ObjectNet3D: A Large Scale Database for 3D Object Recognition

Yu Xiang, Wonhui Kim, Wei Chen, Jingwei Ji, Christopher Choy,
Hao Su, Roozbeh Mottaghi, Leonidas Guibas and Silvio Savarese

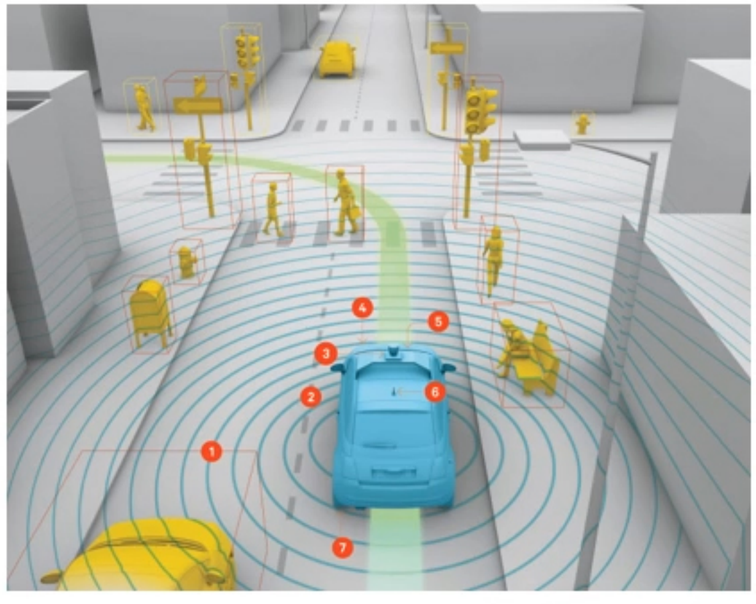
ECCV 2016

Computational Vision
& Geometry Lab

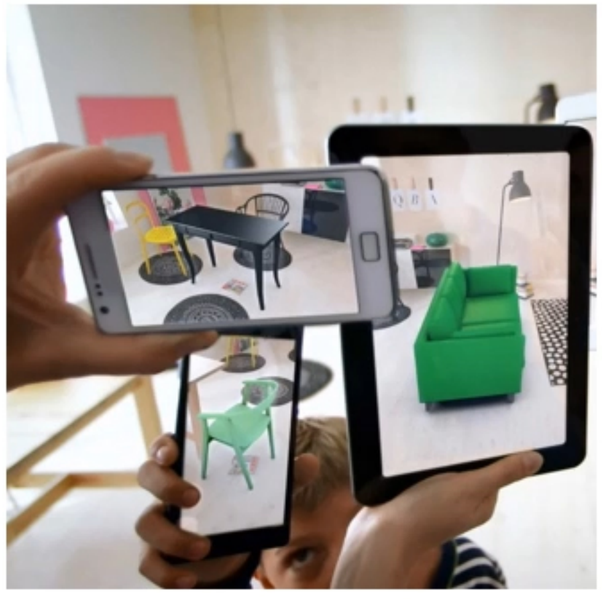
Recognizing the 3D Properties of Objects

- 3D location, 3D pose, 3D shape, etc.
- Applications

Robotics

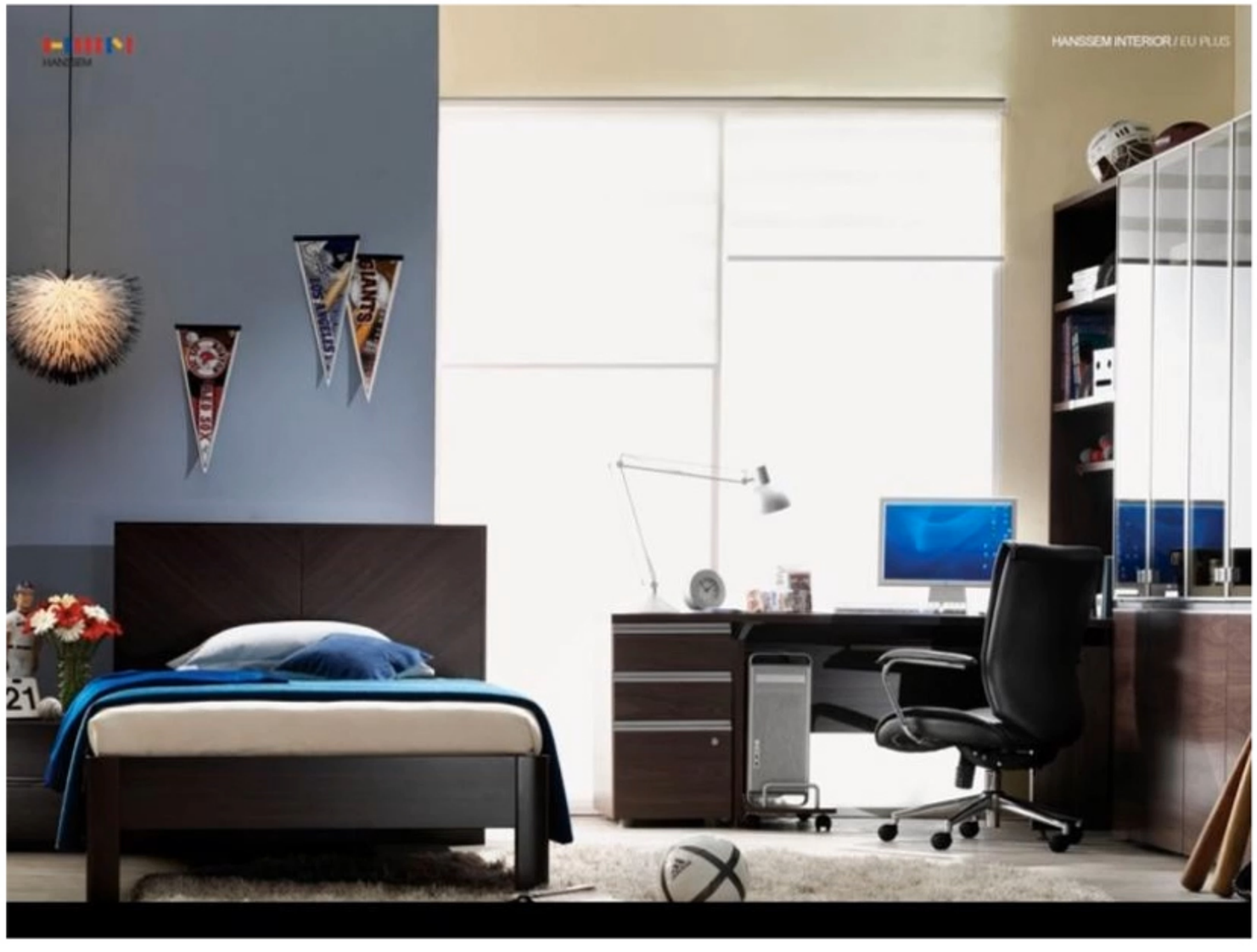


Autonomous Driving

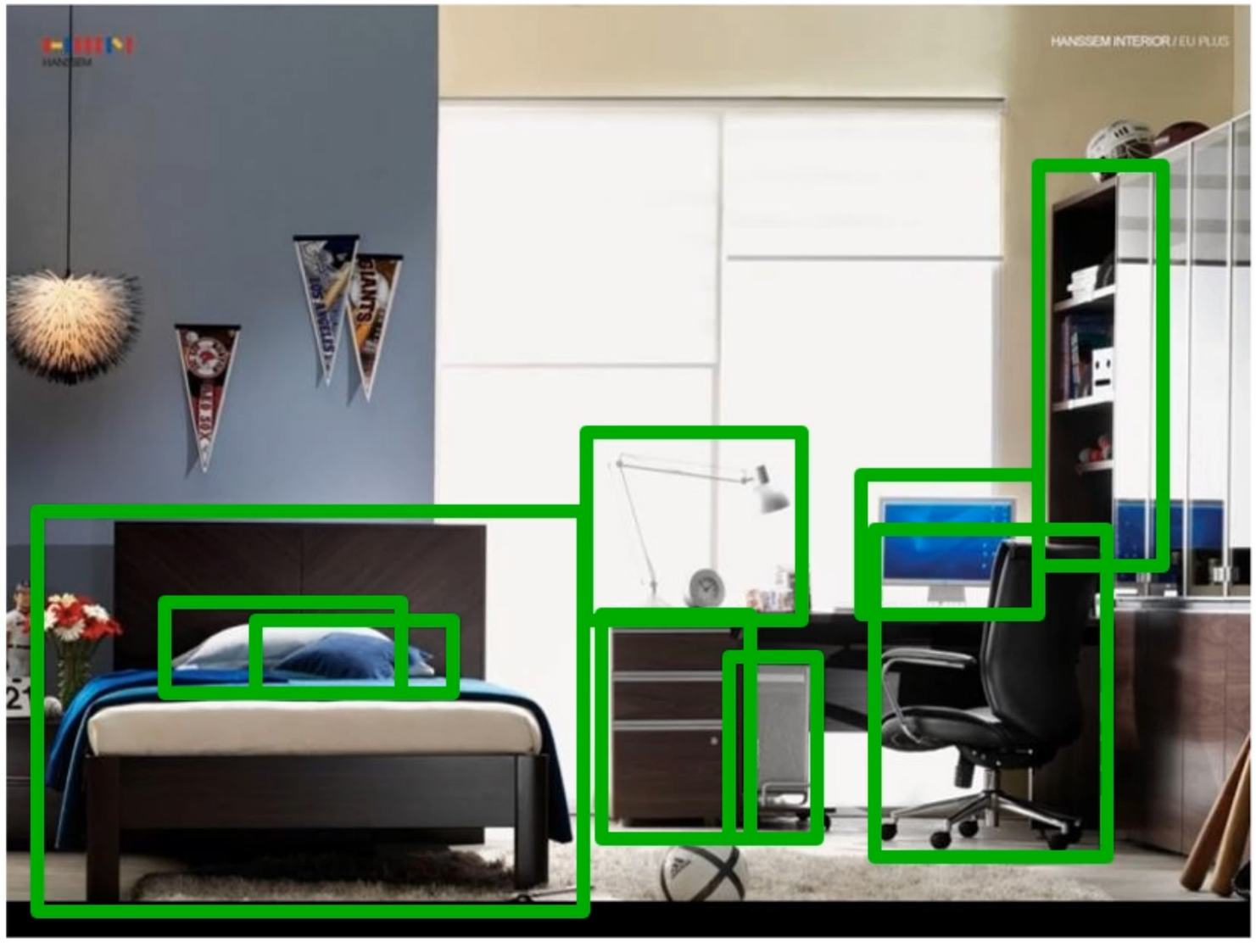


Augmented Reality

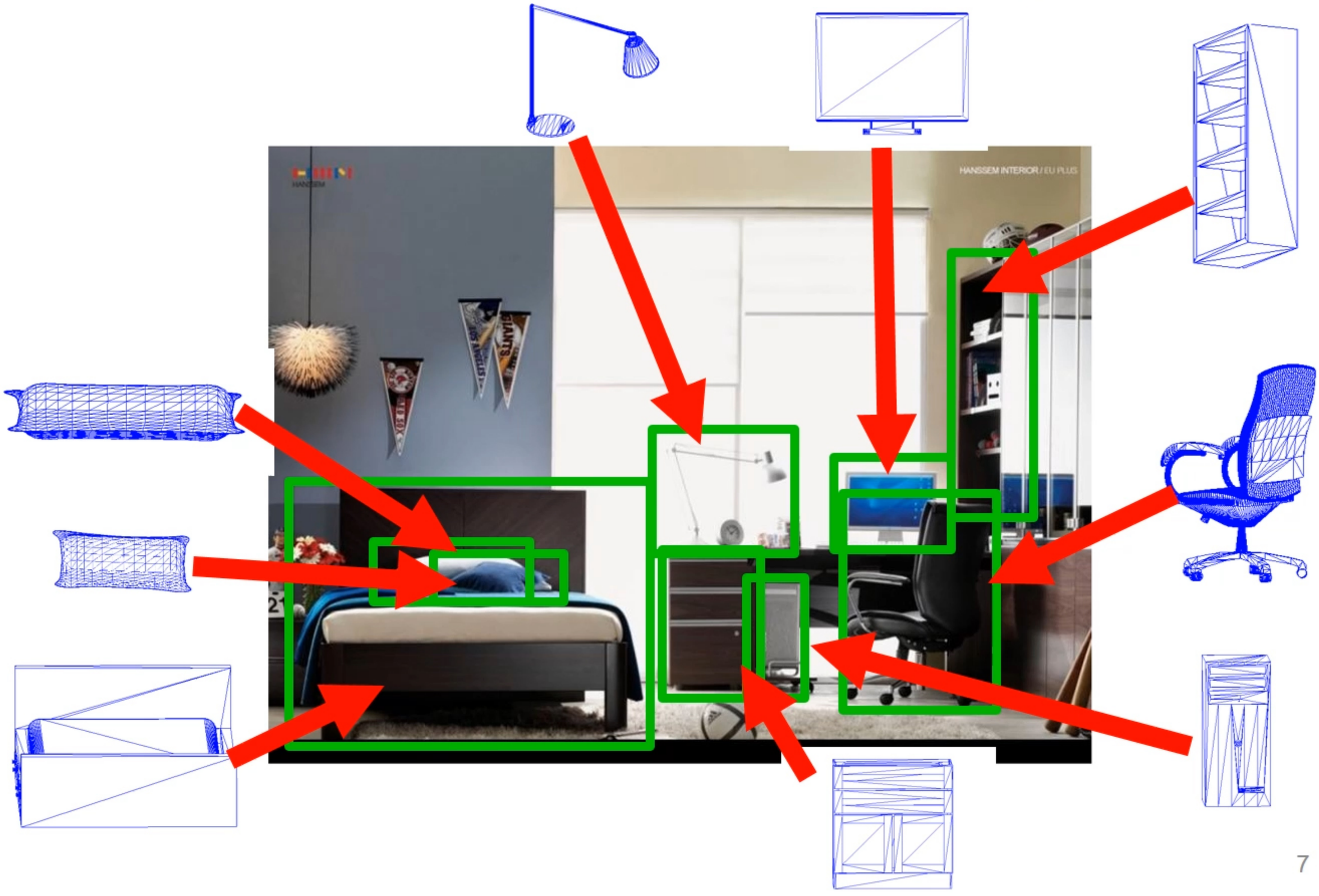
3D Annotation: 2D-3D Alignment



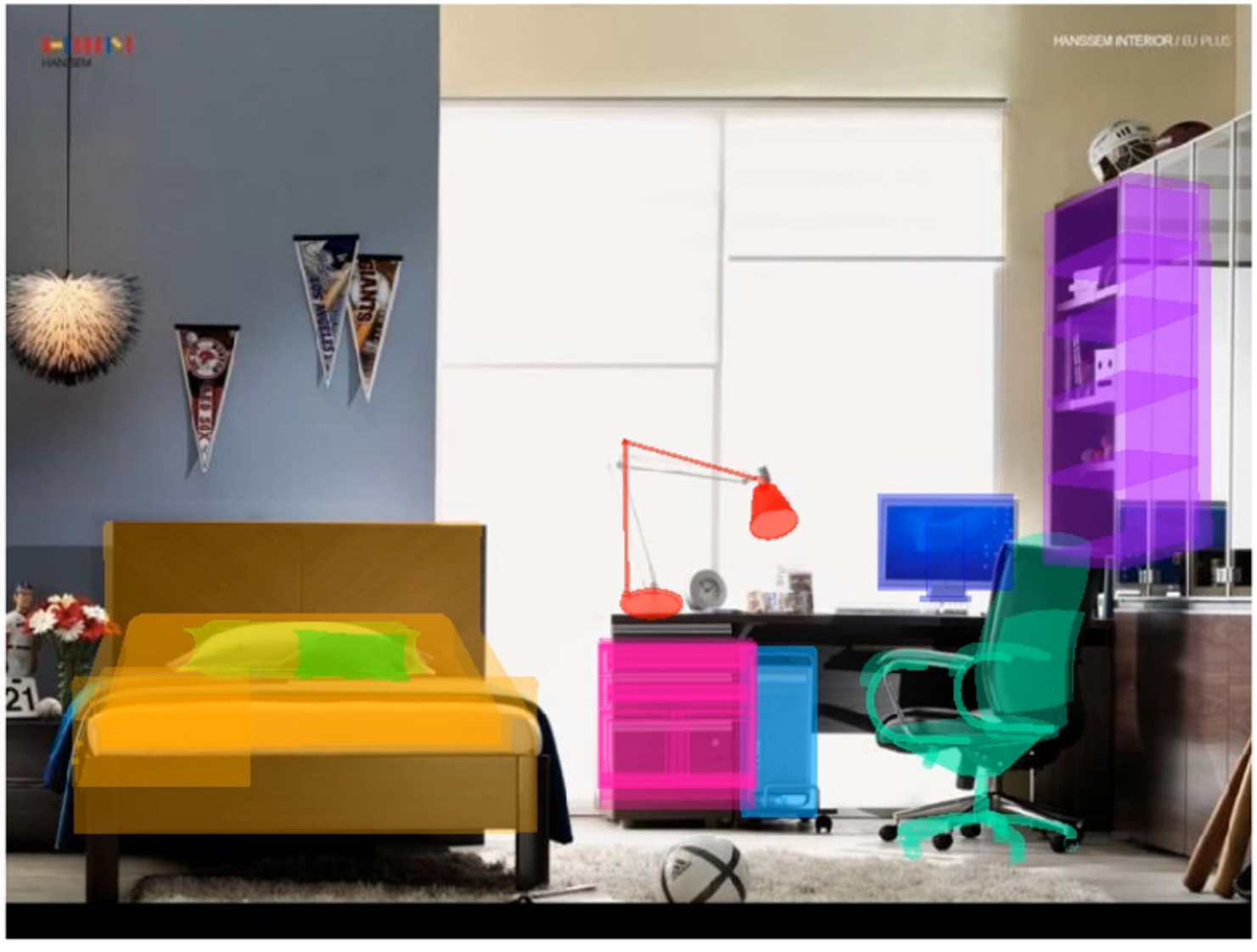
3D Annotation: 2D-3D Alignment



3D Annotation: 2D-3D Alignment



3D Annotation: 2D-3D Alignment



Comparison with Previous Datasets

	#category	#instance	Non-centered objects	Dense viewpoint	3D Shape
3D Object [1]	10	100	✗	✗	✗
EPFL Car [2]	1	20	✗	✓	✗
RGB-D Object [3]	51	300	✗	✓	✗
PASCAL VOC [4]	20	27,450	✓	✗	✗
KITTI [5]	3	80,256	✓	✓	✗
PASCAL3D+ [6]	12	35,672	✓	✓	✓79
ObjectNet3D (Ours)	**100**	**201,888**	✓	✓	✓**44,147**

[1] S. Savarese and L. Fei-Fei. 3d generic object categorization, localization and pose estimation. In ICCV, 2007.
[2] M. Ozuysal, V. Lepetit, and P. Fua. Pose estimation for category specific multiview object localization. In CVPR, 2009.
[3] K. Lai, L. Bo, X. Ren and D. Fox. A large-scale hierarchical multi-view RGB-D object dataset. In ICRA, 2011.
[4] M. Everingham, L. Van Gool, C. K. I.Williams, J.Winn, and A. Zisserman. The pascal visual object classes (voc) challenge. IJCV, 2010.
[5] A. Geiger, P. Lenz, and R. Urtasun. Are we ready for autonomous driving? the kitti vision benchmark suite. In CVPR, 2012.
[6] Y. Xiang, R. Mottaghi and S. Savarese. Beyond PASCAL: A benchmark for 3D object detection in the wild. In WACV, 2014.

Database Construction: Object Categories

- 100 rigid object categories

Aeroplane	Cap	Filing cabinet	Lighter	Remote control	Suitcase
Ashtray	Car	Fire extinguisher	Mailbox	Rifle	Teapot
Backpack	Cellphone	Fish tank	Microphone	Road pole	Telephone
Basket	Chair	Flashlight	Microwave	Satellite dish	Toaster
Bed	Clock	Fork	Motorbike	Scissors	Toilet
Bench	Coffee maker	Guitar	Mouse	Screwdriver	Toothbrush
Bicycle	Comb	Hair dryer	Paintbrush	Shoe	Train
Backboard	Computer	Hammer	Pan	Shovel	Trash bin
Boat	Cup	Headphone	Pen	Sign	Trophy
Bookshelf	Desk lamp	Helmet	Pencil	Skate	Tub
Bottle	Dining table	Iron	Piano	Skateboard	Tvmonitor
Bucket	Dishwasher	Jar	Pillow	Slipper	Vending machine
Bus	Door	Kettle	Plate	Sofa	Washing machine
Cabinet	Eraser	Key	Pot	Speaker	Watch
Calculator	Eyeglasses	Keyboard	Printer	Spoon	Wheelchair
Camera	Fan	Knife	Racket	Stapler	
Can	Faucet	Laptop	Refrigerator	Stove	

Database Construction: Object Categories

- 100 rigid object categories

Aeroplane	Cap	Filing cabinet	Lighter	Remote control	Suitcase
Ashtray	Car	Fire extinguisher	Mailbox	Rifle	Teapot
Backpack	Cellphone	Fish tank	Microphone	Road pole	Telephone
Basket	[illegible]	Flashlight	Microwave	Satellite dish	[illegible]
Bed	Clock	Fork	Motorbike	Scissors	Toilet
Bench	Coffee maker	Guitar	Mouse	Screwdriver	Toothbrush
Bicycle	Comb	Hair dryer	Paintbrush	Shoe	Train
Backboard	Computer	Hammer	Pan	Shovel	Trash bin
Boat	Cup	Headphone	Pen	Sign	Trophy
Bookshelf	Desk lamp	[illegible]	Pencil	[illegible]	[illegible]
Bottle	Dining table	Iron	Piano	Skateboard	Tvmonitor
Bucket	Dishwasher	Jar	Pillow	Slipper	Vending machine
Bus	Door	Kettle	Plate	Sofa	Washing machine
Cabinet	Eraser	Key	Pot	Speaker	Watch
Calculator	Eyeglasses	Keyboard	Printer	Spoon	Wheelchair
Camera	Fan	Knife	Racket	Stapler	
Can	Faucet	Laptop	Refrigerator	Stove	

Vehicles

Furniture

Container

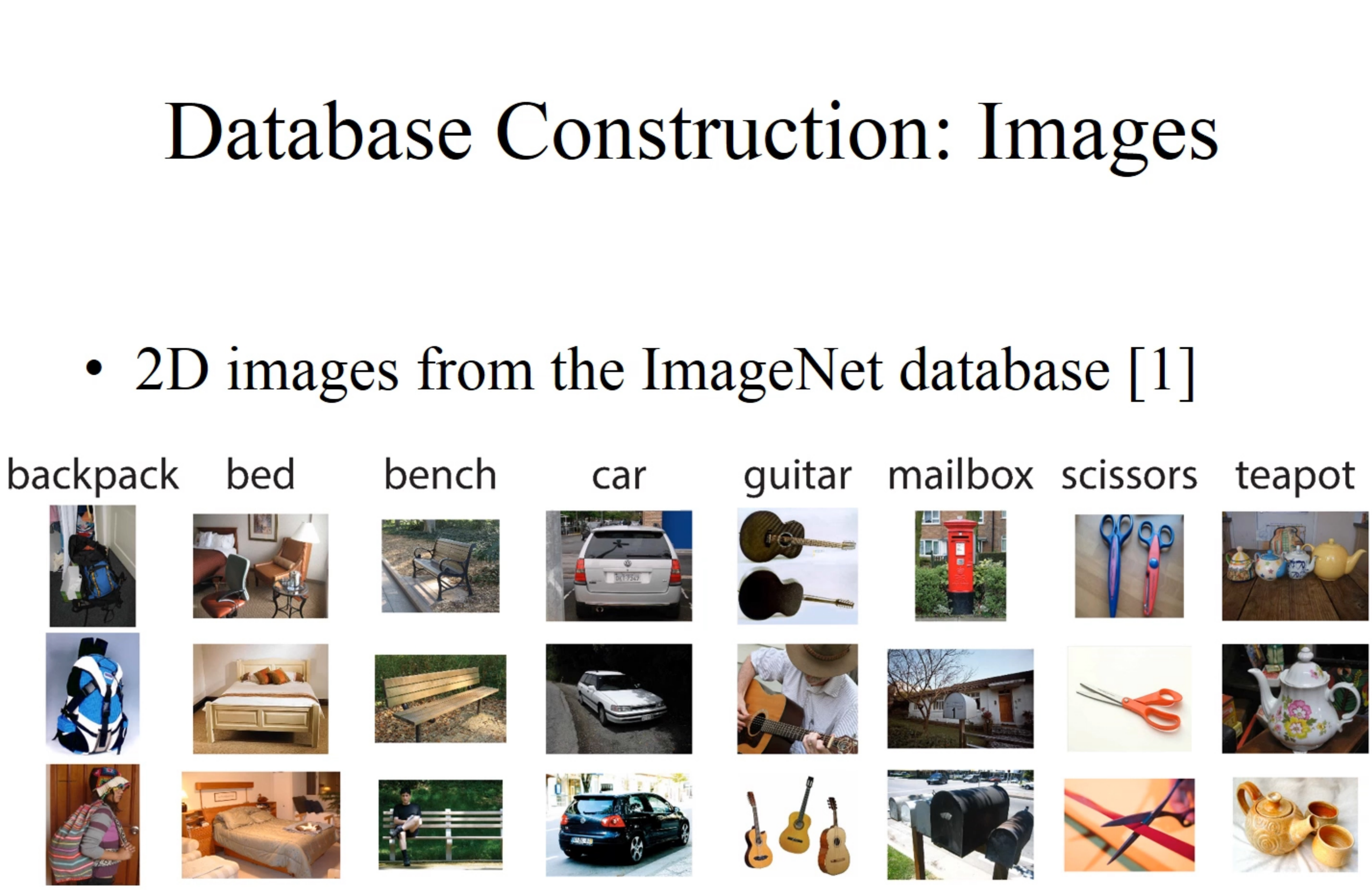
Tools

Electronics

Personal items

Database Construction: Images

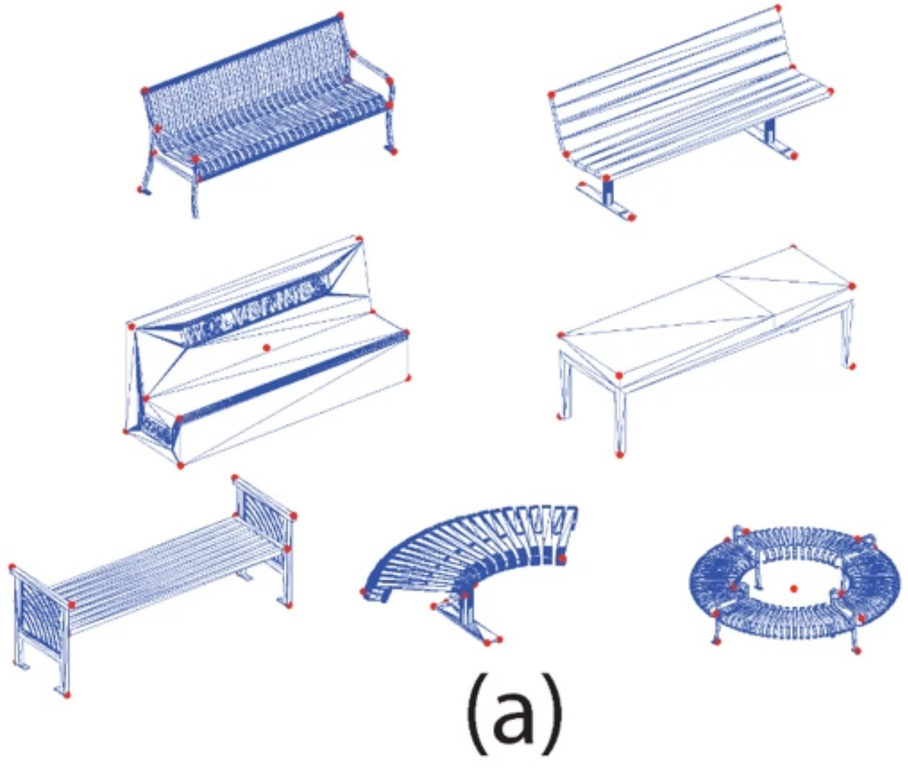
- 2D images from the ImageNet database [1]

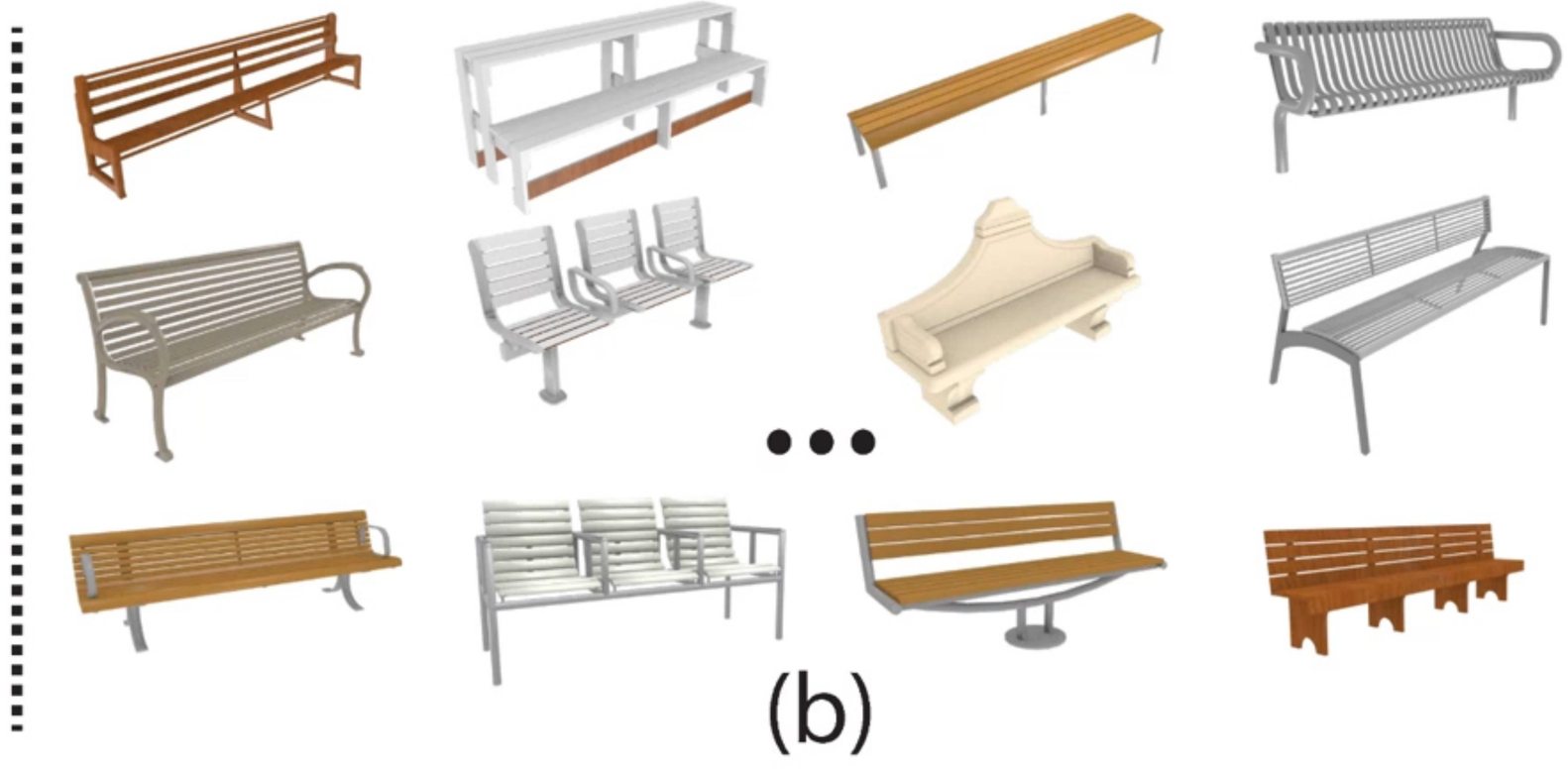


[1] Russakovsky et al. ImageNet Large Scale Visual Recognition Challenge, IJCV 2015

Database Construction: 3D Shapes

- Trimble 3D Warehouse [1]
- ShapeNet database [2]





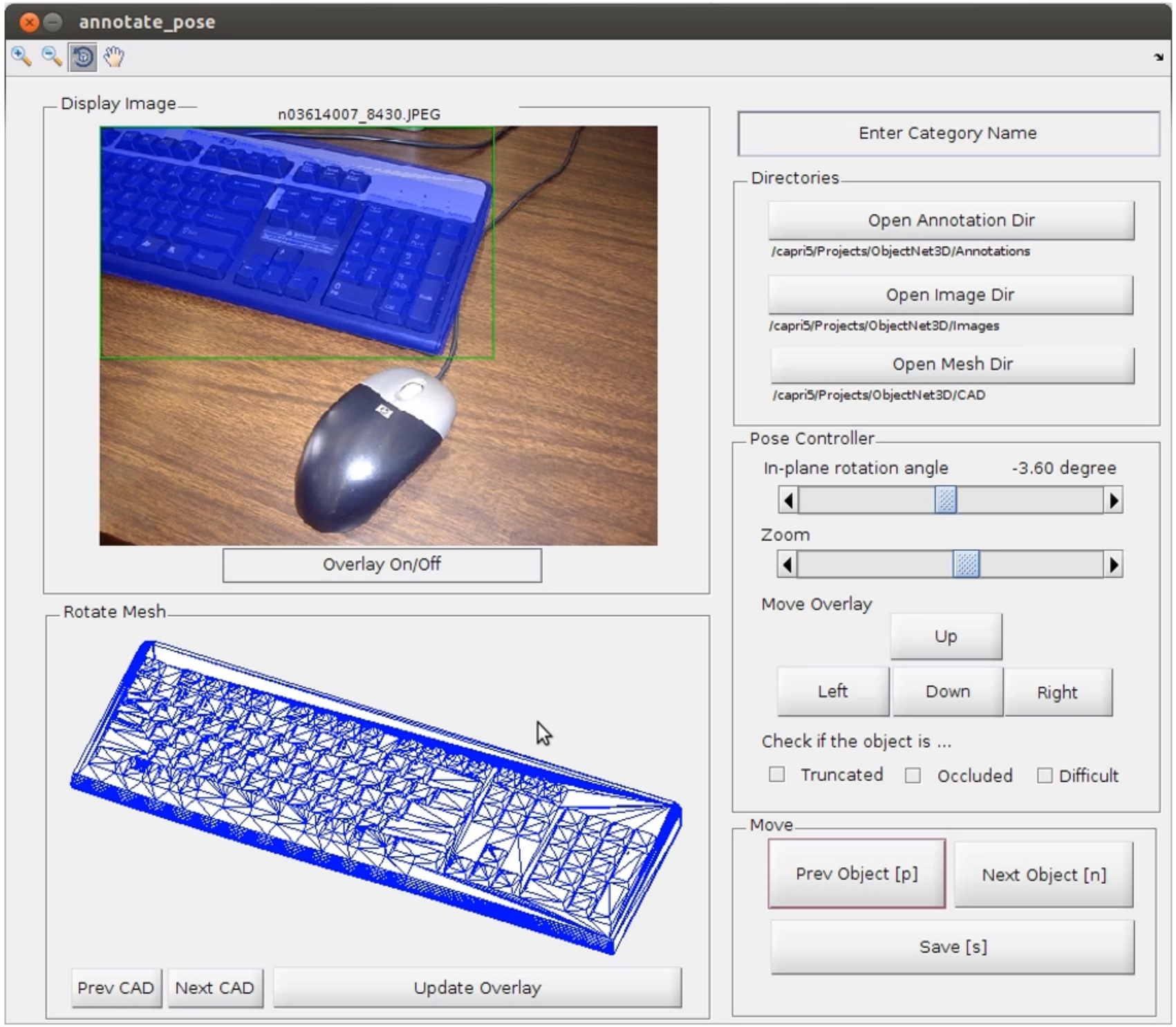
3D Shapes from Trimble 3D Warehouse

3D Shapes from ShapeNet

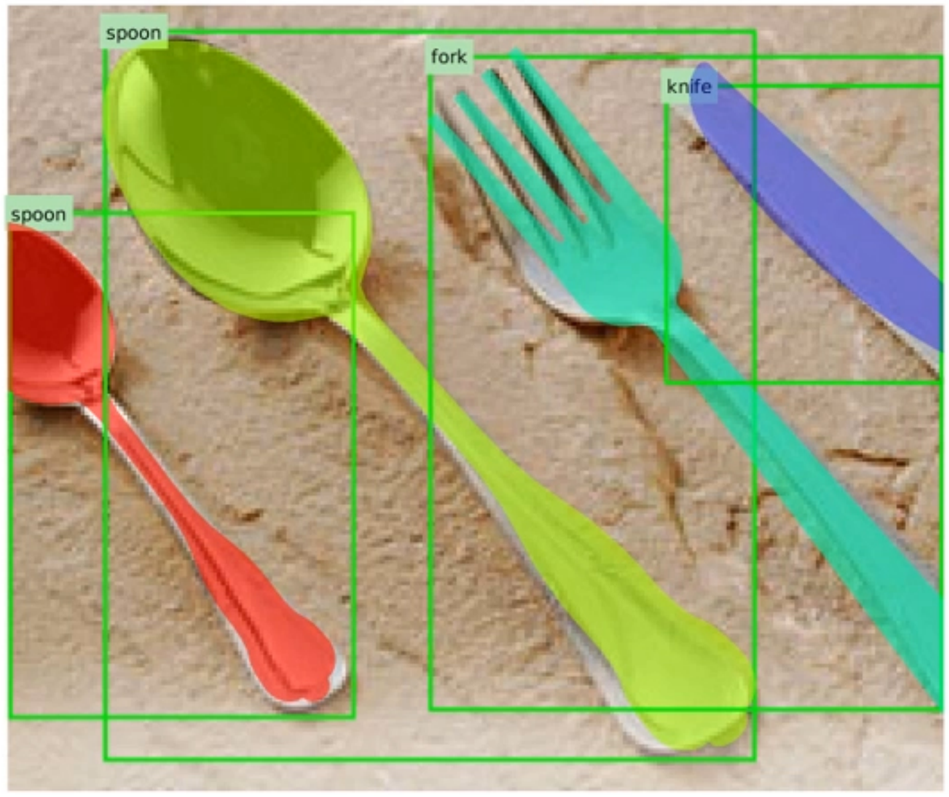
[1] https://3dwarehouse.sketchup.com

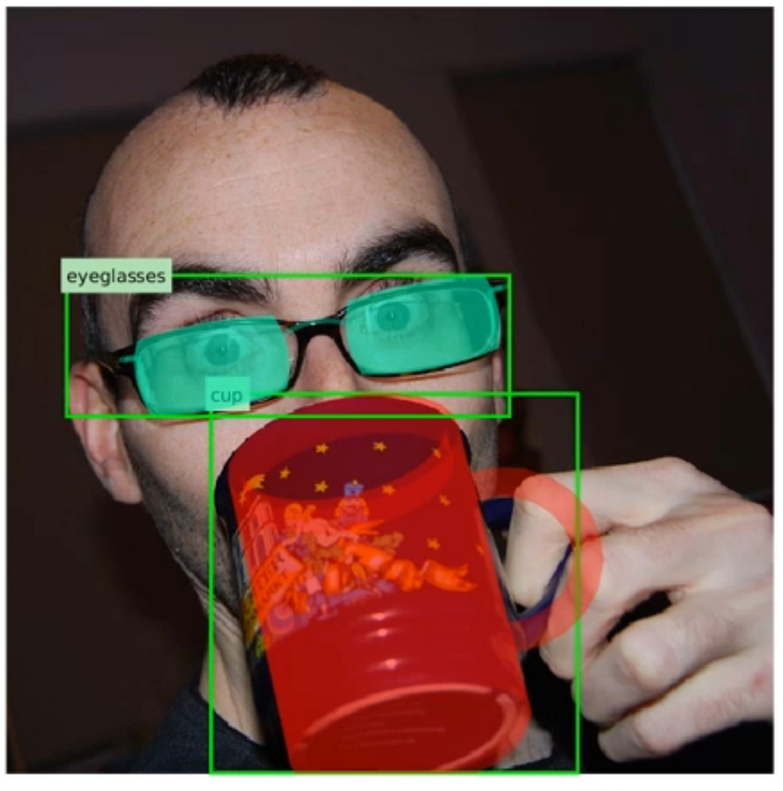
[2] Chang et al. ShapeNet: An Information-Rich 3D Model Repository, arXiv 2015

Database Construction: Annotation Demo

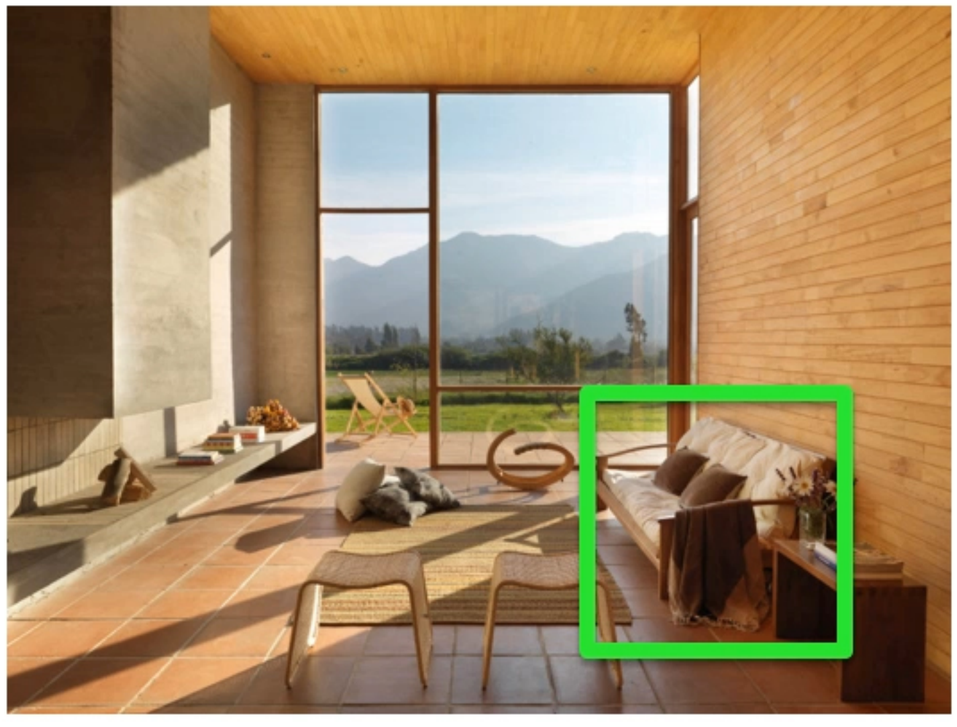


3D Pose Annotation Examples





Database Construction: Image-based 3D Shape Retrieval

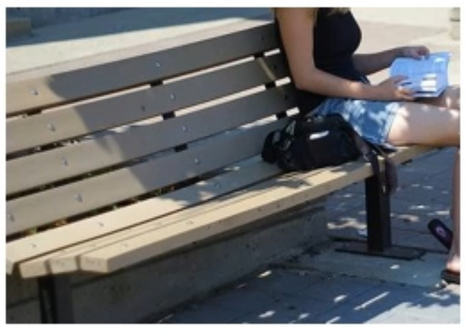


Database Construction: Image-based 3D Shape Retrieval

Test Object | **Rank 1** | **Rank 2** | **Rank 3**




• • •




• • •




• • •

H.O. Song, Y. Xiang, S. Jegelka and S. Savarese. Deep Metric Learning via Lifted Structured Feature Embedding. In CVPR, 2016.

Baseline Experiments

- Object proposal generation
- 2D object detection
- Image-based 3D shape retrieval
- Joint 2D detection and continuous 3D pose estimation

Baseline Experiments

- Object proposal generation

Selective Search: Uijlings et al., IJCV, 2013.
EdgeBoxes: Zitnick et al., ECCV, 2014.
MCG: Arbelaez et al., CVPR, 2014.
RPN: Ren et al., NIPS, 2015.

- 2D object detection

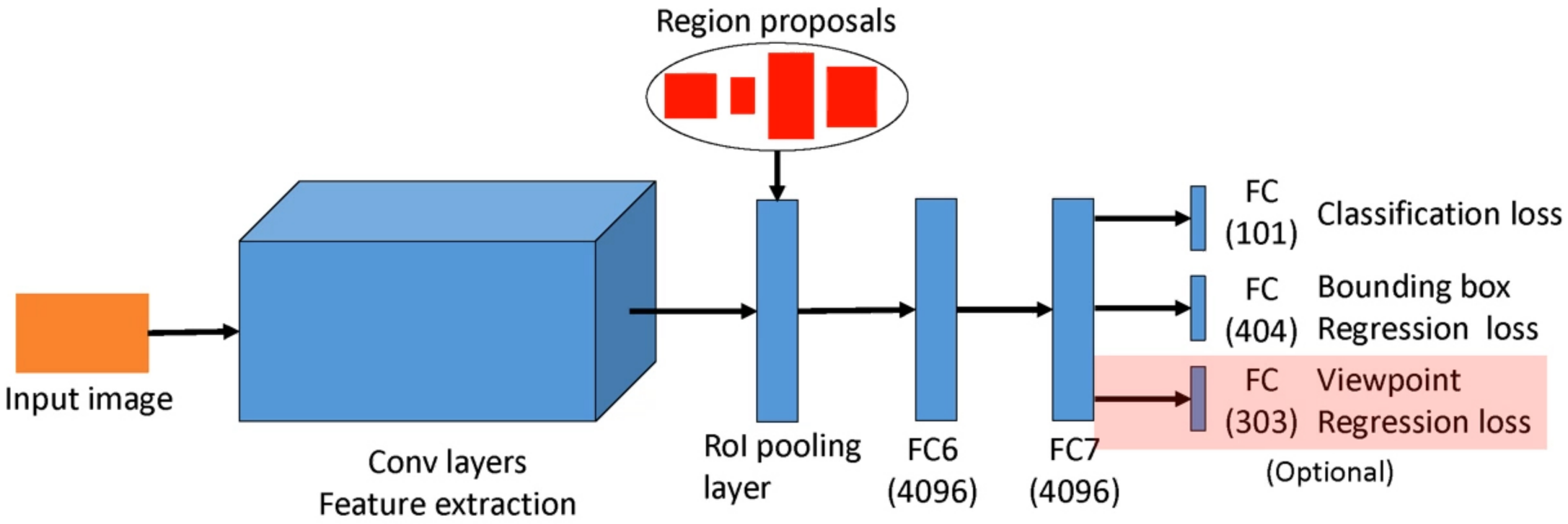
Fast R-CNN: Girshick R., ICCV, 2015.

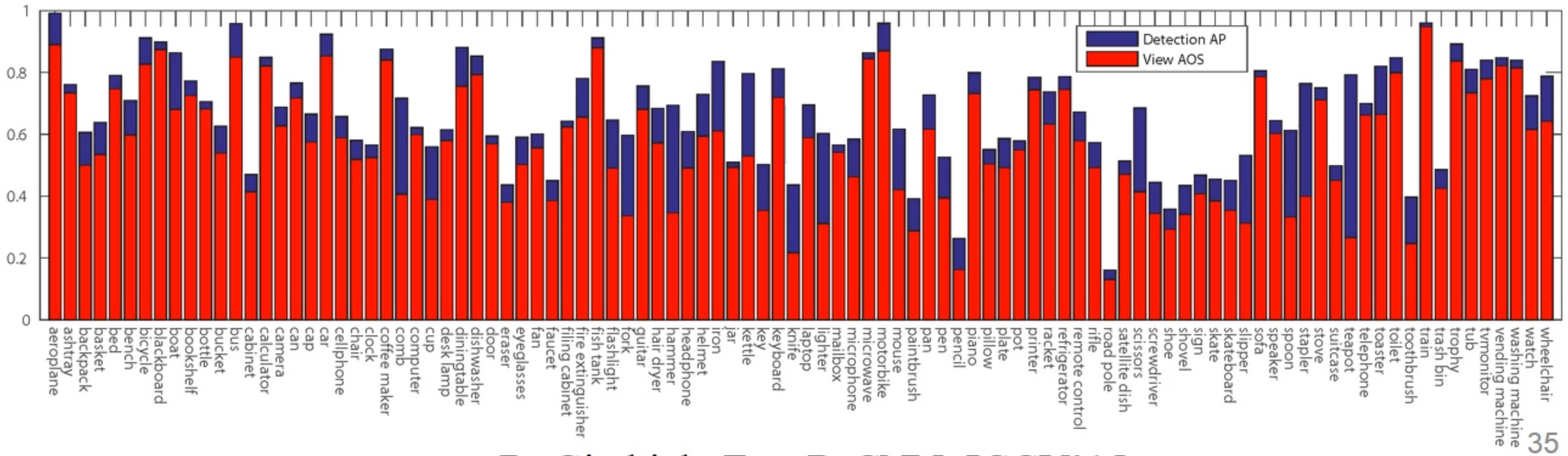
- Image-based 3D shape retrieval

Deep Lifted Structure: Song et al., CVPR, 2016.

- Joint 2D detection and continuous 3D pose estimation

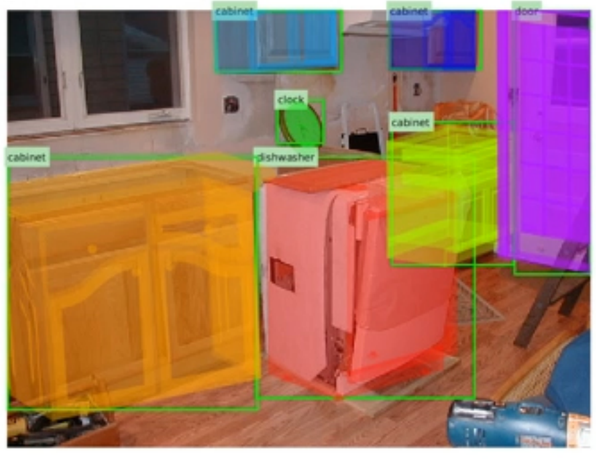
A Network for Object Detection and Pose estimation

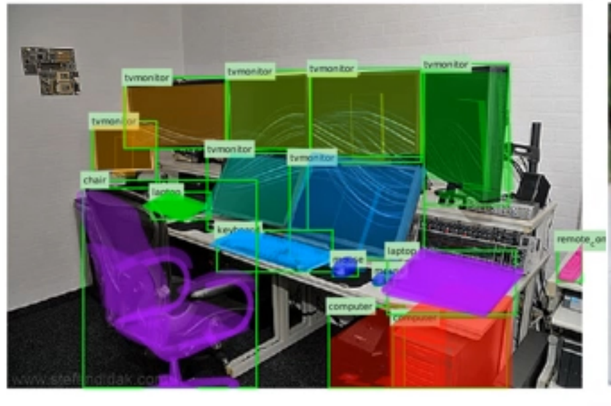




R. Girshick. Fast R-CNN. ICCV'15.

ObjectNet3D





- 100 object categories
- 90,127 images
- 201,888 objects
- 44,147 3D shapes
- 2D-3D alignments
- Baseline experiments on different recognition tasks