

# Share With Thy Neighbors: Single-View Reconstruction by Cross-Instance Consistency

Tom Monnier<sup>①</sup>

Matthew Fisher<sup>②</sup>

Alexei A. Efros<sup>③</sup>

Mathieu Aubry<sup>①</sup>

<https://www.tmonnier.com/UNICORN/>



## Overview and challenges

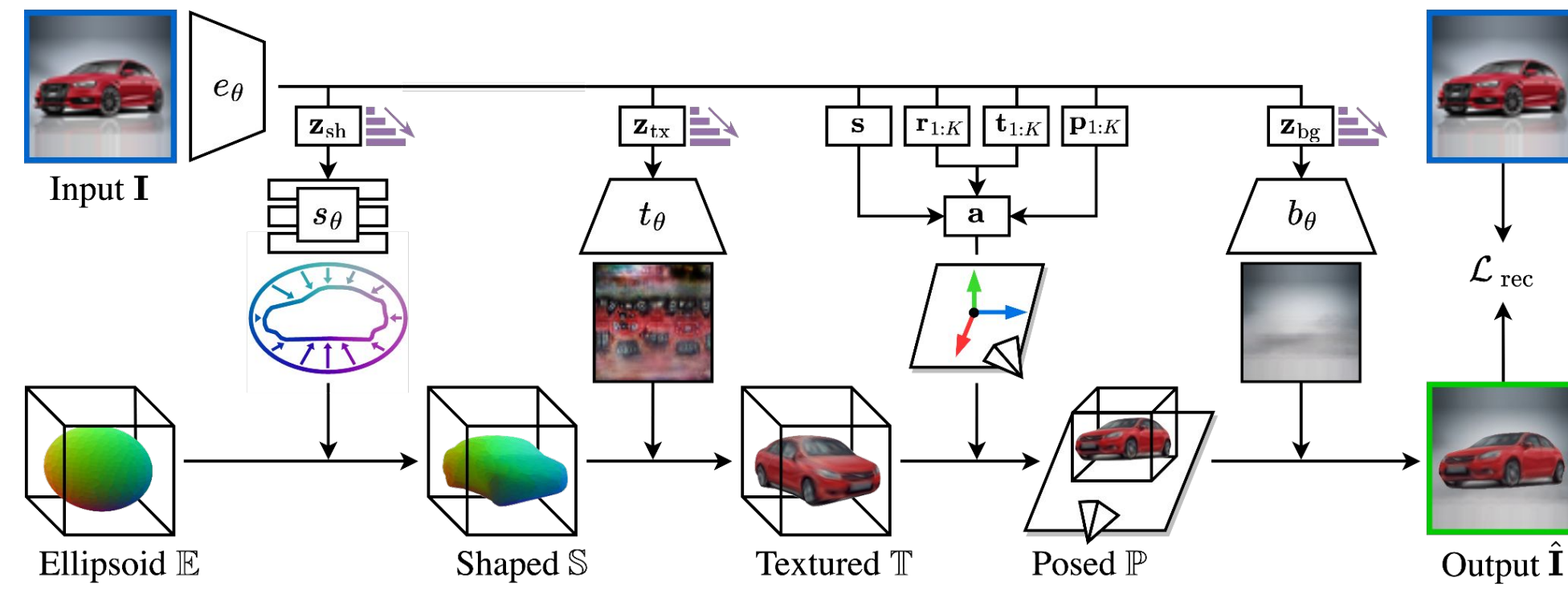
**Goal** → predict 3D from a single image without supervision

**Previous works** → supervision / priors limiting applicability

| Method                              | Supervision | Data           |
|-------------------------------------|-------------|----------------|
| SoftRas [1], DVR [2]                | MV, CK, S   | ShapeNet       |
| CMR [3], SDF-SRN [4], TARS-3D [5]   | CK, S       | ShapeNet, Real |
| IMR [6], UMR [7], UCMR [8], SMR [9] | S, A        | Real           |
| Ours                                | None        | ShapeNet, Real |

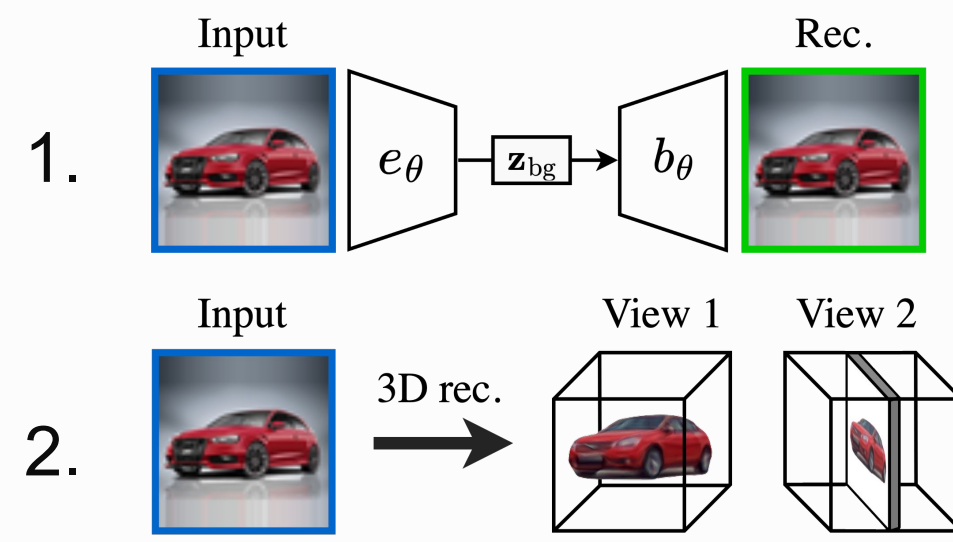
Multi-View, Camera/Keypoints, Silhouette, Assumption (template, symmetry)

**Proposed approach** → img autoencoding into explicit factors



## Dealing w/ overparameterization

**Issue** → degenerate solutions

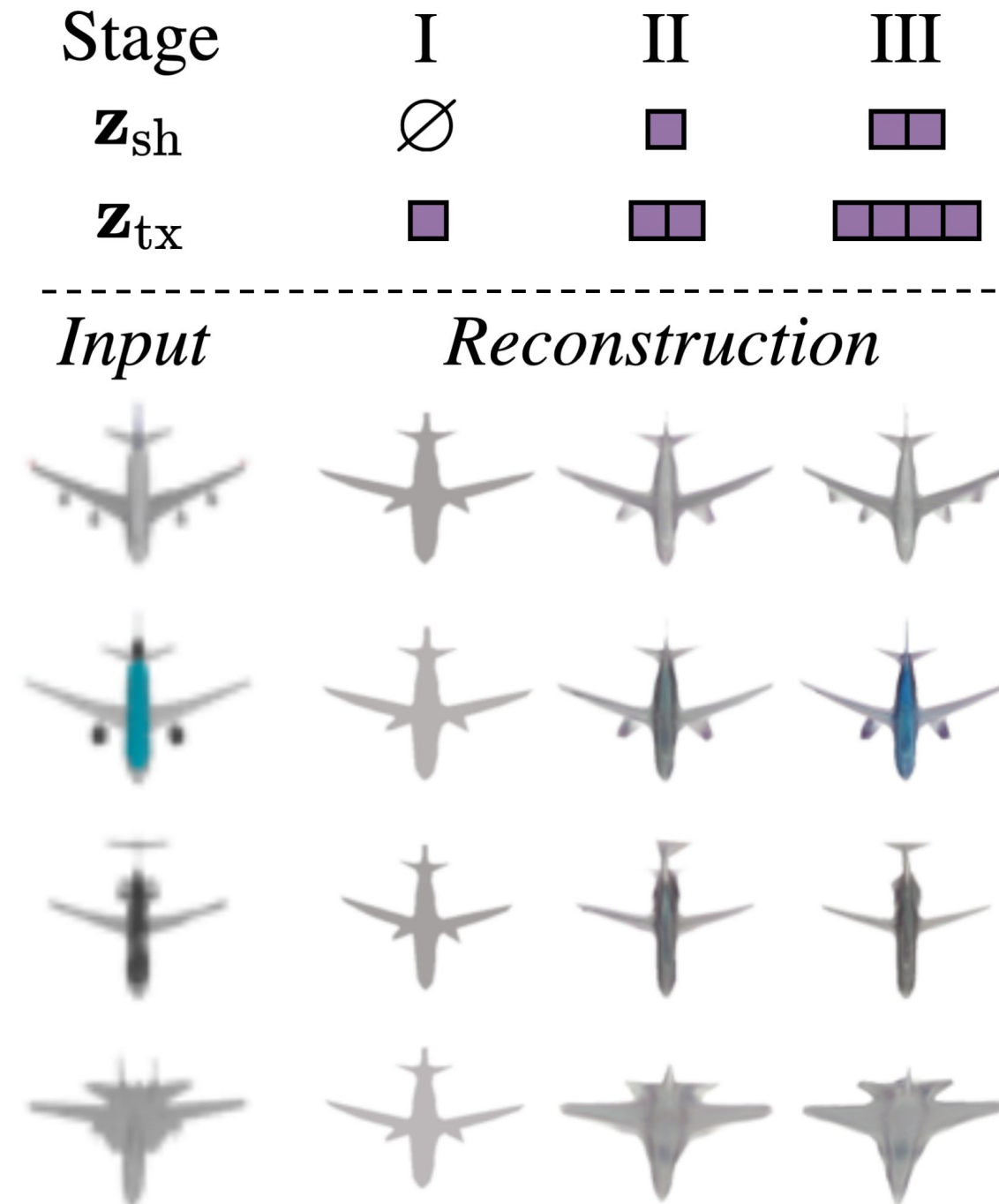


## Cross-instance consistency (SSL)

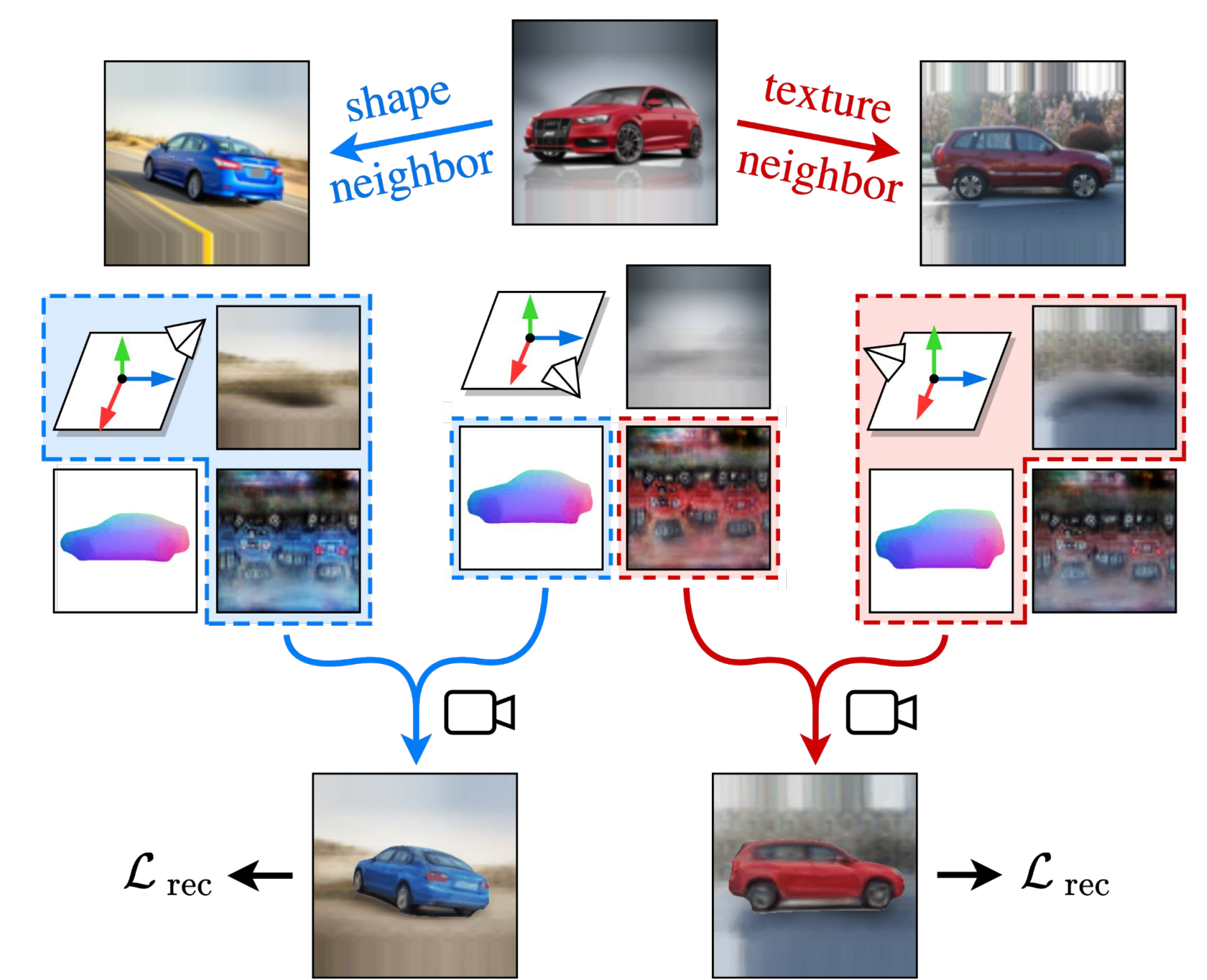
- Progressive conditioning (PC)** = curriculum learning procedure encouraging shared features
- Neighbor reconstruction** = training loss explicitly enforcing consistency between neighbors

## UNICORN: UNsup. Instance COnsistency for 3D Reconstruction

### 1. Progressive conditioning (PC)

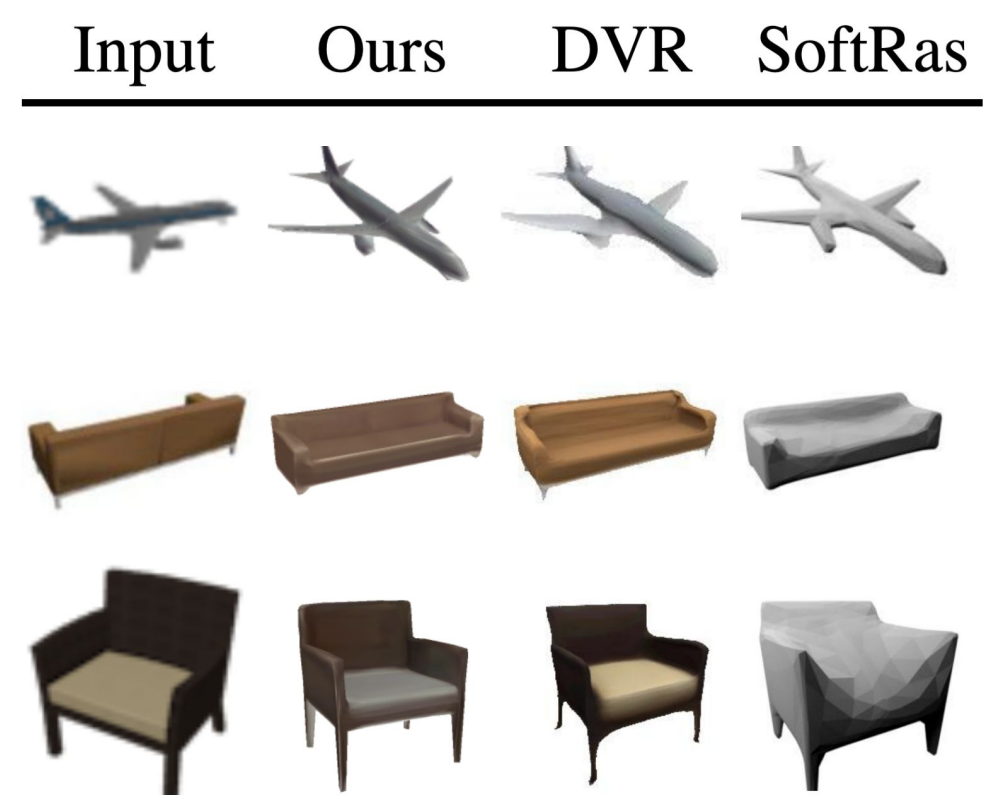


### 2. Neighbor reconstruction $\mathcal{L}_{nbr}$



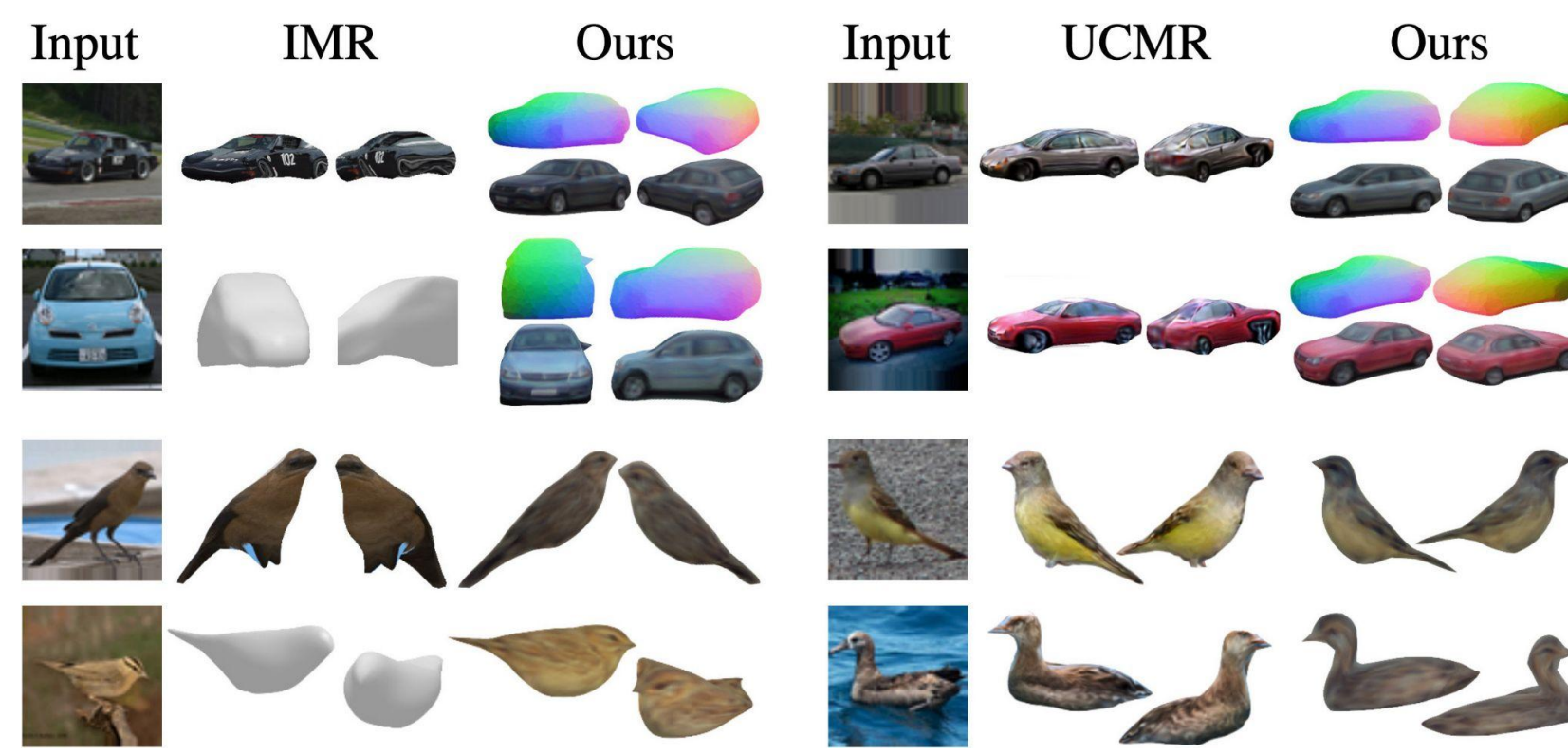
## Results

### Comparison on ShapeNet [A]



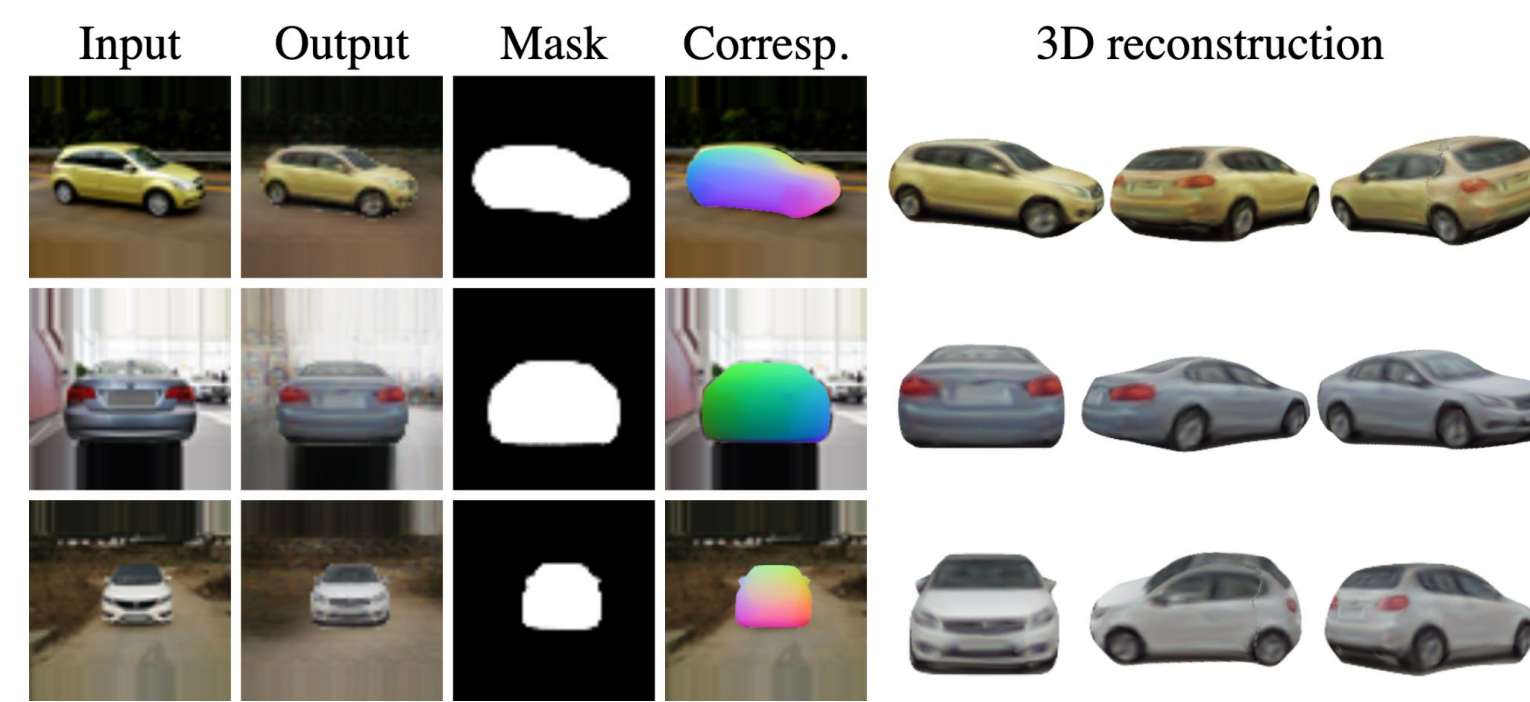
| Method   | Ours  | SDF-SRN | DVR   | DVR   |
|----------|-------|---------|-------|-------|
| MV       |       | ✓       | ✓     | ✓     |
| CK       |       | ✓       | ✓     | ✓     |
| S        |       | ✓       | ✓     | ✓     |
| airplane | 0.110 | 0.128   | 0.114 | 0.111 |
| car      | 0.168 | 0.150   | 0.203 | 0.153 |
| chair    | 0.253 | 0.262   | 0.371 | 0.205 |
| mean     | 0.177 | 0.180   | 0.229 | 0.156 |

### Comparison on Pascal3D+ Car [B] & CUB [C]

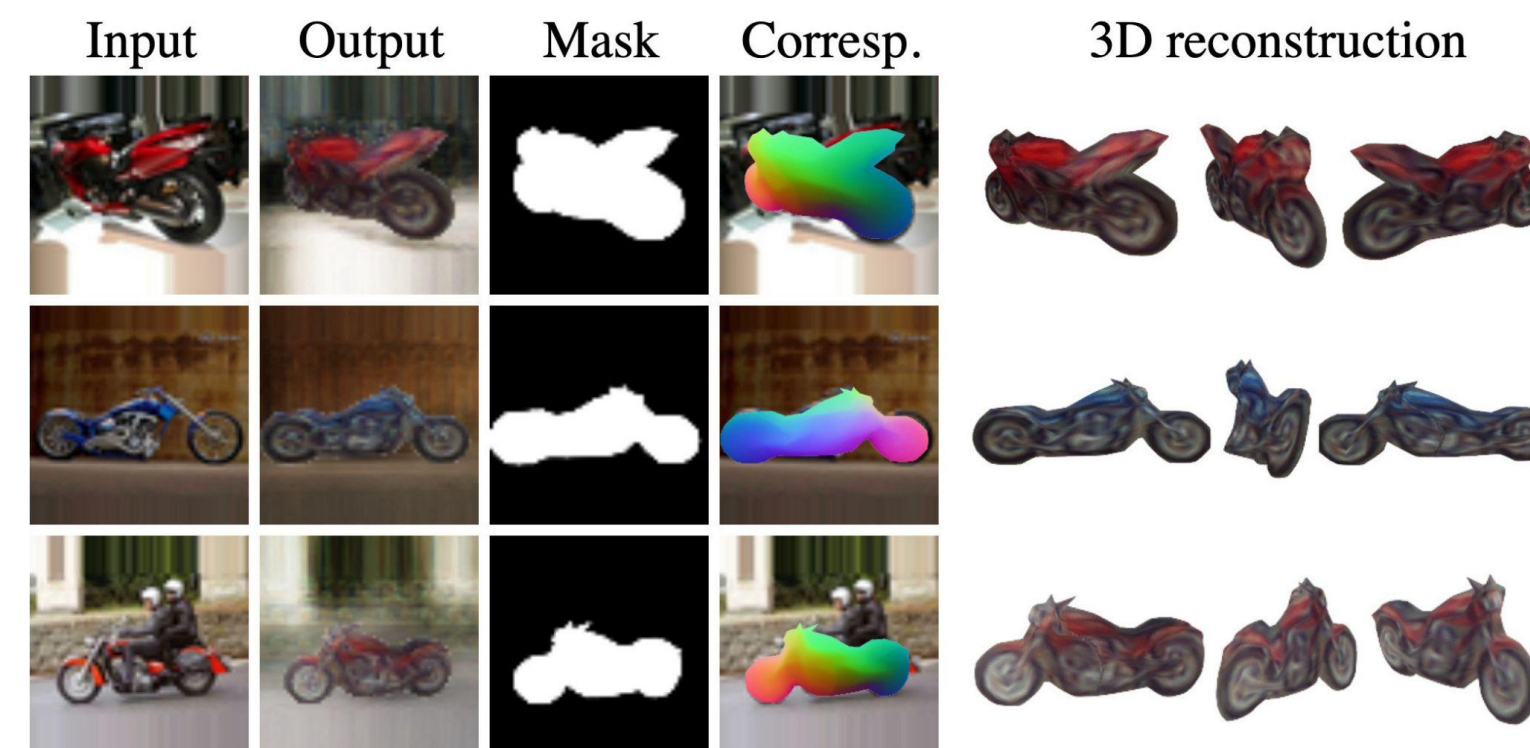


|        | Supervision | Pascal3D+ Car |        |           | CUB-200-2011 |           |
|--------|-------------|---------------|--------|-----------|--------------|-----------|
| Method | CK S A      | 3D IoU↑       | Ch-L1↓ | Mask IoU↑ | PCK@0.1↑     | Mask IoU↑ |
| CMR    | ✓ ✓ ✓       | 64            | -      | -         | 48.3         | 70.6      |
| IMR    | ✓ ✓ ✓       | -             | -      | -         | 53.5         | -         |
| UMR    | ✓ ✓ ✓       | 62            | -      | -         | 58.2         | 73.4      |
| UCMR   | ✓ ✓ ✓       | 67.3          | 0.172  | 73.7      | -            | 63.7      |
| SMR    | ✓ ✓ ✓       | -             | -      | -         | 62.2         | 80.6      |
| Ours   |             | 65.9          | 0.163  | 83.9      | 49.0         | 71.4      |

### Other real images - CompCars [D]

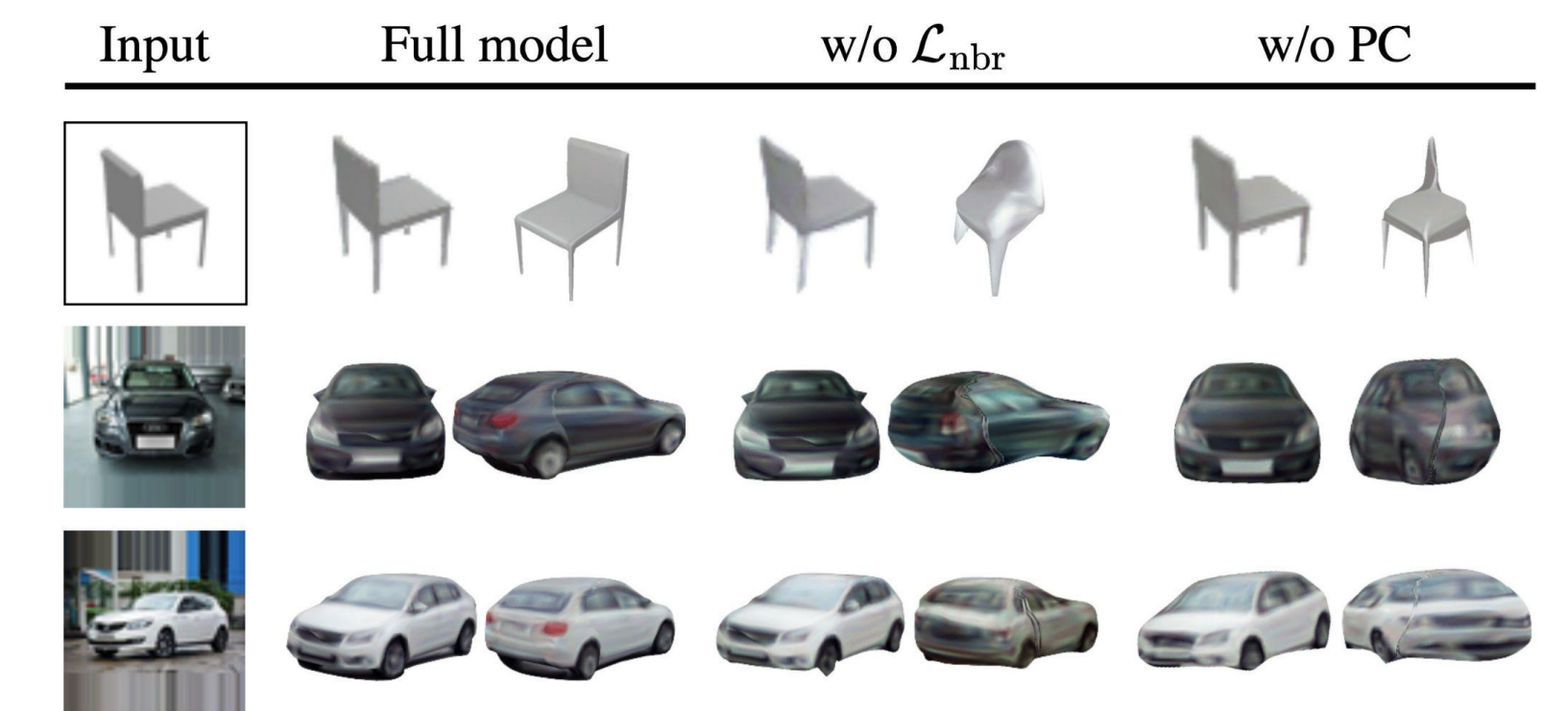


### LSUN Motorbike [E]



## Ablation study

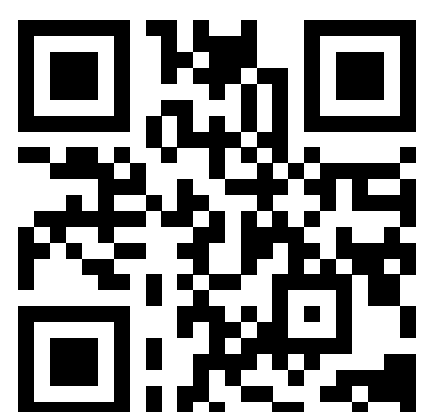
### Qualitative study on SN [A] & CompCars [D]



### Quantitative analysis on ShapeNet [A]

| Model    | Full  | w/o $\mathcal{L}_{nbr}$ | w/o PC |
|----------|-------|-------------------------|--------|
| airplane | 0.110 | 0.124                   | 0.107  |
| bench    | 0.159 | 0.188                   | 0.206  |
| car      | 0.168 | 0.179                   | 0.173  |
| chair    | 0.253 | 0.319                   | 0.527  |
| table    | 0.243 | 0.246                   | 0.598  |
| mean     | 0.187 | 0.211                   | 0.322  |

PyTorch code  
& video results



[A] ShapeNet, Chang et al., arXiv 2015  
 [B] PASCAL, Xiang et al., WACV 2014  
 [C] CUB-200, Welinder et al., Cal. Inst. of Tech. 2010  
 [D] CompCars, Yang et al., CVPR 2015  
 [E] LSUN, Yu et al., arXiv 2016

[1] Soft Rasterizer, Liu et al., ICCV 2019  
 [2] DVR, Niemeyer et al., CVPR 2020  
 [3] CMR, Kanazawa et al., ECCV 2018  
 [4] SDF-SRN, Lin et al., NeurIPS 2020  
 [5] TARS-3D, Duggal and Pathak, CVPR 2022

[6] IMR, Tulsiani et al., arXiv 2020  
 [7] UMR, Li et al., ECCV 2020  
 [8] UCMR, Goel et al., ECCV 2020  
 [9] SMR, Hu et al., CVPR 2021