SAGNet: Structure-aware Generative Network for 3D-Shape Modeling

Zhijie Wu	Shenzhen University
Xiang Wang	Shenzhen University
Di Lin	Shenzhen University
Dani Lischinski	The Hebrew University of Jerusalem
Daniel Cohen-Or	Shenzhen University
Hui Huang	Shenzhen University

Outline

01 Motivation

02 Method

03 Result

04 Conclusion

Outline

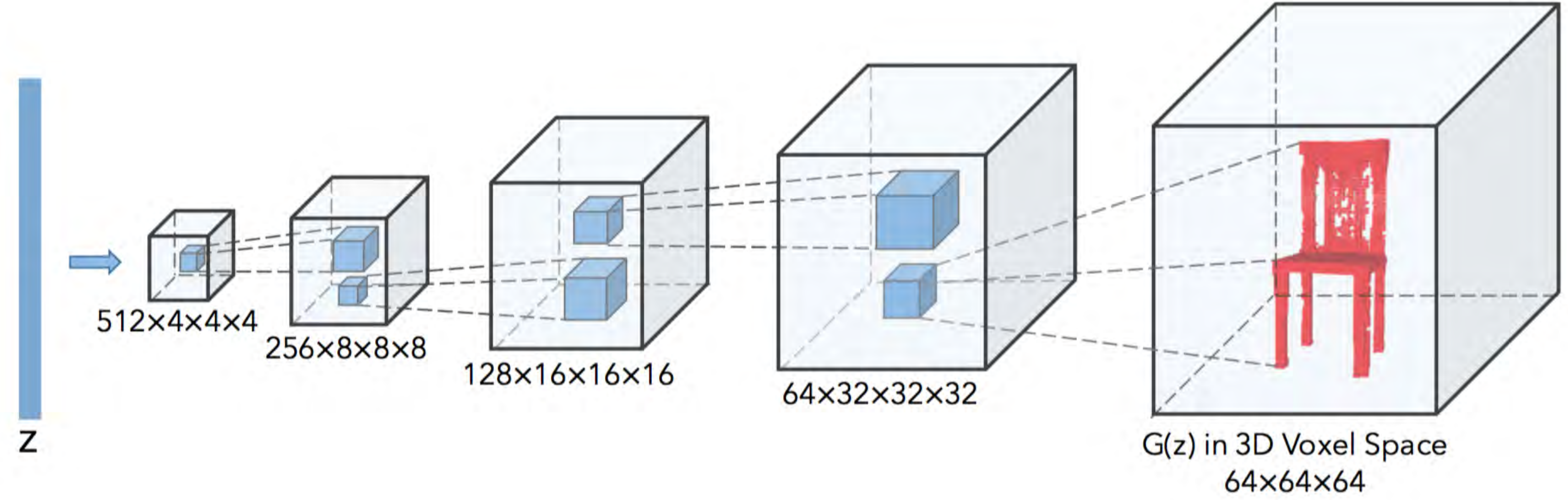
01 Motivation

02 Method

03 Result

04 Conclusion

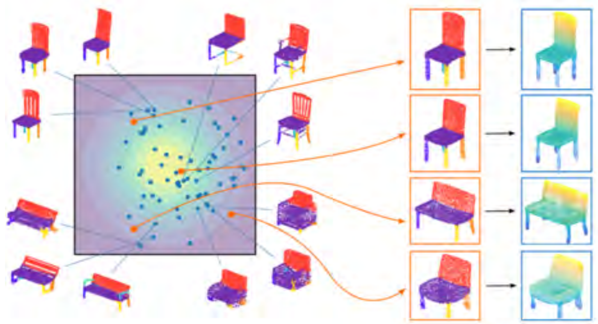
Background



[Wu et al. NIPS 2016. "Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling"]

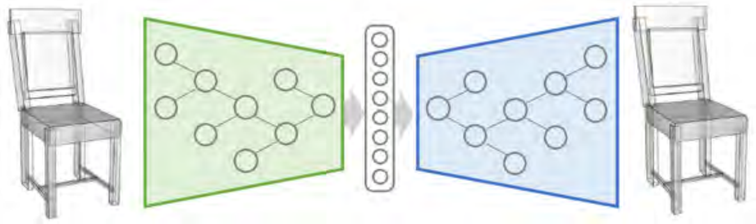
Related Work

work 01



[Nash et.al 2017]

work 02

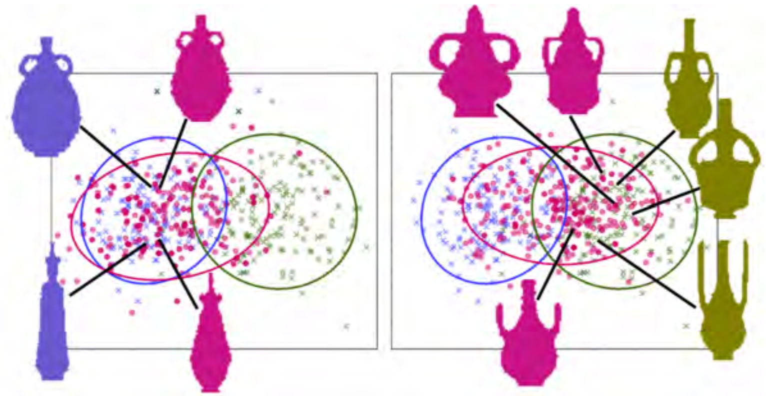


[Li et.al 2017]

work 03

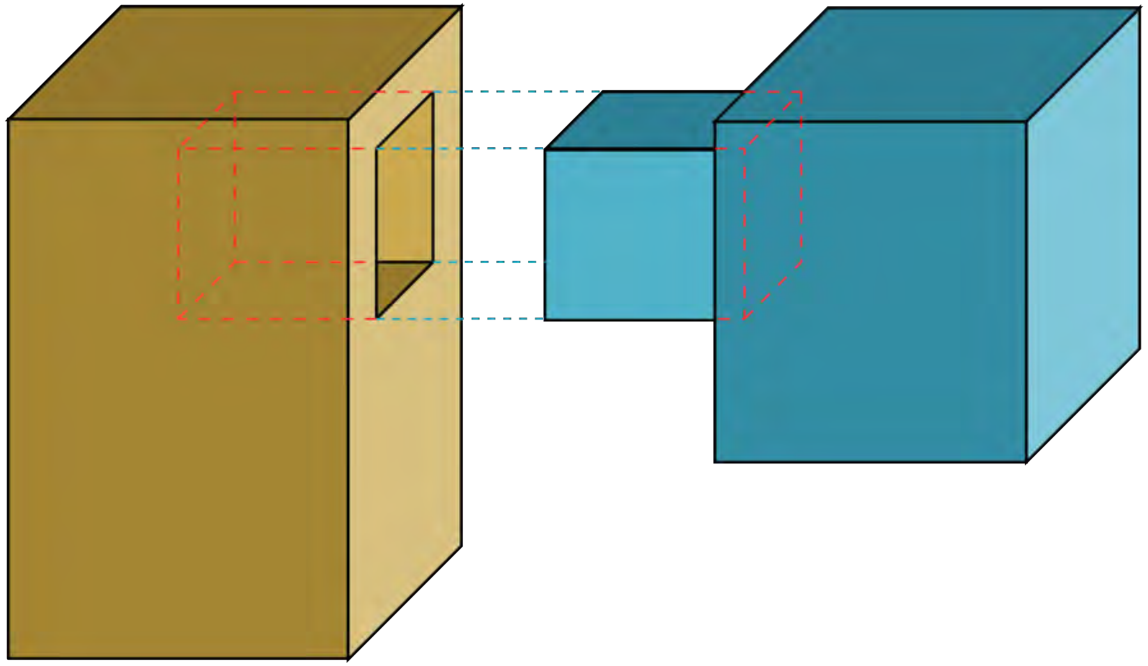
[Wang et.al 2018]

work 04



[Schor et.al 2018]

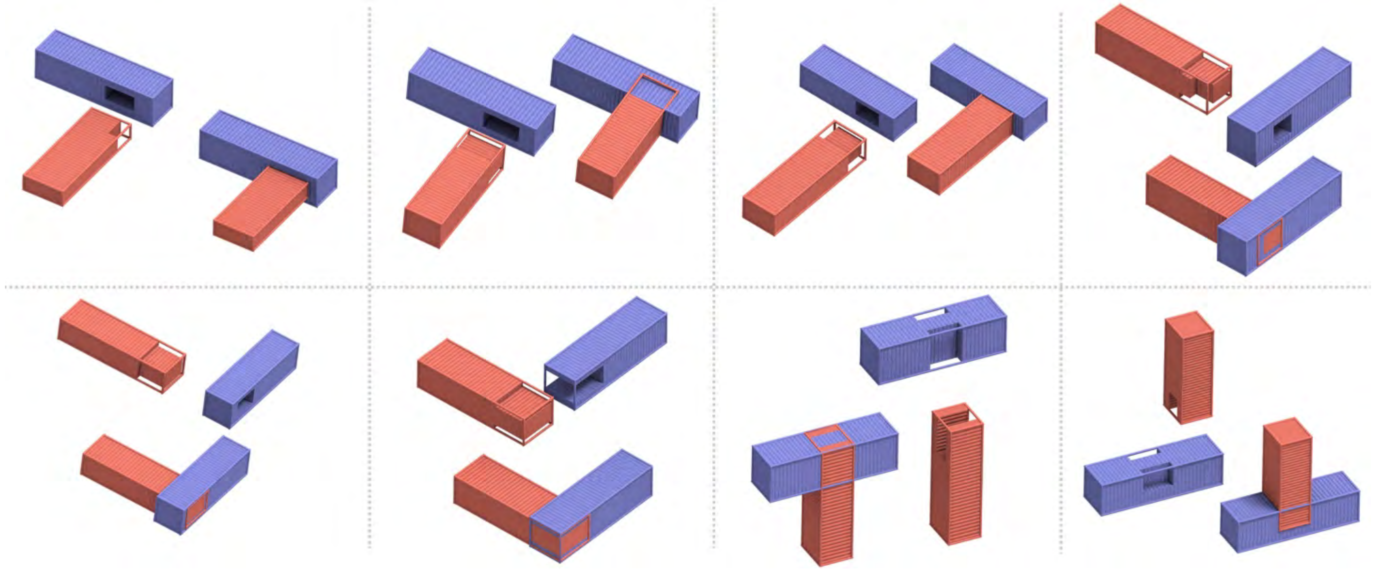
Motivations



Tenon-mortise joint

Each tenon-mortise joint consists of two parts. One part, in orange, has a cavity into which the second part, in blue, exactly fits. Thus, if the relation between the relative position of the parts and their geometry is not learned well by the network, it is unlikely that the network would succeed in generating the orange parts with a correctly sized and placed cavity.

Motivations



Tenon-mortise joint

Each tenon-mortise joint consists of two parts. One part, in orange, has a cavity into which the second part, in blue, exactly fits. Thus, if the relation between the relative position of the parts and their geometry is not learned well by the network, it is unlikely that the network would succeed in generating the orange parts with a correctly sized and placed cavity.

Outline

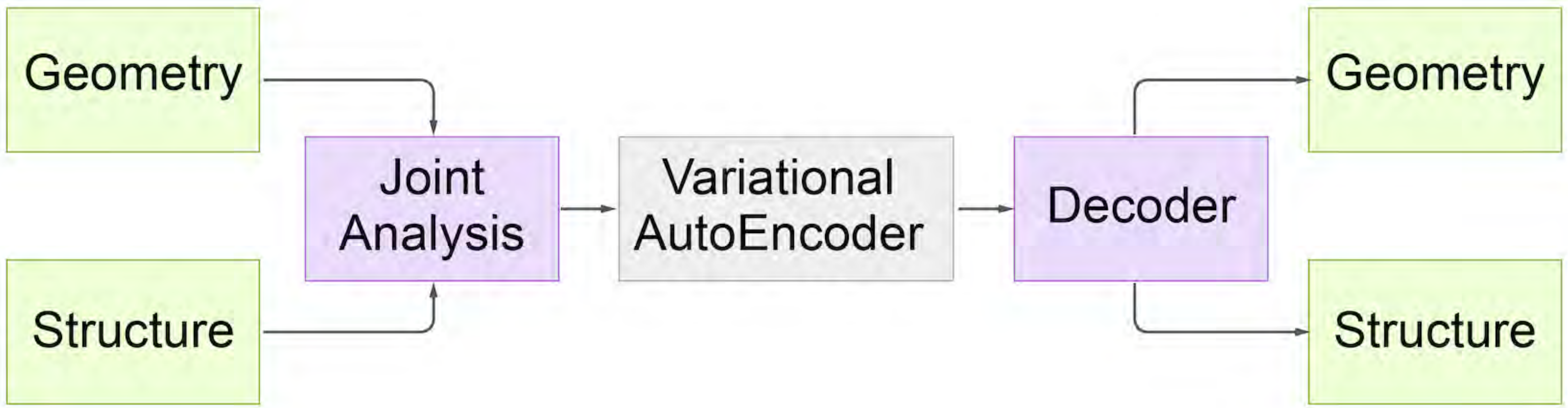
01 Introduction

02 Method

03 Results

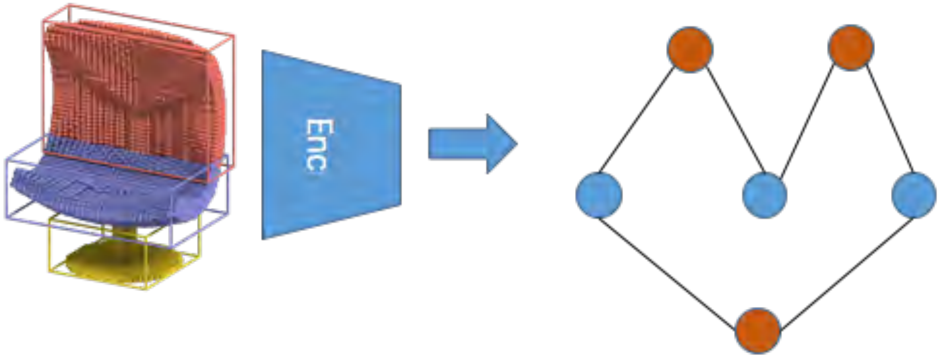
04 Conclusion

Framework Pipeline



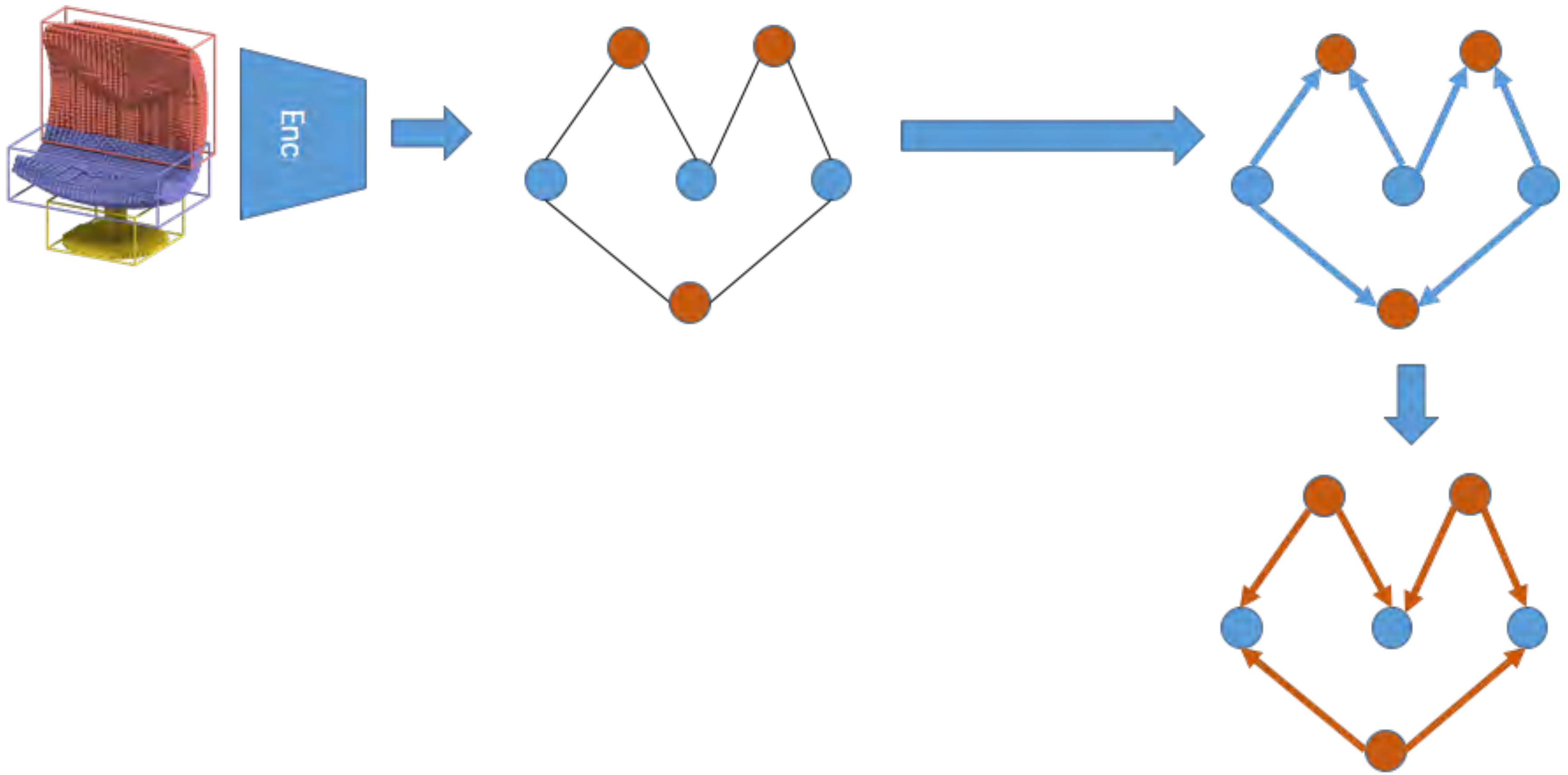
Framework Pipeline

To better jointly analyse the geometry and structure information, we follow the message passing strategy to update these two information iteratively, share similar spirit of [Xu et al. 2017].



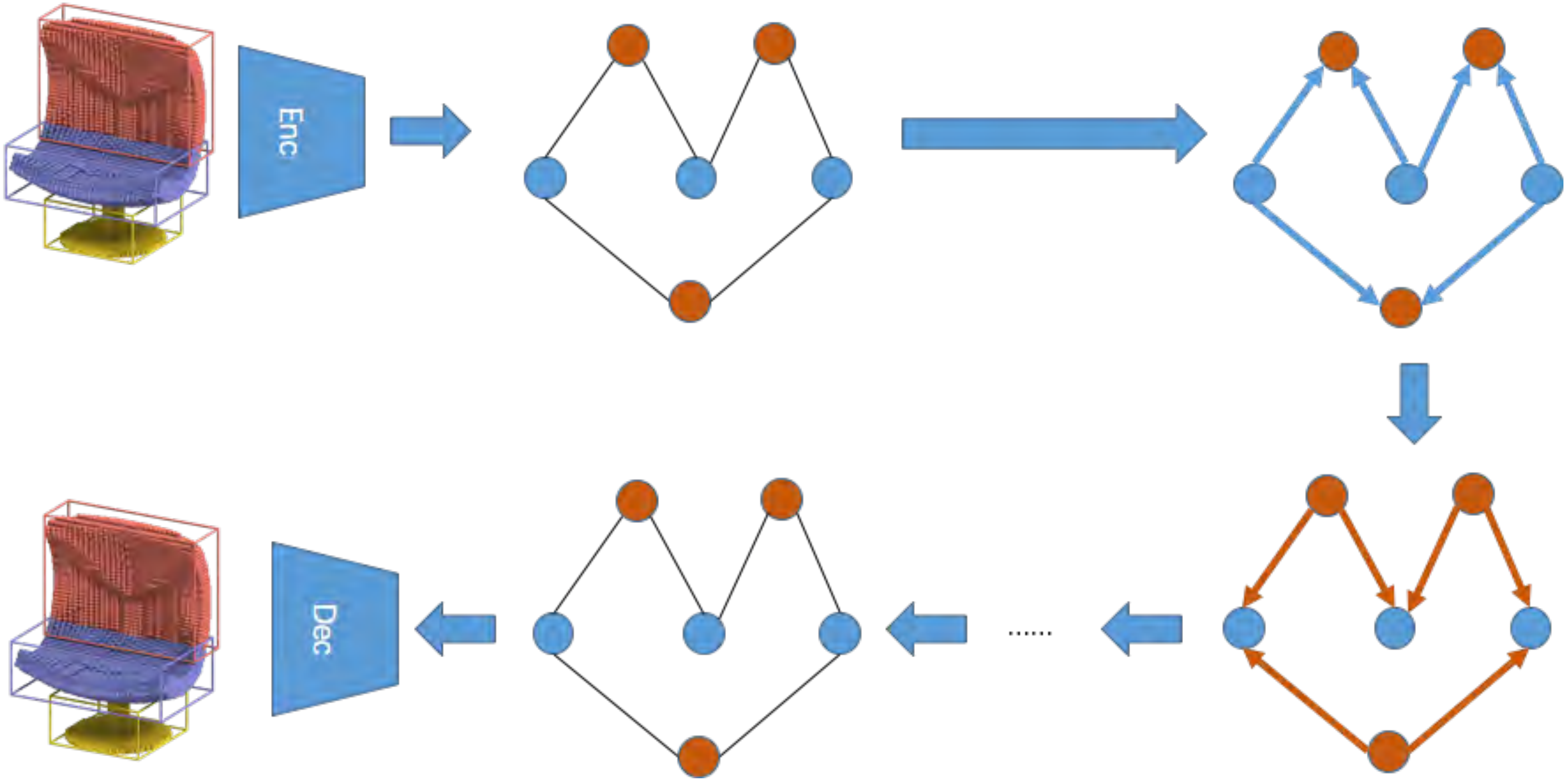
Framework Pipeline

To better jointly analyse the geometry and structure information, we follow the message passing strategy to update these two information iteratively, share similar spirit of [Xu et al. 2017].

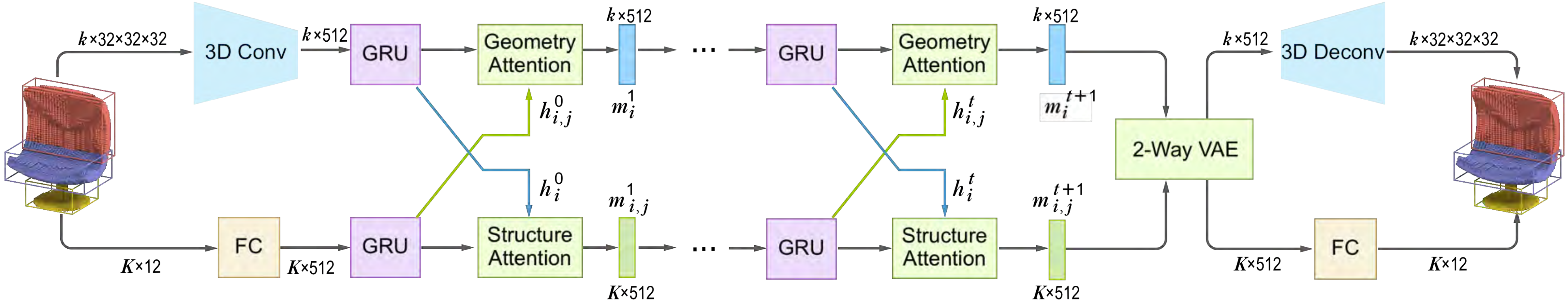


Framework Pipeline

To better jointly analyse the geometry and structure information, we follow the message passing strategy to update these two information iteratively, share similar spirit of [Xu et al. 2017].



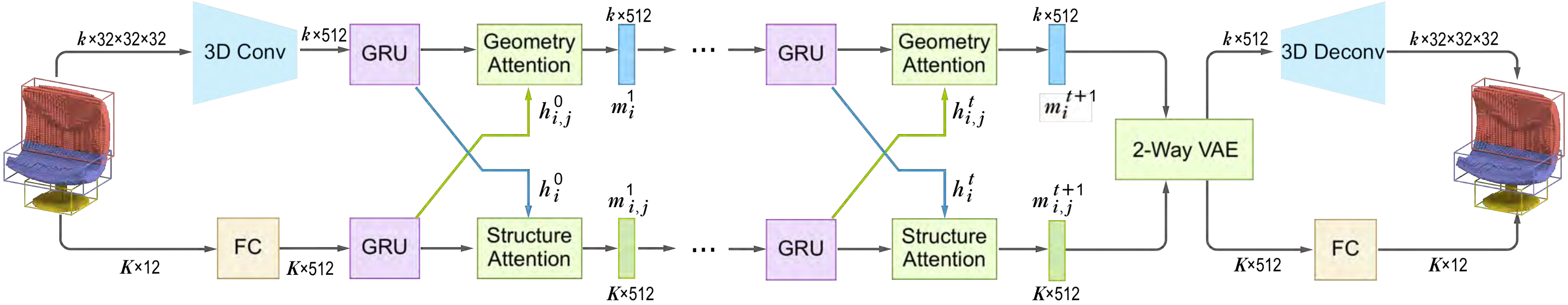
Architecture

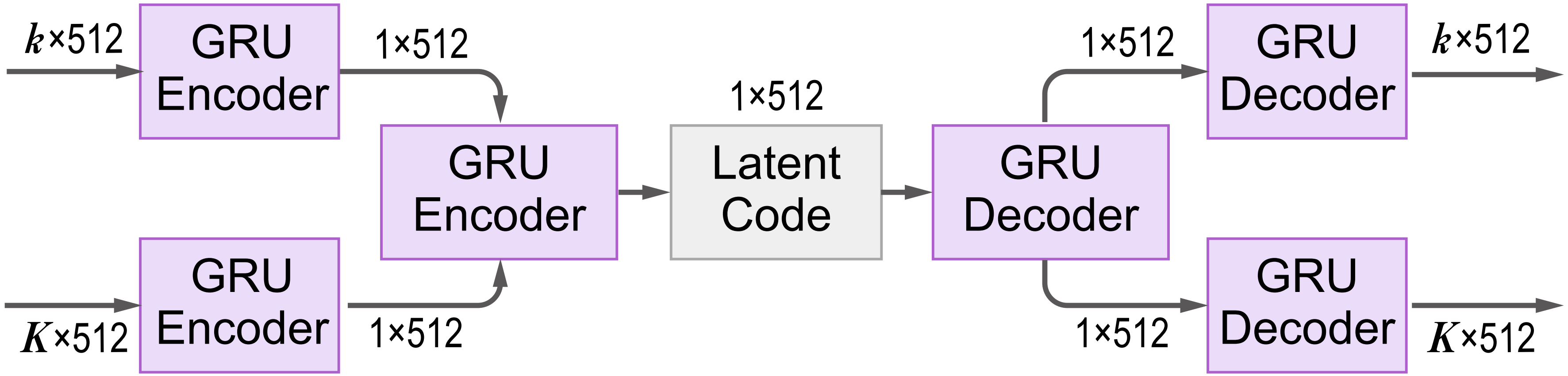


$$m_i^{t+1} = \sum_{j \neq i} f([h_i^t, h_{i,j}^t])h_{i,j}^t,$$

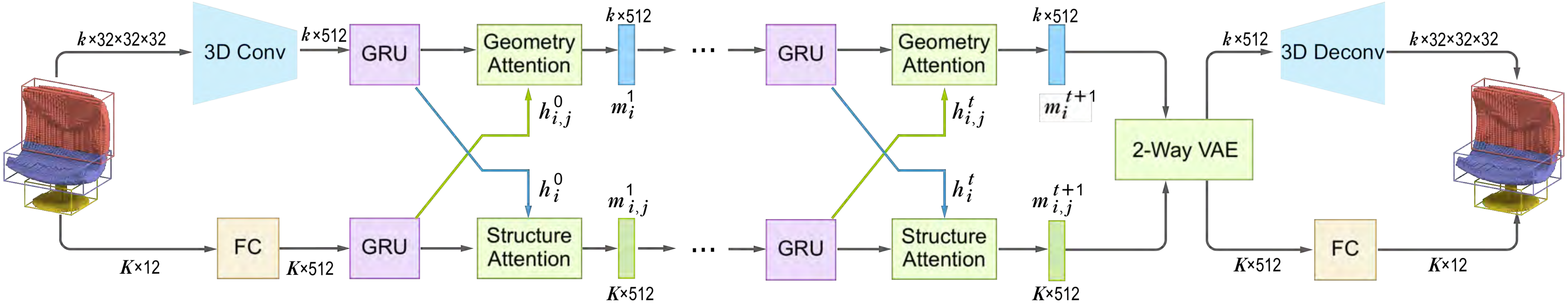
$$m_{i,j}^{t+1} = f([h_{i,j}^t, h_i^t])h_i^t + f([h_{i,j}^t, h_j^t])h_j^t$$

Architecture





Training strategy



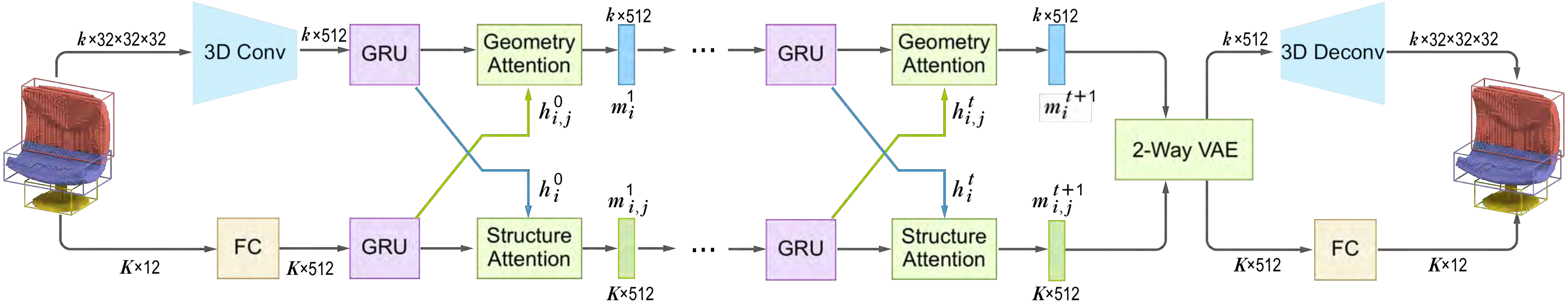
$$L_s = L_f + \lambda L_{KL}$$

where

$$L_{KL} = KL(q_\phi(z|x,y,c)||p_\phi(z|c)),$$

$$L_f = -E_{q_\phi(z|v,b,c)}[log(p_\varphi(v,b|z,c))]$$

Training strategy



$$L_s = L_f + \lambda L_{KL}$$

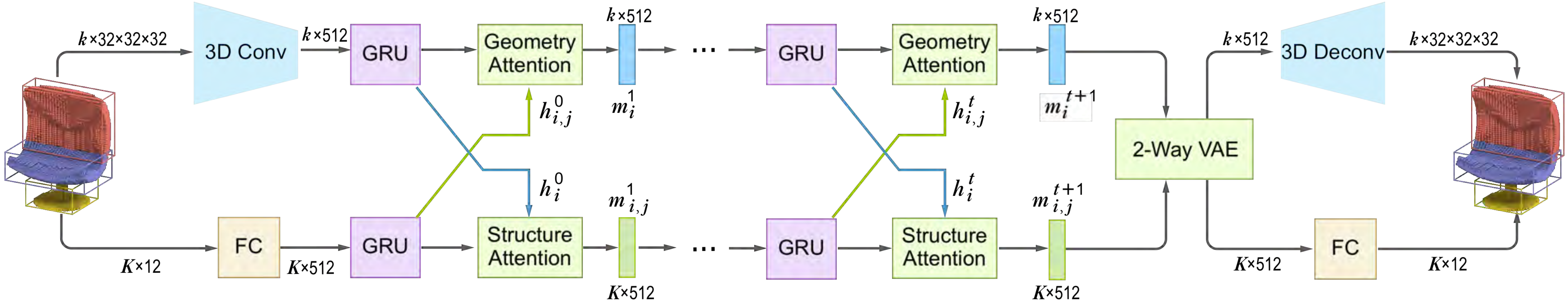
where

$$L_{KL} = KL(q_\phi(z|x,y,c)||p_\phi(z|c)),$$

$$L_f = -E_{q_\phi(z|v,b,c)}[log(p_\varphi(v,b|z,c))]$$

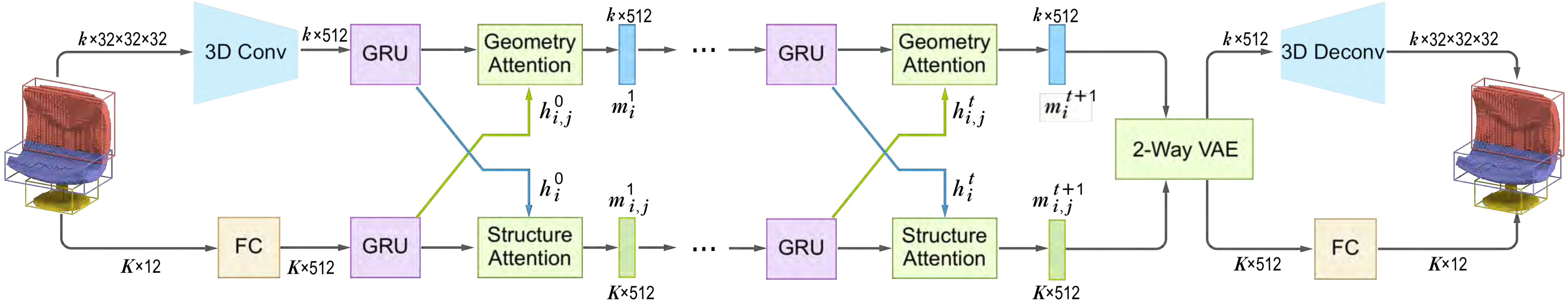
Posterior Collapse Problems!

Training strategy



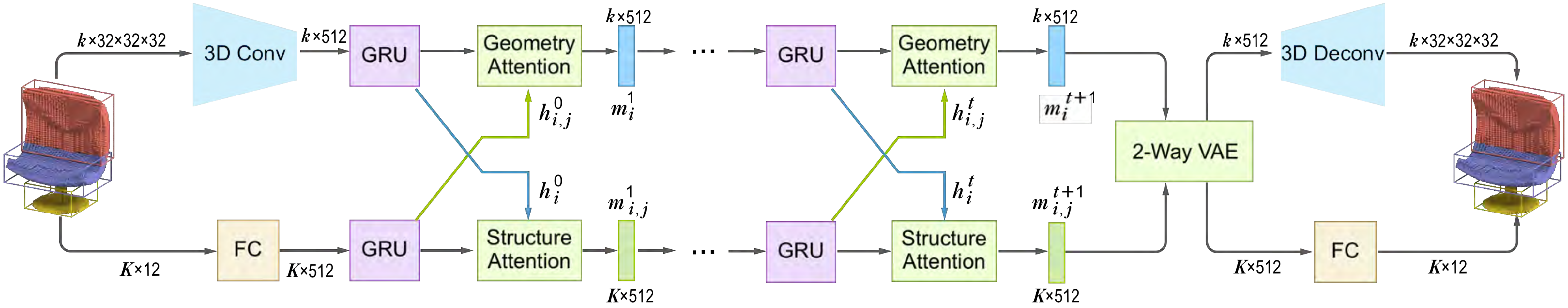
VAEs are hard to train when combined with powerful autoregressive decoders or RNNs. This is due to the "posterior collapse" problem: the model ends up relying solely on the properties of the decoder while ignoring the latent variables, which become uninformative.
[Bowman et al. Generating Sentences from a Continuous Space]

Training strategy(First phase)



$$L_f = -E_{q_\phi(z|v,b,c)}[log(p_\varphi(v,b|z,c))]$$

Training strategy(Second phase)



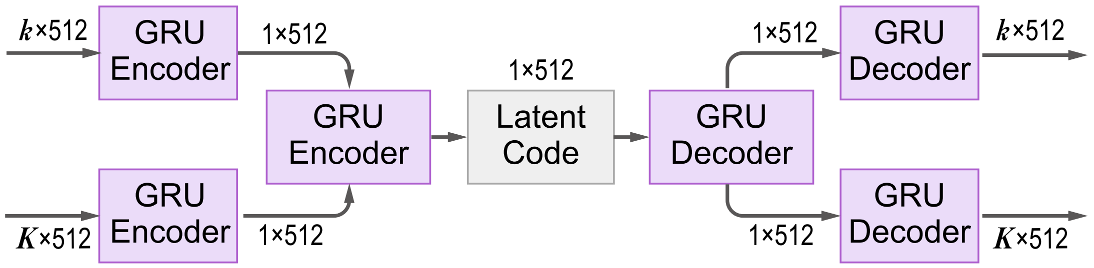
$$L_s = L_f + \lambda L_{KL}$$

where

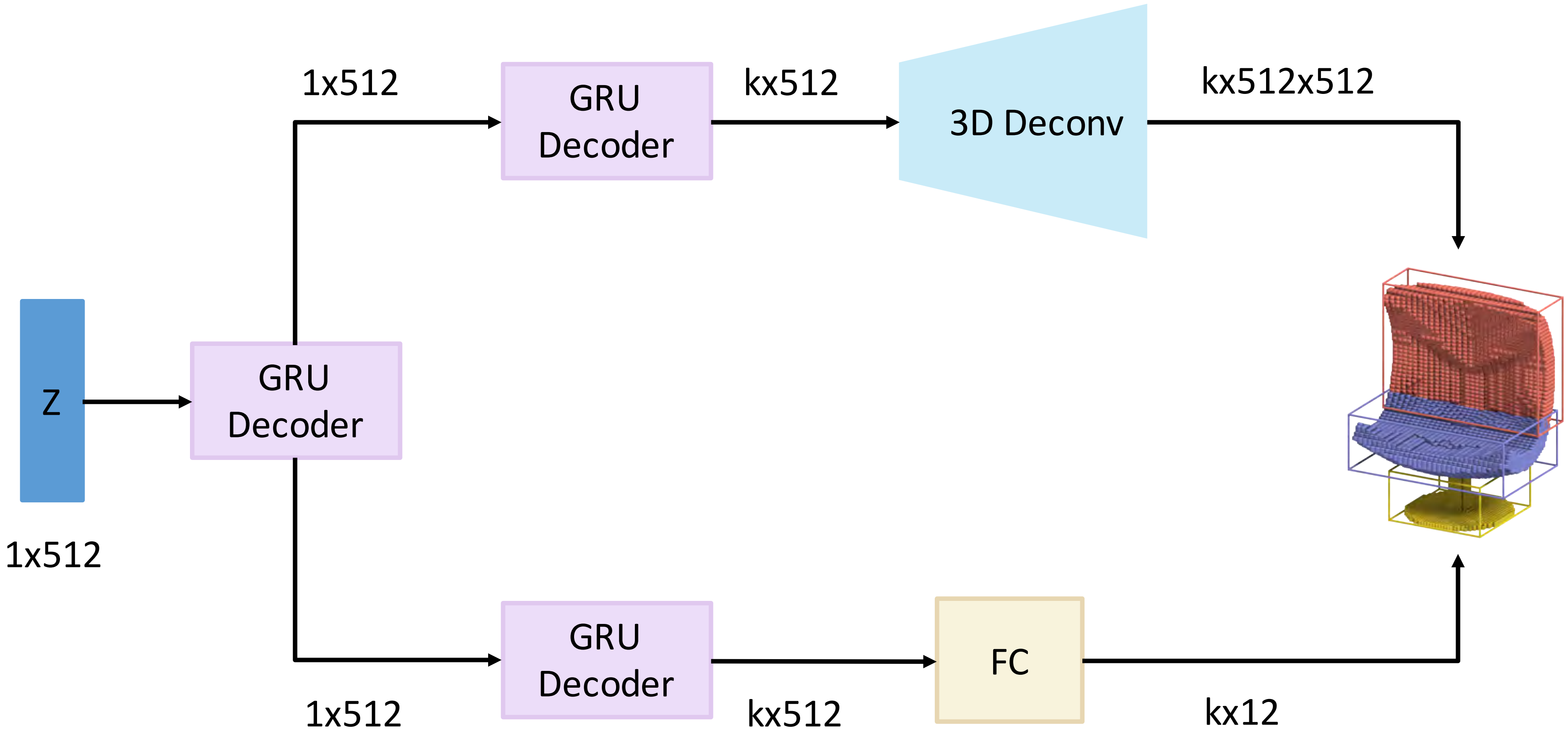
$$L_{KL} = KL(q_\phi(z|x,y,c)||p_\phi(z|c)),$$

$$\boxed{R = \sum_{i=1}^{k} ||h'_i - h_i||_2^2 + \sum_{i=1}^{k} \sum_{j=i+1}^{k} ||h'_{i,j} - h_{i,j}||_2^2,}$$

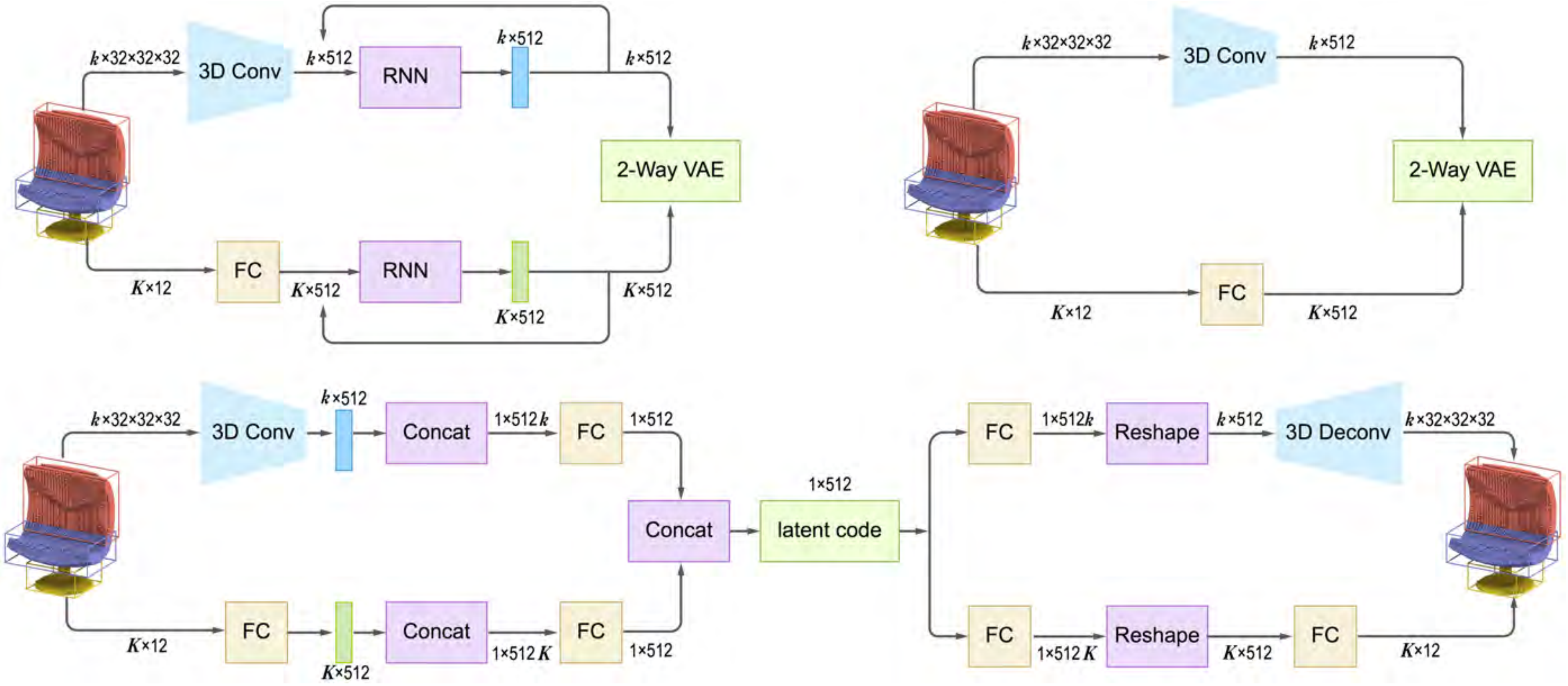
$$L_f = -E_{q_\phi(z|v,b,c)}[log(p_\varphi(v,b|z,c))]$$



Testing Procedure



Ablation study frameworks



The architectures of three ablation study baseline models. The top left diagram denotes the **No-attention** baseline model. The top right diagram corresponds to **No-GRU** baseline model. The diagram that lies at bottom indicates the baseline model of **Concatenation**.

Outline

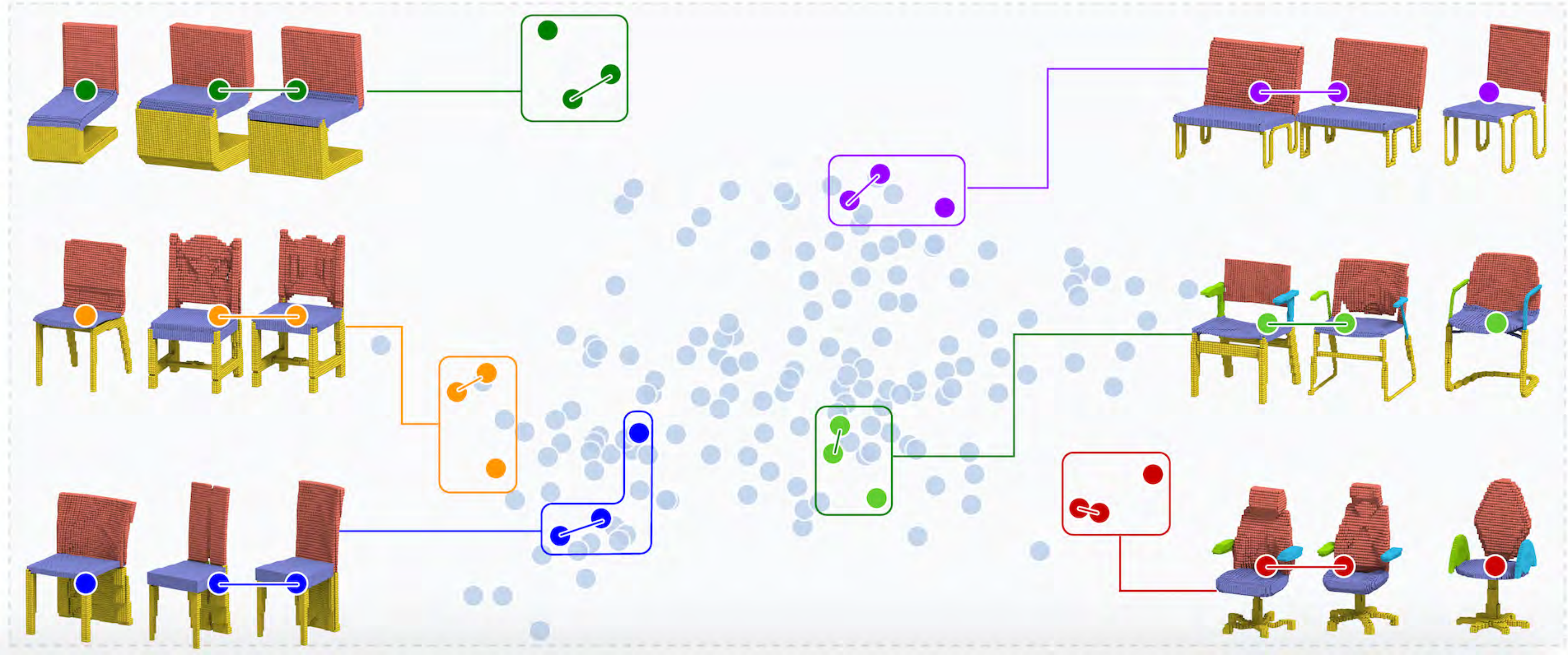
01 Introduction

02 Method

03 Results

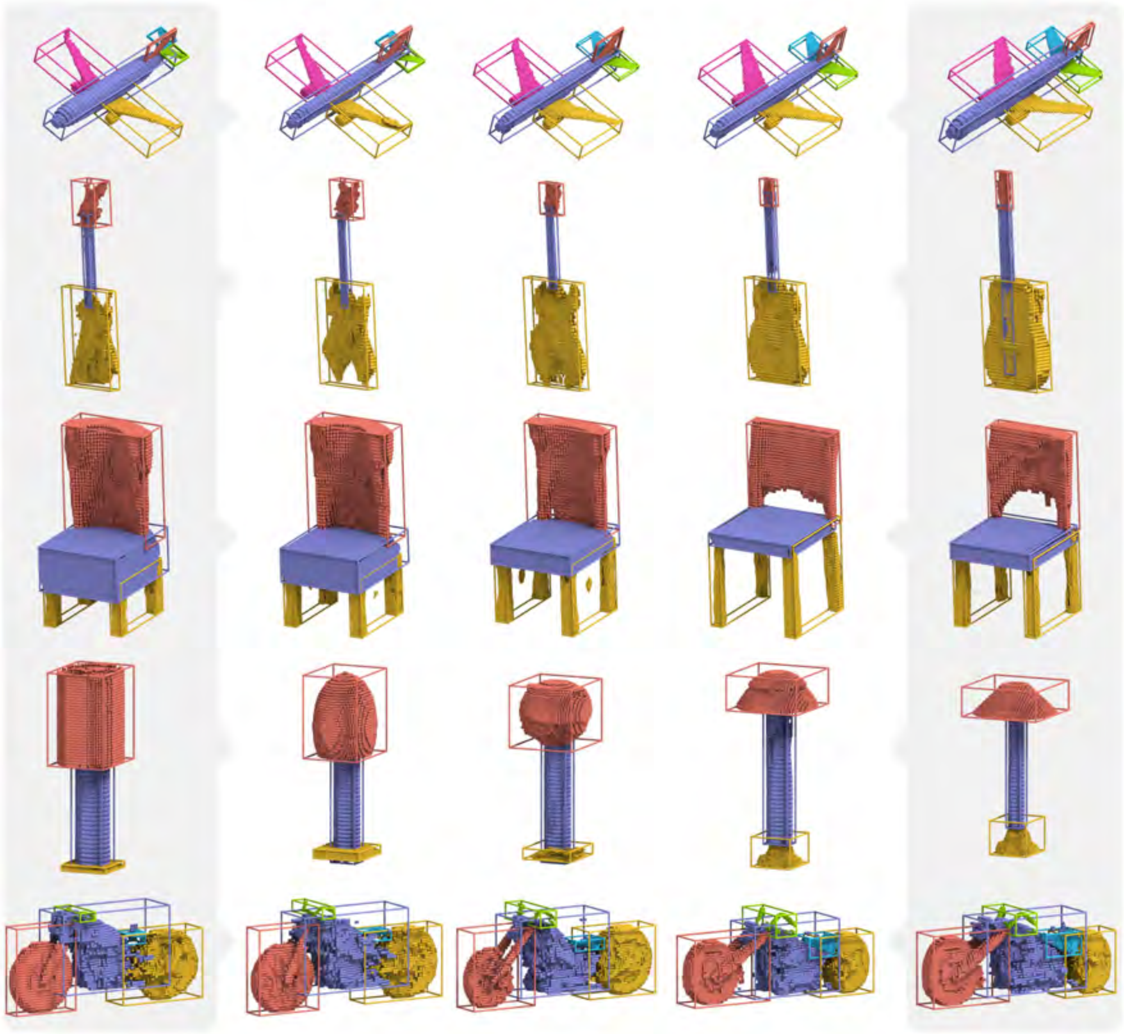
04 Conclusion

Embedding space



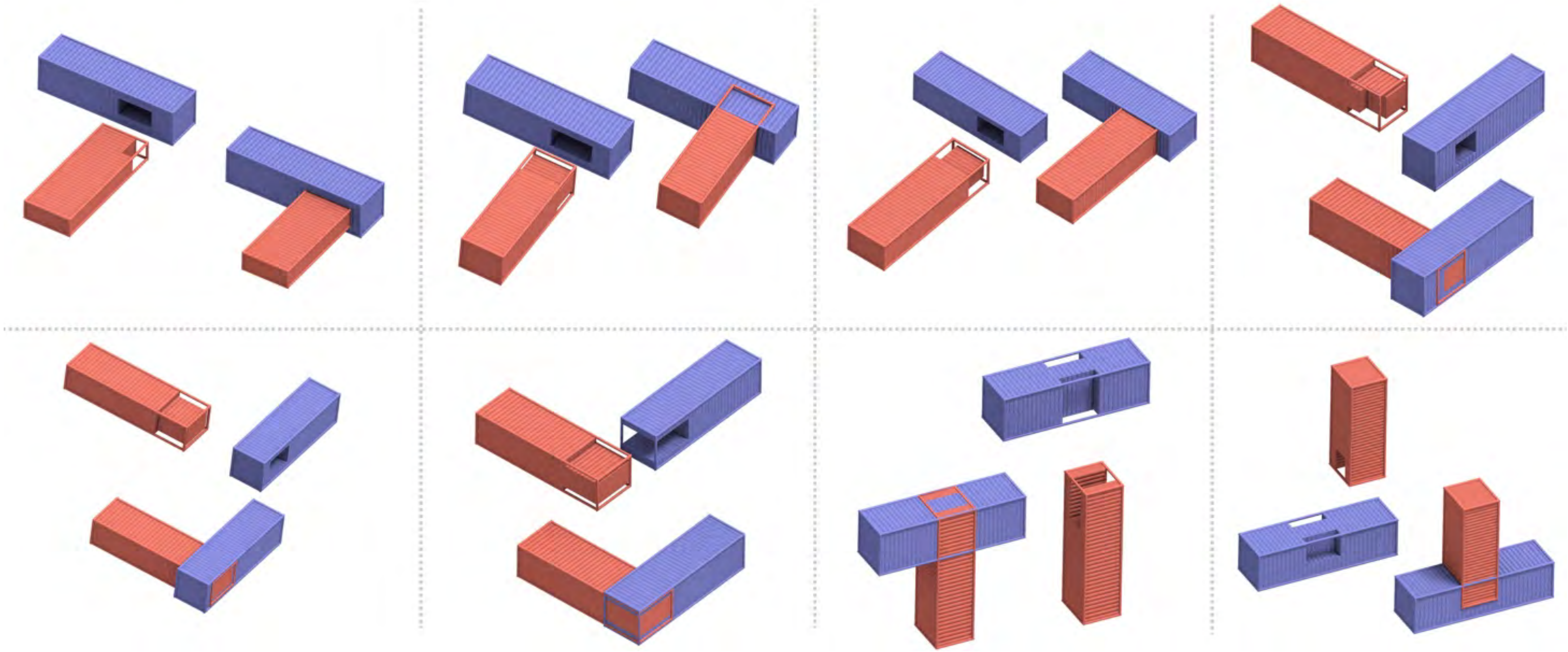
kNN results

Interpolation results

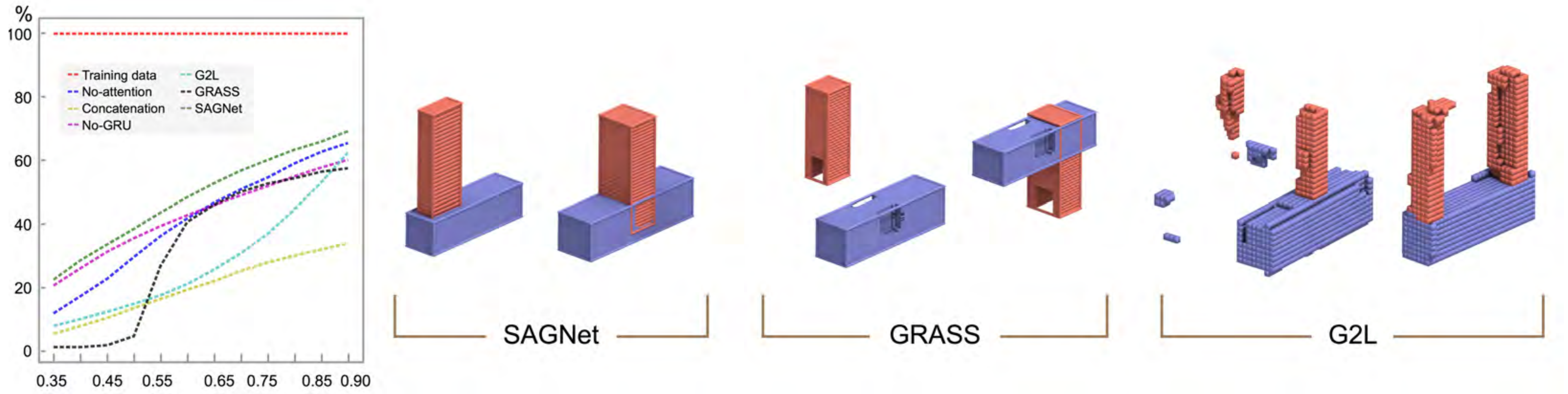


The synthesis result comparison

Tenon-mortise joint



Results on tenon-mortise joints

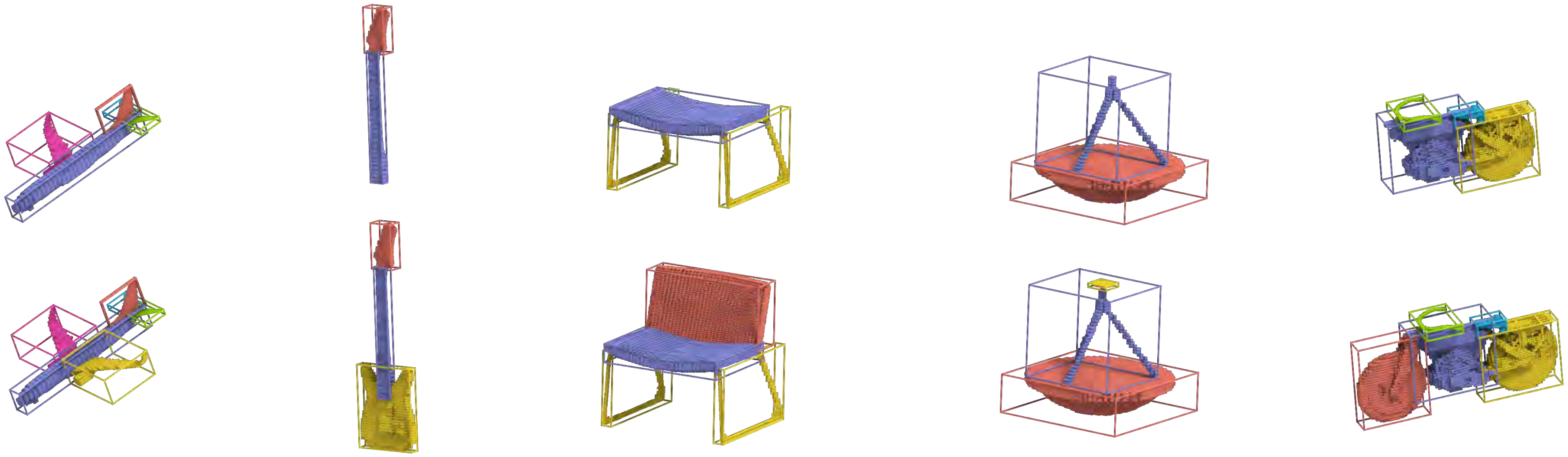


We randomly generate 1000 test samples with the trained network and measure how well the convex parts fit into the cavity of the non-convex ones. To quantitatively measure the fitting accuracy, we calculate, Ro (Re), the portion of occupancy(empty) voxels of the non-convex part that are overlapped with occupancy voxels of the convex part. Then the smaller score R = 1 − (Re −Ro) indicates better fitting status between the two parts.

Results on tenon-mortise joints

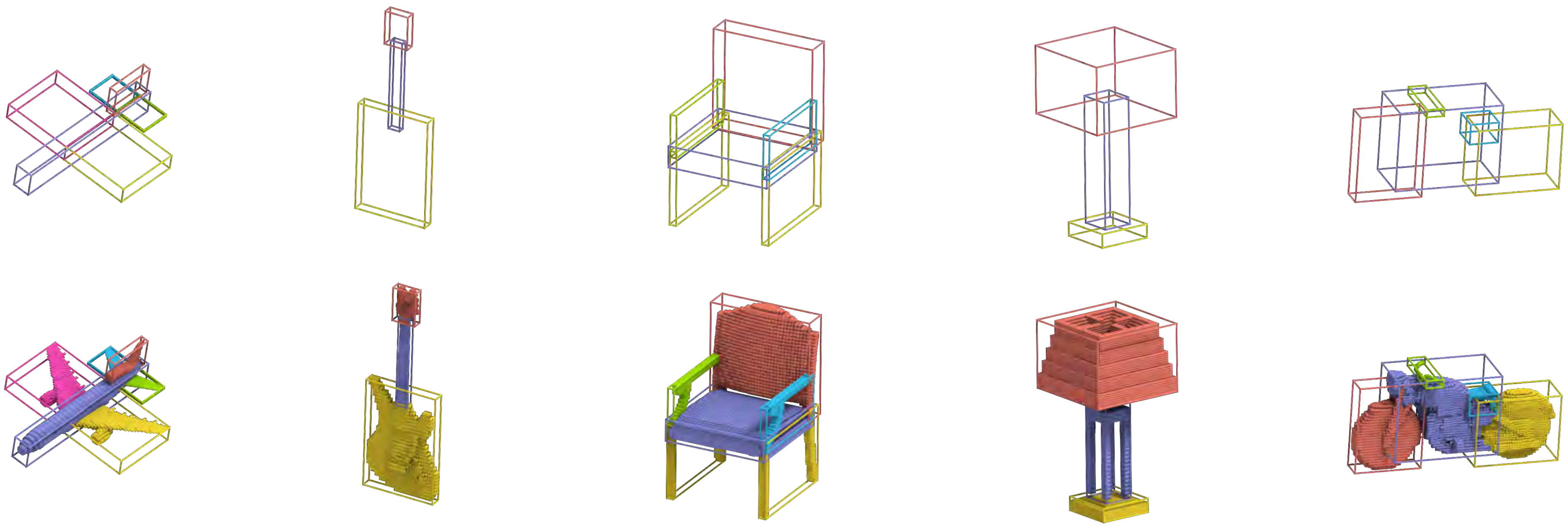
	Scores on Synthetic Data						
	SAGNet	No-attention	No-GRU	Concatenation	G2L	GRASS	Training data
R_o	0.291	0.343	0.301	0.307	0.086	0.554	0.0
R_e	0.585	0.593	0.544	0.321	0.298	0.683	1.0
$R_{over} = R_e - R_o$	**0.294**	0.250	0.243	0.013	0.211	0.129	1.0
Inception Score	**6.26**	6.01	5.95	5.32	5.44	1.95	7.98

Application Results



Shape completion

Application Results



Structure-geometry Translation

Application Results

Geometry-structure Translation

Outline

01 Introduction

02 Method

03 Results

04 Conclusion

Conclusion

1

We use the semantically segmented training set to learn the implicit dependencies between geometry of parts and their spatial arrangement.

2

The presented network allows us to generate 3D shapes with separate control over their geometry and structure.

3

The designed tenon-mortise joints can quantitatively measure the learning ability to capture the dependencies between geometry and structure.

Several avenues for future work

work 01

Extend the two-branch auto-encoder to process other properties or develop a k-way auto-encoder where k properties are learned in parallel.

work 02

Learn more about the geometry of the parts by employing part-level generators to potentially generate finer details.

work 03

Strengthen the flexibility to model various 3D shapes with different numbers of parts.

work 04

Explore the possibility to apply the idea of joint analysis to other areas and capture more complex geometry and structure details.

Thank you!