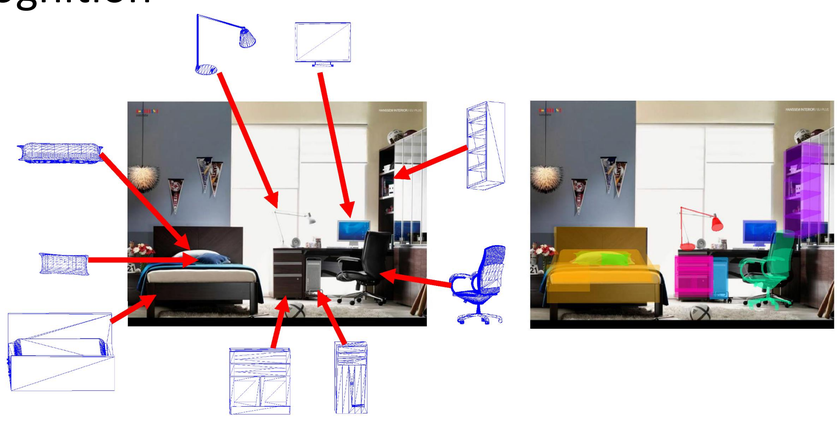
ObjectNet3D: A Large Scale Database for 3D Object Recognition

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

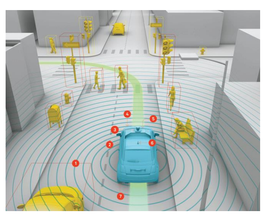
❖ Introduction

❑ Goal: build a large scale database for 3D object recognition

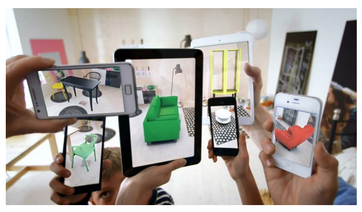


❑ Application: recognizing the 3D properties of objects such as 3D location, 3D pose, 3D shape, etc.

Robotics



Autonomous Driving



Augmented Reality

❑ 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] Savarese & Fei-Fei. ICCV, 2007.
[2] M. Ozuysal, et al. CVPR, 2009.
[3] K. Lai et al. ICRA, 2011.
[4] M. Everingham, et al. IJCV, 2010.
[5] A. Geiger, et al. CVPR, 2012.
[6] Y. Xiang, et al. WACV, 2014.

❖ Database Construction

❑ 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	

❑ 2D image from the ImageNet database

backpack bed bench car guitar mailbox scissors teapot

Deng et al. CVPR, 2009.

Acknowledgements. We acknowledge the support of NSF grants IIS-1528025 and DMS-1546206, a Google Focused Research award, and grant SPO # 124316 and 1191689-1-UDAWF from the Stanford AI Lab-Toyota Center for Artificial Intelligence Research.

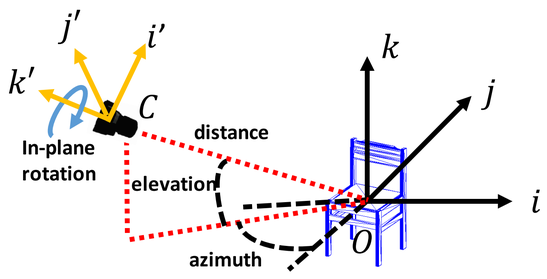
❖ Database Construction

❑ 3D shapes from Trimble 3D Warehouse and ShapeNet

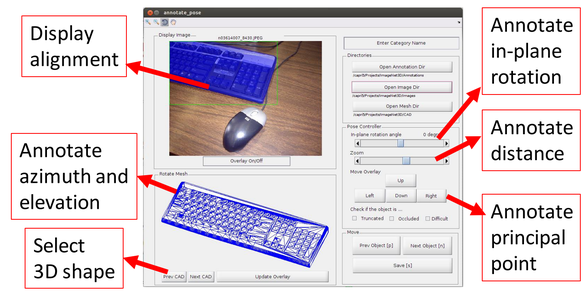
https://3dwarehouse.sketchup.com

Chang et al. ShapeNet, arXiv 2015

❑ Camera model and annotation tool

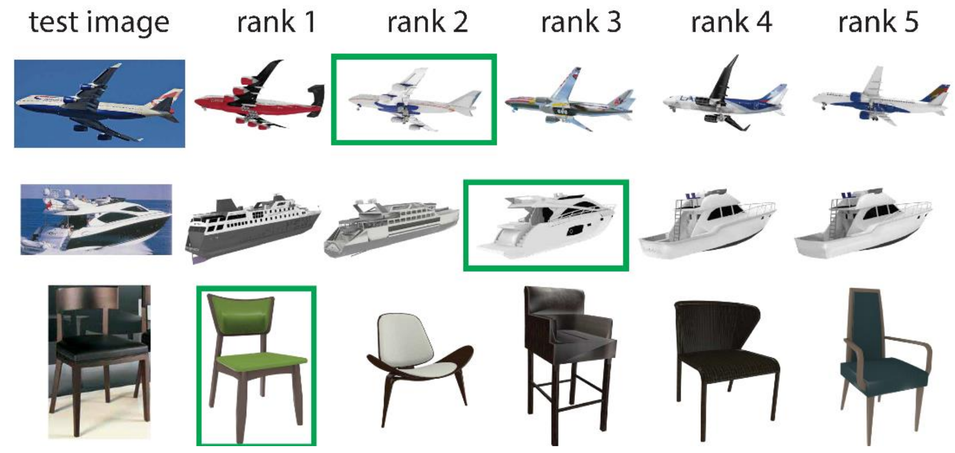


Camera model



Annotation interface

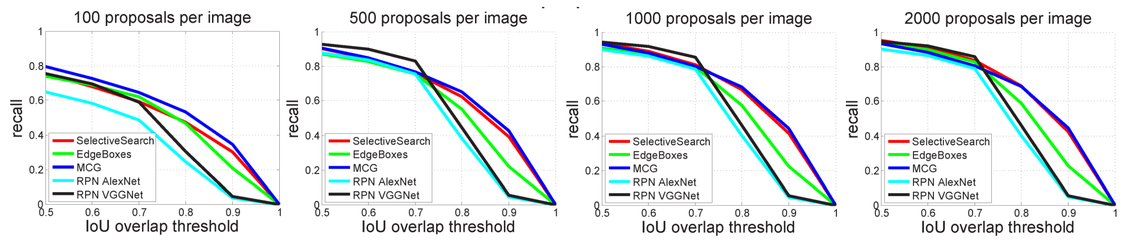
❑ Image-based 3D shape retrieval



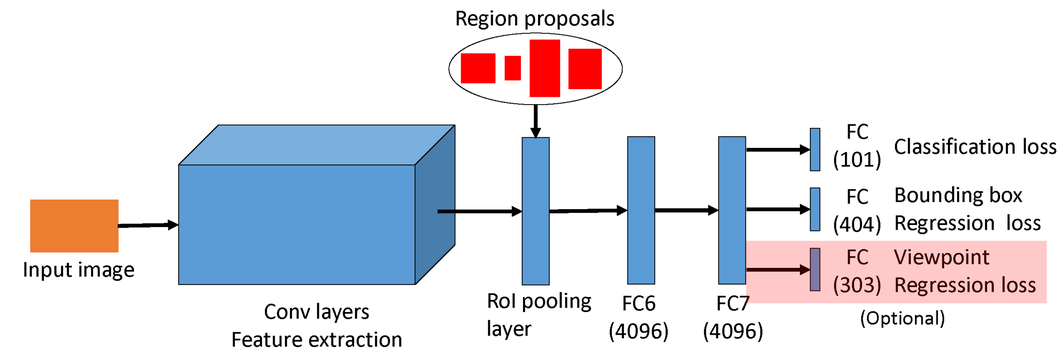
H.O. Song, et al. Deep Metric Learning via Lifted Structured Feature Embedding. In CVPR, 2016.

❖ Baseline Experiments

❑ Object proposal generation



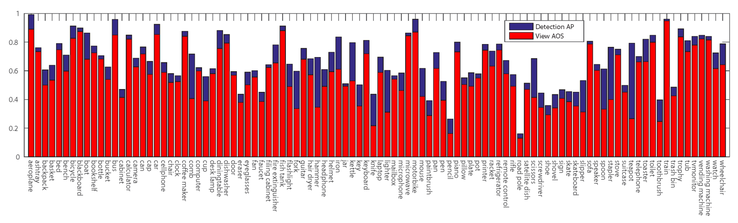
❑ Object detection and pose estimation



mAP	Selective Search	EdgeBox	MCG	RPN
AlexNet	**56.3**	52.5	56.0	54.2
VGGNet	67.3	64.5	67.0	**67.5**

AOS/AVP	Selective Search	EdgeBox	MCG	RPN
AlexNet	**48.1 / 37.3**	44.5 / 33.6	47.8 / 37.0	46.1 / 35.4
VGGNet	**57.1 / 43.0**	54.2 / 39.6	56.8 / 42.9	57.0 / 42.6

2D object detection



Joint 2D object detection and pose estimation

AlexNet: Krizhevsky et al., NIPS, 2012.
VGGNet: Simonyan., arXiv:1409.1556, 2014.
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.

❑ Image-based 3D shape retrieval

The mean Recall@20 is 69.2%.