

PaperReading:
[CVPR 2017 Oral]

A point Set Generation Network for 3D
Object Reconstruction from a Single
Image



Input

Reconstructed 3D point cloud

Contributions

- We are the first to study the point set generation problem by deep learning;
- On the task of 3D reconstruction from a single image, we apply our point set generation network and significantly outperform state of the art;
- We systematically explore issues in the architecture and loss function design for point generation network;
- We propose a principled formulation and solution to address the groundtruth ambiguity issue for the 3D reconstruction from single image task.

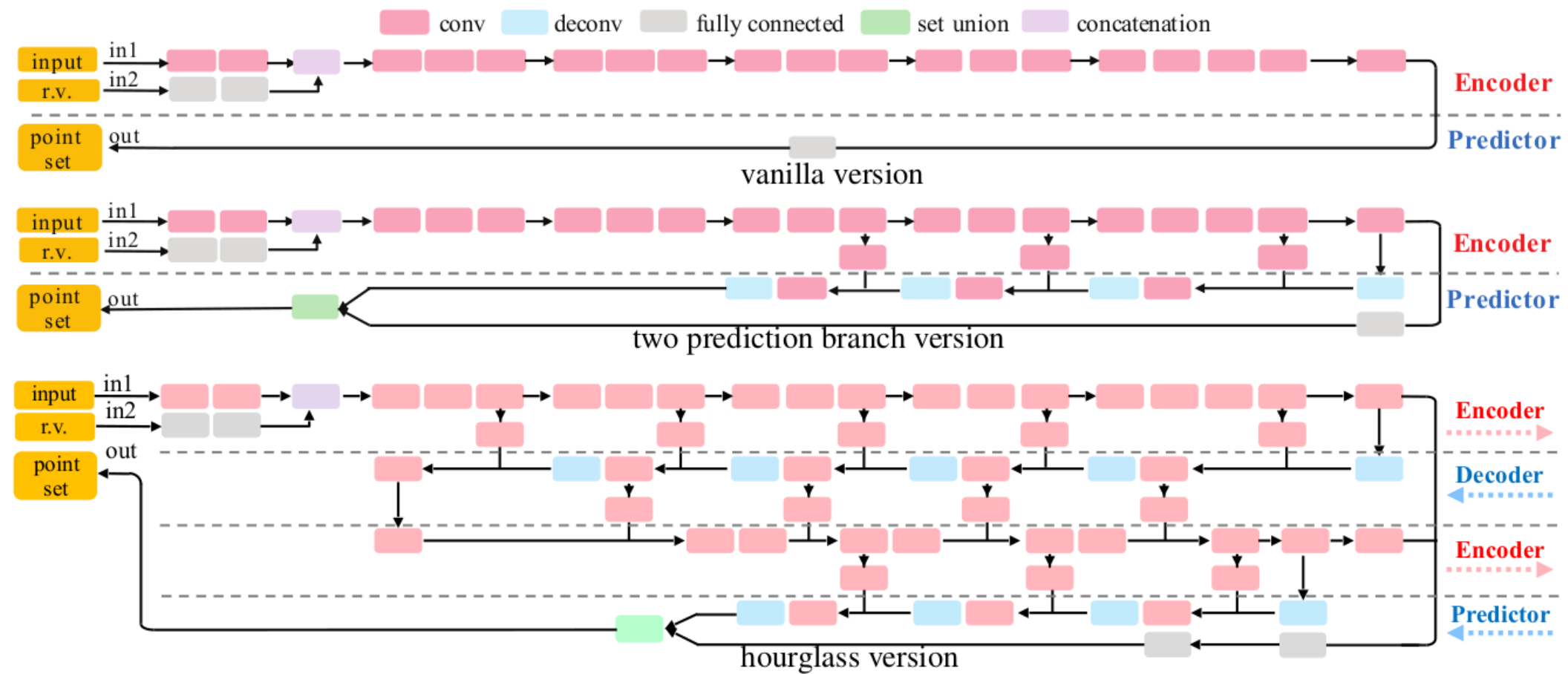
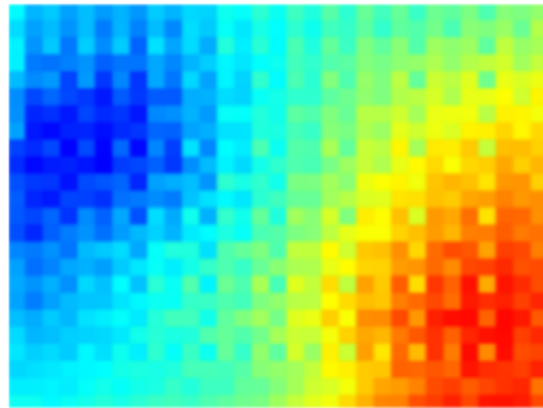
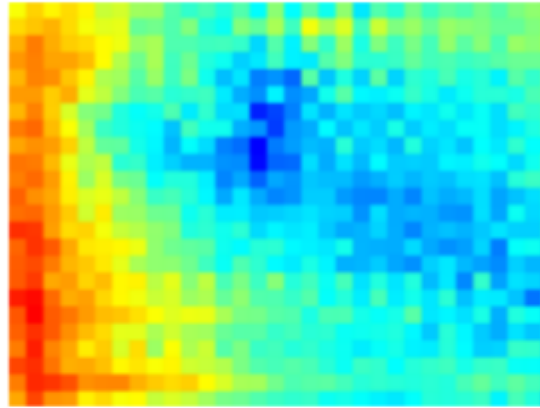


Figure 2. PointOutNet structure

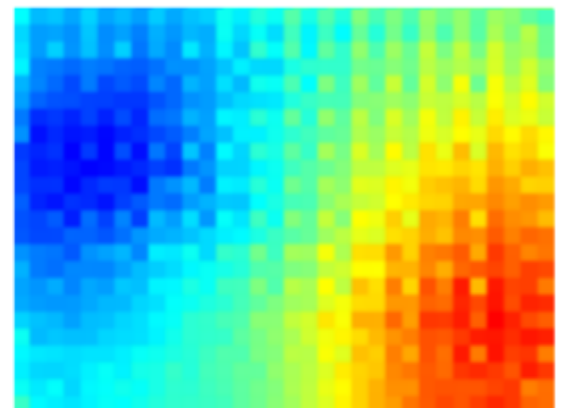
deconv branch



x-channel

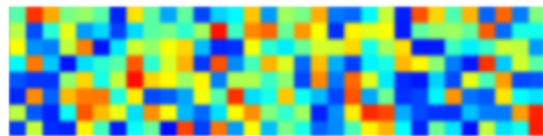


y-channel

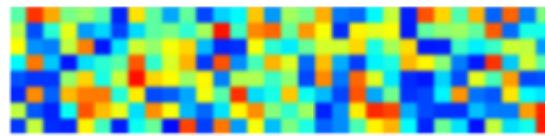


z-channel

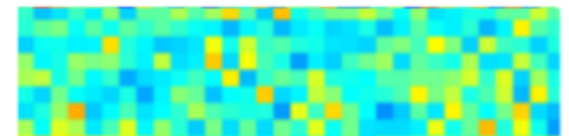
fully connected branch



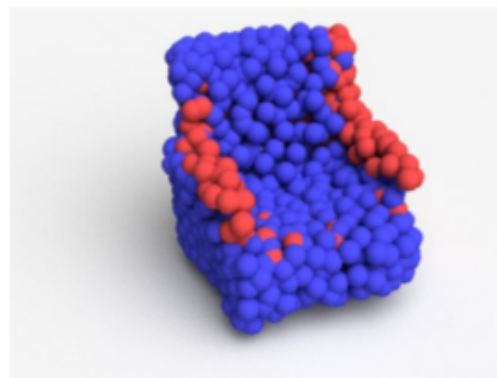
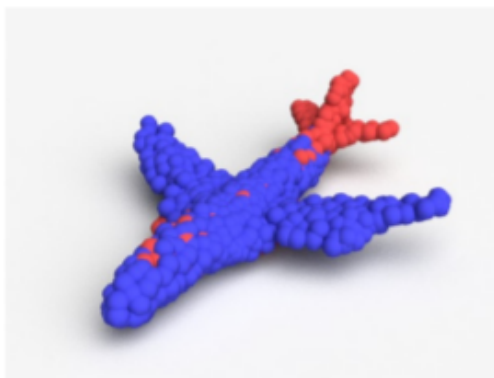
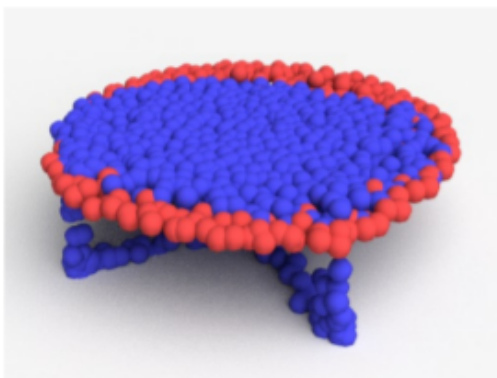
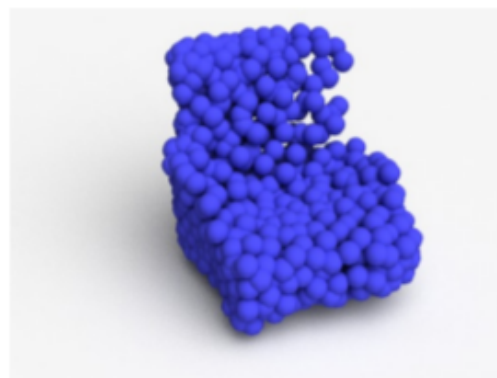
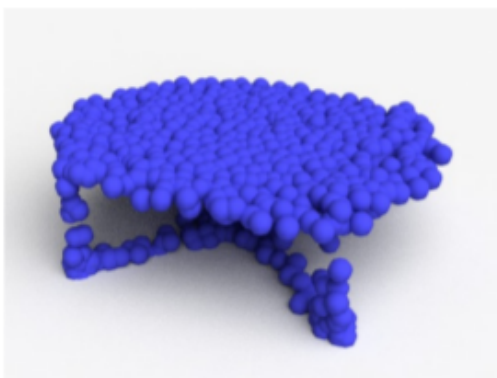
x-channel



y-channel



z-channel



Loss

$$L(\{S_i^{pred}\}, \{S_i^{gt}\}) = \sum d(S_i^{pred}, S_i^{gt})$$

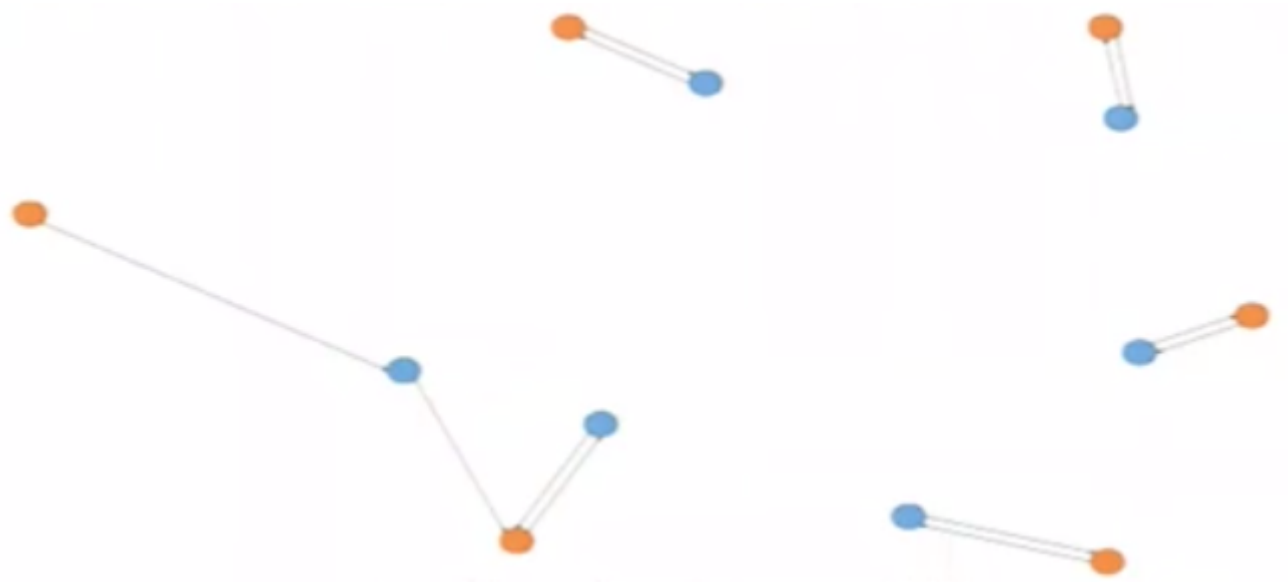
Earth Mover's Distance

$$d_{EMD}(S_1, S_2) = \min_{\phi: S_1 \rightarrow S_2} \sum_{x \in S_1} \|x - \phi(x)\|_2$$



Chamfer Distance

$$d_{CD}(S_1, S_2) = \sum_{x \in S_1} \min_{y \in S_2} \|x - y\|_2^2 + \sum_{y \in S_2} \min_{x \in S_1} \|x - y\|_2^2$$



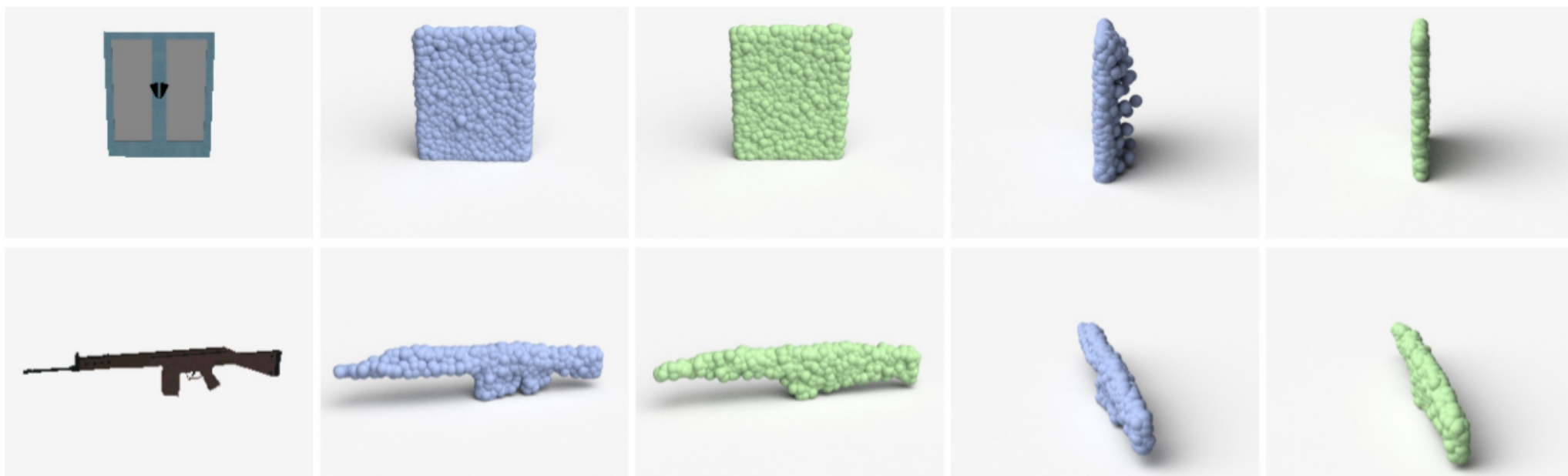

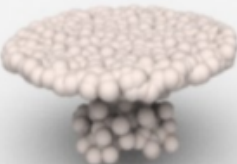




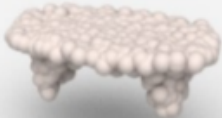









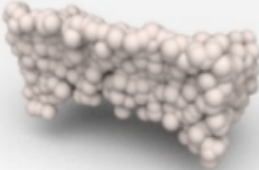
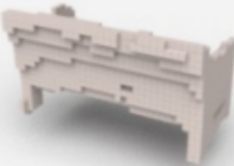
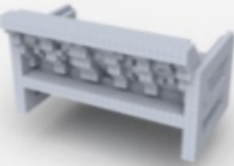



Figure 15. Comparison of predictions of networks trained by CD (blue, on the left) and EMD (green, on the right).

input image	ours	ours (post-processed)	ground truth	3D-R2N2
				
				
				
				

category	Ours	3D-R2N2		
	1 view	1 view	3 views	5 views
plane	0.601	0.513	0.549	0.561
bench	0.550	0.421	0.502	0.527
cabinet	0.771	0.716	0.763	0.772
car	0.831	0.798	0.829	0.836
chair	0.544	0.466	0.533	0.550
monitor	0.552	0.468	0.545	0.565
lamp	0.462	0.381	0.415	0.421
speaker	0.737	0.662	0.708	0.717
firearm	0.604	0.544	0.593	0.600
couch	0.708	0.628	0.690	0.706
table	0.606	0.513	0.564	0.580
cellphone	0.749	0.661	0.732	0.754
watercraft	0.611	0.513	0.596	0.610
mean	0.640	0.560	0.617	0.631

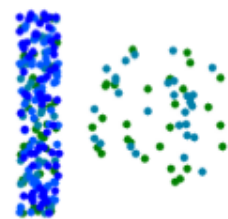
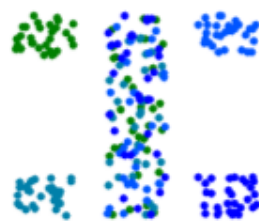
Generation of Multiple Plausible Shapes



Generation of Multiple Plausible Shapes



Input



EMD
mean



CD
mean



(a)

(b)

(c)

(d)

