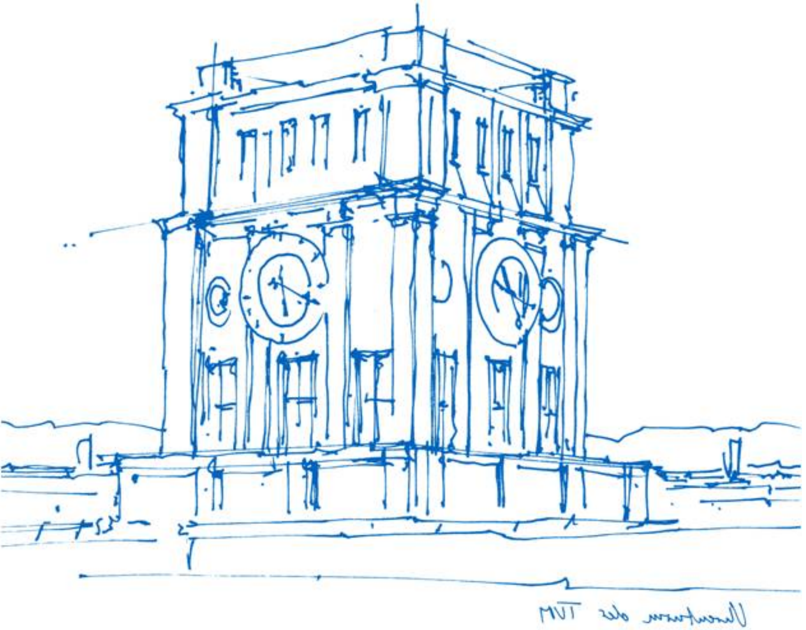
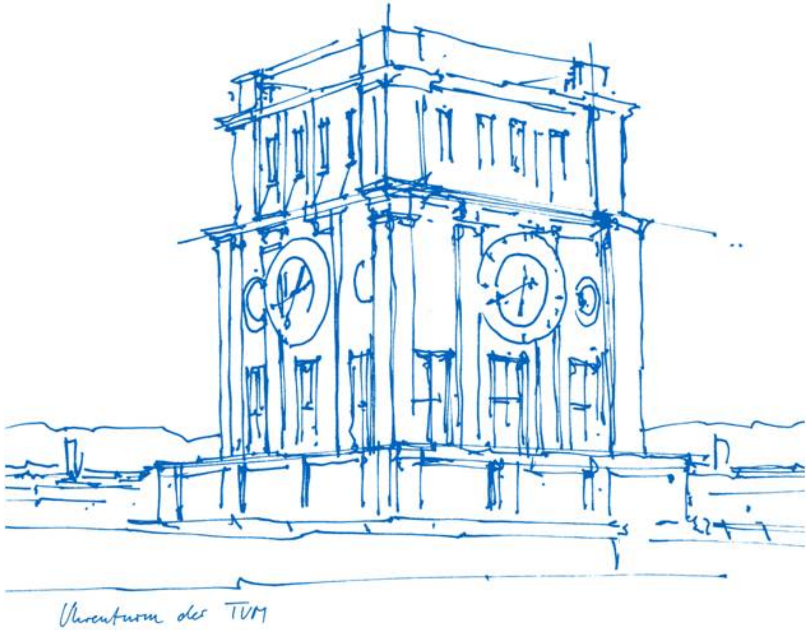
Master’s Thesis

# 3D Shape Completion from Sparse Point Clouds

Christian Diller



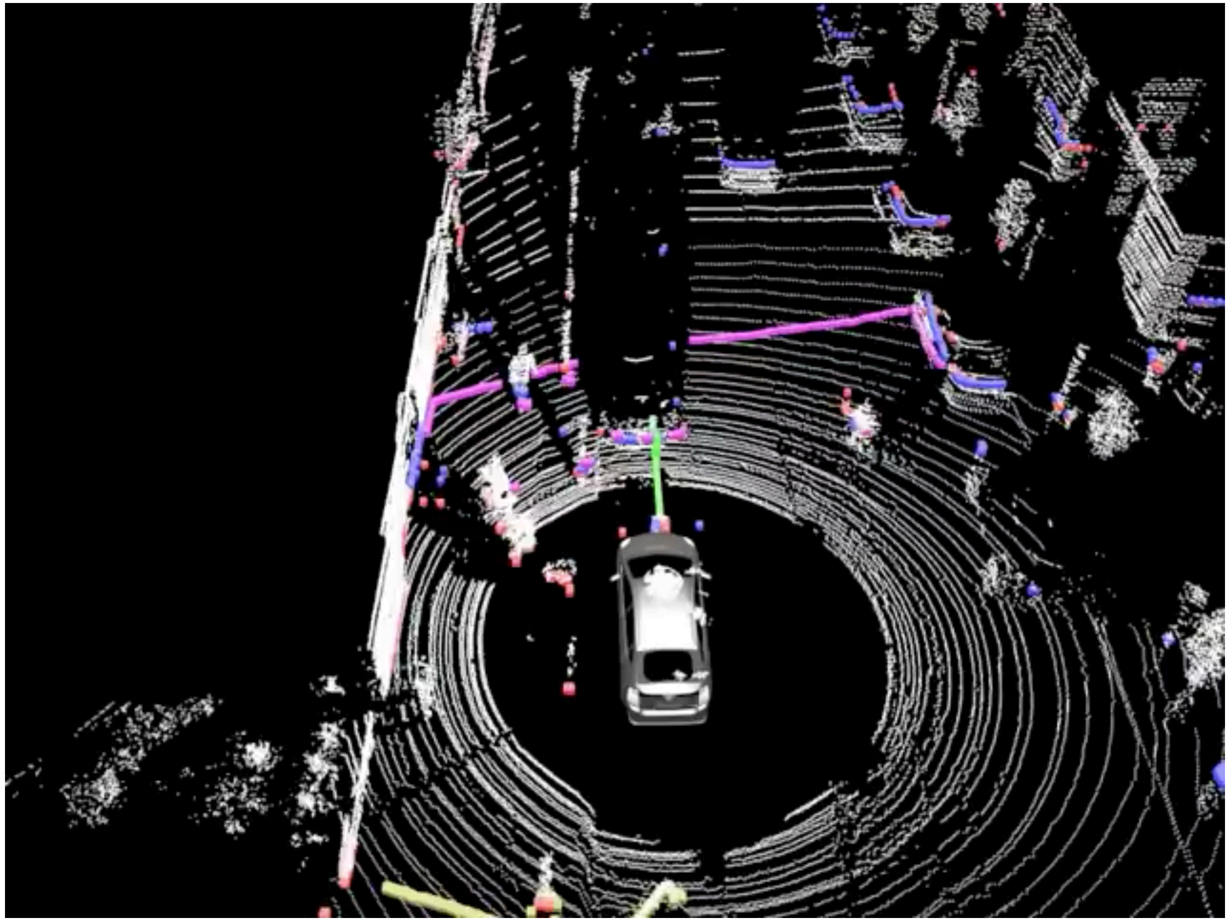



26th July 2019

# Motivation

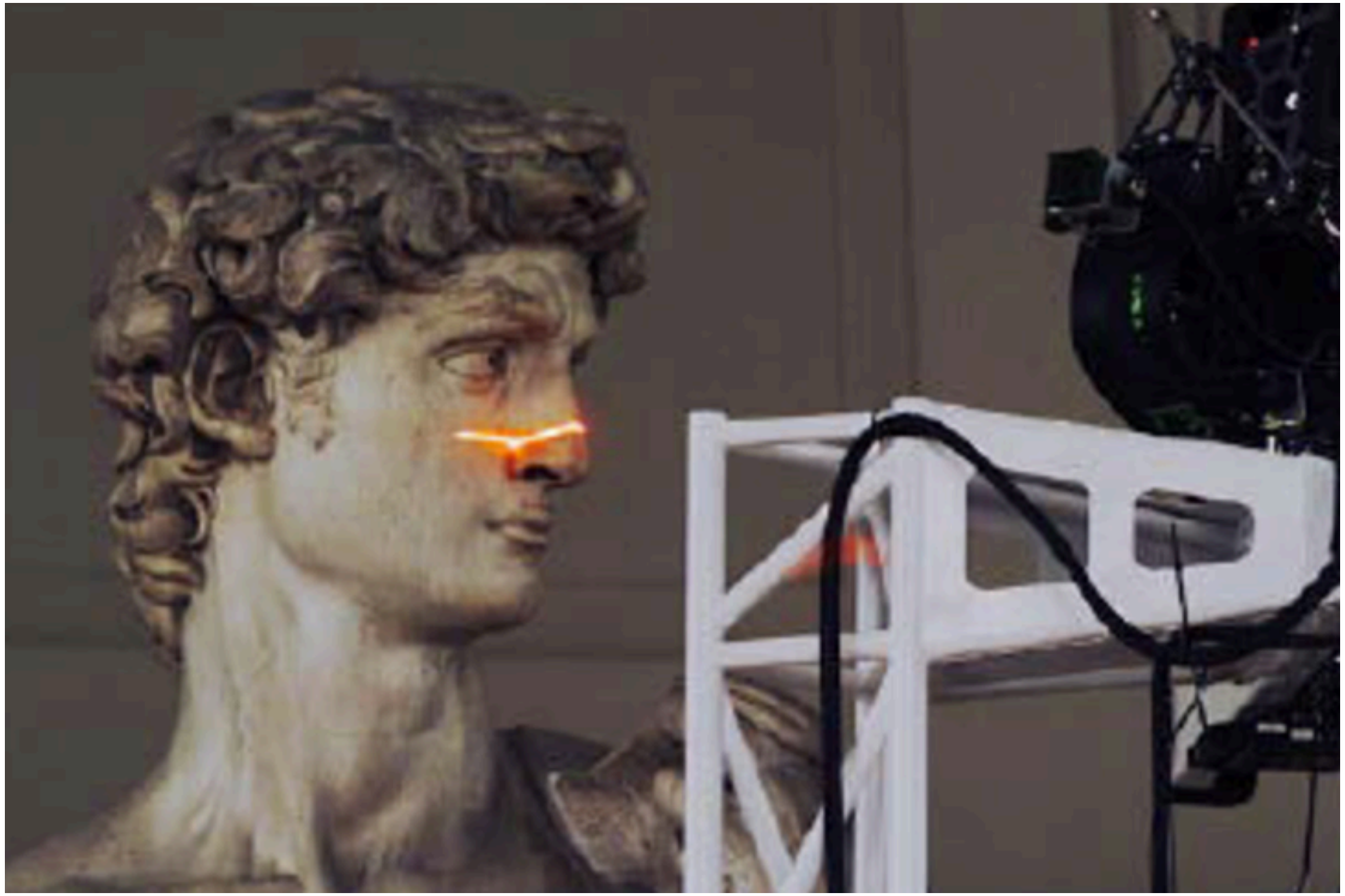
# LiDAR Scanning

- Sensing the environment
- Sparse point measurements
- Only partially visible objects



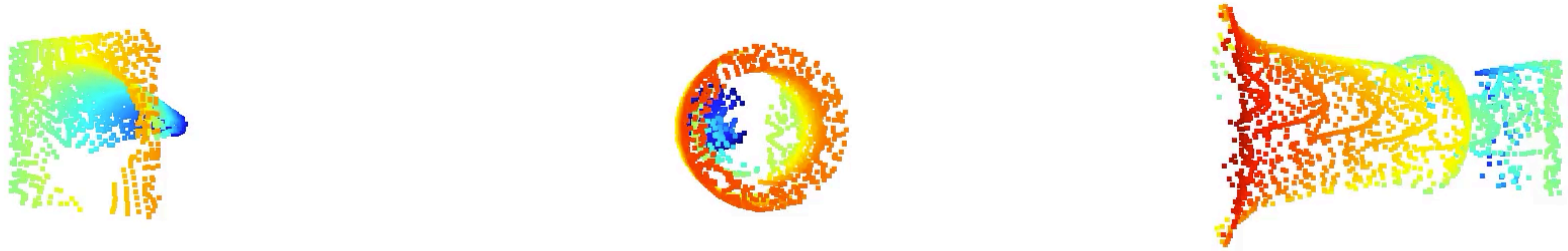
# Laser Scanning

- Precise point locations
- Scanline approach
- Requires controlled environment



# Problem Statement

Sparse and Partial Point Clouds

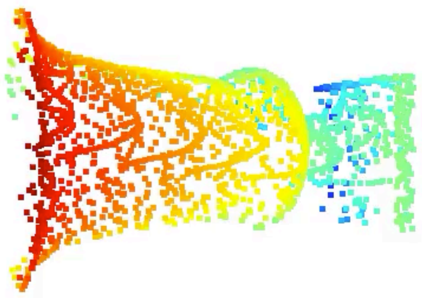


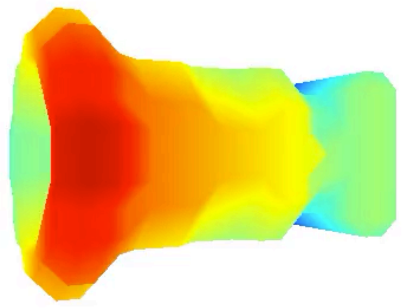
3D Shape Completion

Dense Surface Meshes

# 3D Shape Completion from Sparse Point Clouds

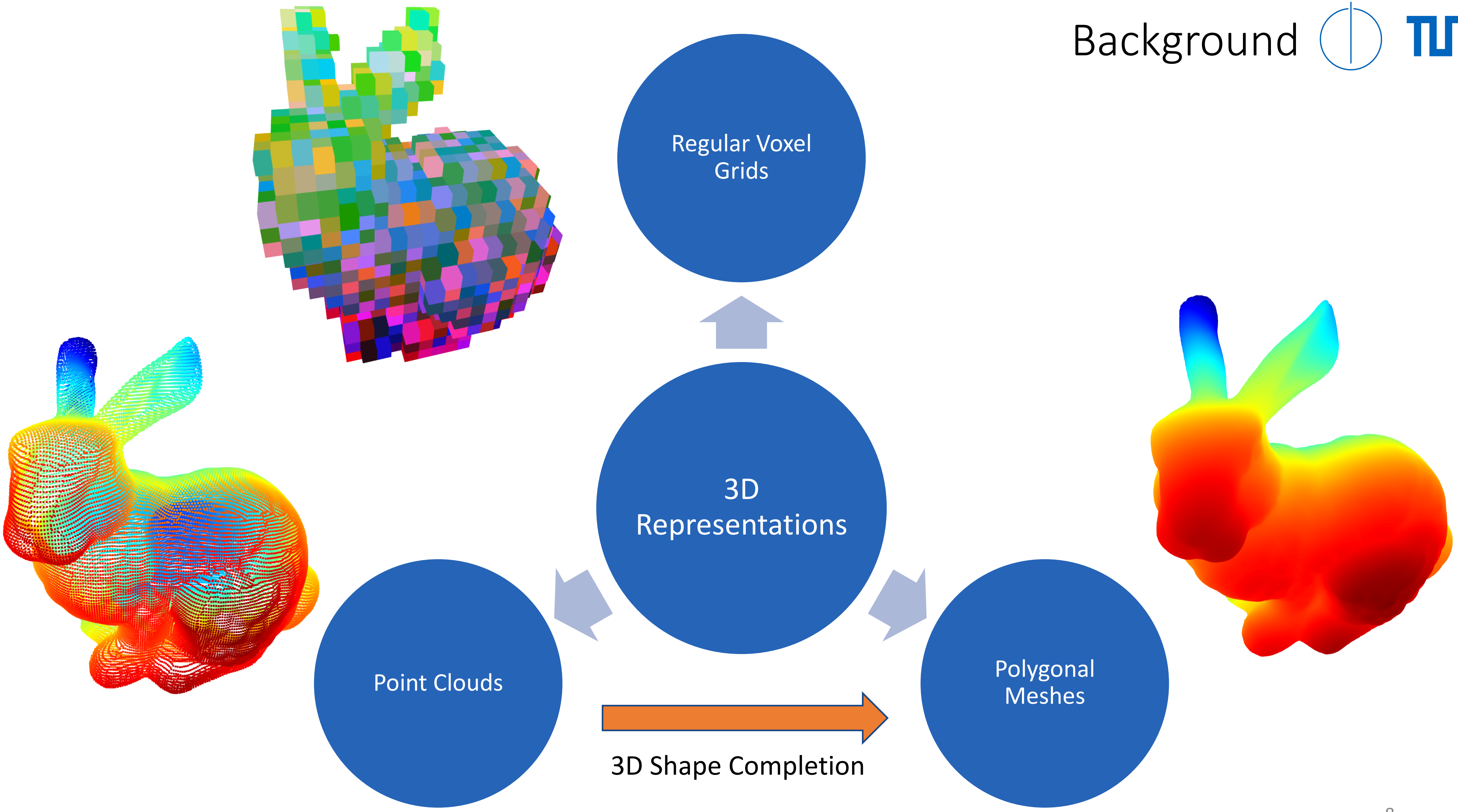
- Background
- Data Generation
- Network Architecture
- Evaluation Results

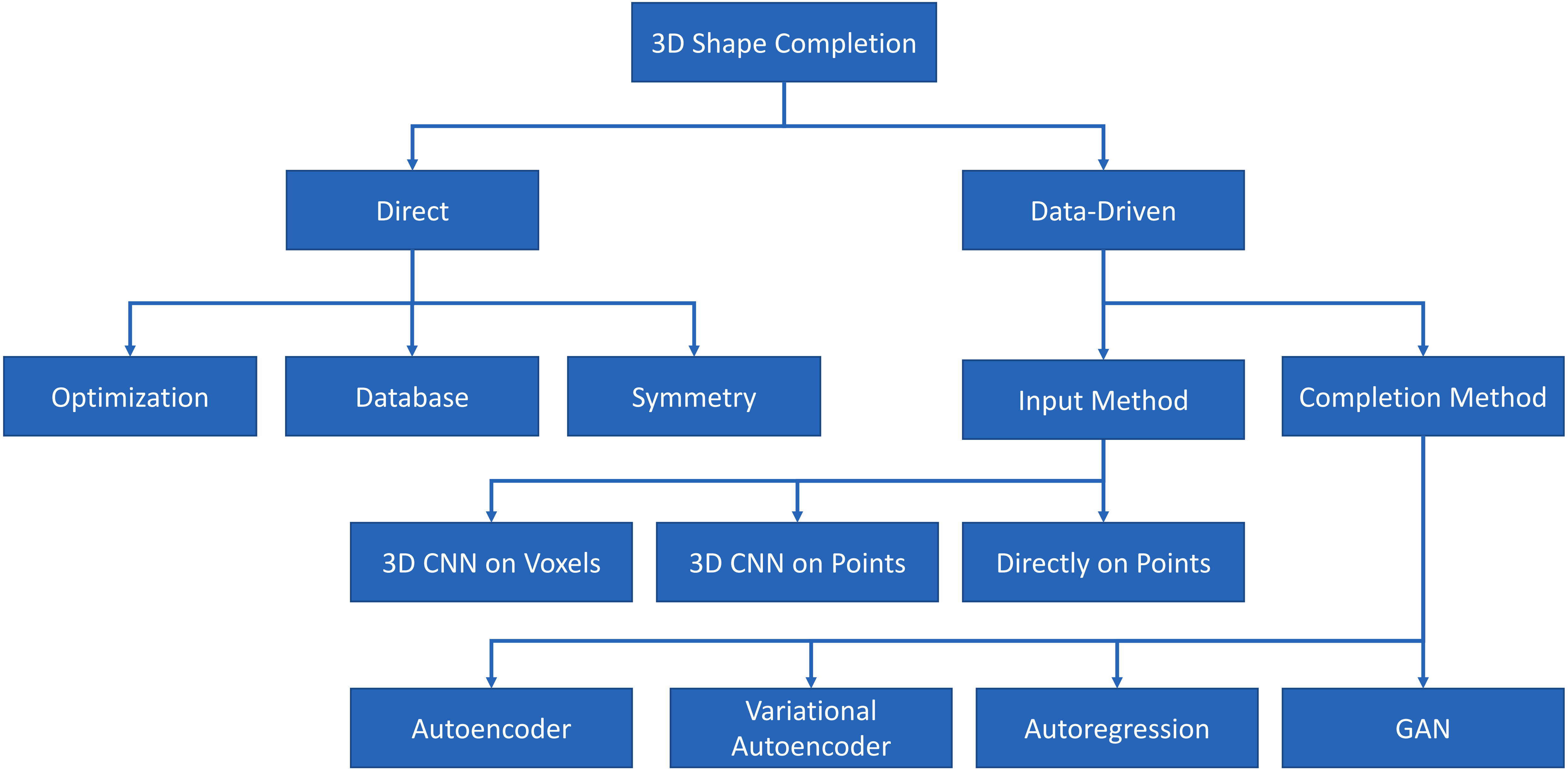




# Background

# Background

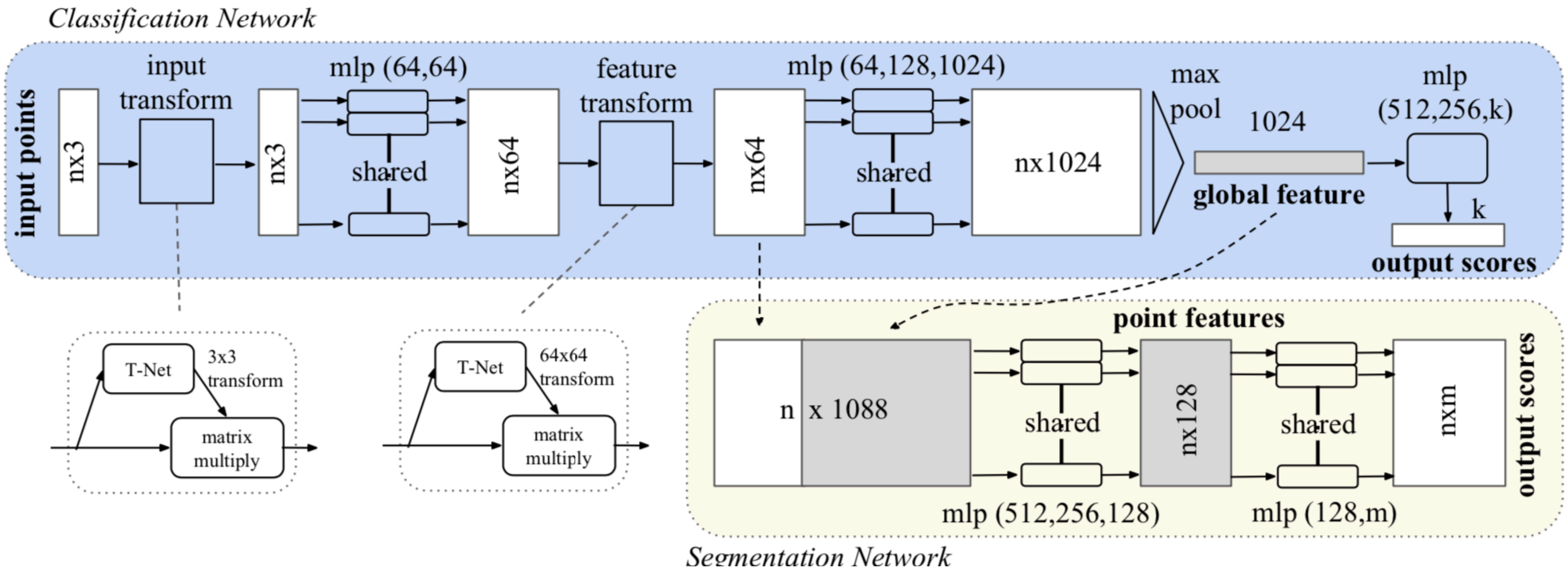

Regular Voxel Grids
3D Representations
Point Clouds
Polygonal Meshes
3D Shape Completion


3D Shape Completion
Direct
Data-Driven
Optimization
Database
Symmetry
Input Method
Completion Method
3D CNN on Voxels
3D CNN on Points
Directly on Points
Autoencoder
Variational Autoencoder
Autoregression
GAN

# PointNet



- Classification and Segmentation
- Operating directly on unstructured point clouds
- Uses symmetric max pooling operation

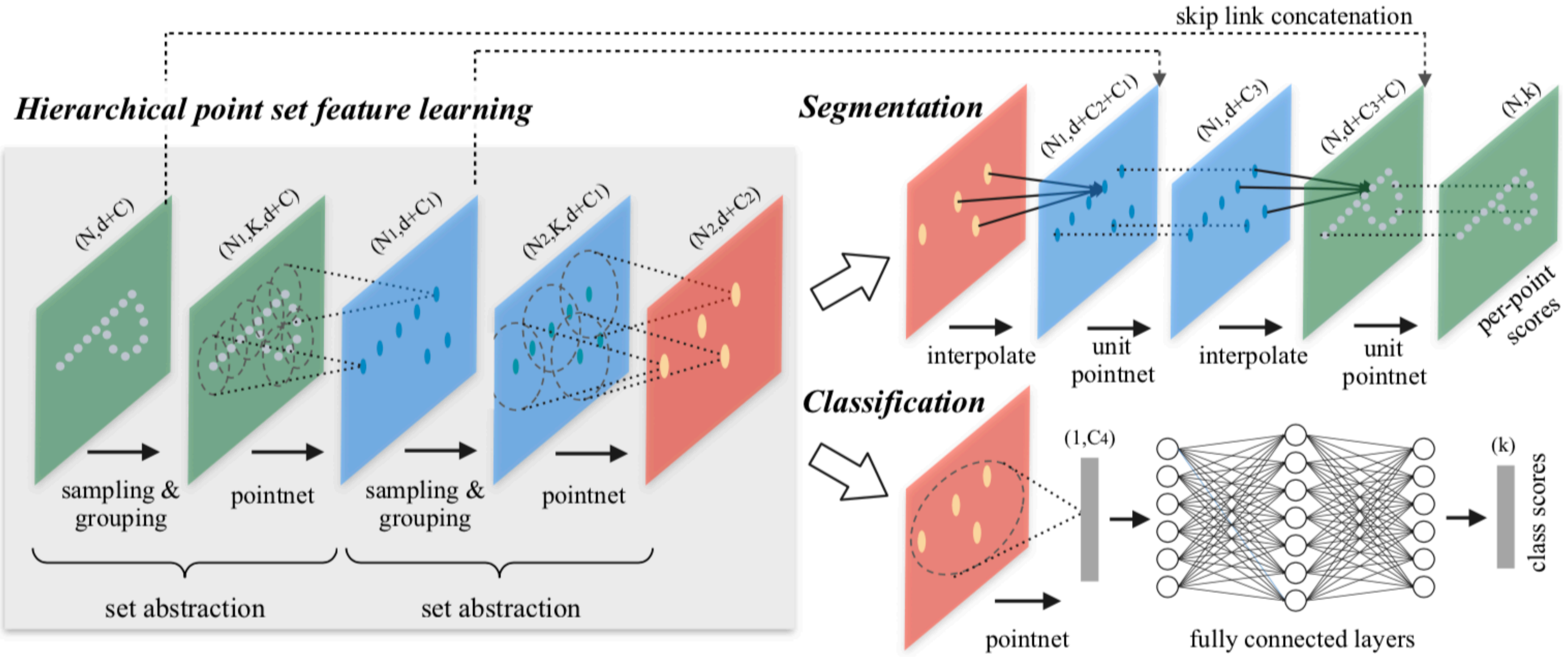





32.C. R. Qi, H. Su, K. Mo, and L. J. Guibas. “PointNet - Deep Learning on Point Sets for 3D Classification and Segmentation.” In: *CVPR* (2017), pp. 77–85.

# PointNet++

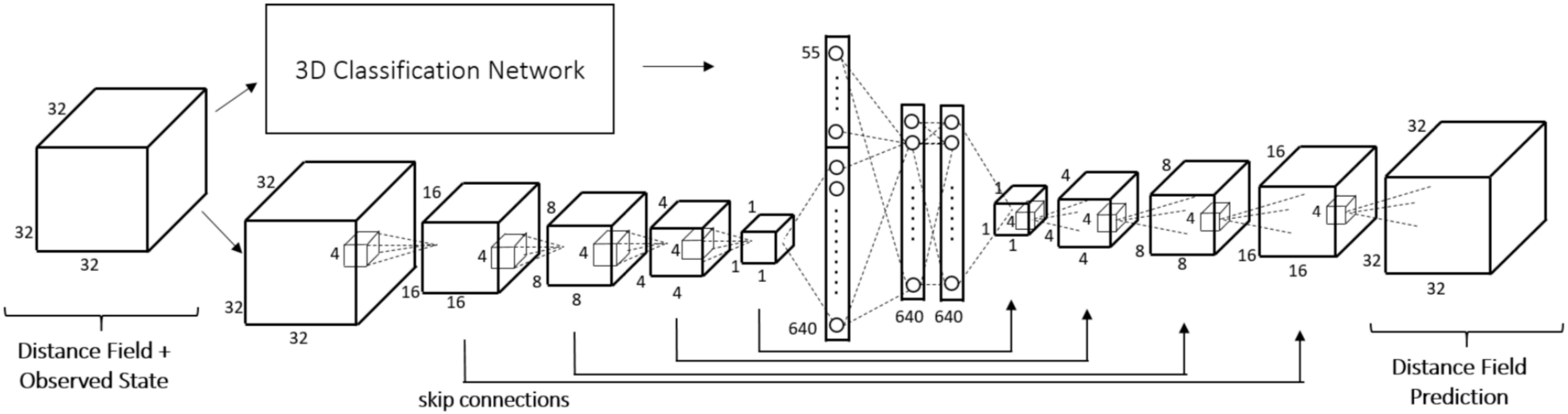
- Improvements by stacking multiple PointNets
- Captures local point neighborhoods




49.C. R. Qi, L. Yi, H. Su, and L. J. Guibas. “PointNet++ - Deep Hierarchical Feature Learning on Point Sets in a Metric Space.” In: *NIPS* (2017).

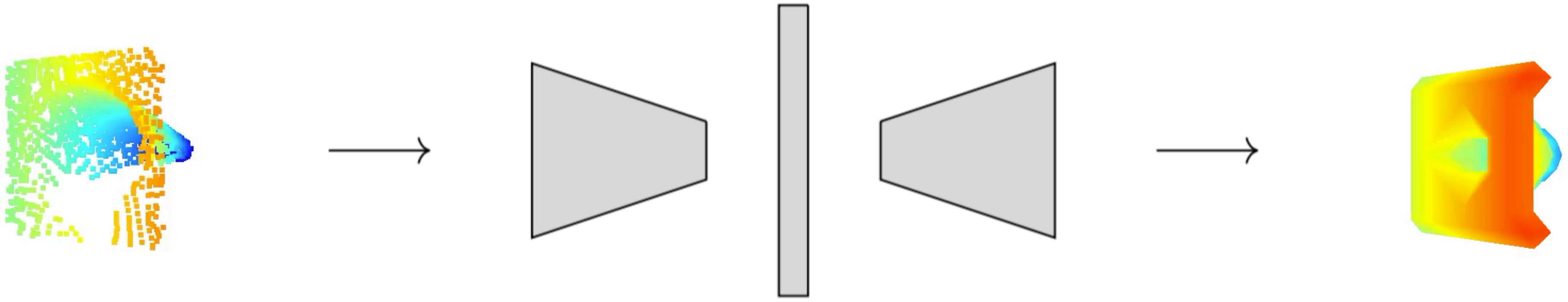
# 3D Encoder Predictor CNN

- Learns shape completion with autoencoder-like architecture
- Operates on regular voxel grids




A. Dai, C. R. Qi, and M. Nießner. "Shape Completion Using 3D-Encoder-Predictor CNNs and Shape Synthesis." In: *CVPR* (2017), pp. 6545–6554.

# Method Overview



**Input:**

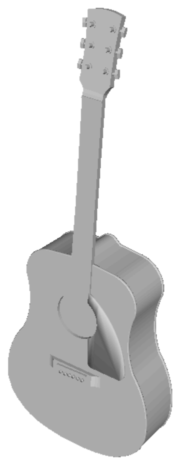
- Sparse and Partial Point Clouds
- Unstructured list of xyz coordinates in 3D space

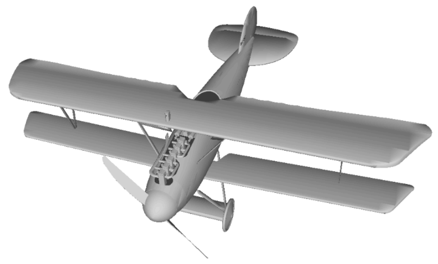
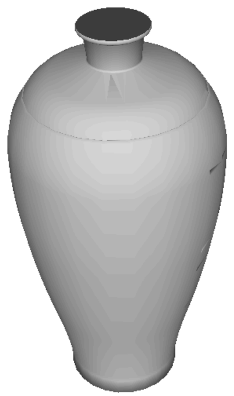
**Output:**

- Dense Surface Meshes
  - Vertices and faces
- Unsigned Distance Field as intermediary representation

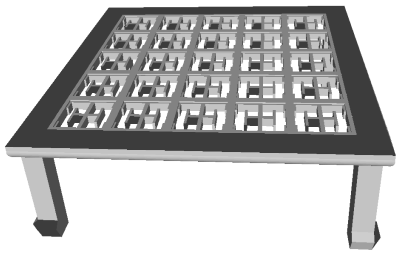
# Data Generation

# ModelNet40



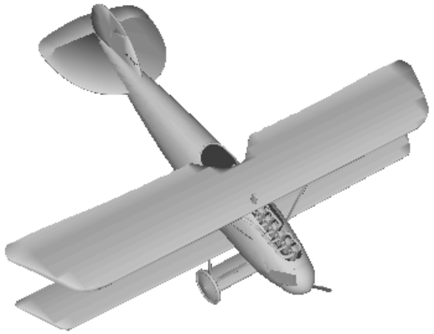



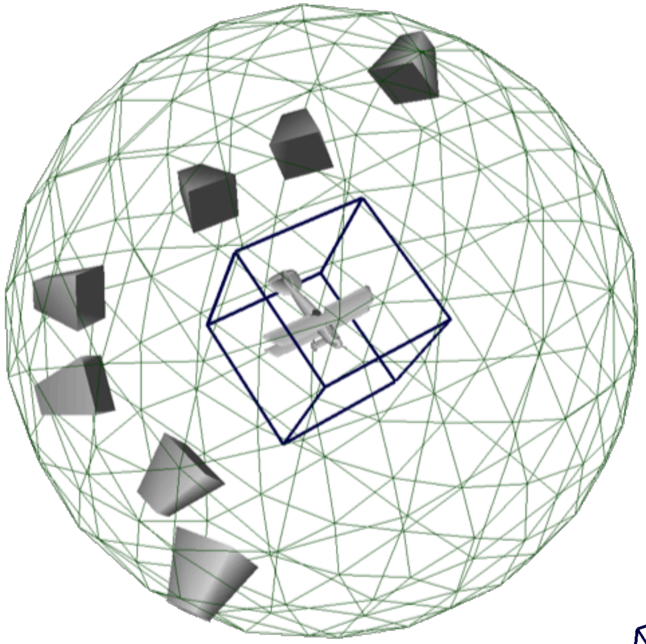
- High-Quality 3D CAD models
- 40 object classes
- 9843 train and 2468 test models

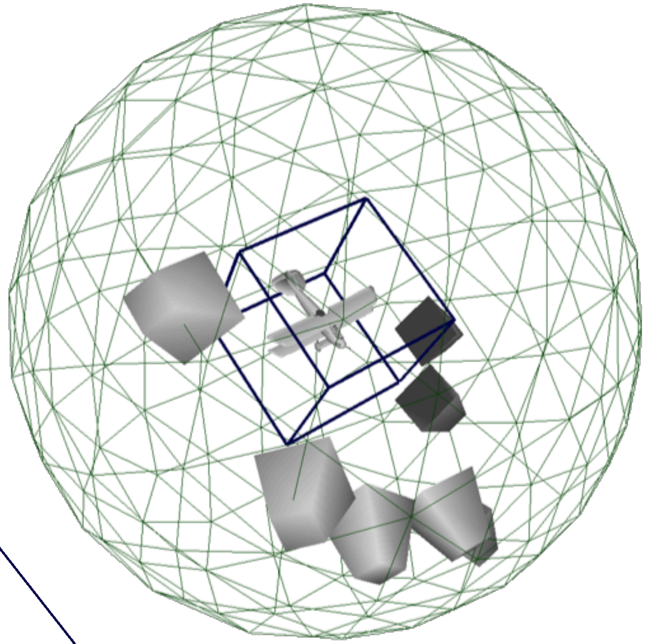


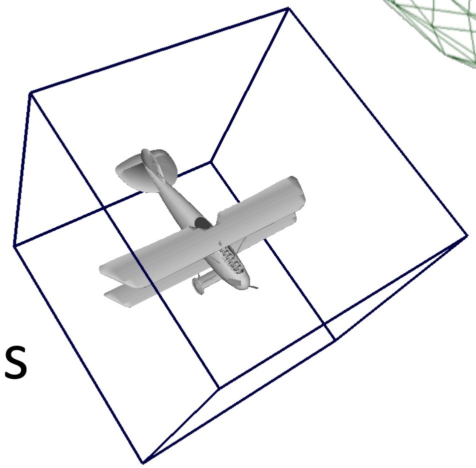
66.Z. Wu, S. Song, A. Khosla, F. Yu, L. Zhang, X. Tang, and J. Xiao. “3D ShapeNets - A deep representation for volumetric shapes.” In: *CVPR* (2015), pp. 1912–1920.

# Trajectory Sampling





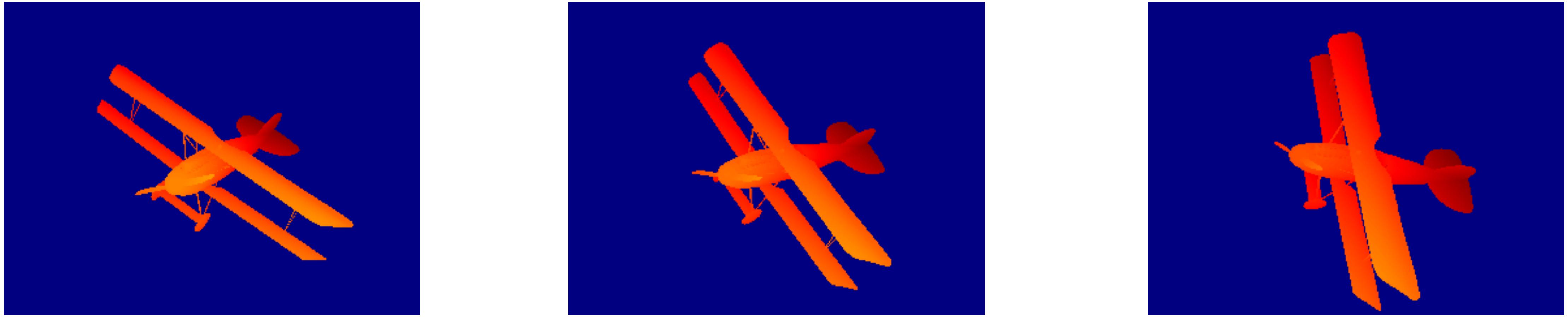




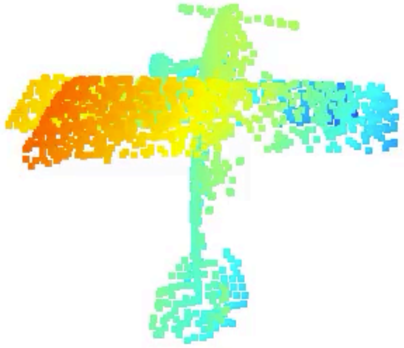
1. Normalization to unit cube
2. Trajectory sampling with jitter
3. Virtual rendering from generated cameras

Augmentations: 6 trajectories and up to 6 rotations per object

# Virtual Rendering

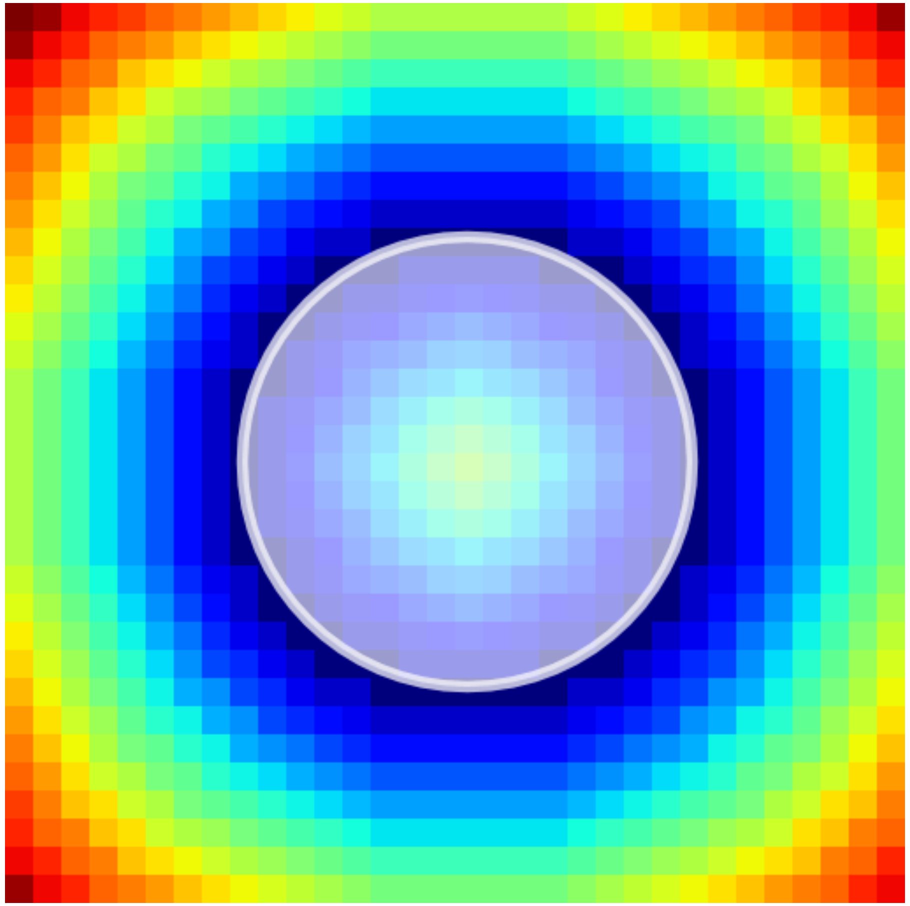


1. Virtual rendering from each camera
2. Backprojecting into common 3D space
3. Subsampling to get exactly 2048 points



# Data Formats

- Input Data
  - Partial Point Clouds: 2048 points
  - Signed Distance Fields: $32^3$ voxels
- Target Data
  - Unsigned Distance Fields: $32^3$ voxels
  - Complete Point Cloud: 4096 points

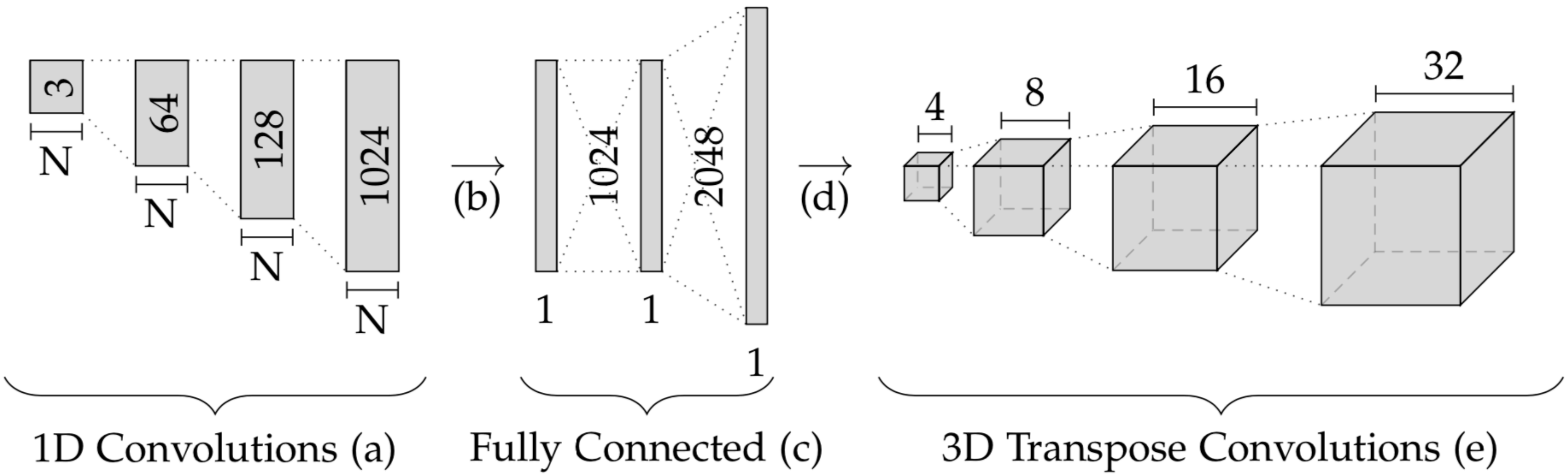


2D slice through a distance field volume

# Network Architecture

# Distance Field Generation



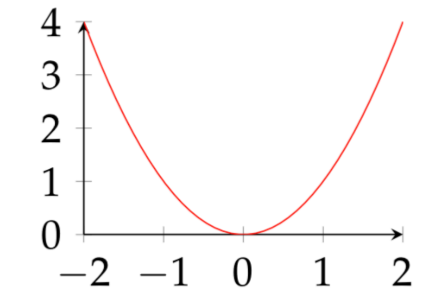



- Ingesting 3D point cloud with 1D convolutional layers
- Symmetric Max Pooling layer
- Fully Connected layers on latent vector
- Reshape and 3D Transpose Convolutions for volume generation

# Loss on Distance Fields

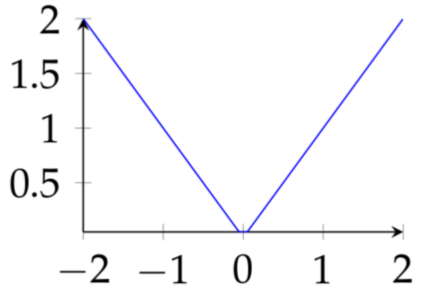
- $l_2$ Loss: Lacking Robustness

$$d_{l_2}(x,y) = \frac{1}{n}\sum_i (x_i - y_i)^2$$



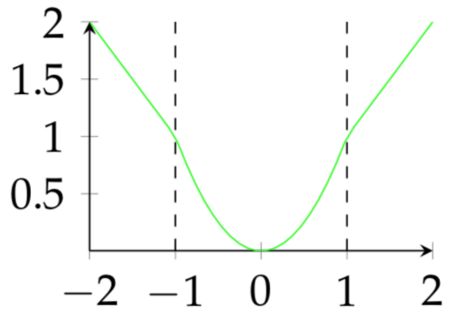

- $l_1$ Loss: Lacking Stability

$$d_{l_1}(x,y) = \frac{1}{n}\sum_i |x_i - y_i|$$



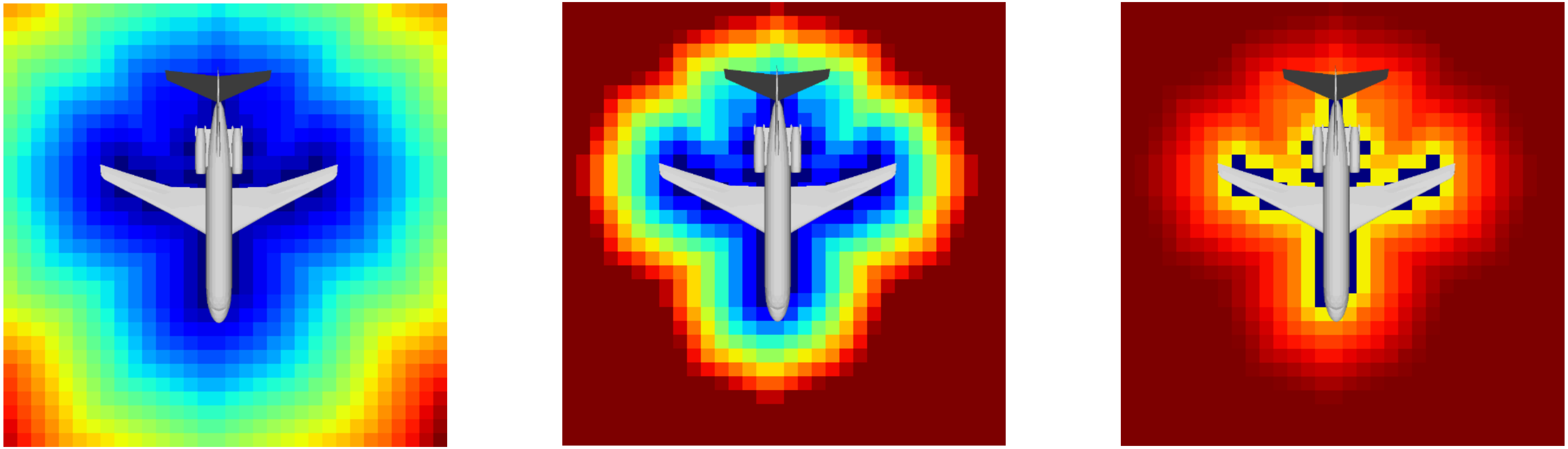

- Huber (smooth $l_1$) Loss

$$d_{huber}(x,y) = \frac{1}{n}\sum_i z_i, \qquad \text{with} \qquad z_i = \begin{cases} \frac{1}{2}(x_i - y_i)^2, & |x_i - y_i| < 1 \\ |x_i - y_i|, & \text{otherwise} \end{cases}$$



# Loss on Distance Fields





- Voxels far away contribute less to shape but influence loss more
- Truncation removes high-value voxels from volume
- Additional log scaling emphasizes changes in surface-near voxels

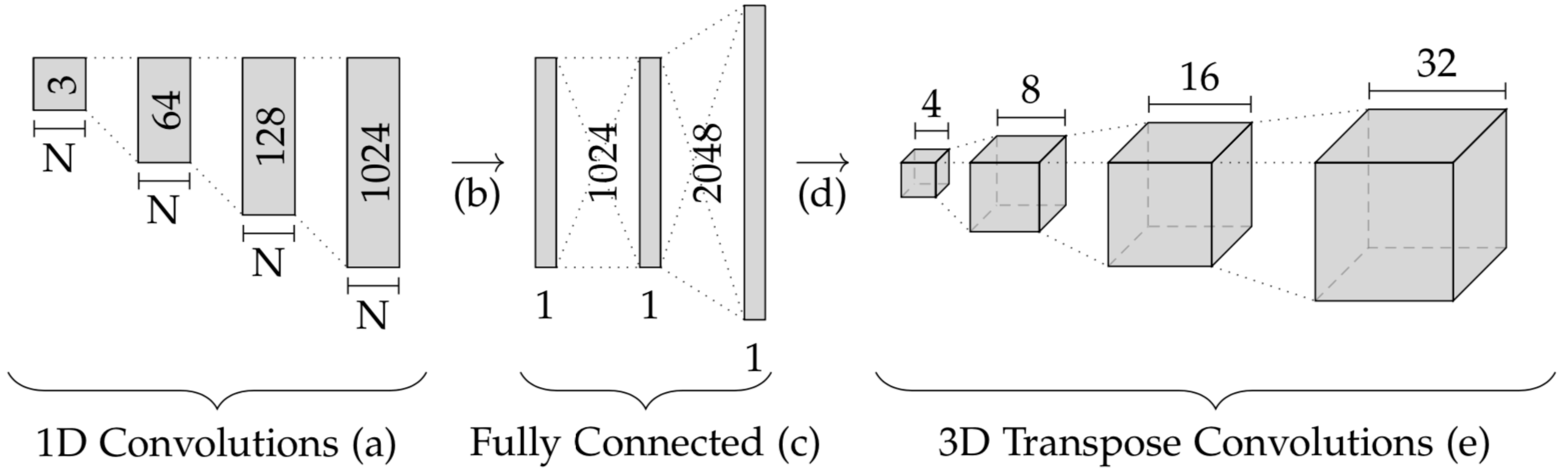
# Evaluation: Loss on Distance Fields

| Rotations | Truncation and Log Scaling | $l_1$ loss | Truncation and Log Scaling | $l_1$ loss |
|---|---|---|---|---|
| 1 | No | 0.016105 | Yes | **0.000881** |
| 3 | No | 0.016066 | Yes | 0.000996 |
| 6 | No | 0.016103 | Yes | 0.001123 |

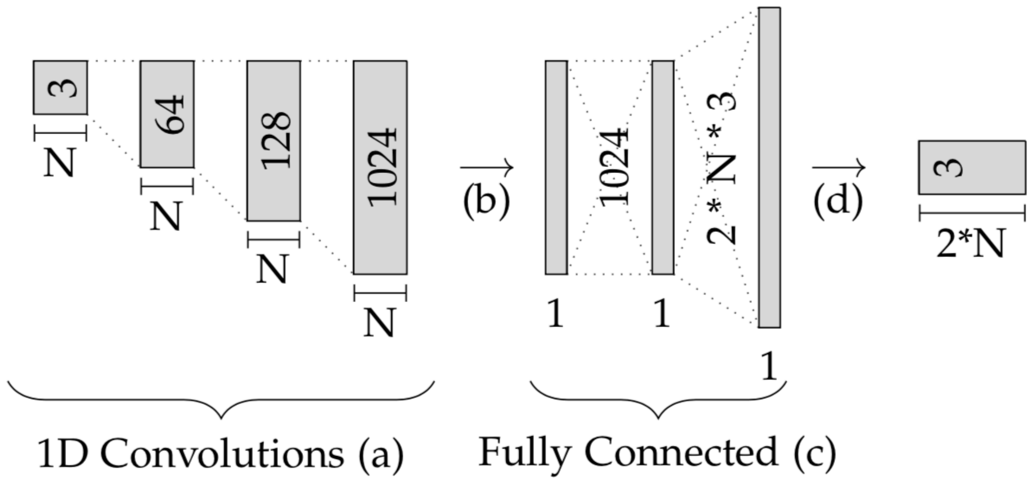
Evaluating how truncation and logarithmic scaling impacts the resulting $l_1$ distance

# Design Studies






- Point Cloud Generation



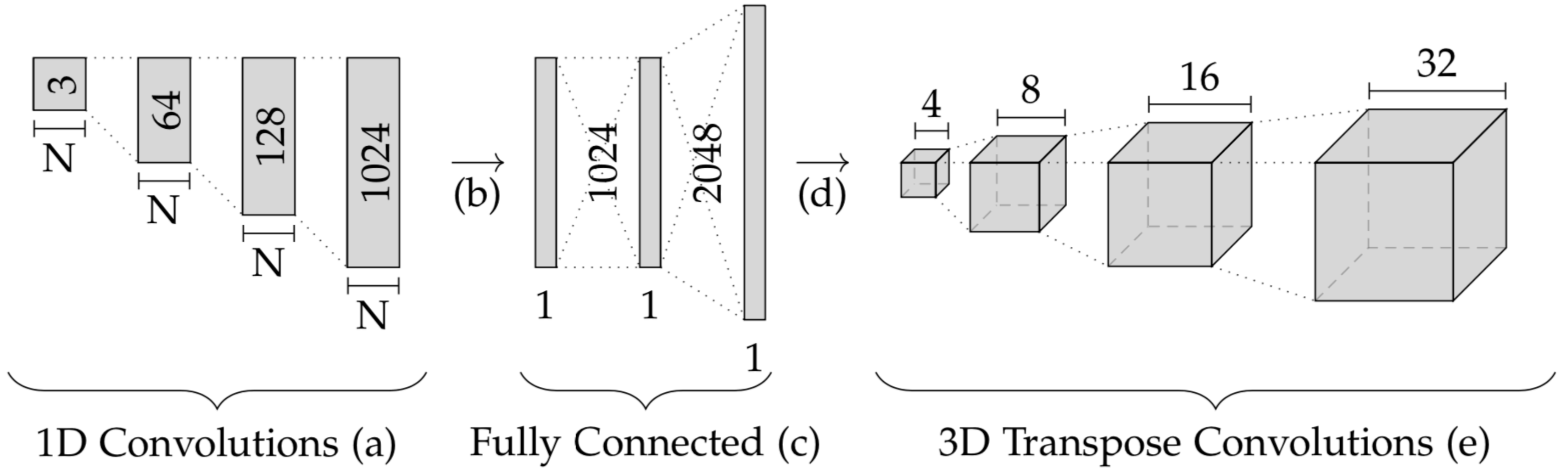
# Evaluation: Point Cloud Generation

| Rotations | Accuracy | Completeness |
|---|---|---|
| 1 | 82.2% | **79.2%** |
| 3 | 87.2% | 69.3% |
| 6 | **87.8%** | 67.1% |

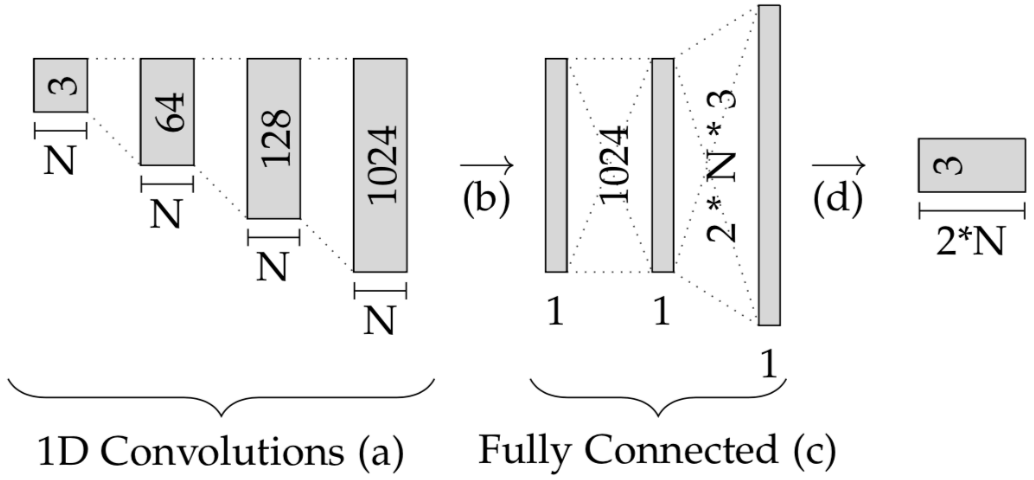
Comparing accuracy and completeness across different number of rotational augmentations during training

- Accuracy: Percentage of points with distance to their ground-truth correspondence above a threshold
- Completeness: Percentage of points with distance to their prediction correspondence above a threshold

# Design Studies




- Hybrid Decoder

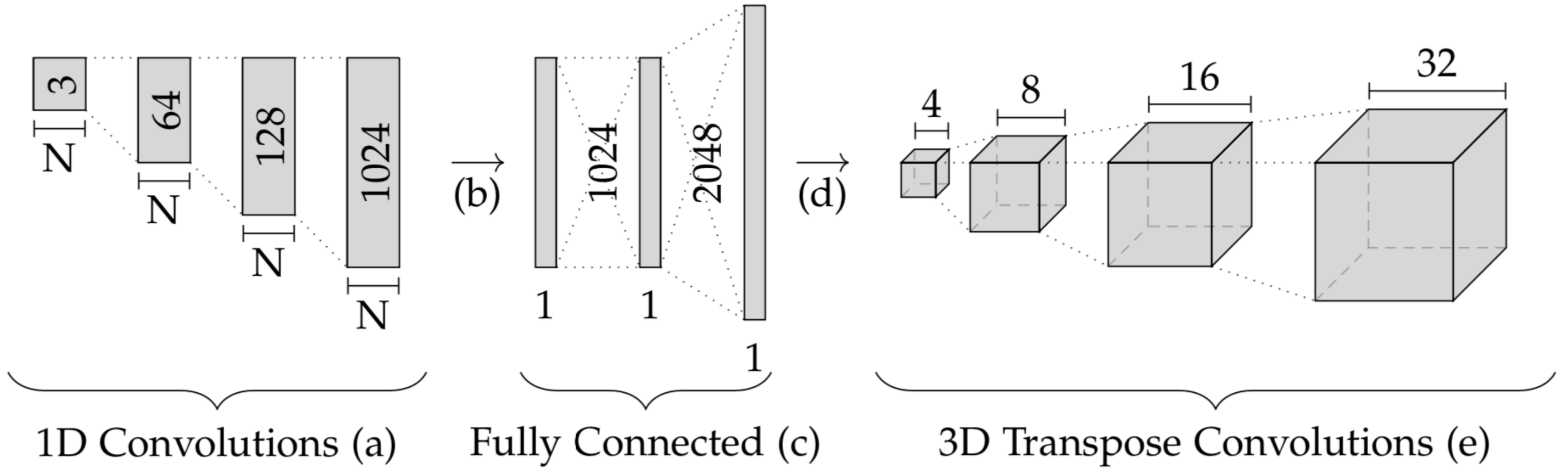


# Evaluation: Hybrid Decoder

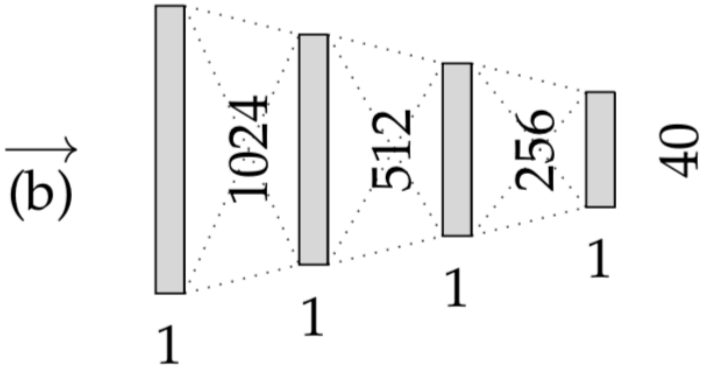
| Decoder | Rotations | $l_1$ Distance | Accuracy | Completeness |
|---|---|---|---|---|
| Point Cloud | 1 | - | 82.2% | 79.2% |
| Point Cloud | 3 | - | 87.2% | 69.3% |
| Point Cloud | 6 | - | 87.8% | 67.1% |
| Distance Field | 1 | **0.000881** | - | - |
| Distance Field | 3 | 0.000996 | - | - |
| Distance Field | 6 | 0.001123 | - | - |
| Hybrid | 1 | 0.000970 | 82.7% | **79.3%** |
| Hybrid | 3 | 0.000988 | **88.4%** | 69.7% |
| Hybrid | 6 | 0.001266 | 87.3% | 66.4% |

Comparing the impact of using both distance field and point cloud decoders vs. only one of them

# Design Studies




- Classification Pseudo-Loss



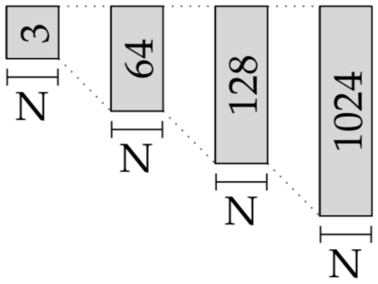
# Evaluation: Classification

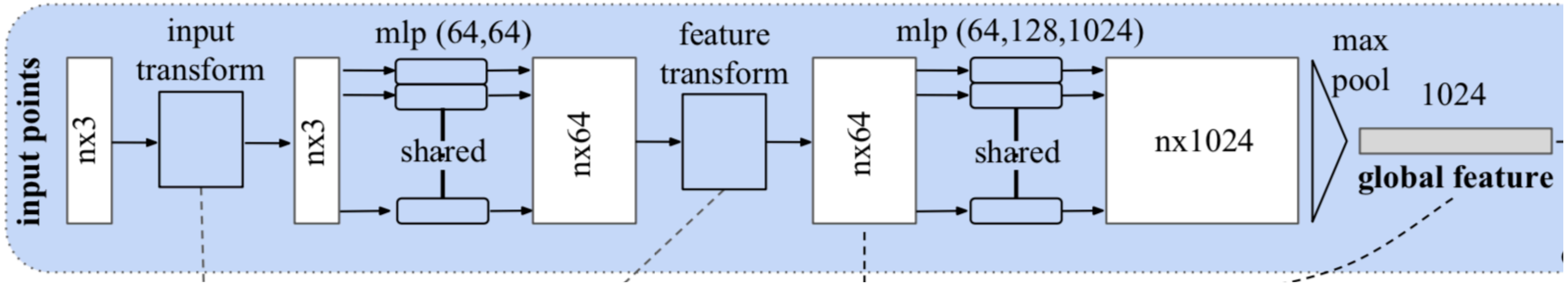
| Decoder | Classification Branch | $l_1$ loss | Classification Branch | $l_1$ loss |
|---|---|---|---|---|
| Distance Field | No | **0.000881** | Yes | 0.001272 |
| Point Cloud | No | **82.2% / 79.2%** | Yes | 75.8% / 65.1% |

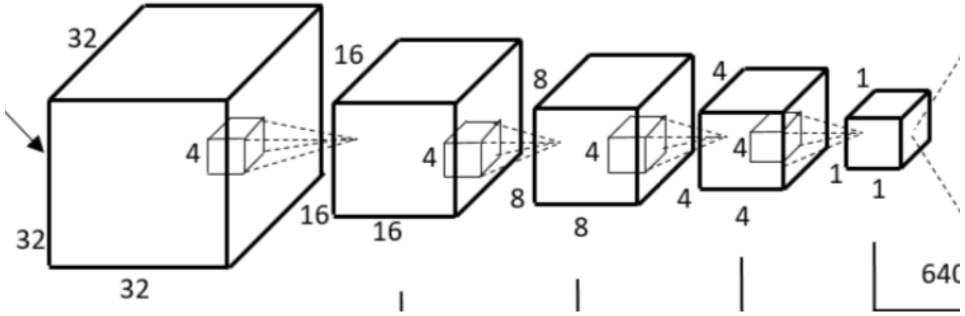
Evaluating how adding a classification branch impacts prediction performance

# Design Studies

- Encoders
  - PointNet
  - PointNet++
  - 3D-EPN



# Evaluation: Encoders



| Encoder | $l_1$ loss (1 rotation) | $l_1$ loss (3 rotations) | $l_1$ loss (6 rotations) |
|---|---|---|---|
| PointNet | 0.001145 | **0.000854** | 0.001212 |
| PointNet++ | 0.001126 | - | - |
| Point Cloud | **0.000881** | 0.000996 | **0.001123** |
| 3D-EPN | 0.00967 | 0.000907 | 0.001150 |

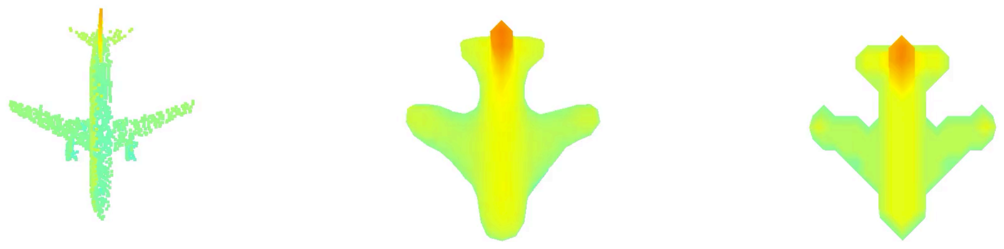
Comparing the impact of using different encoders on completion performance

# Results

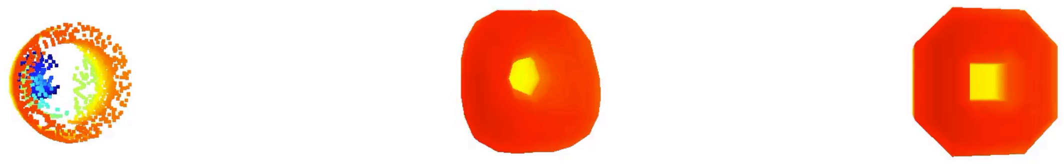
# Evaluation: Overall

| Encoder | Rotations | Classification | Truncation & Log | $l_1$ Distance |
|---|---|---|---|---|
| Point Cloud | 1 | | | 0.016105 |
| Point Cloud | 1 | | ✓ | **0.000881** |
| Point Cloud | 1 | ✓ | ✓ | 0.001272 |
| PointNet | 1 | | ✓ | 0.001145 |
| PointNet++ | 1 | | ✓ | 0.001126 |
| 3D-EPN | 1 | | ✓ | 0.000967 |
| Point Cloud | 3 | | | 0.016066 |
| Point Cloud | 3 | | ✓ | 0.000996 |
| PointNet | 3 | | ✓ | **0.000854** |
| 3D-EPN | 3 | | ✓ | 0.000907 |
| Point Cloud | 6 | | | 0.016103 |
| Point Cloud | 6 | | ✓ | **0.001123** |
| PointNet | 6 | | ✓ | 0.001212 |
| 3D-EPN | 6 | | ✓ | 0.001150 |

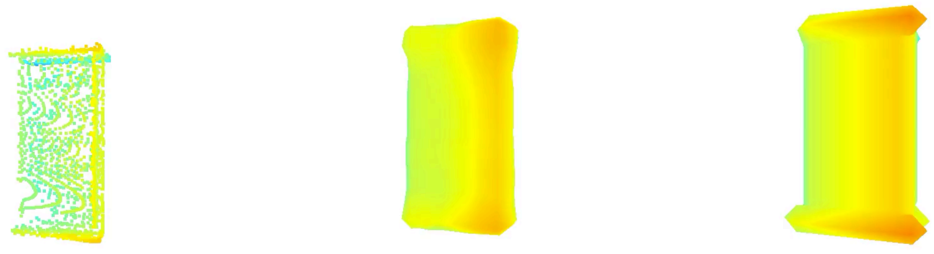
# Qualitative: Mesh



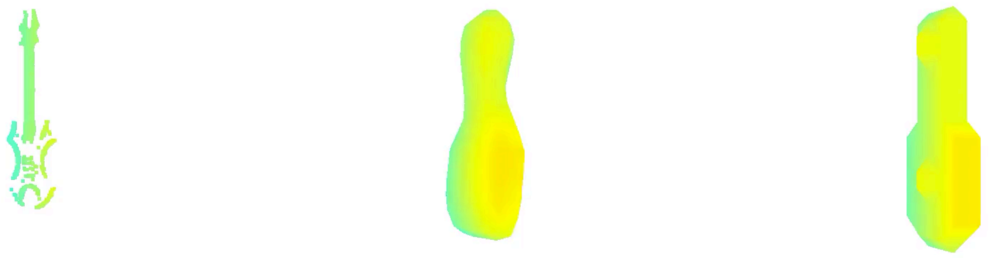
airplane



cup



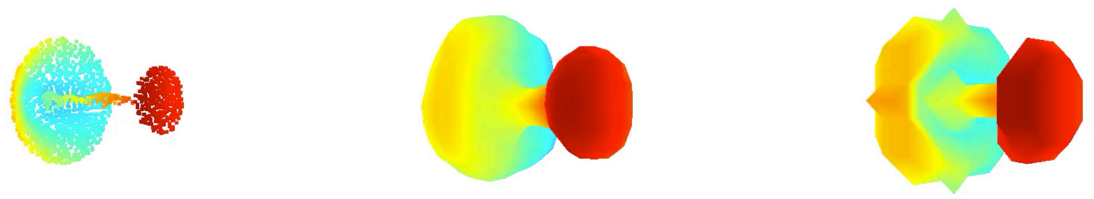
bed



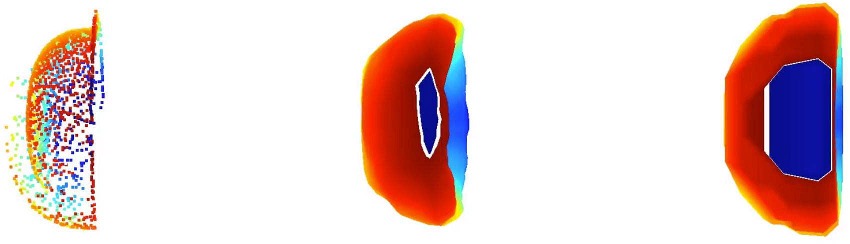
guitar

# Qualitative: Mesh

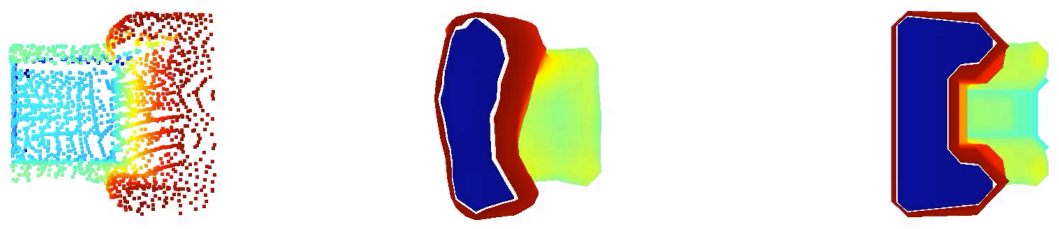
bottle



lamp

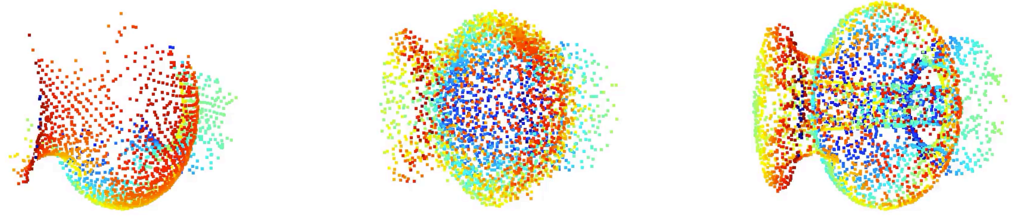


bowl

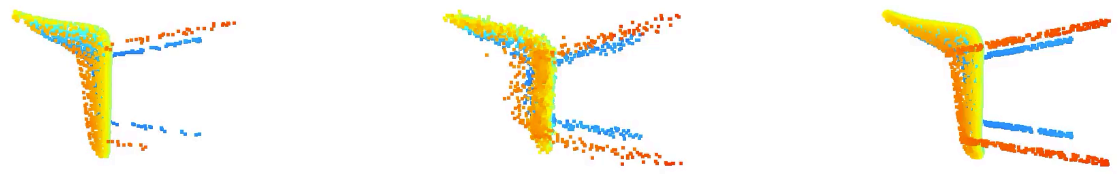


sofa

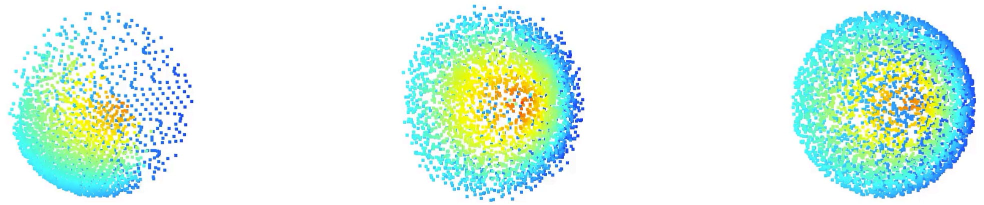
# Qualitative: Point Cloud



vase



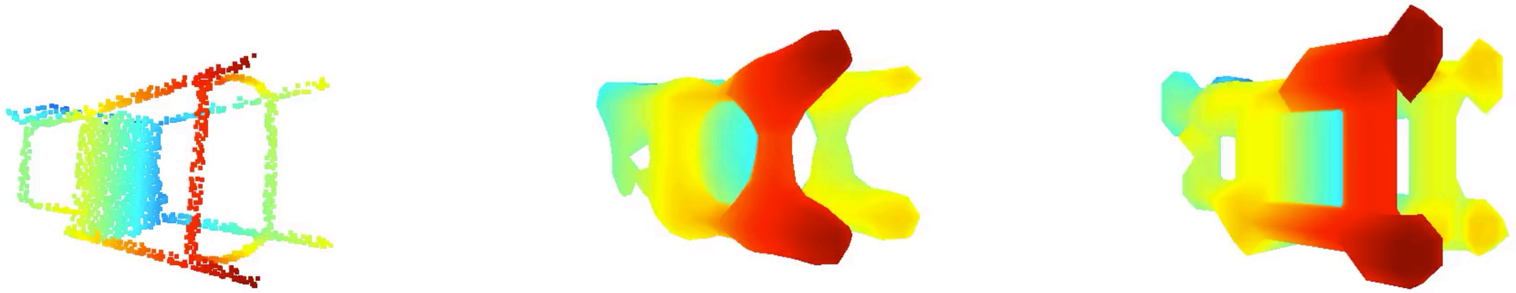
chair



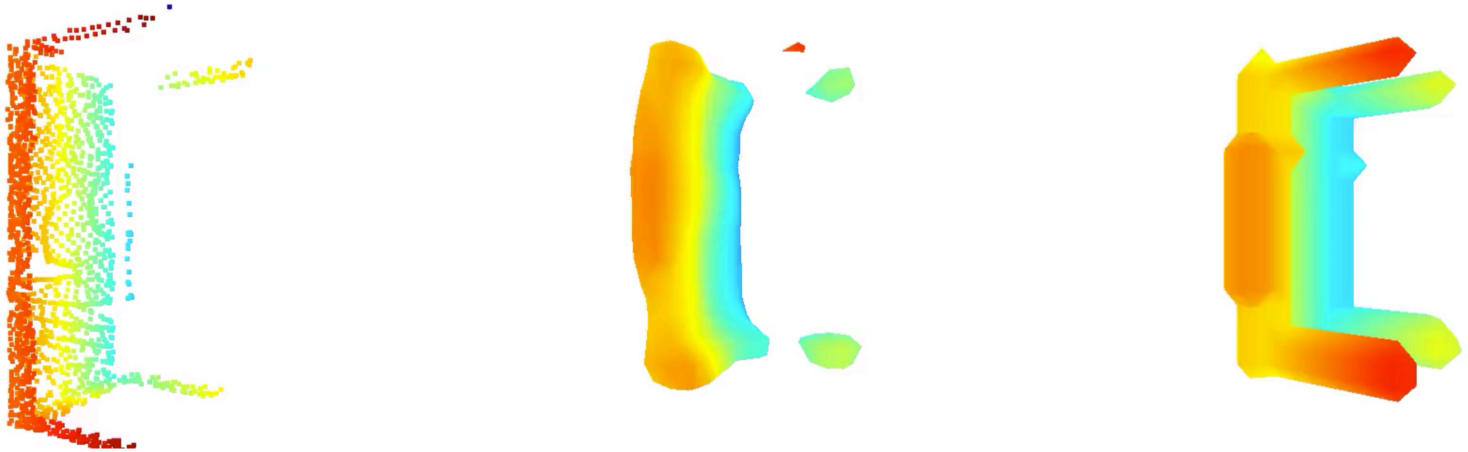
cone

airplane

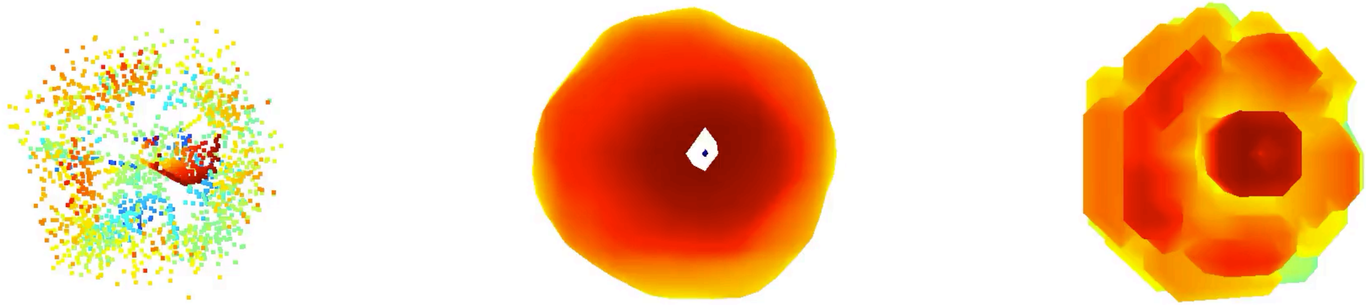
# Limitations



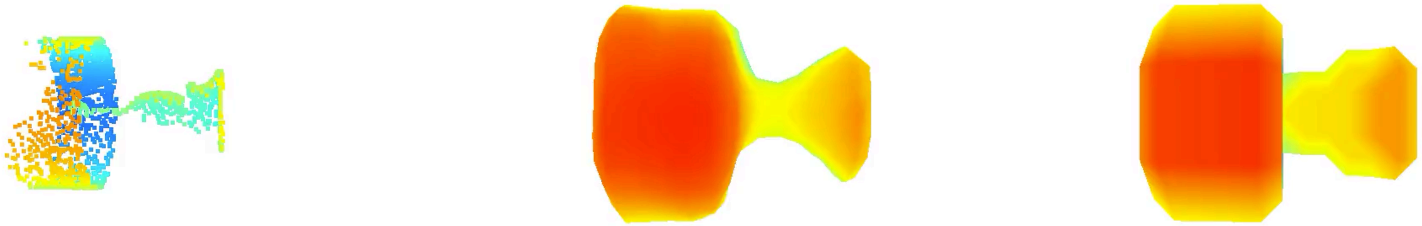
Missing geometry for fine structures



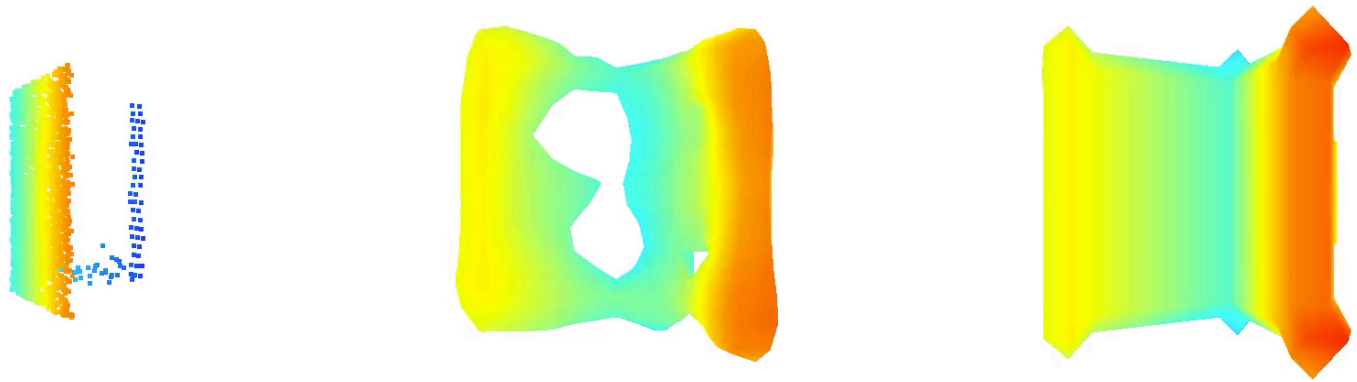
# Limitations



Fused geometry for fine structures



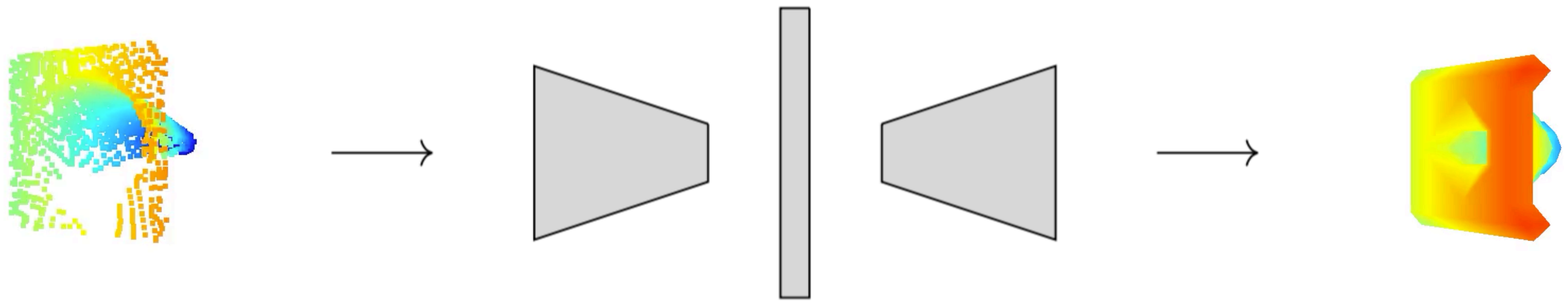
# Limitations



Missing geometry for areas with little information

# Conclusion

- Taking sparse and partial point clouds as input
- Data-driven shape completion using an autoencoder-like architecture
- Outputting dense surface mesh



# Thank you