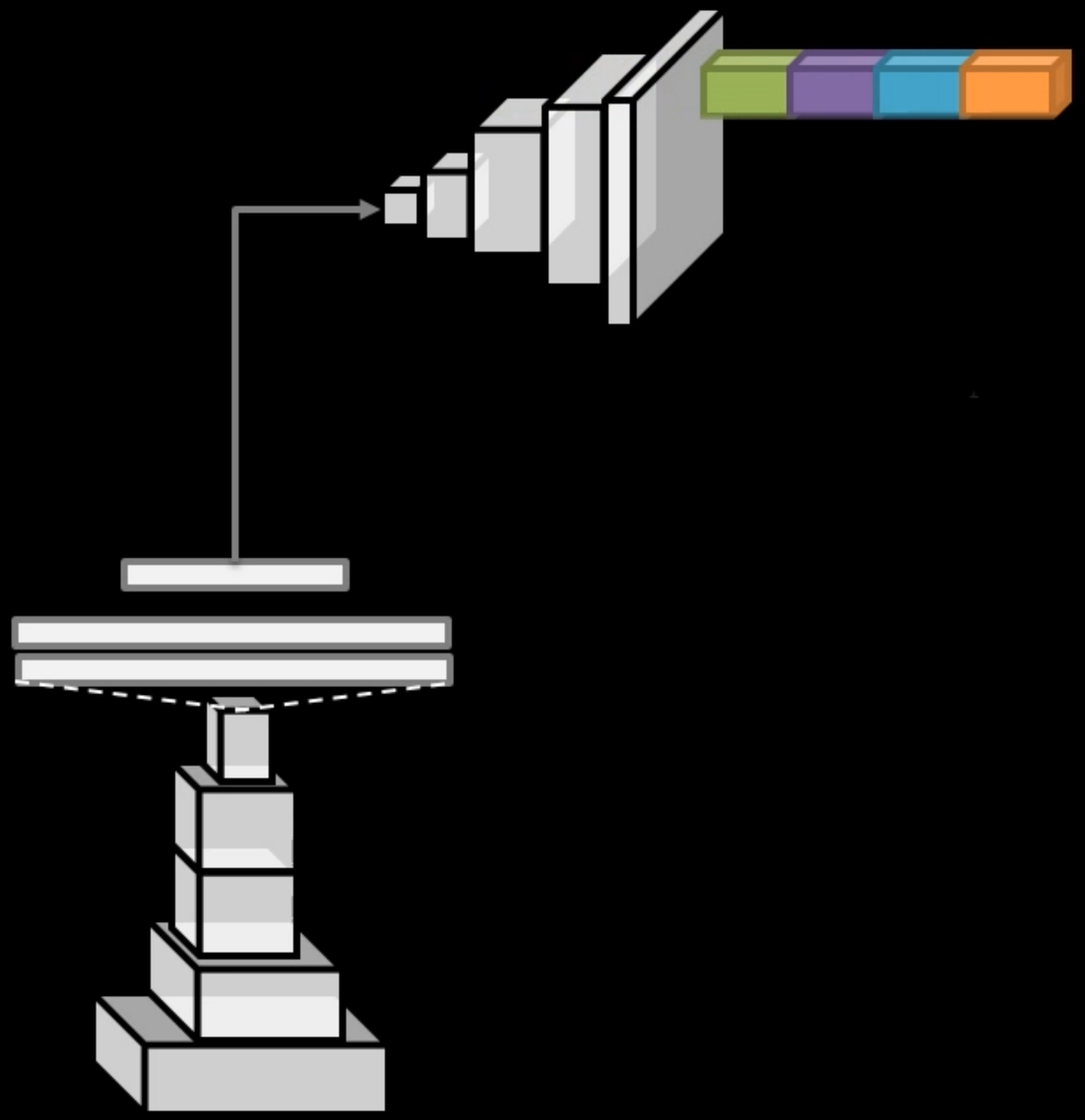
Step #2: Train the **Image-Network** to regress for Auto-Encoder Coefficients

Training the TL-Embedding

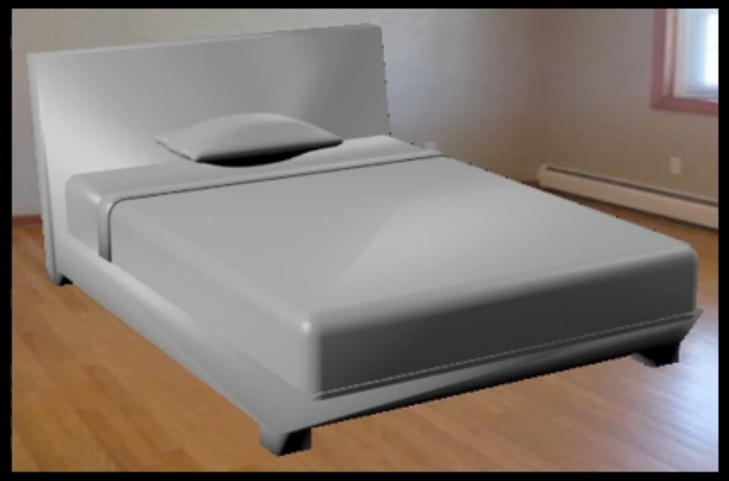


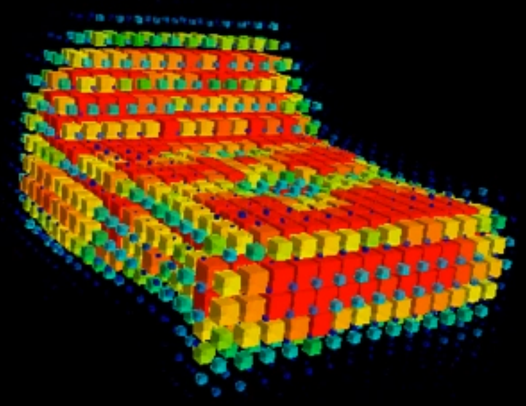
Step #3: Jointly, with both the losses

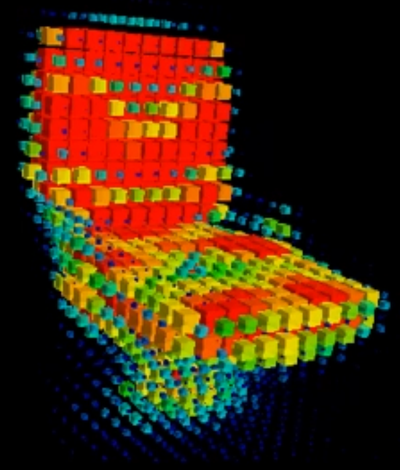
Testing the TL-Embedding

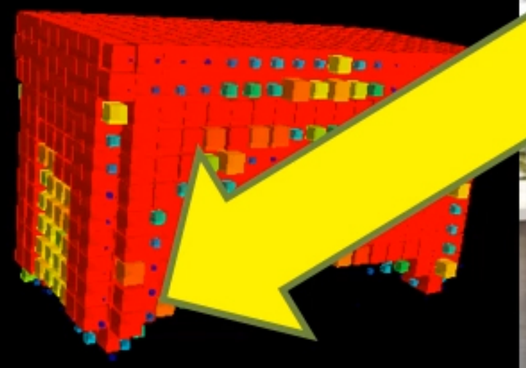


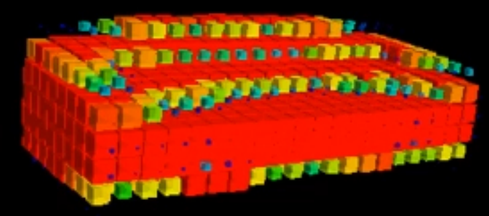
Reconstructing Test Models

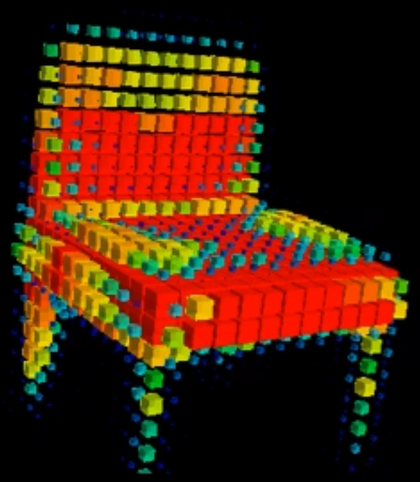


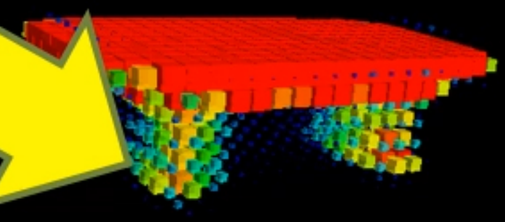




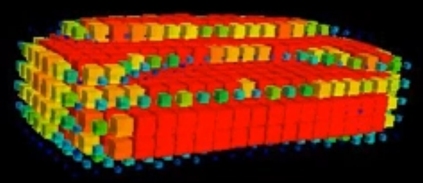


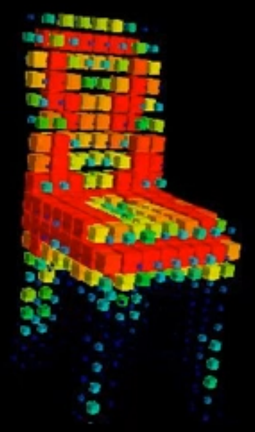


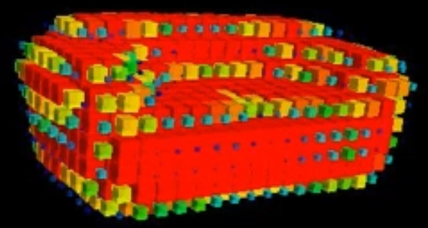


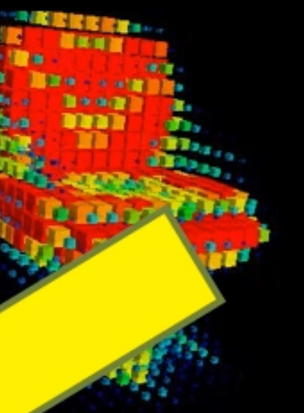


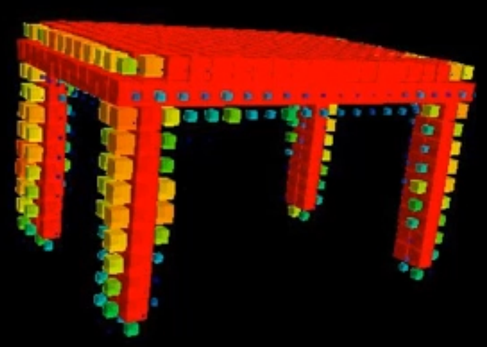
Reconstructing IKEA

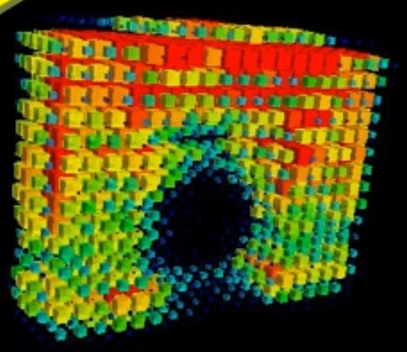









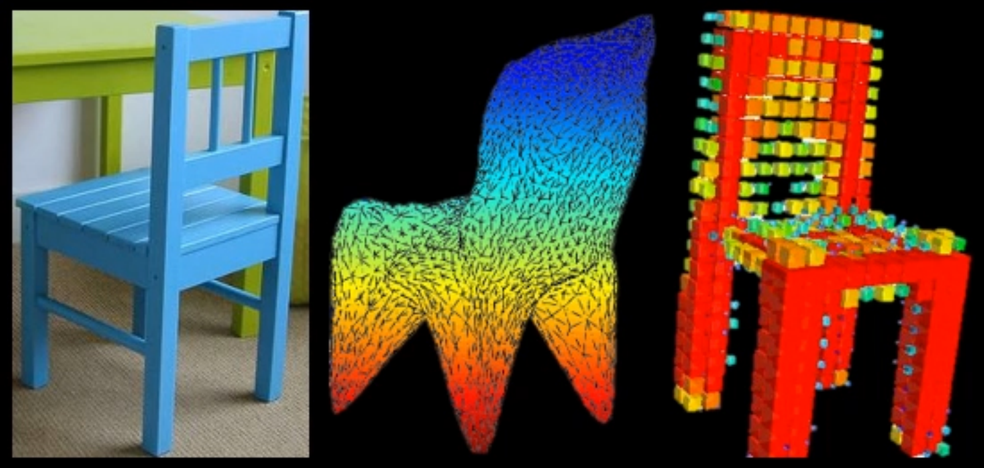




Data from Lim *et al.* Parsing ikea objects: Fine pose estimation (ICCV '13)

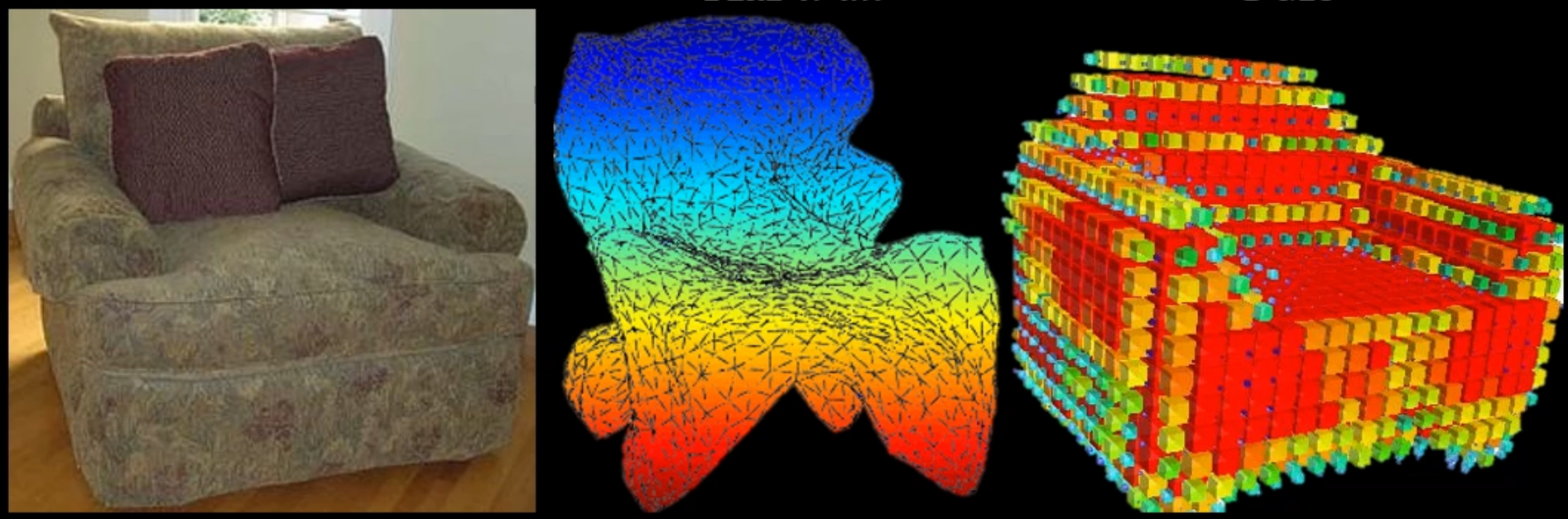
Compared to the Kar *et al.* (CVPR'15)

Kar *et al.* Ours



Kar *et al.* Ours

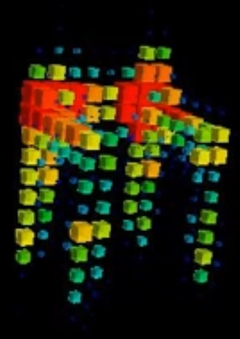
Kar *et al.* Ours

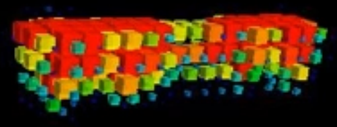


A. Kar, S. Tulsiani, J. Carreira and J. Malik. Category-Specific Object Reconstruction from a Single Image (CVPR'15)
Data from Xiang *et al.* Beyond pascal: A benchmark for 3d object detection in the wild. (WACV'14)

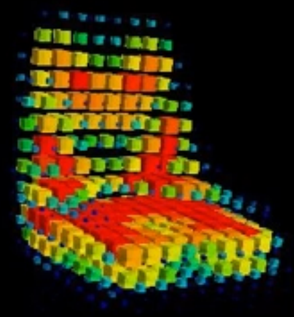
Common Failure Cases

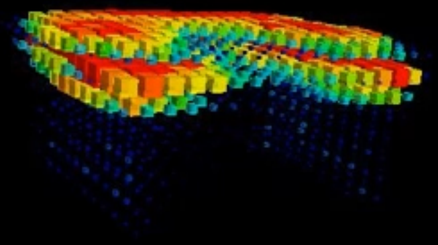
Truncation, occlusions and multiple objects



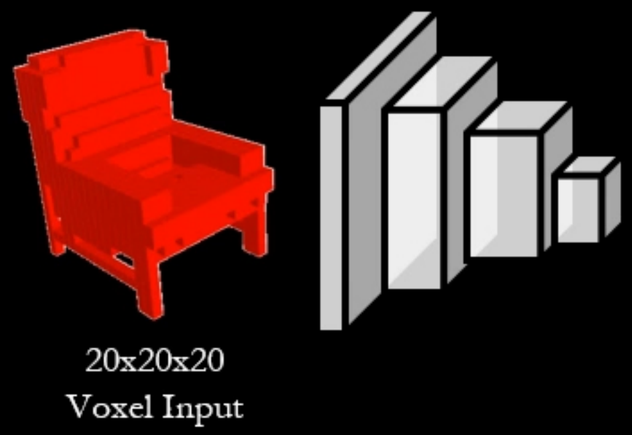


Non-canonical viewpoints

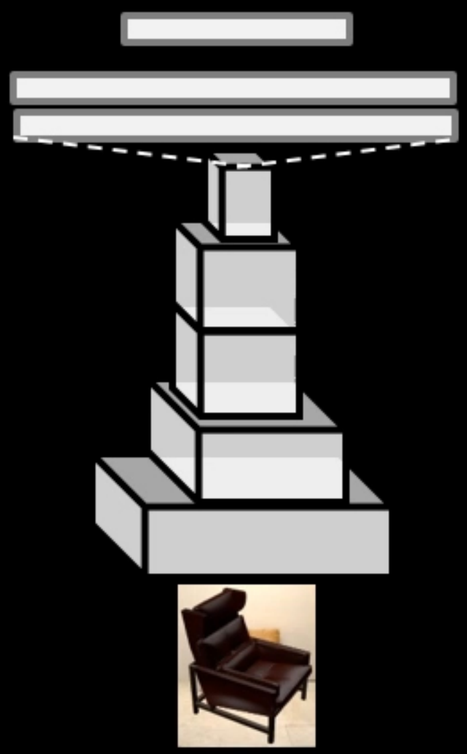




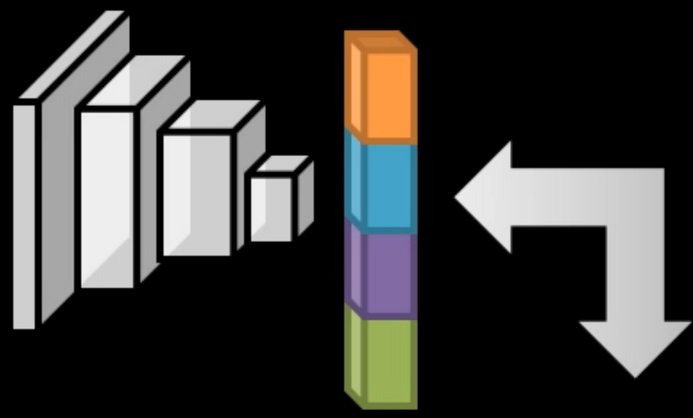
Application: 3D Model Retrieval



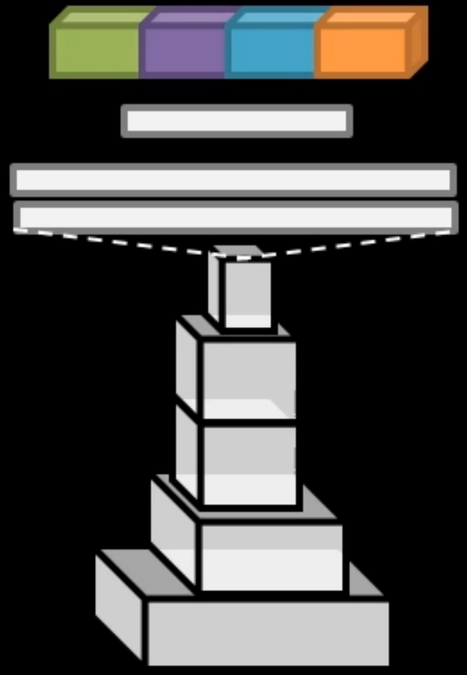
20x20x20
Voxel Input



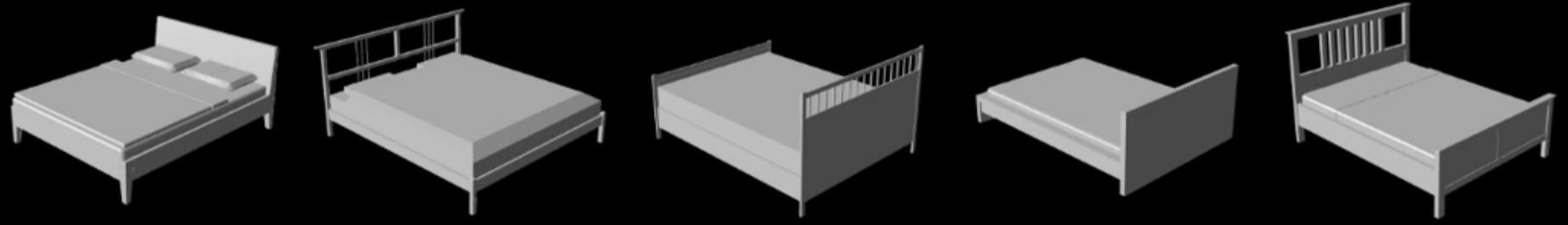
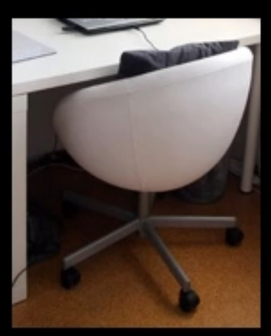
Application: 3D Model Retrieval



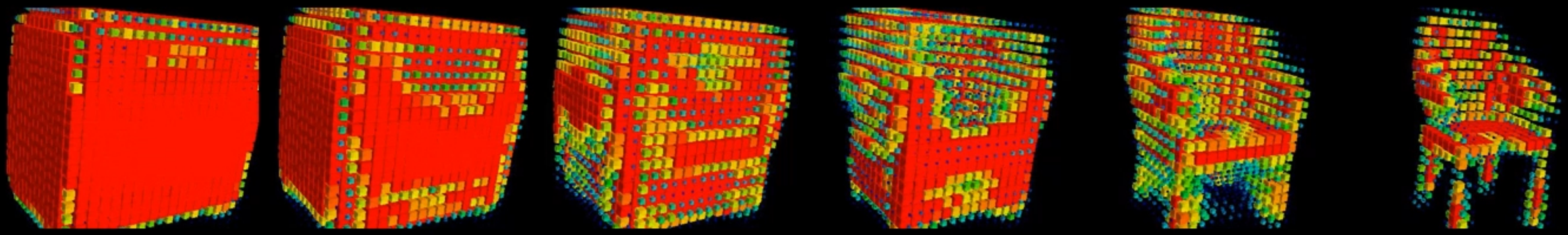
Nearest Neighbor Search

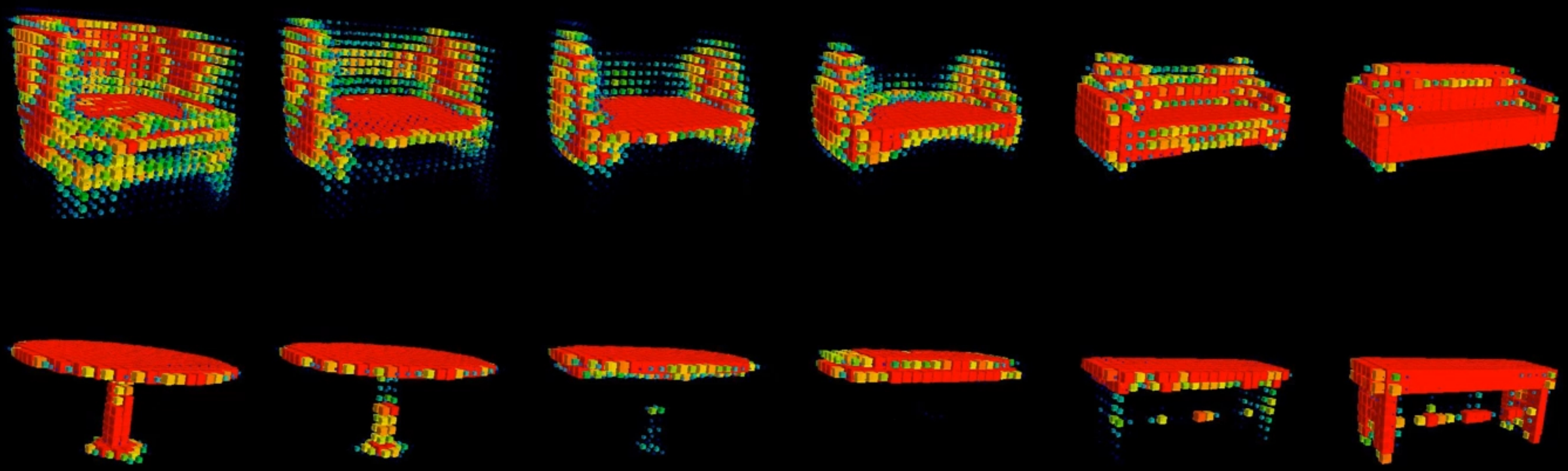


Results: Fast 3D Model Retrieval

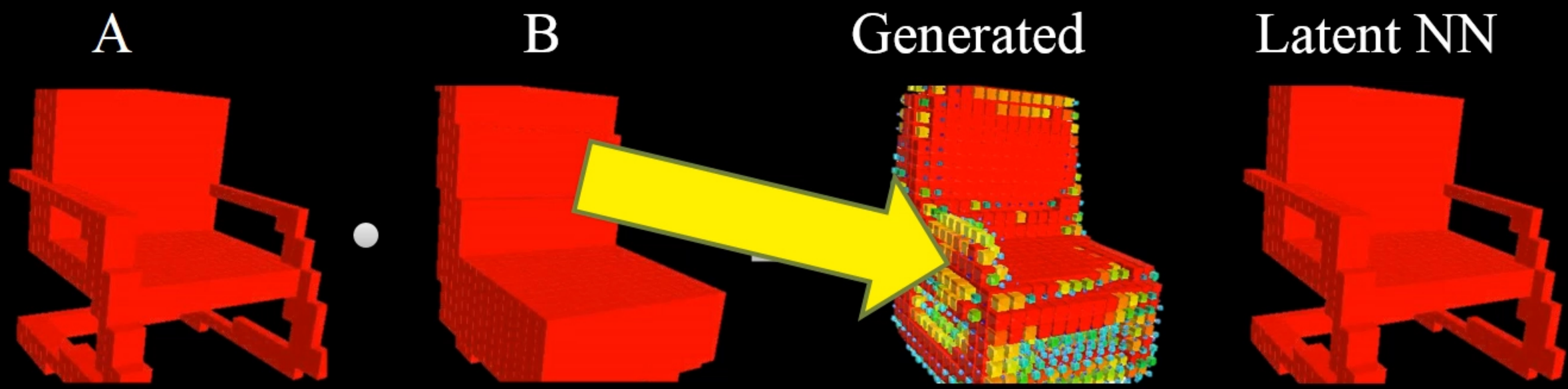





A Latent Space that is smooth…





… generative



… and allows for 3D arithmetic!

Code and Models available online!

rohitgirdhar / **GenerativePredictableVoxels**

Watch 5 | Star 19 | Fork 11

Code | Issues 0 | Pull requests 0 | Projects 0 | Pulse | Graphs

Code release for ECCV16 paper "Learning a Predictable and Generative Vector Representation for Objects"
https://rohitgirdhar.github.io/GenerativePredictableVoxels/

7 commits | 2 branches | 0 releases | 1 contributor

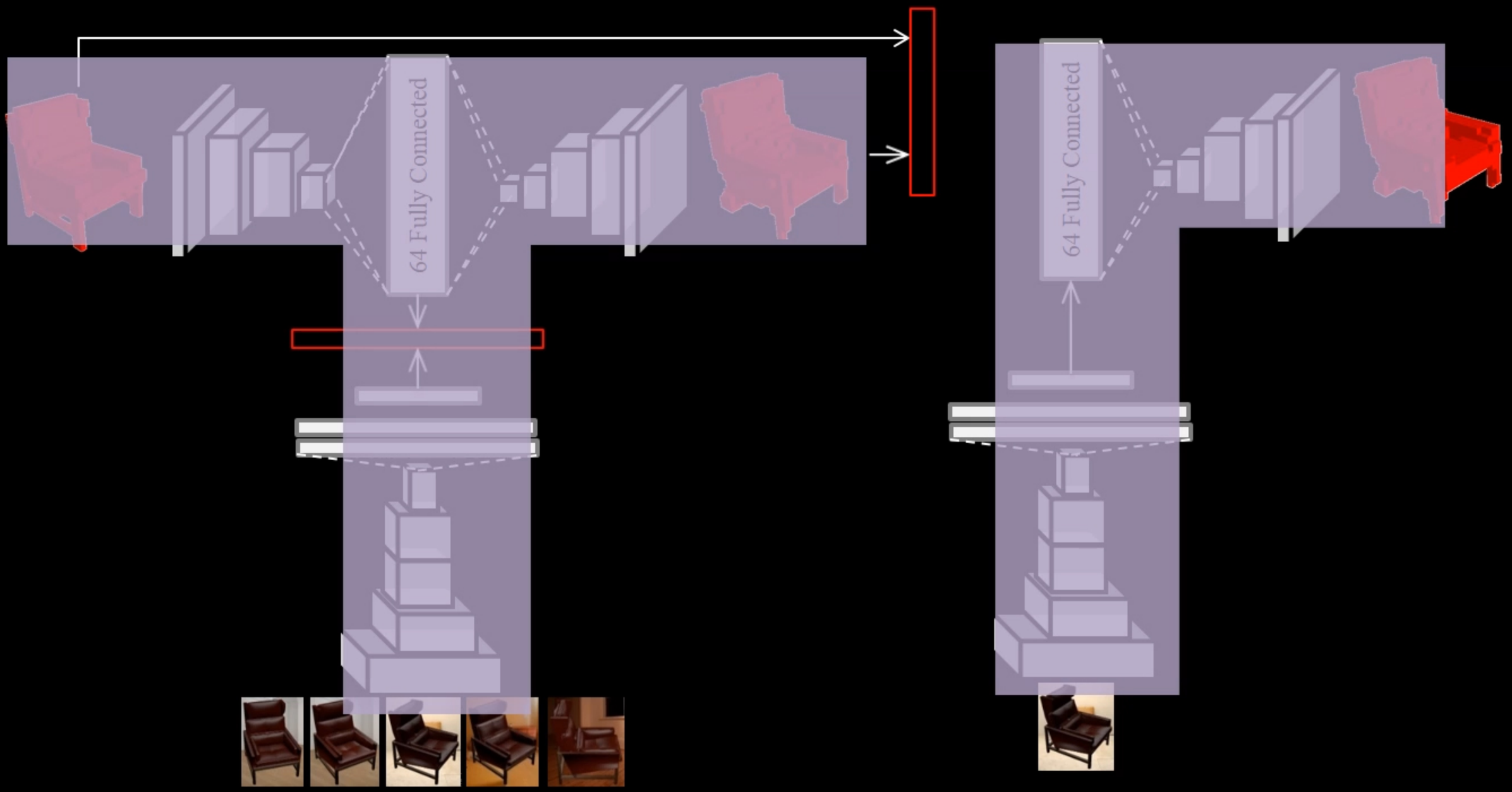
Branch: master | New pull request | Find file | Clone or download

rohitgirdhar Merge branch 'master' of github.com:rohitgirdhar/GenerativePredictabl... | Latest commit d4f4254 a day ago

data	testing output	21 days ago
dataset	splits	a day ago
models	testing output	21 days ago
src	readme for testing	21 days ago
README.md	bibtex	11 days ago

git.io/viym3

Thank You!



Visit us at Poster S-4A-09