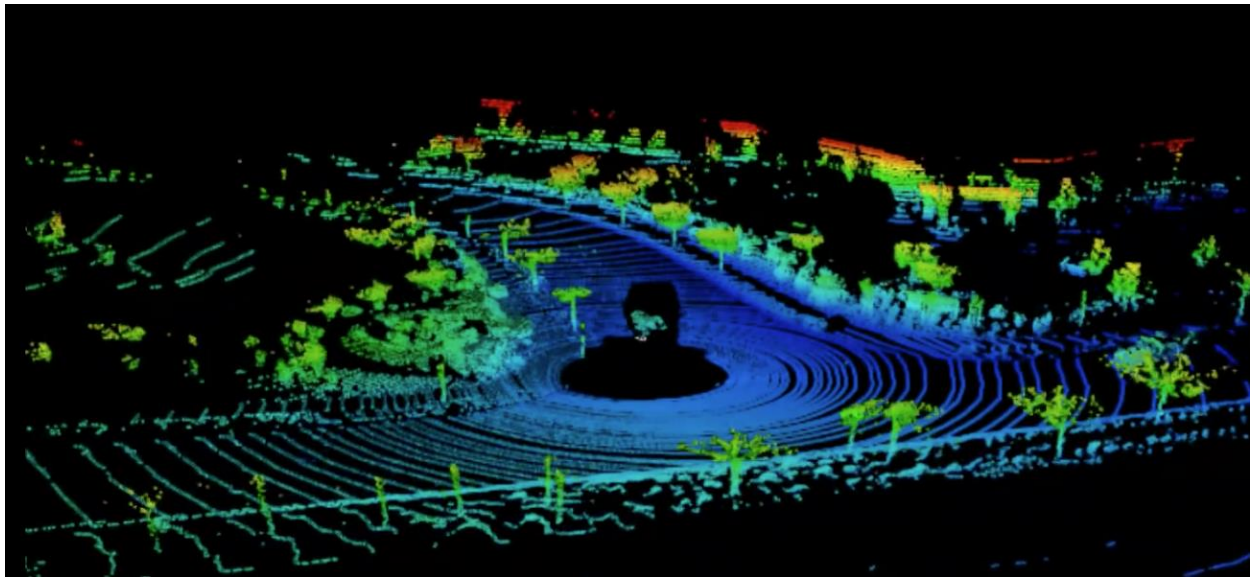
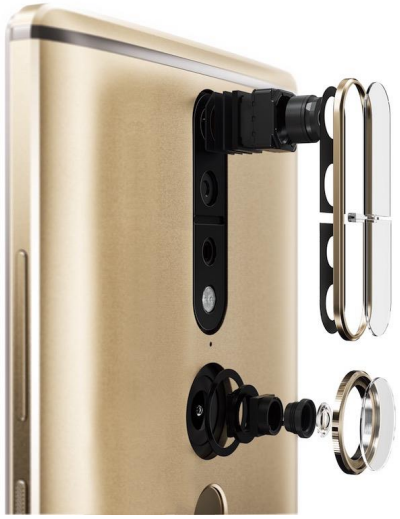
Volumetric and Multi-View CNNs for Object Classification on 3D Data

Charles R. Qi*, Hao Su*, Matthias Nießner,
Angela Dai, Mengyuan Yan, Leonidas J. Guibas

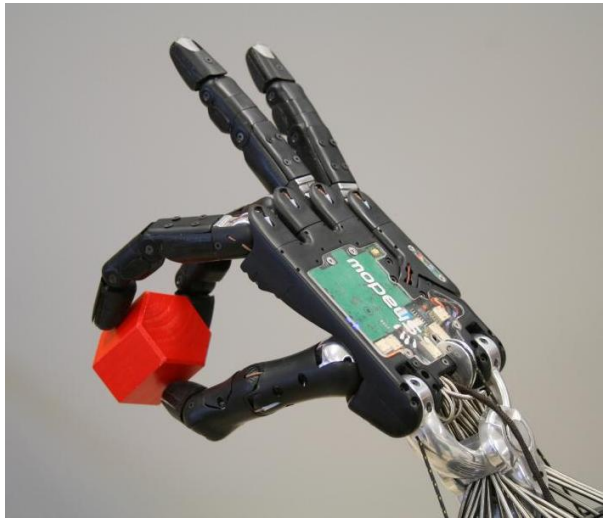
Stanford University

Rich Applications of 3D





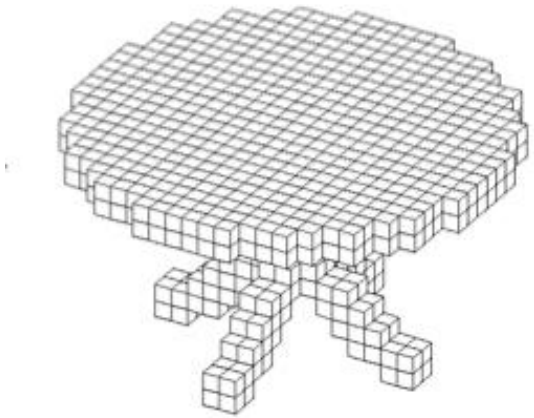
Augmented Reality



Robot Perception

3D Representations for Generic Object Classification

Volumetric



3DShapeNets by Z. Wu et al. CVPR 15

VoxNet by D. Maturana et al. IEEE/RSJ 15

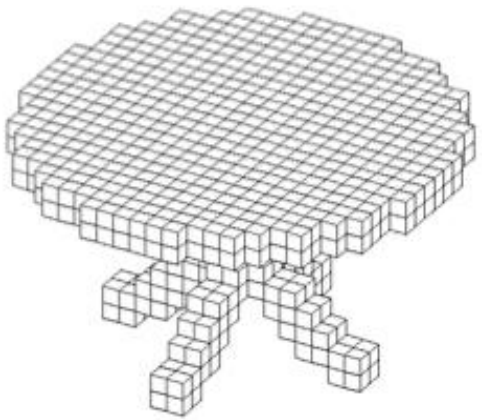
Multi-Views

MVCNN by H. Su et al. ICCV 15

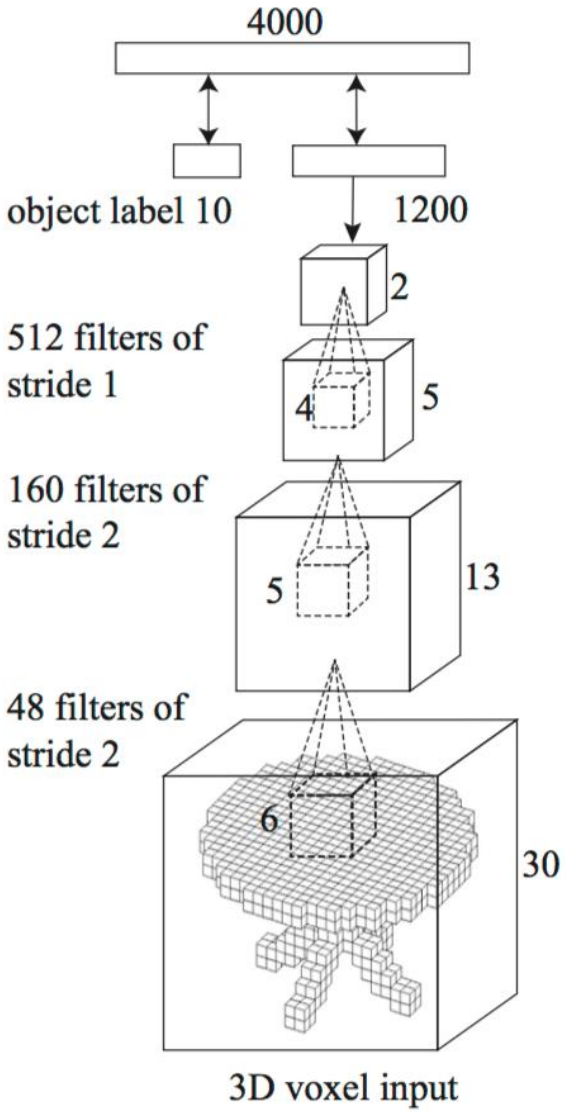
DeepPano by B. Shi et al. IEEE/SPL 15

Volumetric CNNs Revisited

Volumetric CNNs



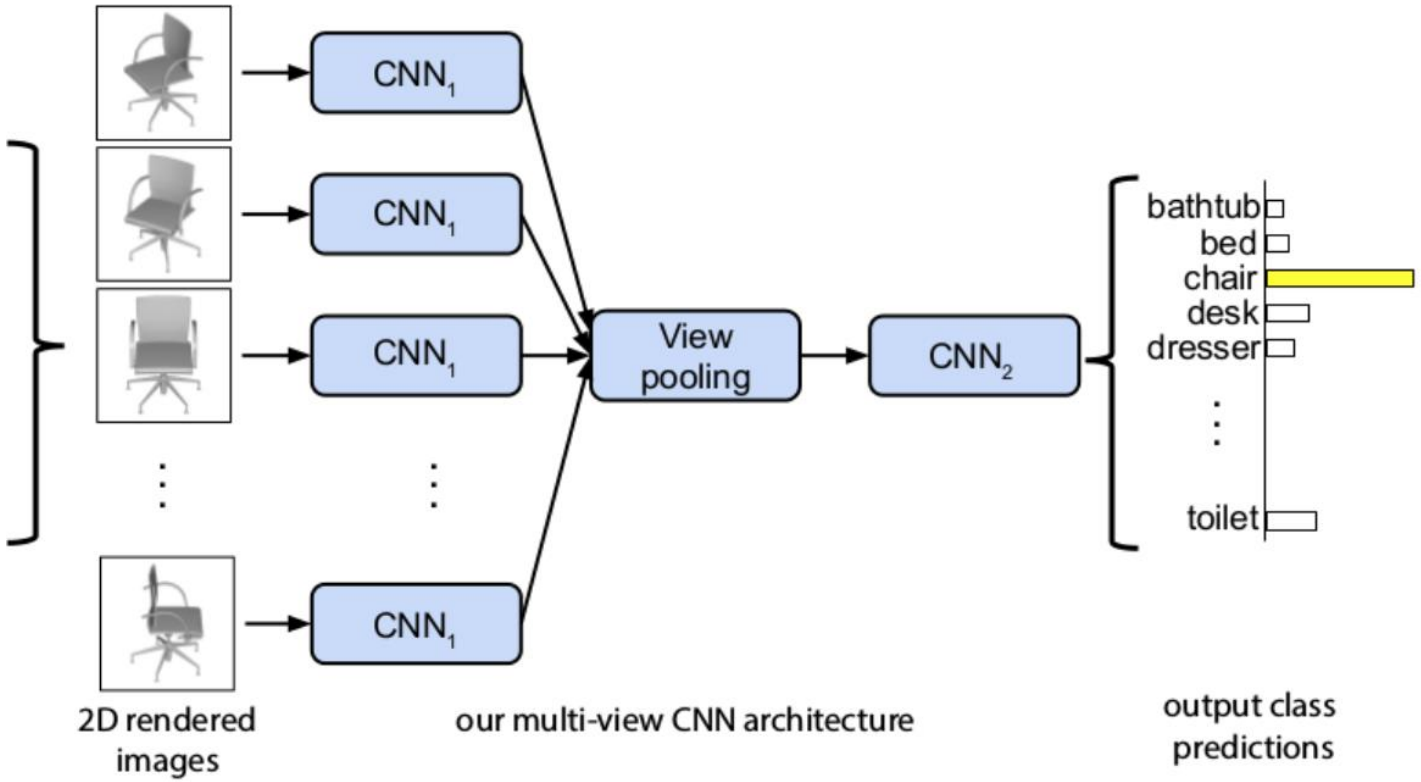
3DShapeNets by Z. Wu et al. CVPR 15



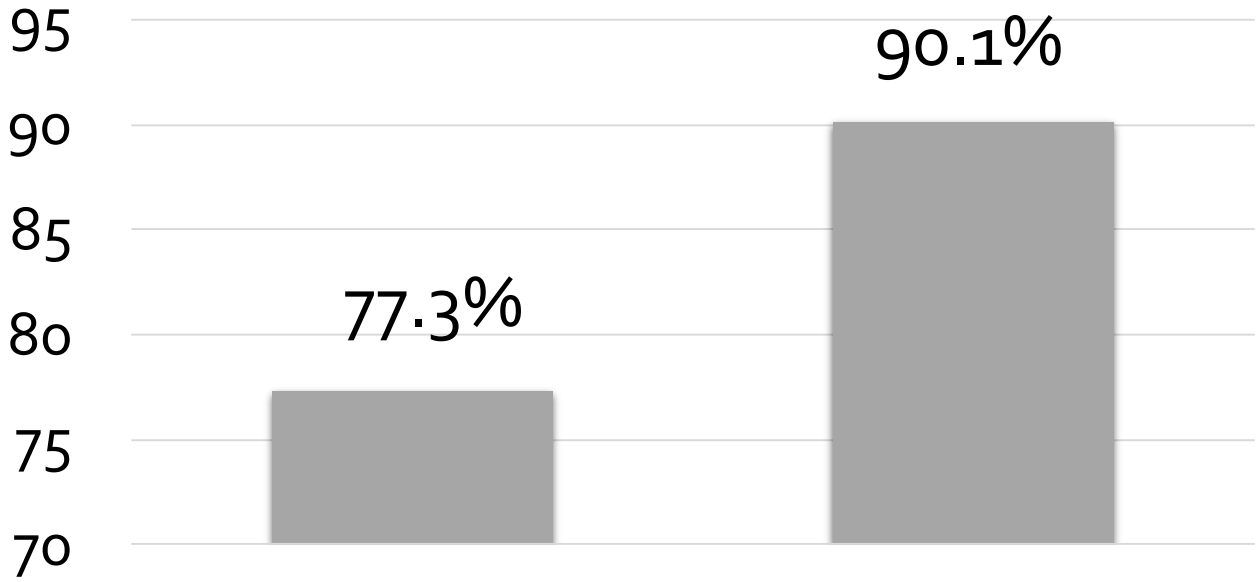
Multi-View CNNs Revisited

Multi-View CNNs

MVCNN by H. Su et al.
ICCV 15



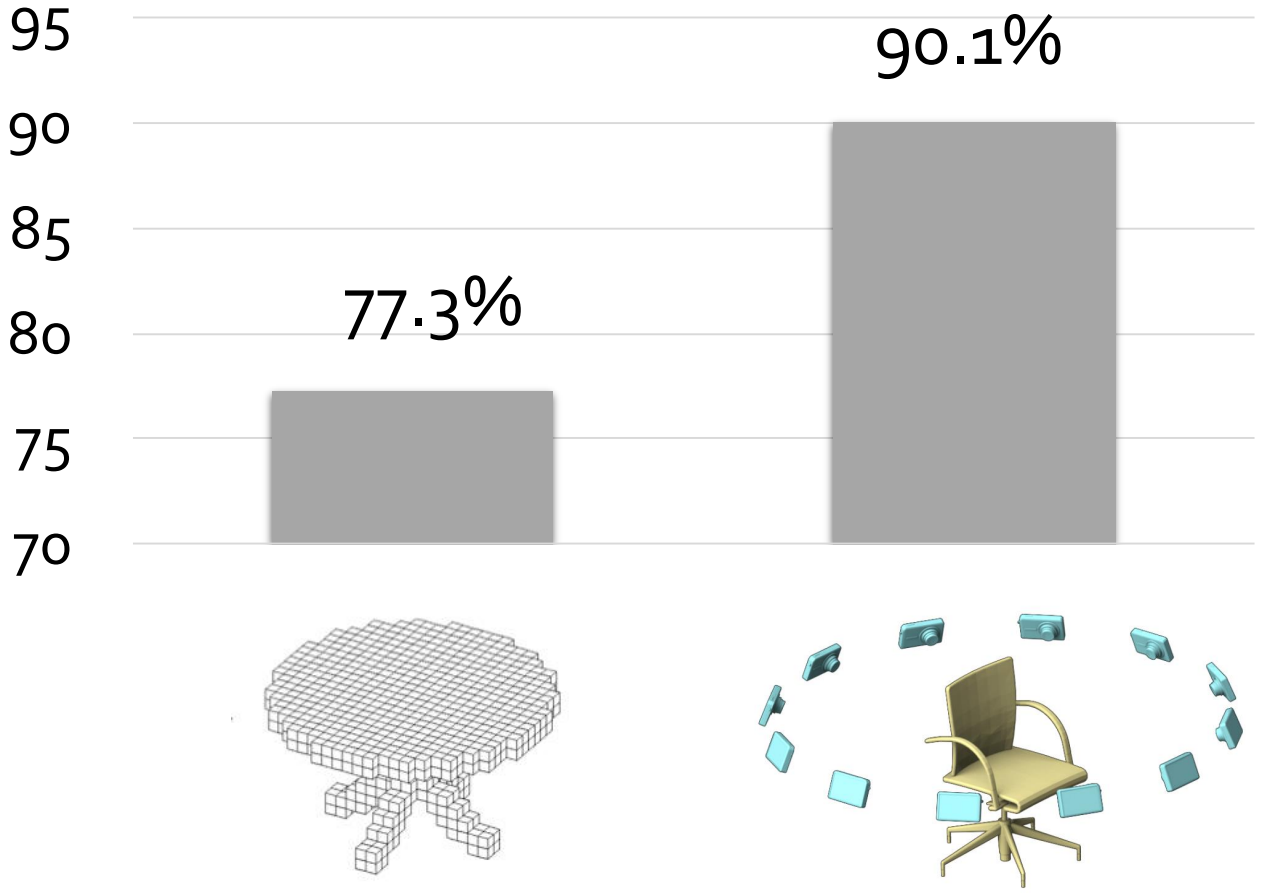
Shape Classification Results Revisited



3DShapeNets
Wu et al.

MVCNN
Su et al.

Shape Classification Results Revisited



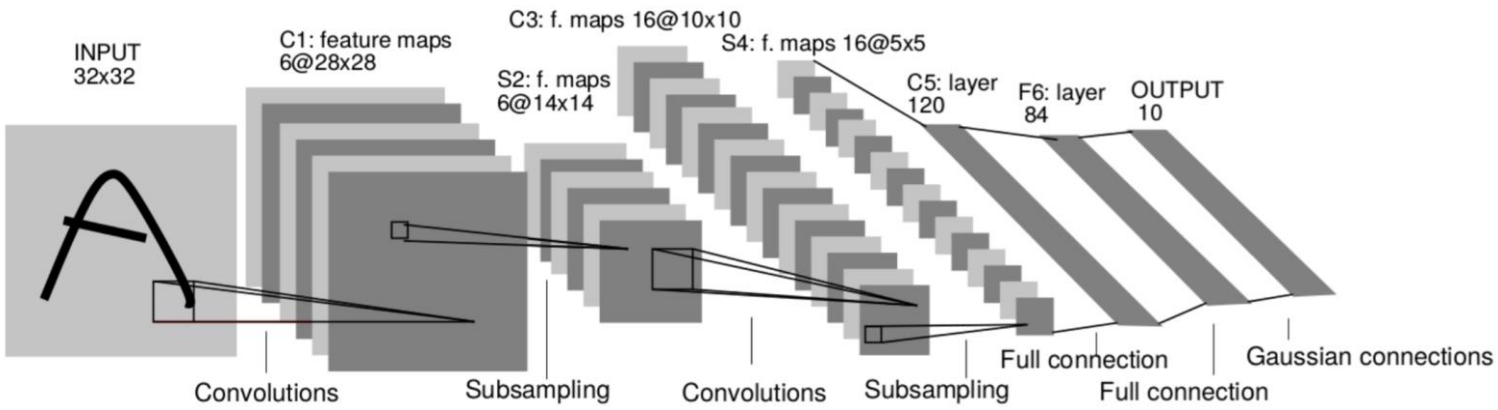
3DShapeNets
Wu et al.

MVCNN
Su et al.

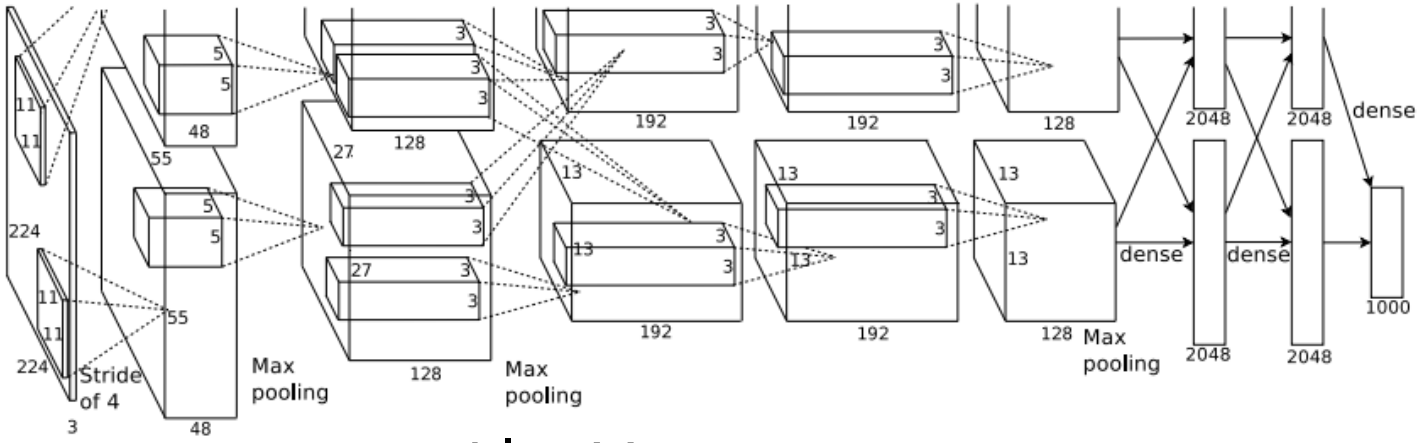
Big gap between volumetric and multi-view based methods

Why?

Cause 1: Architecture and Engineering

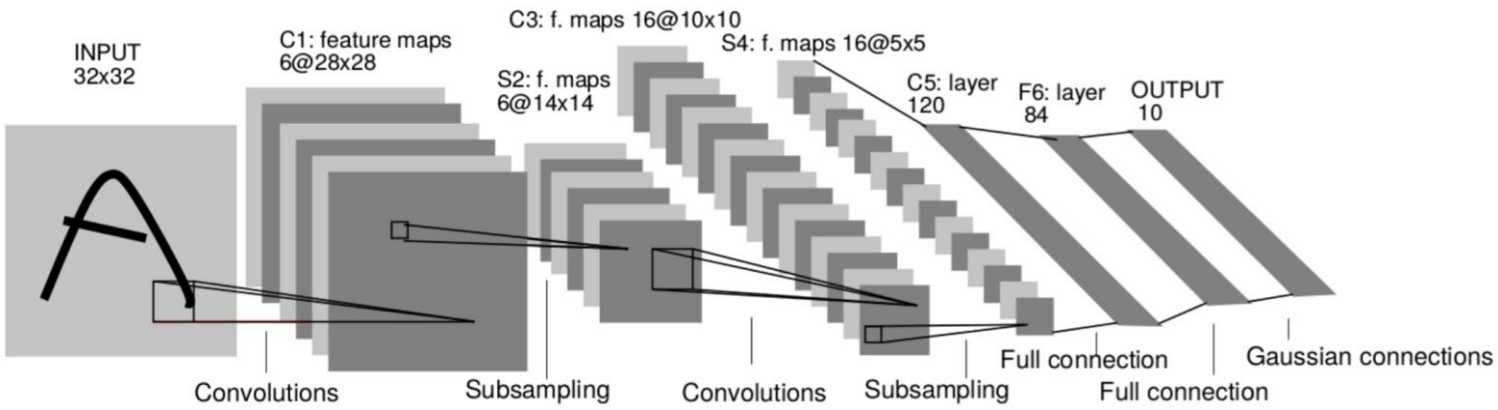


LeNet, **1998**

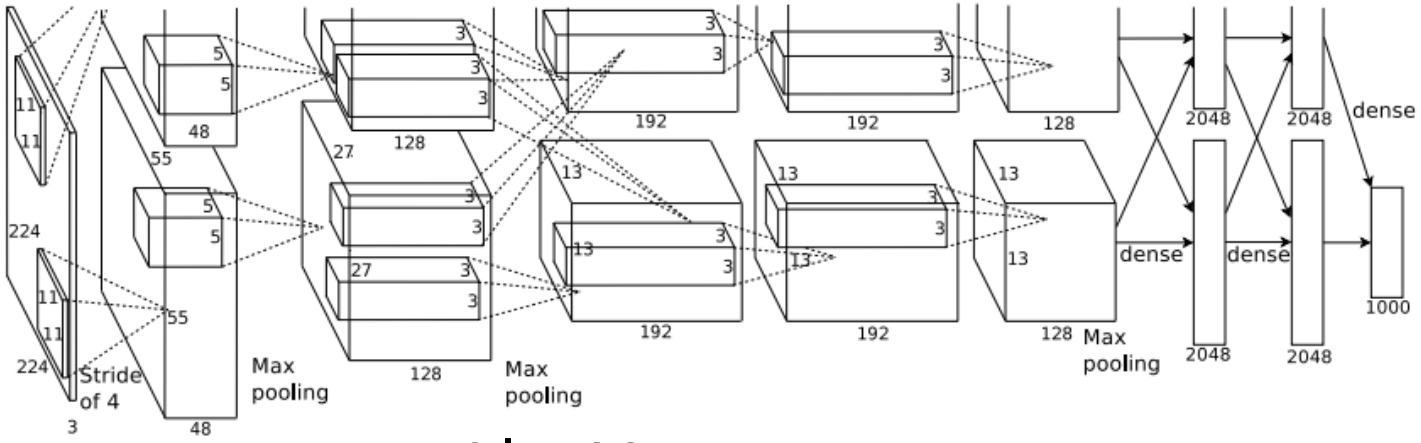


AlexNet, **2012**

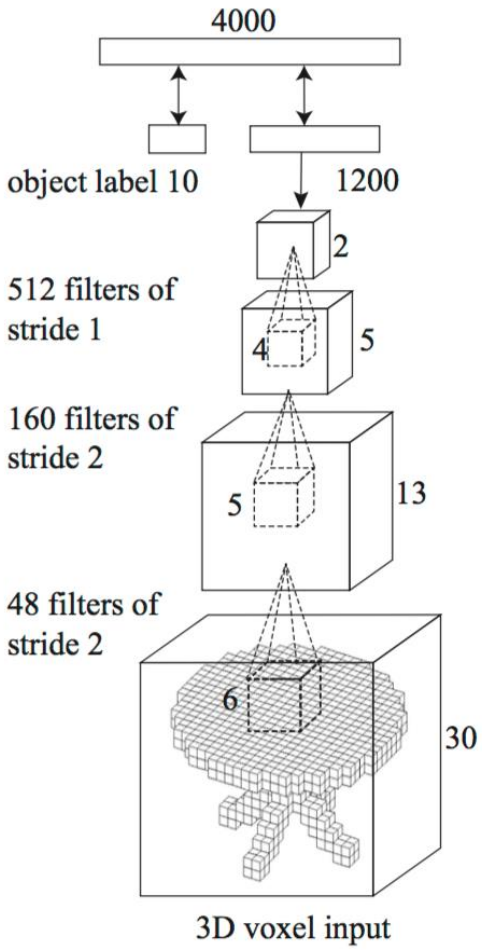
Cause 1: Architecture and Engineering



LeNet, **1998**



AlexNet, **2012**

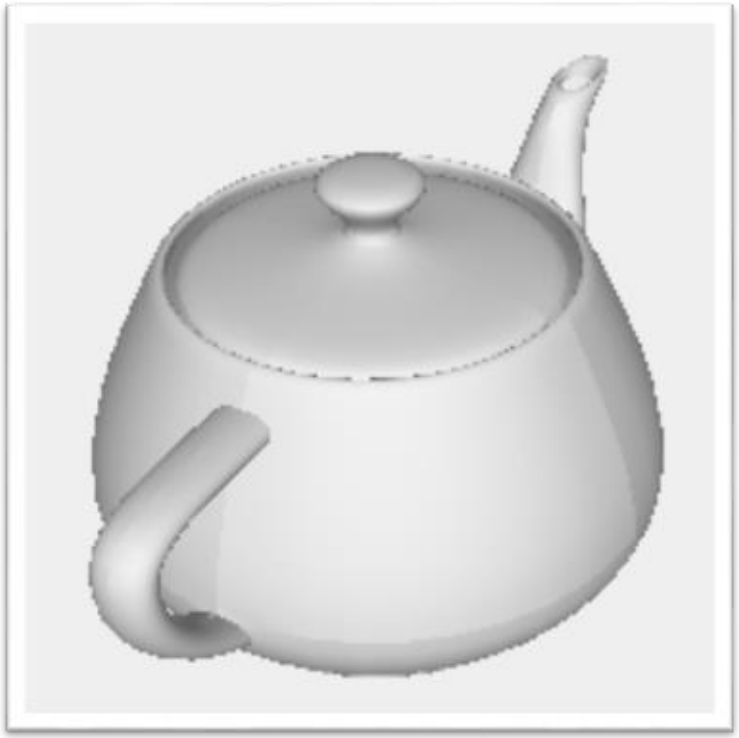


3DShapeNets, **2015**

Cause 2: Resolution

Multi-View CNNs

MVCNN Su et al.

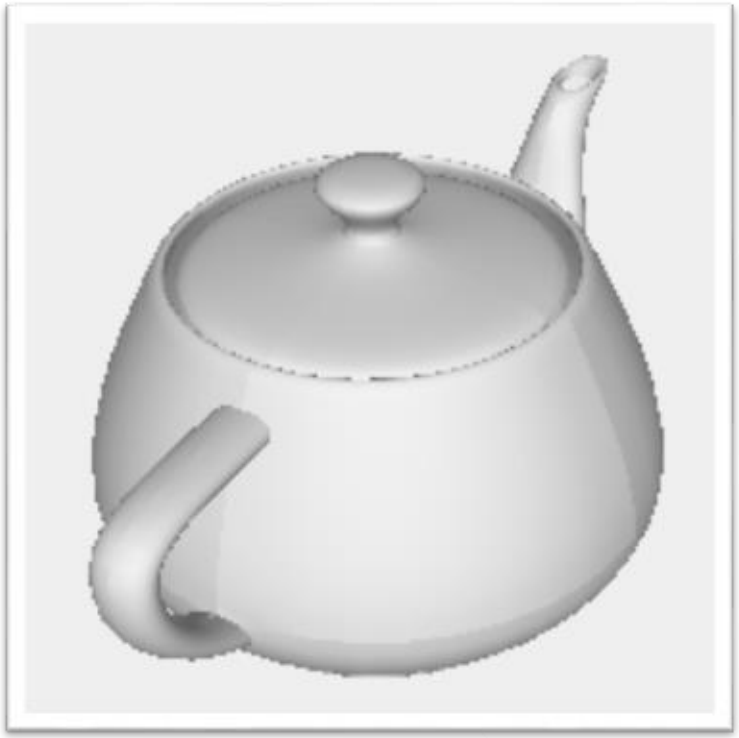


224x224 Images

Cause 2: Resolution

Multi-View CNNs

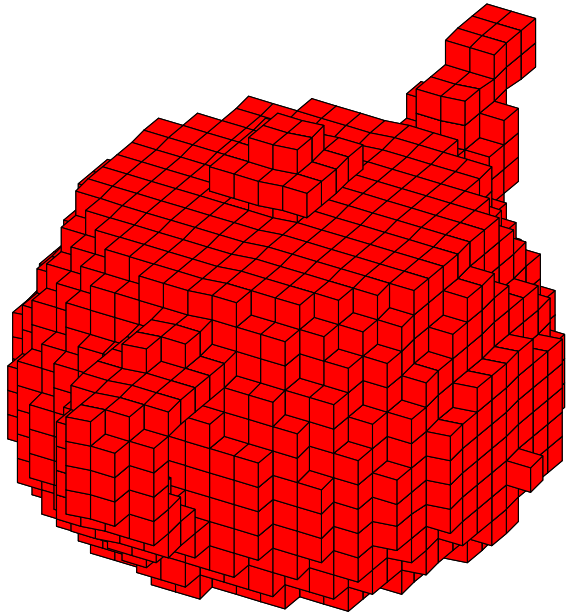
MVCNN Su et al.



224x224 Images

Volumetric CNNs

3DShapeNets Wu et al.

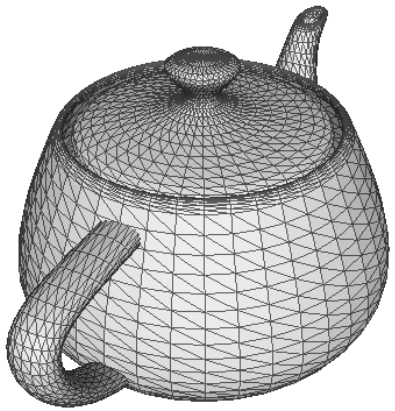


30x30x30 Volumes

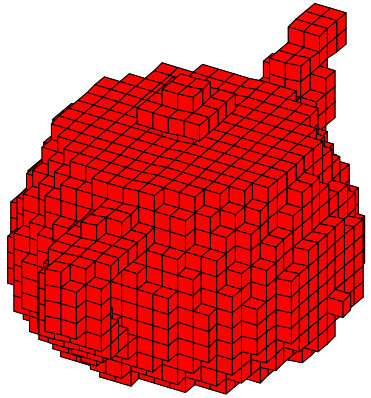
Diagnosis of Causes: Variable Control

- Same resolution, study architectures
- Same architecture, look into resolutions

Sphere Rendering

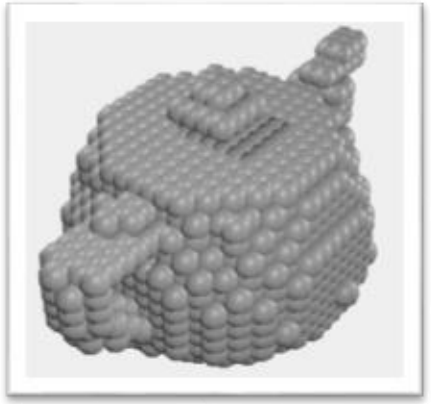


Polygon Mesh



Occupancy Grid

30x30x30

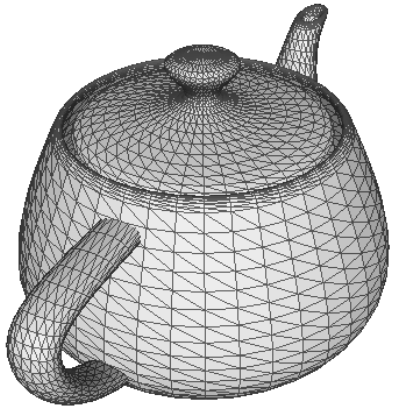


Image

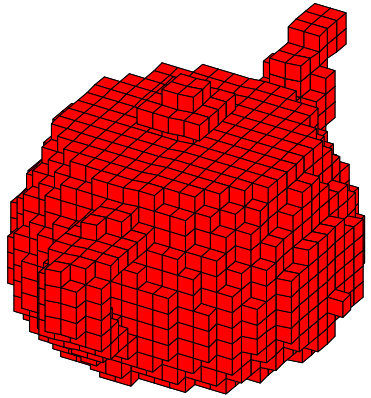
224x224

Sphere Rendering

Same "3D Resolution"

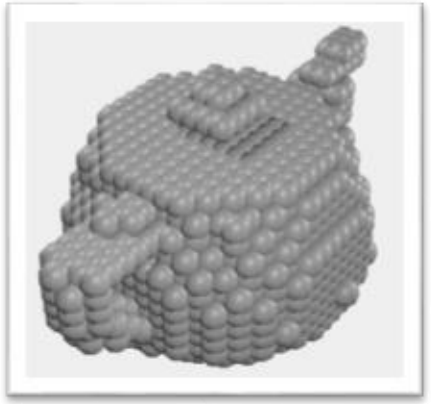


Polygon Mesh



Occupancy Grid

30x30x30



Image

224x224

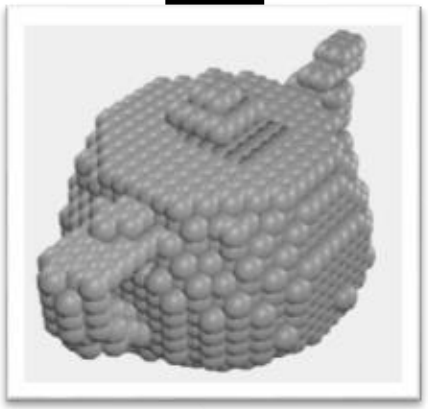
Investigation into Architecture

Different Architecture

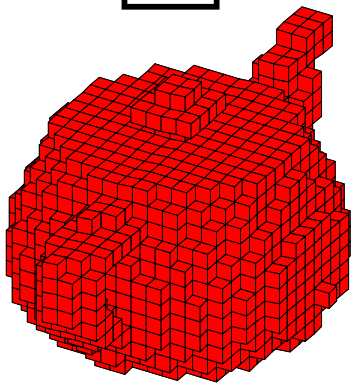
Multi-View Image CNN

3D CNN

Same 3D Resolution (30x30x30)

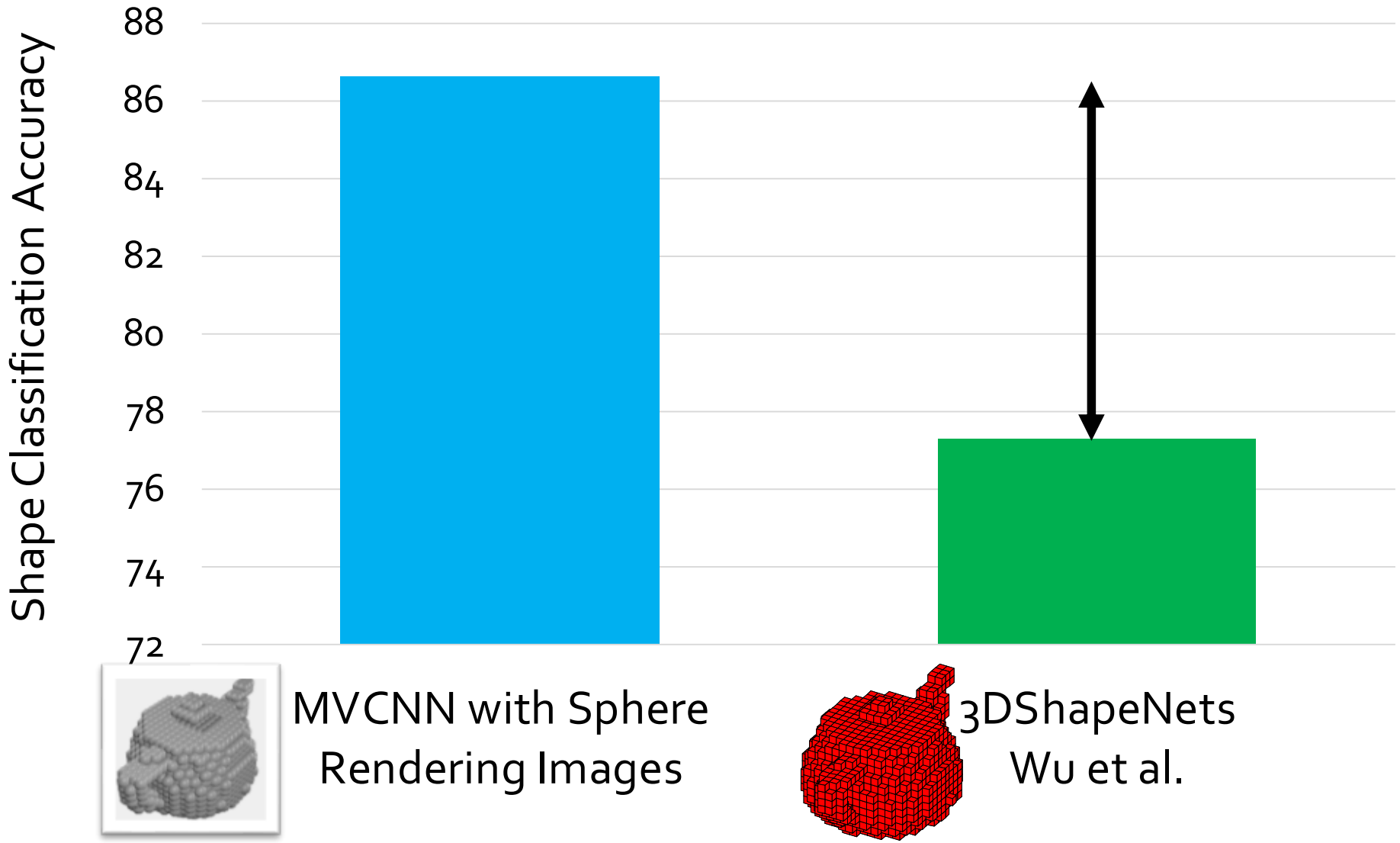


Sphere Rendering Images



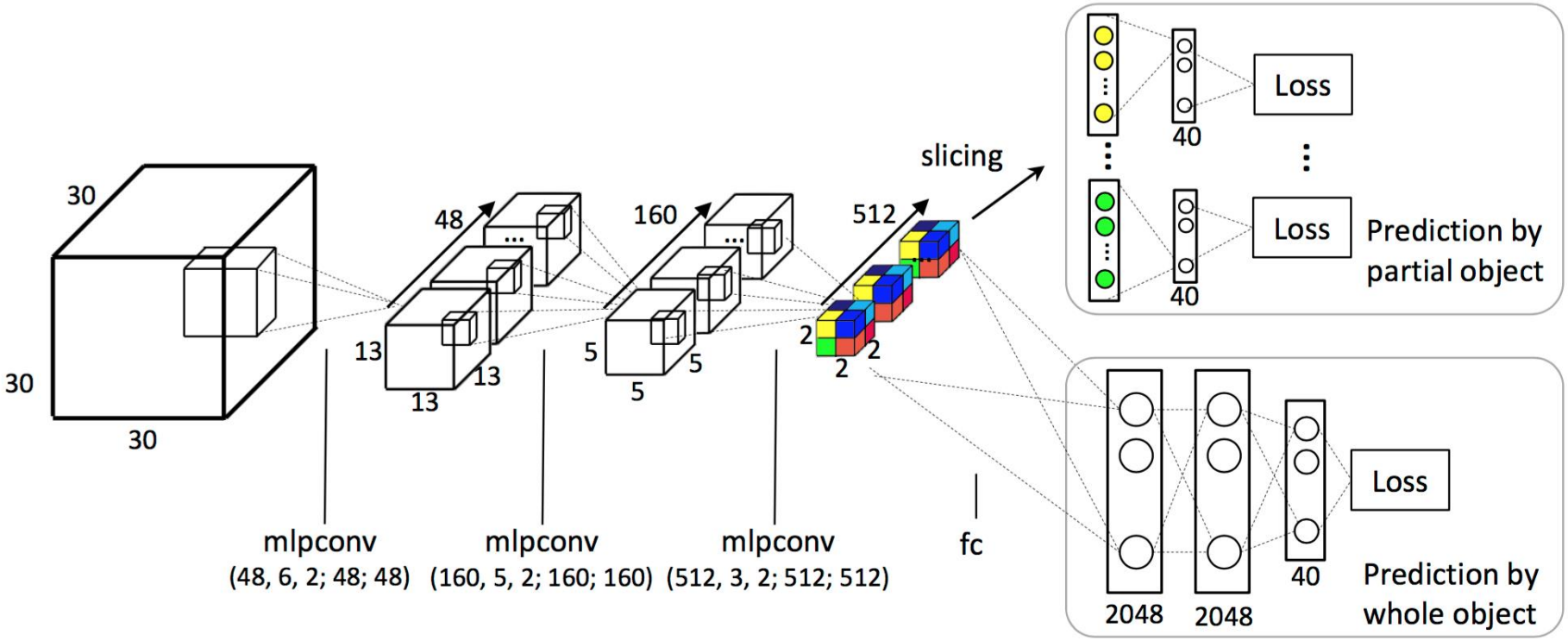
Occupancy Grid Volumes

CNNs with Same 3D Resolution Inputs



Novel 3D CNN Architectures

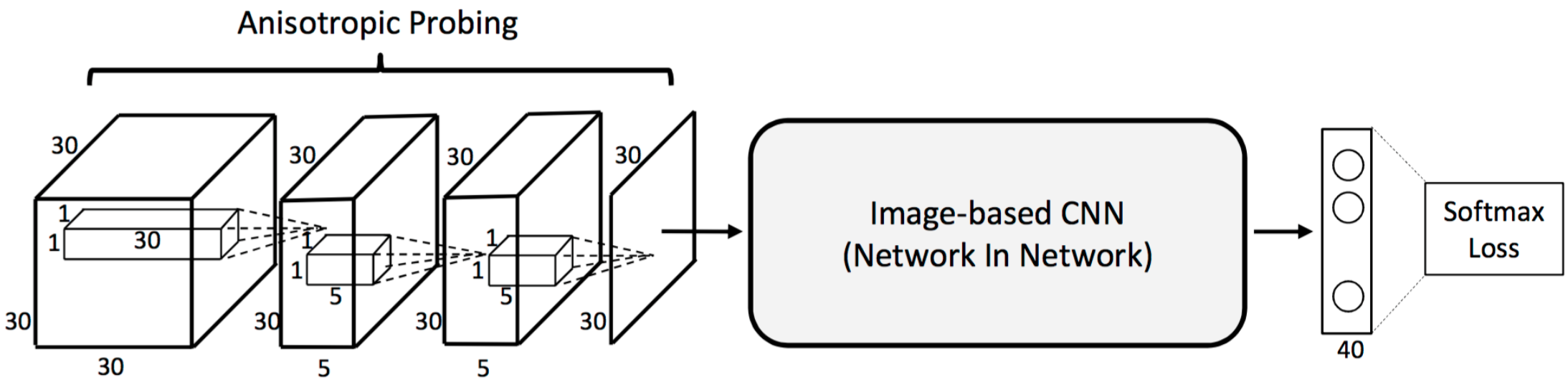
❑ 3D NIN with Subvolume Supervision



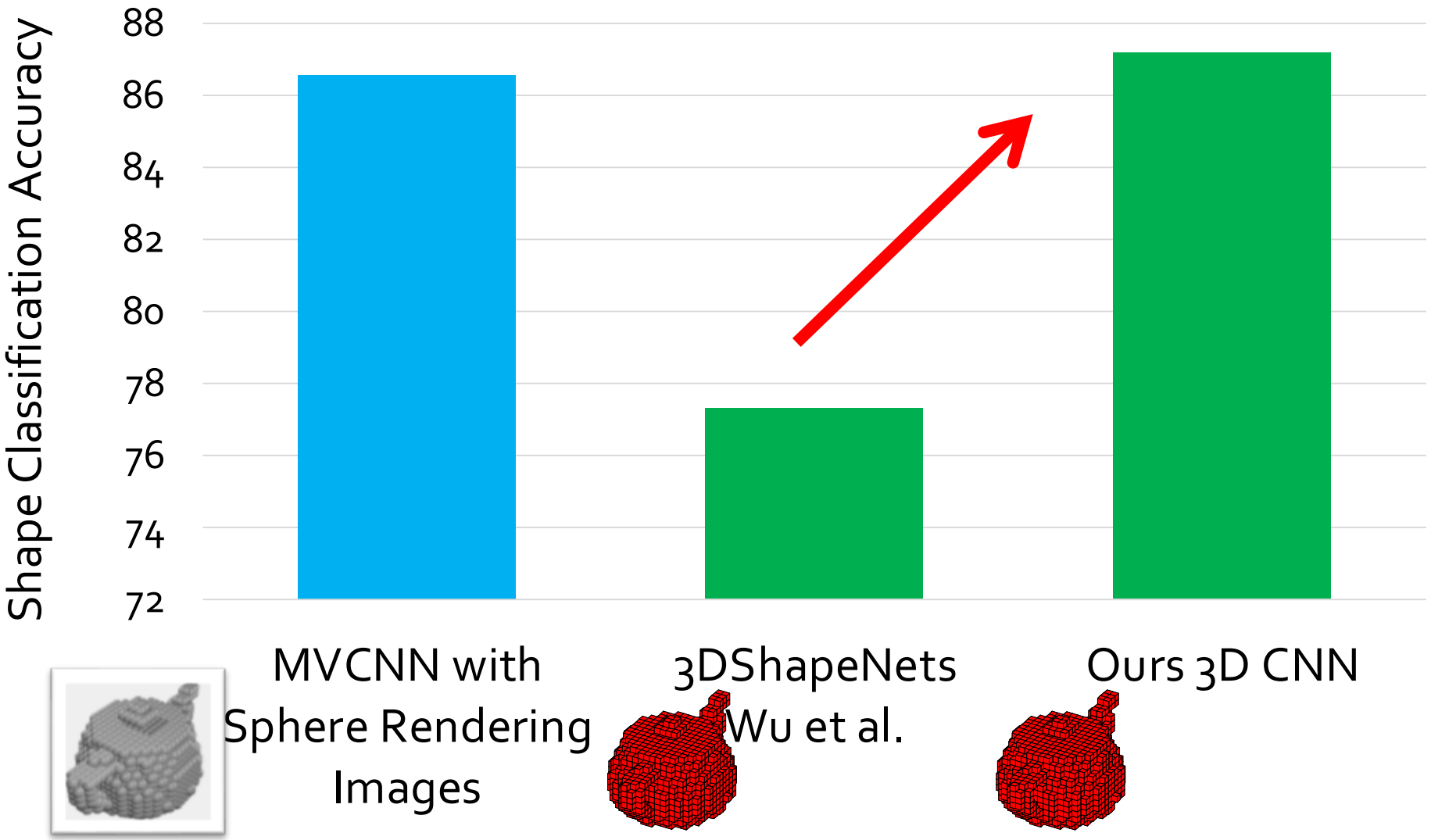
Push Harder for Learning Better!

Novel 3D CNN Architectures

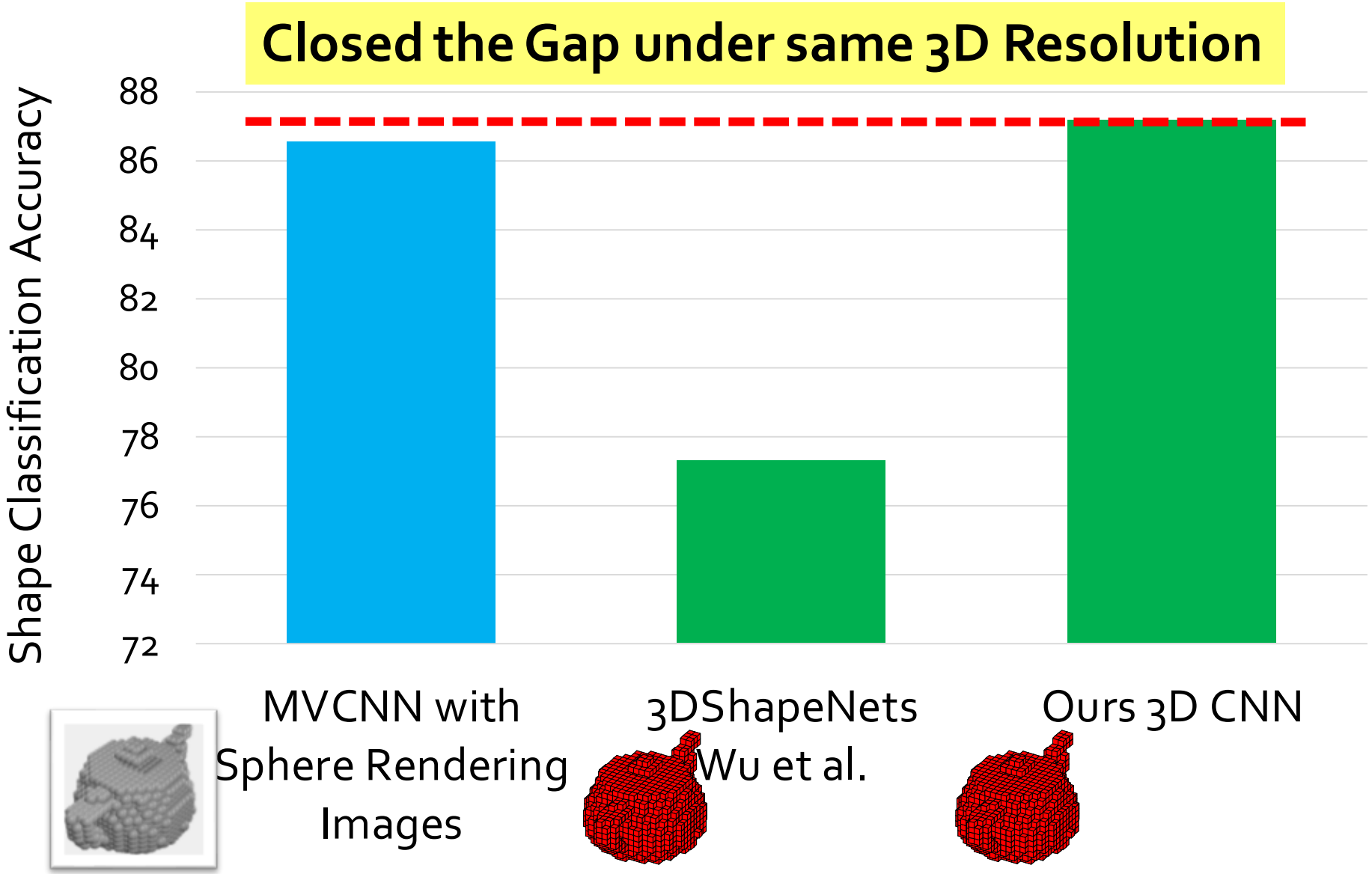
❑Anisotropic Probing Network



Results of Our Novel 3D CNNs



Results of Our Novel 3D CNNs



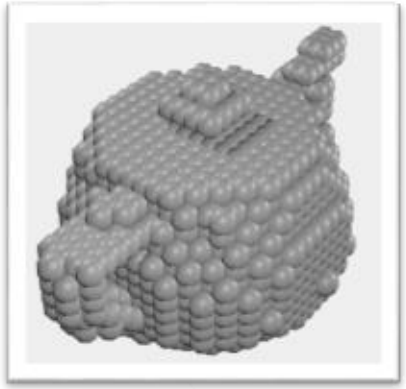
Investigation into Resolution

Multi-View Image CNN

Multi-View Image CNN

Same Architecture

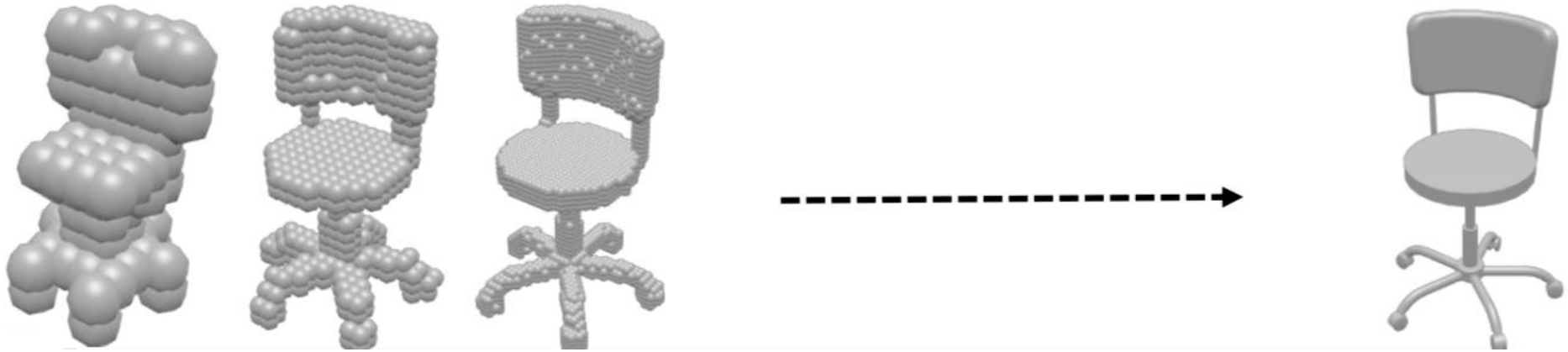
Standard Rendering Images

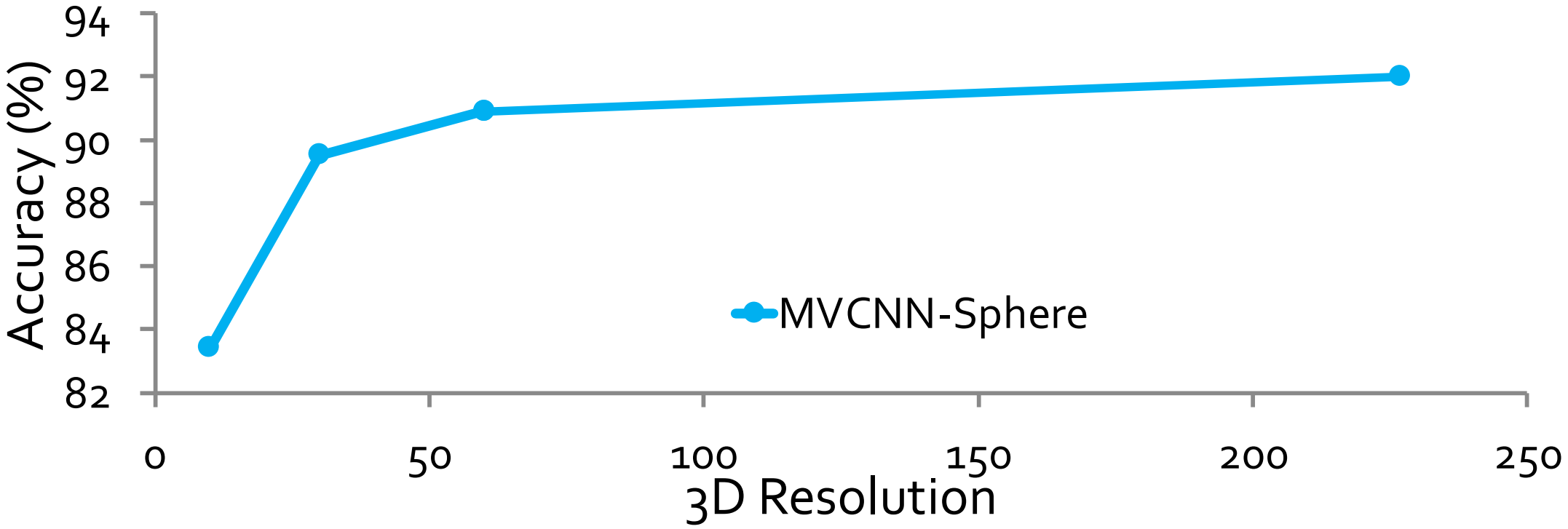


Sphere Rendering Images

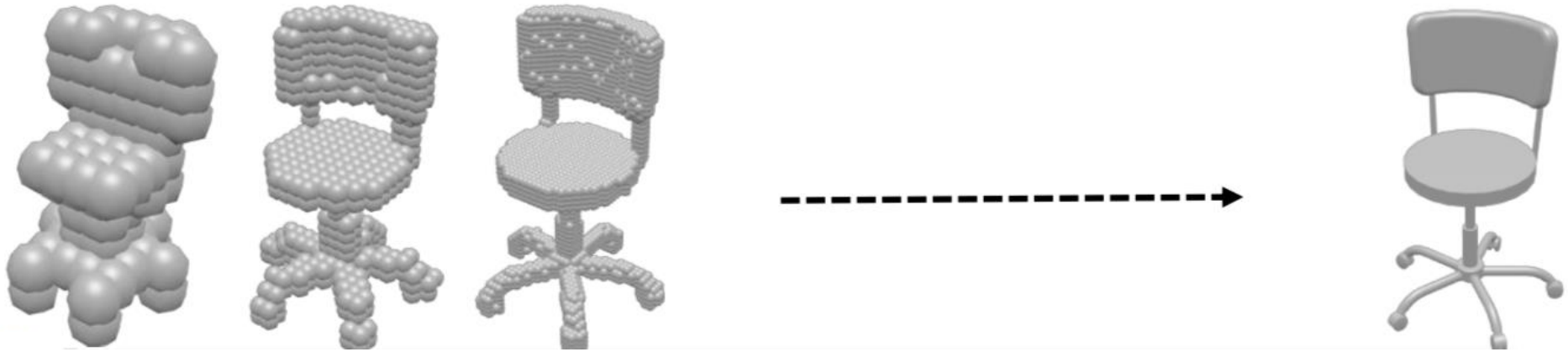
Different 3D Resolution

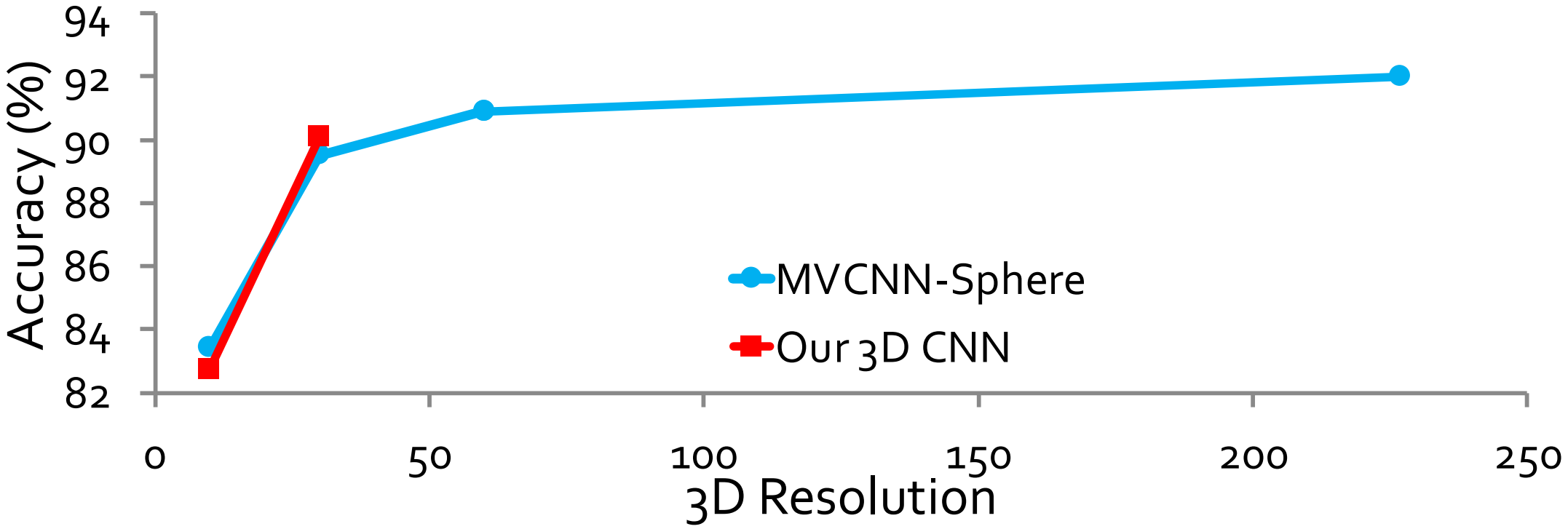
Performance Trend wrt 3D Resolution





Performance Trend wrt 3D Resolution

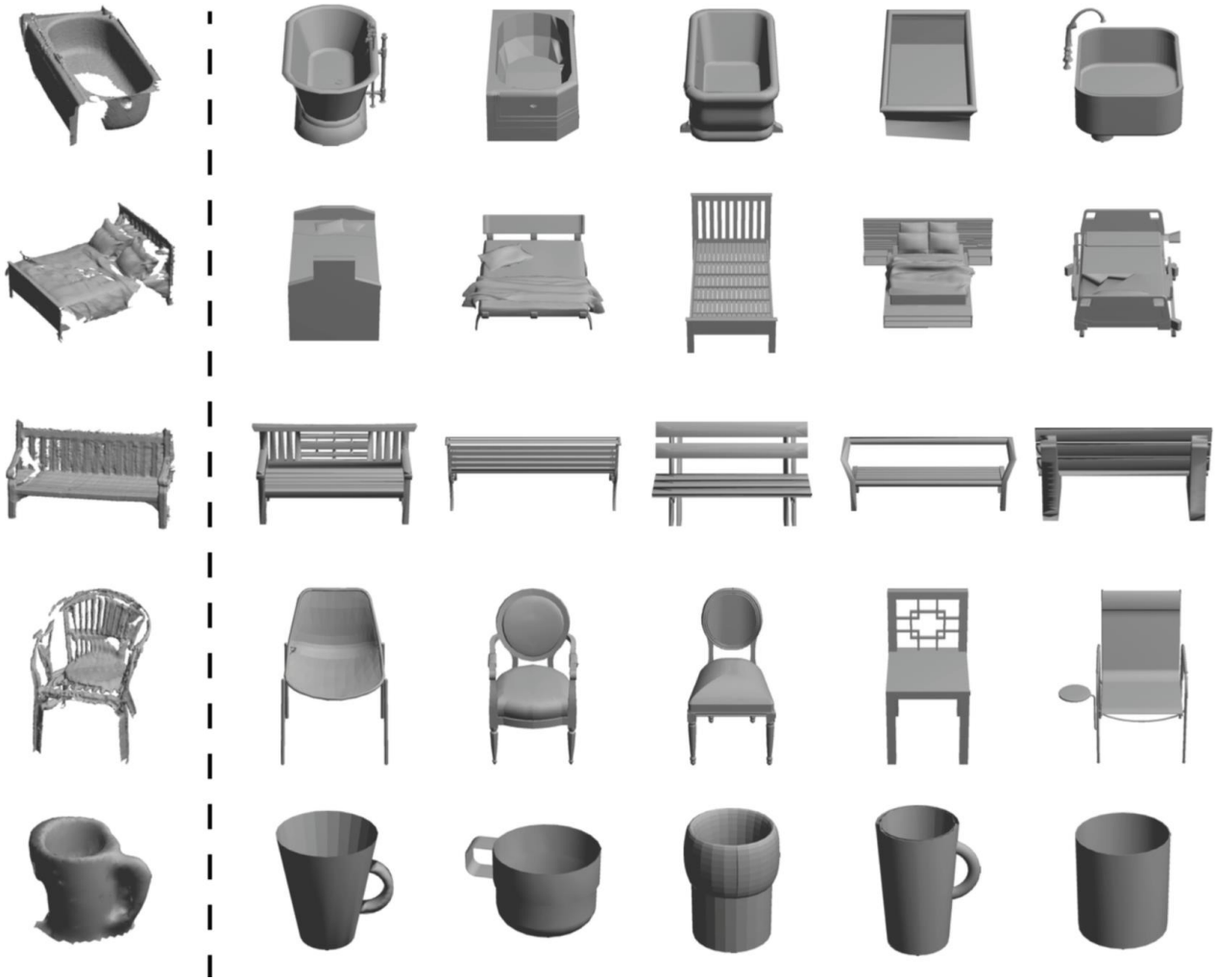




Generalization to Real Scans

Real Scan Dataset
243 objects
12 categories

Shape retrieval on scan data



Volumetric and Multi-View CNNs for Object Classification on 3D Data

Code and Data Available Online!

http://graphics.stanford.edu/projects/3dcnn/

Welcome to Our Poster #38!