

## National University of Defense Technology

## Introduction

**Goal:** Recover a cuboid structure of the object parts and their mutual relations from a single RGB image.



## > Contribution:

- **Problem**: Faithful and detailed recovery of 3D part structure of an object from 2D images.
- Architecture: Integrating a *convolutional structure* masking network for object discerning and a recursive *structure recovery network* for structure recovery.
- **Prototype applications**: Recovered box structures are used: 1) to refine volumetric 3D shapes reconstructed from single images and 2) to assist structure-aware editing of 2D images.



# Im2Struct: Recovering 3D Shape Structure from a Single RGB Image Chengjie Niu Jun Li Kai Xu





An overview of our network architecture. The structure masking network is a two-scale CNN which is trained to produce a contour mask for the object of interest. The structure recovery network first fuses the feature map of the masking network and the CNN feature of the original image, and then decode the fused feature recursively into a cuboid structure. The red arrows in the resulting chair structure to the right most indicate recovered reflectional symmetries between chair legs.



entropy loss for node classification.

### **Evaluation:**

• Hausdorff Error:  $\frac{1}{2T} \sum_{i=1}^{T} (D(S_i, S_i^{gt}) + D(S_i^{gt}, S_i))$ 

 $D(S_1, S_2) = \frac{1}{n} \sum_{B_j^1 \in S_1} \min_{B_k^2 \in S_2} H(B_j^1, B_k^2) \quad H(B^1, B^2) = \max_{p \in B^1} \min_{q \in B^2} ||p - p|| = 0$ 

• Thresholded Accuracy:  $\delta = H(B_i, B_i^*)/L(B_i^*) < thresholded$ 

	Method	Overall Pi	xel	Per-Class	
	single-scale	0.953		0.917 0.964	
	two-scale (w/o jump)	0.982			
	two-scale (with jump)	0.988		0.983	
	Mathad	Hausdorff	Thre	esholded Acc.	
	Method	Error	$\delta <$	0.2	$\delta < 0.1$
-q	Vanilla VGG-16	0.0980	96.8%		67.8%
<b>1</b>	Structure masking (VGG-16)	0.0894	97.8%		75.3%
	Vanilla VGG-19	0.0922	96.4%		72.2%
shold	Structure masking (VGG-19)	0.0846	97.6	5%	78.5%





A comparison of part-based 3D shape reconstruction with two alternatives from single view images. Note, however, our method recovers only bounding boxes of shape parts.



A Google image search challenge to our method for 3D shape structure recovery. Images are returned from Google image search. The recovered box structure is shown under each image. Failure cases are marked with red boxes.