# 3D SHAPE REPRESENTATION AND ITS APPLICATIONS

2017-9.6

# CONTENTS/OUTLINE

3D Shape Representation&Applications

- 1 Introduction
- 2 Traditional Shape Representation
- 3 3D Deep Learning
- Conclusion

Growing market of crowd-sourcing for 3D modeling

yobi 3D

Clara.io

Google SketchUp

Sketchfab

TINKERCAD

……

# 1. INTRODUCTION

3D Visual Data in Physical World

Augmented Reality

3D vs. 2D images?

- ✓Real
- ✓Informative
- ✓Expensive to Capture: 3D Scanner
- ✓Uneasy to perform feature extraction

# 1. INTRODUCTION

3D Visual Data in Physical World

Augmented Reality

3D vs. 2D images?

✓Real

✓Informative

✓Expensive to Capture: 3D Scanner

✓Uneasy to perform feature extraction

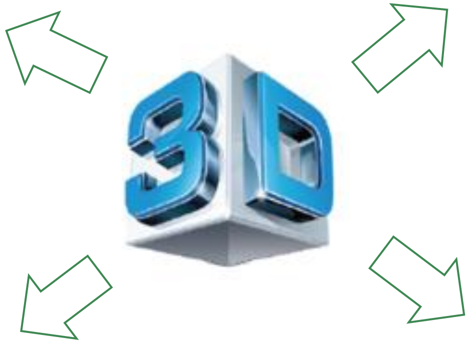
Why study 3D Feature?:  e.g. Reuse

# 1. INTRODUCTION


Robotics

## Applications Fields

- Industry
- Entainment
- Medical
- Geology
- Cultrual Relic Protection
- E-commerce



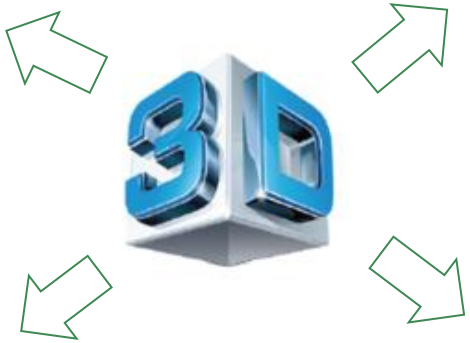
# 1. INTRODUCTION

## Applications Fields

- Industry
- Entainment
- Medical
- Geology
- Cultrual Relic Protection
- E-commerce


Robotics


3D Printer



# 1. INTRODUCTION

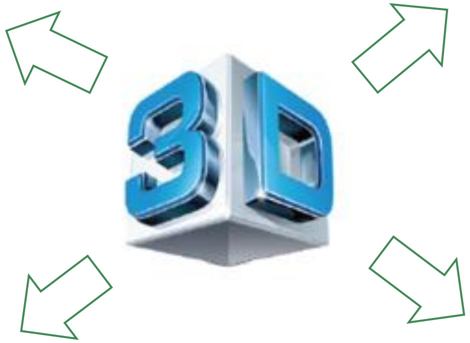
## Applications Fields

- Industry
- Entainment
- Medical
- Geology
- Cultrual Relic Protection
- E-commerce

Robotics

3D Printer

3D Film/Game/Design



# 1. INTRODUCTION

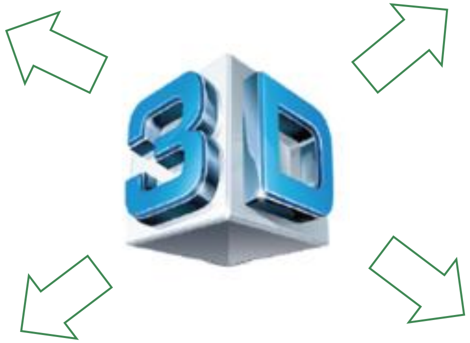
## Applications Fields

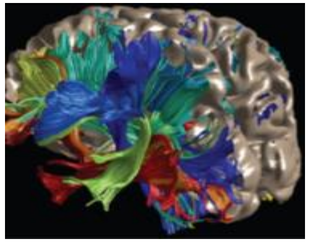
- Industry
- Entainment
- Medical
- Geology
- Cultrual Relic Protection
- E-commerce


Robotics


3D Printer


3D Film/Game/Design




Medical Image Processing

# 1. INTRODUCTION

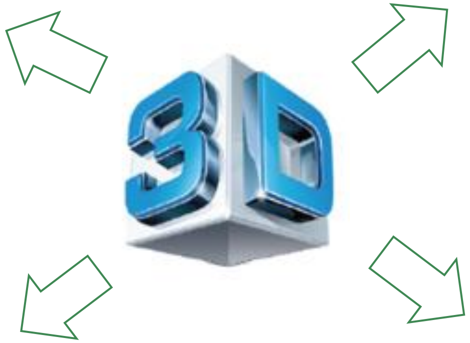
## Applications Fields

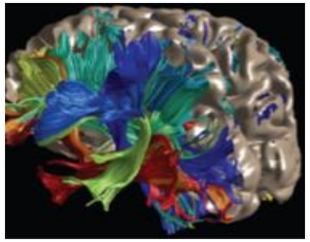
- Industry
- Entainment
- Medical
- Geology
- Cultrual Relic Protection
- E-commerce

Robotics

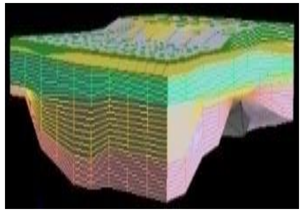
3D Printer

3D Film/Game/Design





Medical Image Processing



Geology

# 1. INTRODUCTION

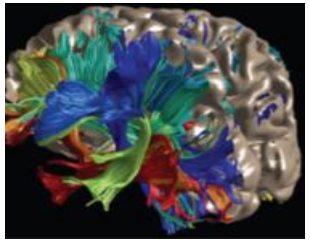
## Applications Fields

- Industry
- Entainment
- Medical
- Geology
- Cultrual Relic Protection
- E-commerce

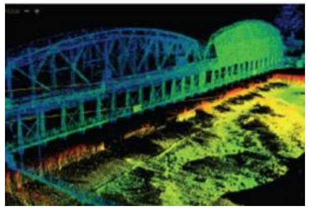
Robotics

3D Printer

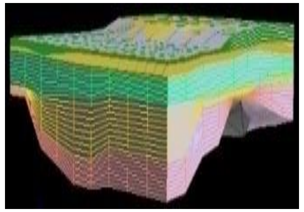
3D Film/Game/Design



Medical Image Processing



Automatic Driving



Geology

# 1. INTRODUCTION

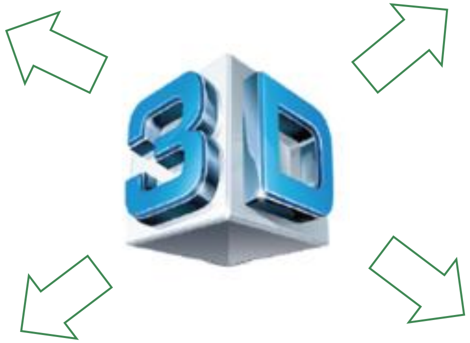
## Applications Fields

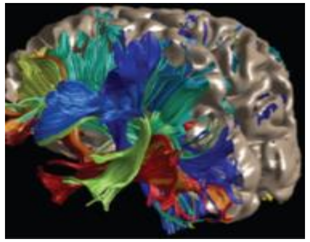
- Industry
- Entainment
- Medical
- Geology
- Cultrual Relic Protection
- E-commerce

Robotics

3D Printer

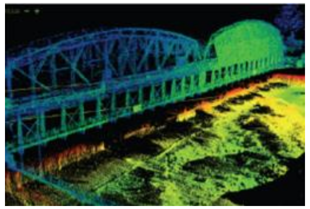
3D Film/Game/Design



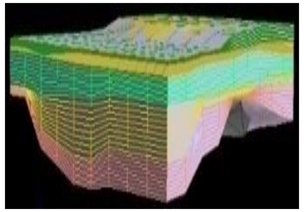


Medical Image Processing

Cultrual Relic Protection



Automatic Driving

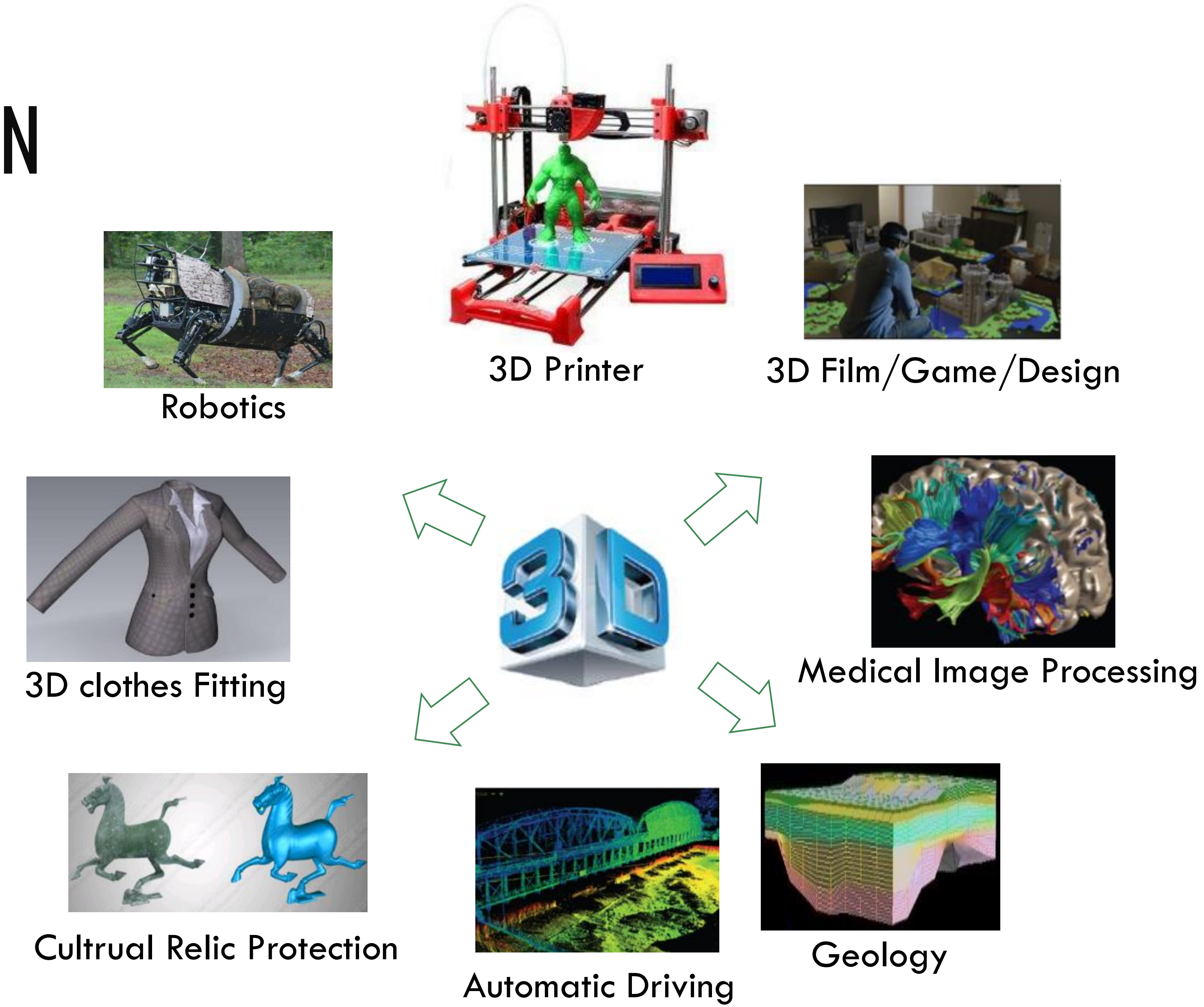


Geology

# 1. INTRODUCTION

## Applications Fields

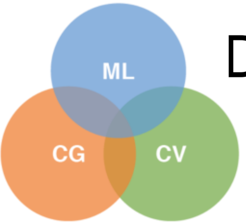
- Industry
- Entainment
- Medical
- Geology
- Cultrual Relic Protection
- E-commerce



# 1. INTRODUCTION

## Fundermental Research

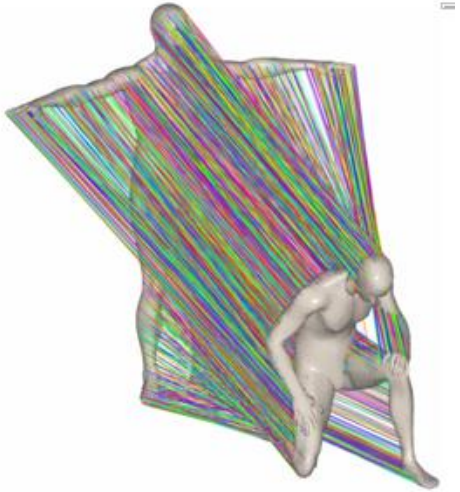
- Object Classification/Retrieval
- Segmentation
- Matching&Correspondence
- Reconstruction



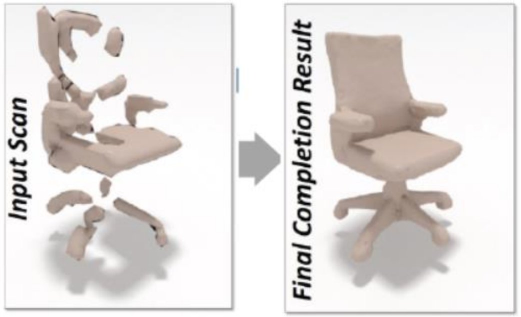

## Depends on Feature Representation

- Global Feature
- Local Feature (keypoint based)





Monocular
3D reconstruction




Shape completion

Shape modeling

# 1. INTRODUCTION

## 3D-assisted image analysis


Cross-view image retrieval

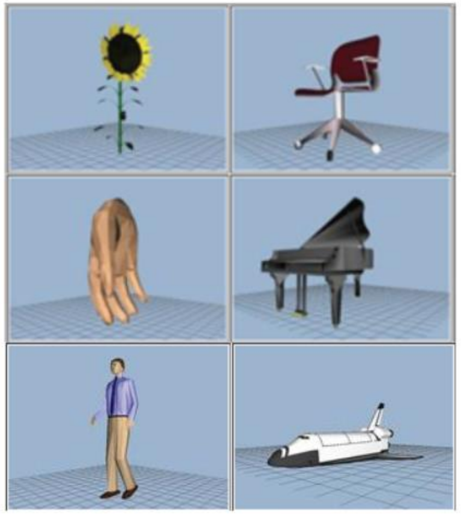
Intrinsic decomposition

# 1. INTRODUCTION

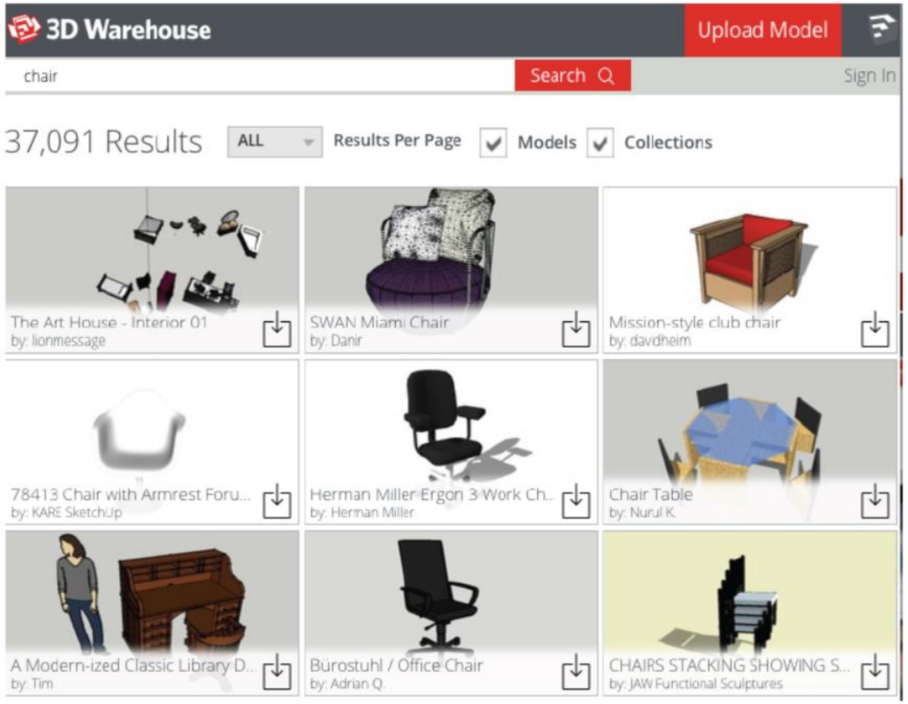
Benchmarks&Datasets: Princeton Shape Benchmark, 3D Warehouse, SHREC Contest, ...

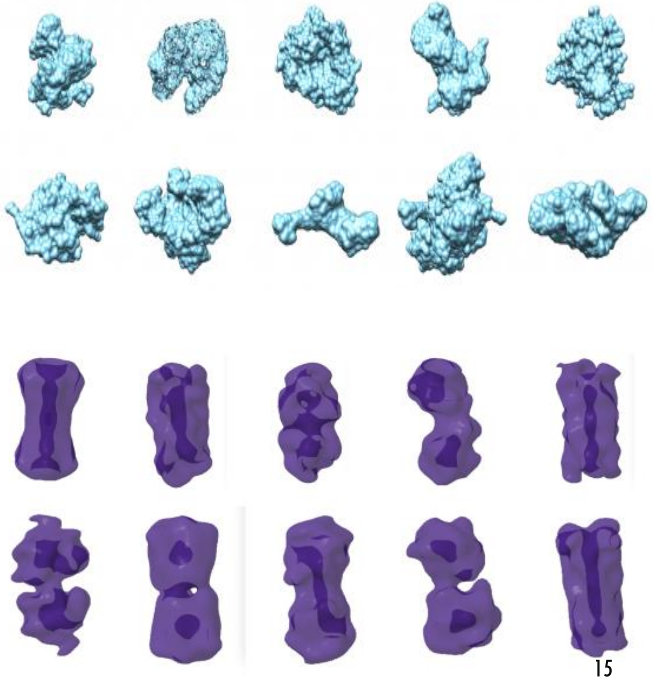
Focuses: General Shapes/Non-Rigid Shapes/Scene Shapes

1800 models in 90 categories



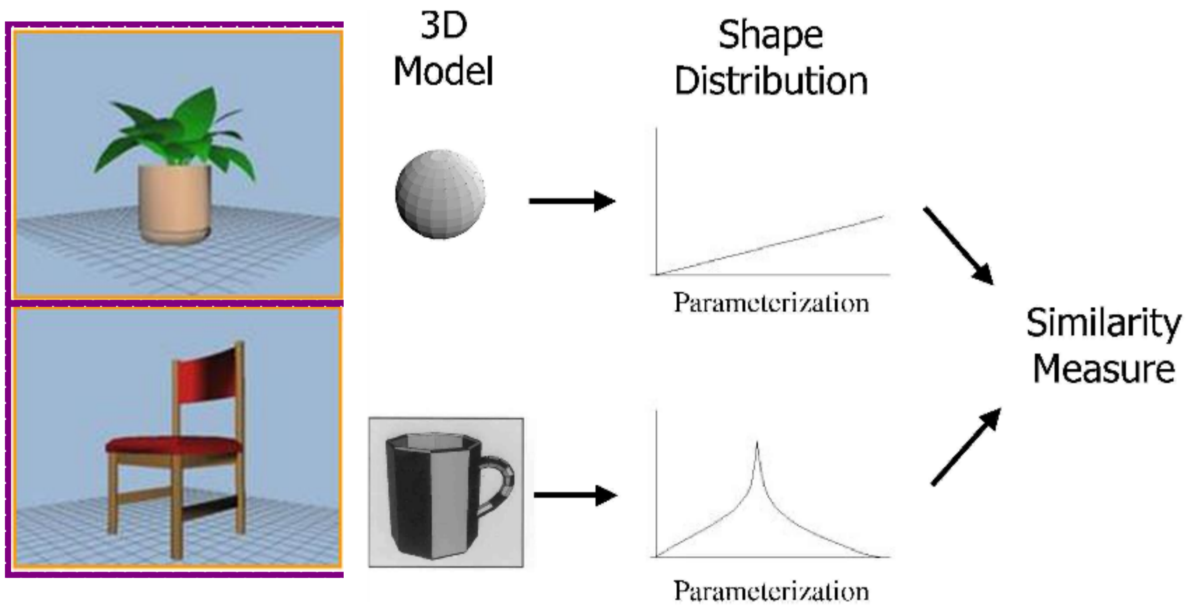
Princeton shape benchmark
[Shilane et al. 04]



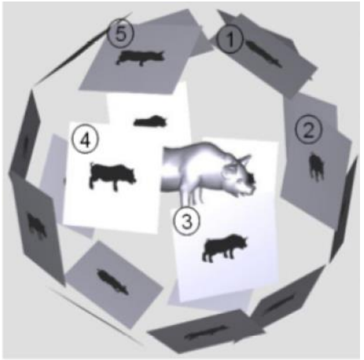



# 2 TRADITIONAL APPROACH

Shape Distributions, Princeton 2002

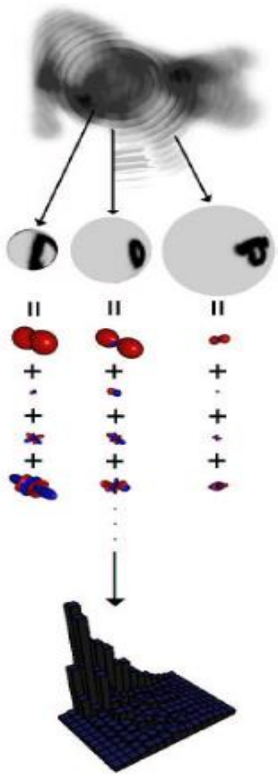



Light Field Descriptor, Chen et al.2003

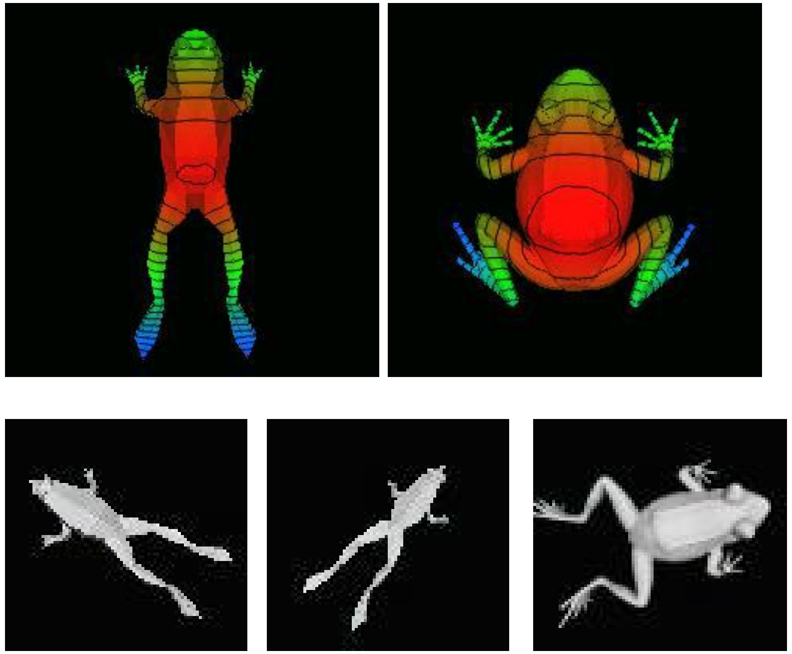


# 2 TRADITIONAL APPROACH

Sphere Harmonic, Princeton 2003



Topology Skeleton 2003

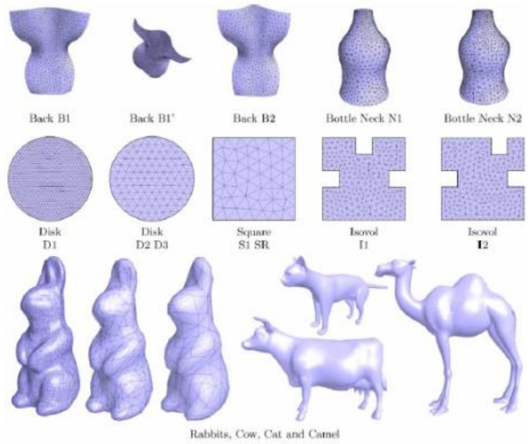


# 2 TRADITIONAL APPROACH

Features Defined on Laplace Spectral Domain

$$(\Delta_X - \frac{\partial}{\partial t})u(x,t) = 0$$

$$\boxed{h_t(x,y) = \sum_{l=1}^{\theta} e^{-\lambda_l t}\phi_l(x)\phi_l(y)}$$




ShapeDNA 2006

Global Point Signature(GPS) 2007

Heat Kernel Signature/ShapeGoogle 2009-2011

$$F_{DNA}(X) = (\lambda_1, \lambda_2, \cdots, \lambda_{\theta-1}, \lambda_\theta)$$

$$GPS(x) = (\frac{\phi_1(x)}{\sqrt{\lambda_1}}, \frac{\phi_2(x)}{\sqrt{\lambda_2}}, \cdots \frac{\phi_\theta(x)}{\sqrt{\lambda_\theta}})$$

$$HKS(x,t) = \sum_{l=1}^{\theta} e^{-\lambda_l t}\phi_l^2(x)$$

# 2 TRADITIONAL APPROACH

Basic Procedure:
Step 1： Model Sampling
Step 2： Pairwise Distances
Step 3： Histogram Vectors

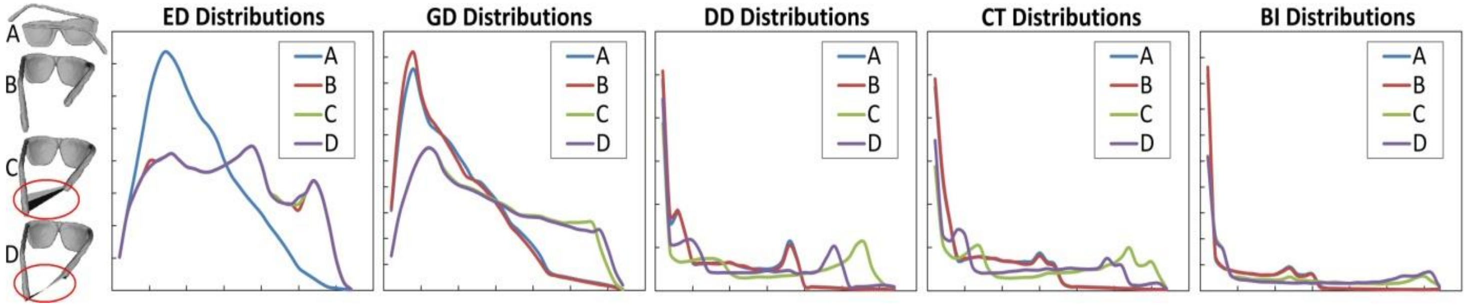
## Intrinsic Shape Contexts

- Geodesic Distance 2012

## Distance Feature

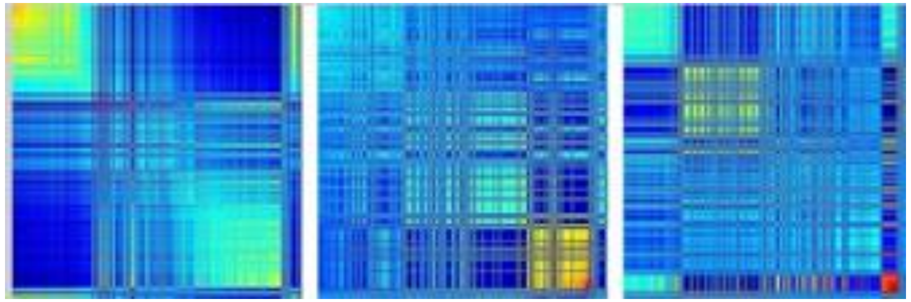
- Euclidean Distance
- Geodesic Distance
- Spectral Distance 2010-2015

$$d_S(x,y) = \| \boldsymbol{\psi}(x) - \boldsymbol{\psi}(y) \|_{L2}^2$$
$$= \begin{cases} d_D(t,x,y), & \sigma(\lambda_l) = e^{-\lambda_l t} \\ d_C(t,x,y), & \sigma(\lambda_l) = 1/\sqrt{\lambda_l} \\ d_B(t,x,y), & \sigma(\lambda_l) = 1/\lambda_l \end{cases}$$



# 2 TRADITIONAL APPROACH

## Modal Transformation Feature, 2015



Step1： Intrinsic function $H: V \otimes V \rightarrow \mathbb{R}$

$$H(x_i, x_j) = g(x_i) \otimes g(x_j), 1 \leq i, j \leq n \qquad g(x_i) \otimes g(x_j) = g(x_j) \otimes g(x_i)$$

Step 2： Define Modal feature $MHF(X) = \dfrac{\mu}{\| \mu \|_{L2}}$

$$\mu = (\mu_1, \mu_2, \cdots, \mu_i, \cdots, \mu_k), \qquad k \ll n \qquad \mu_1 > \cdots > \mu_i > \cdots > \mu_k, \qquad H\Phi_i = \mu_i \Phi_i$$

## Integrate Heat Kernel Signature, 2015

$$IHKS(x) = \sum_l \phi_l^2(x) / \lambda_l$$

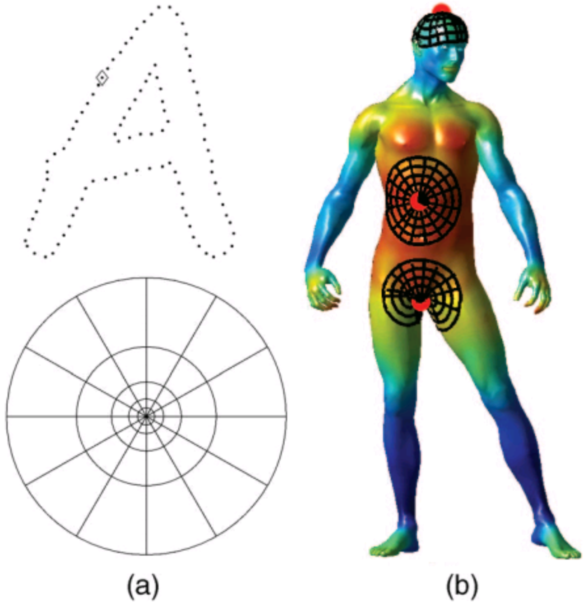
$$= \int_t HKS(x, t) dt$$

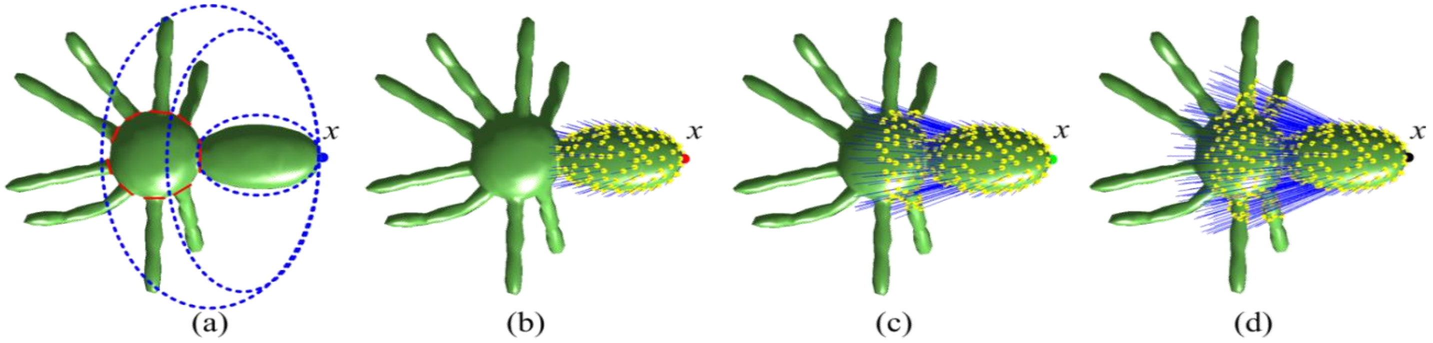
$$h(x, y, T) = \sum_l \phi_l(x) \phi_l(y) \mathbb{k}(\lambda_l, T) / \pi(\lambda_l, T)$$

# 2 TRADITIONAL APPROACH

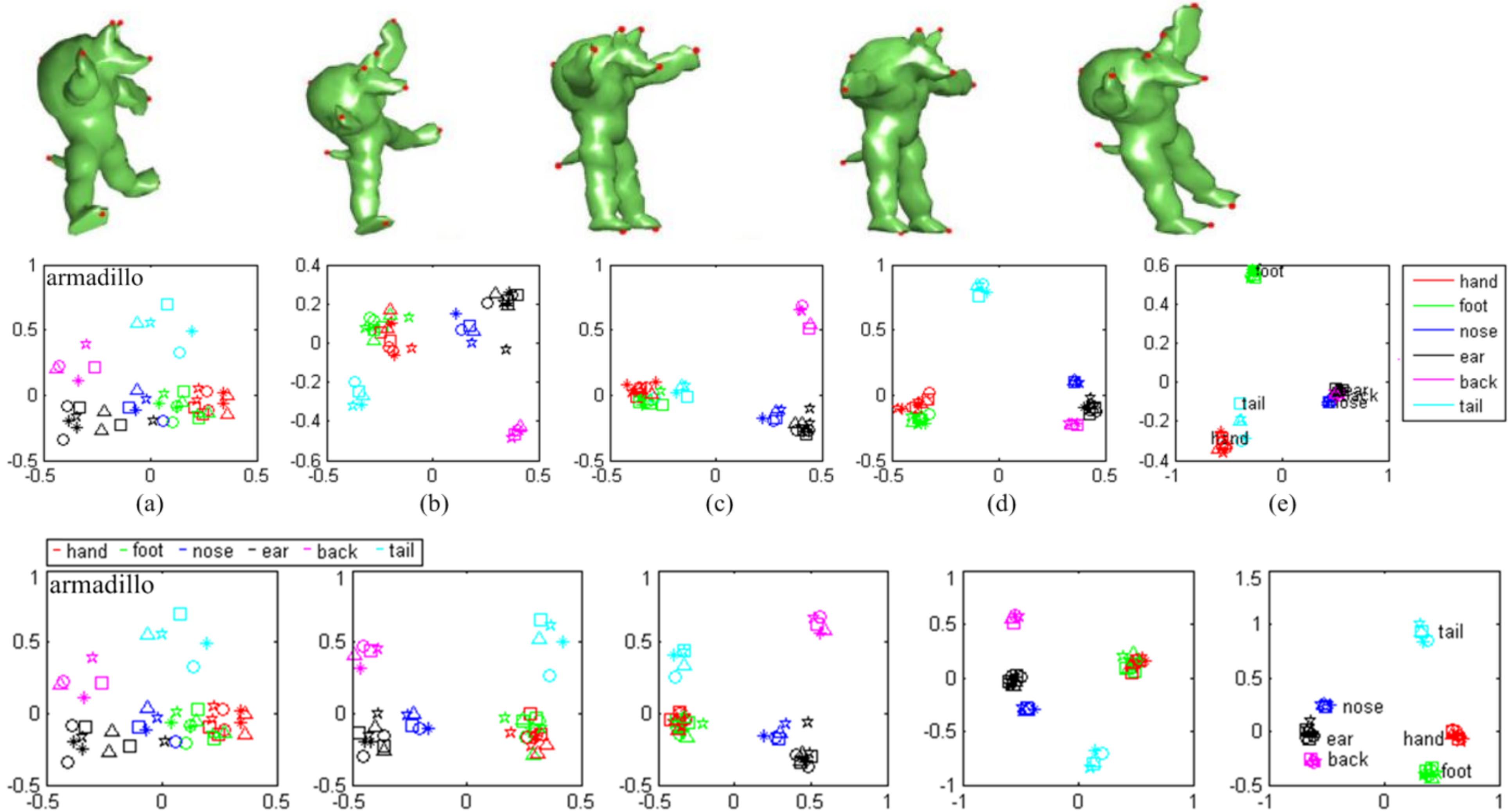
## Multiscale Shape Context 2016

Critical Point Detection=>Feature Extraction => Global Vector Quantization



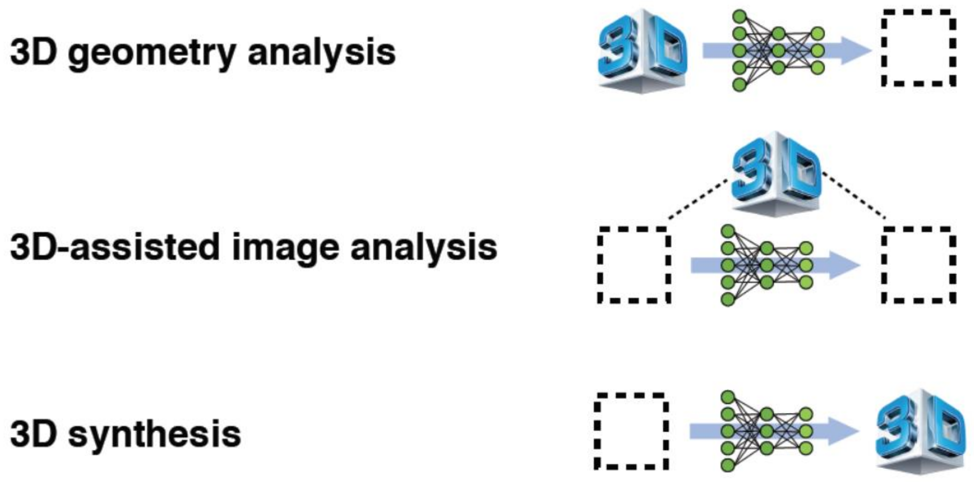
# 2 TRADITIONAL APPROACH



# 3 DEEP LEARNING APPROACH

## 3D Deep Learning Issues

- Different Applications
- Different 3D Data Representations



# 3 DEEP LEARNING APPROACH

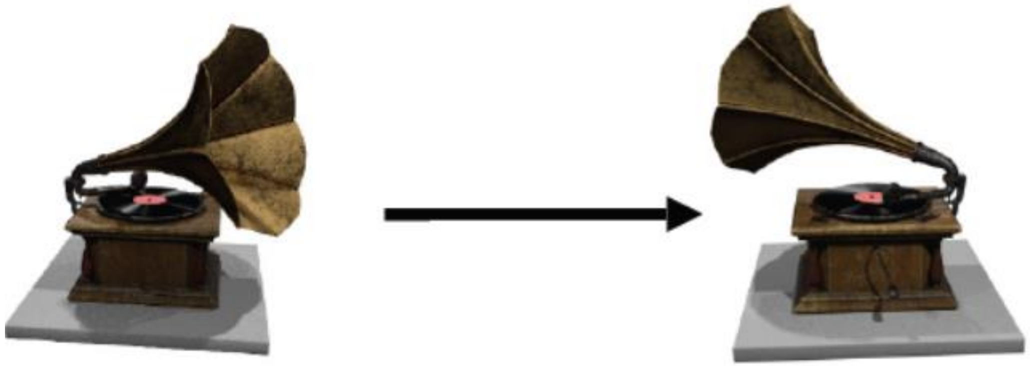
Multi-view RGB-D images

Volumetric

Polygonal Mesh

Point Cloud

Primitive-based CAD



Novel view image synthesis

# 3 DEEP LEARNING APPROACH

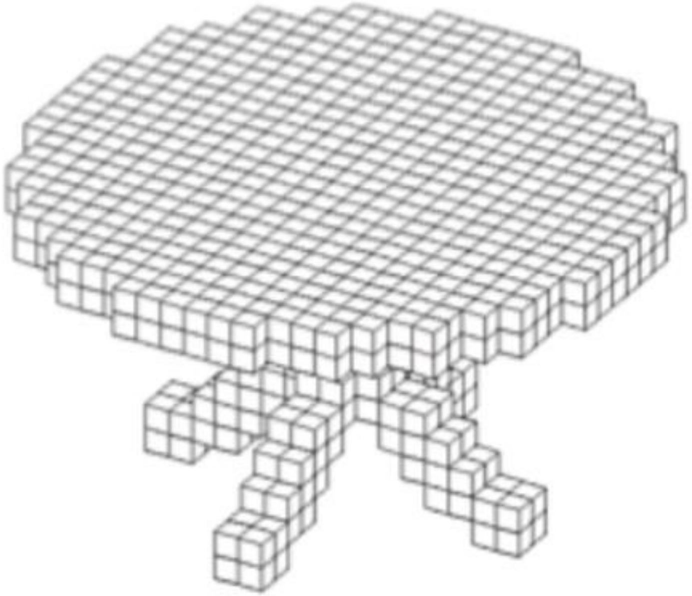
Multi-view RGB-D images

Volumetric

Polygonal Mesh

Point Cloud

Primitive-based CAD



# 3 DEEP LEARNING APPROACH

Multi-view RGB-D images

Volumetric

Polygonal Mesh

Point Cloud

Primitive-based CAD

# 3 DEEP LEARNING APPROACH

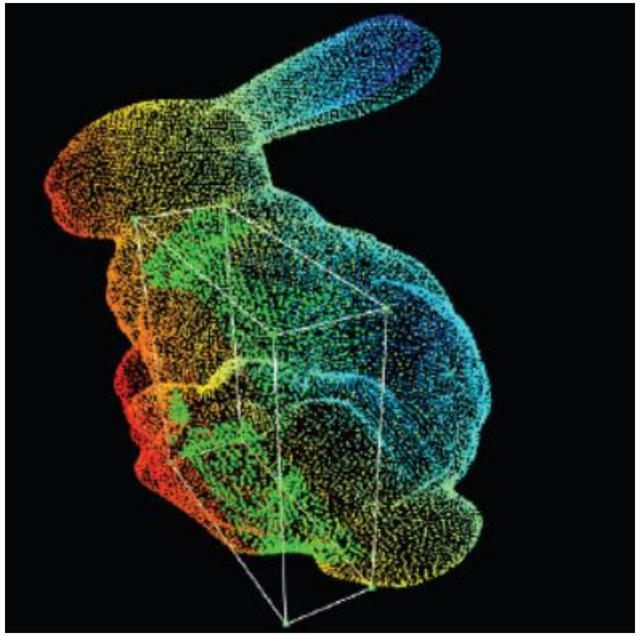
Multi-view RGB-D images

Volumetric

Polygonal Mesh

Point Cloud

Primitive-based CAD model



# 3 DEEP LEARNING APPROACH

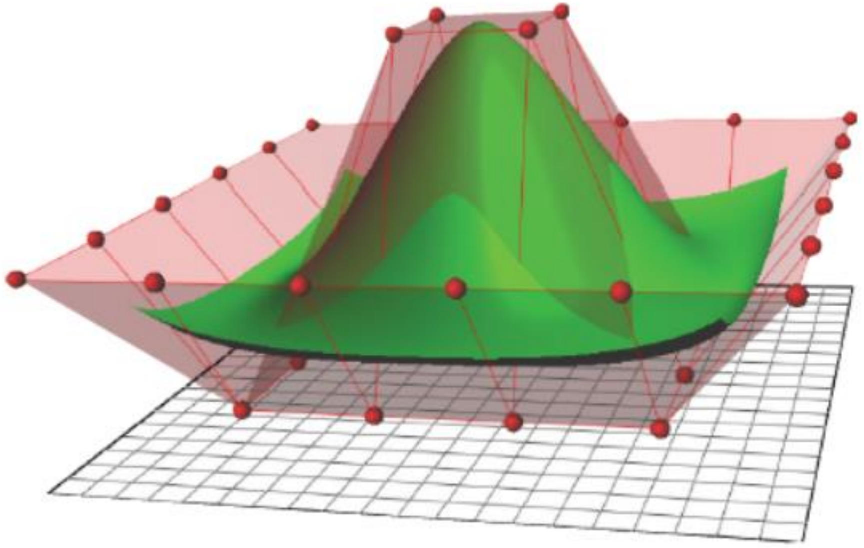
Multi-view RGB-D images

Volumetric

Polygonal Mesh

Point Cloud

Primitive-based CAD model



# 3 DEEP LEARNING APPROACH

Multi-view RGB-D images

Volumetric

Rasterized form (regular grids)

Polygonal Mesh

Point Cloud

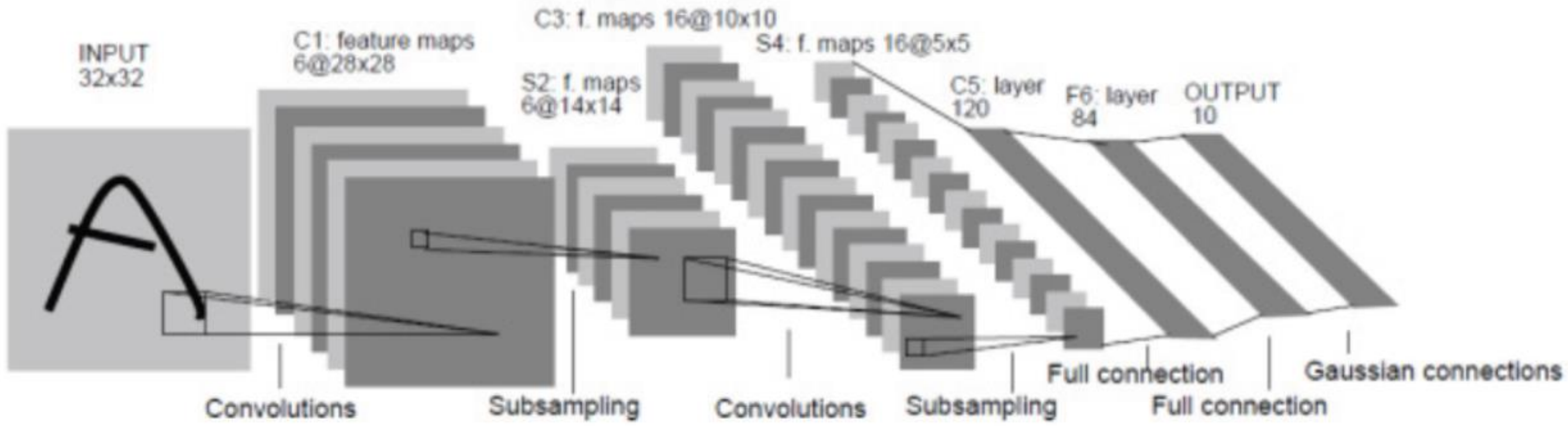
Primitive-based CAD model

Geometric form (irregular grids)

# 3 DEEP LEARNING APPROACH

## Fundamental Challenges

- Can we directly apply CNN on 3D data?



# 3 DEEP LEARNING APPROACH

- Convolution needs an underlying structure

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| 1 | 44 | 33 | 12 | 20 | 23 | 35 | 14 |
| 51 | 16 | 40 | 32 | 46 | 48 | 28 | 17 |
| 29 | 60 | 3 | 63 | 49 | 55 | 36 | 7 |
| 52 | 22 | 26 | 41 | 38 | 10 | 61 | 53 |
| 2 | 24 | 19 | 11 | 34 | 43 | 5 | 8 |
| 57 | 9 | 37 | 42 | 25 | 21 | 27 | 18 |
| 30 | 56 | 50 | 64 | 4 | 59 | 6 | 13 |
| 58 | 47 | 45 | 31 | 39 | 15 | 62 | 54 |

$$(f * g)[n] = \sum_{m=-M}^{M} f[n-m]g[m]$$

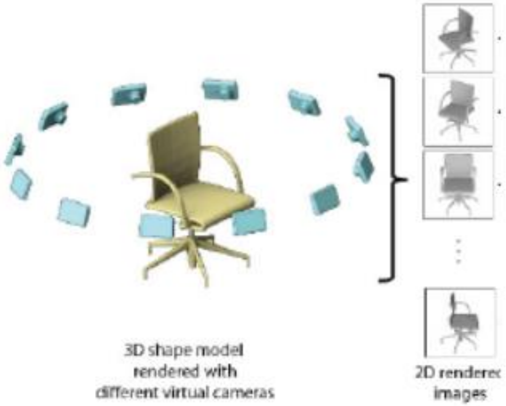
**Rasterized form**

- Can directly apply CNN
- Has other challenges

**Geometric form**

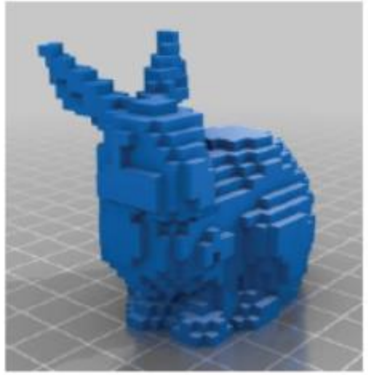
- Cannot directly apply CNN

# 3 DEEP LEARNING APPROACH




Multi-view

[Su et al. 2015]
[Kalogerakis et al. 2016]
...

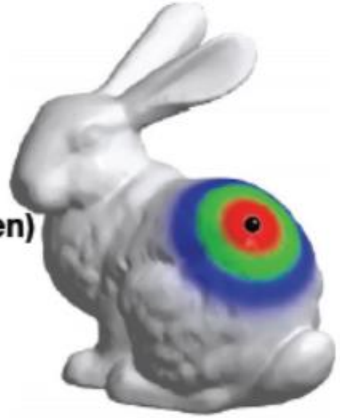


Volumetric

[Maturana et al. 2015]
[Wu et al. 2015] (GAN)
[Qi et al. 2016]
[Liu et al. 2016]
**[Wang et al. 2017] (O-Net)**
**[Tatarchenko et al. 2017] (OGN)**
...

Point cloud

**[Qi et al. 2017] (PointNet)**
**[Fan et al. 2017] (PointSetGen)**



Mesh (Graph CNN)

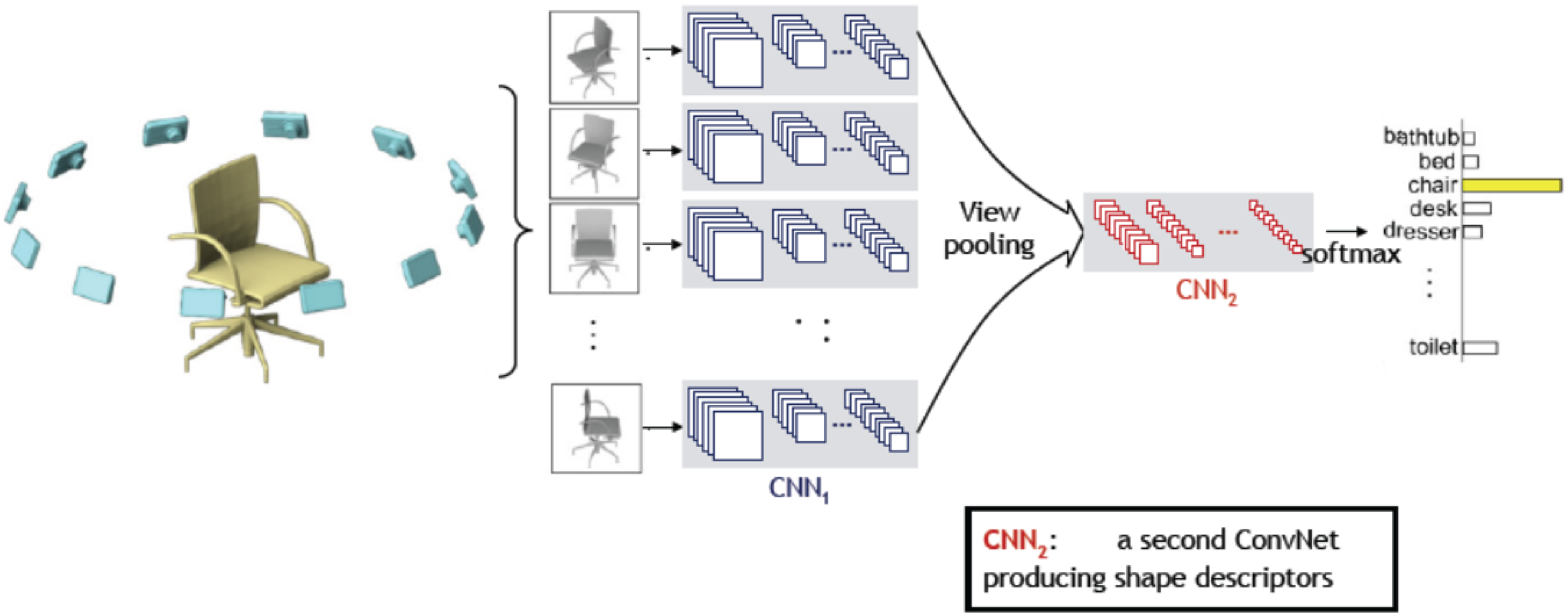
[Defferard et al. 2016]
[Henaff et al. 2015]
**[Yi et al. 2017] (SyncSpecCNN)**
...

Part assembly

**[Tulsiani et al. 2017]**
**[Li et al. 2017] (GRASS)**

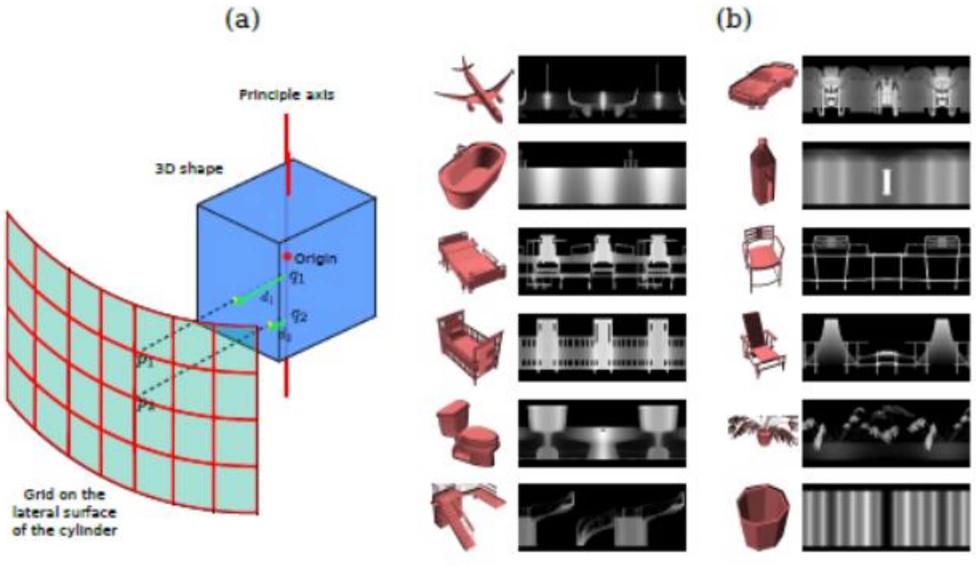
# DEEP LEARNING ON MULTIVIEW REPRESENTATION

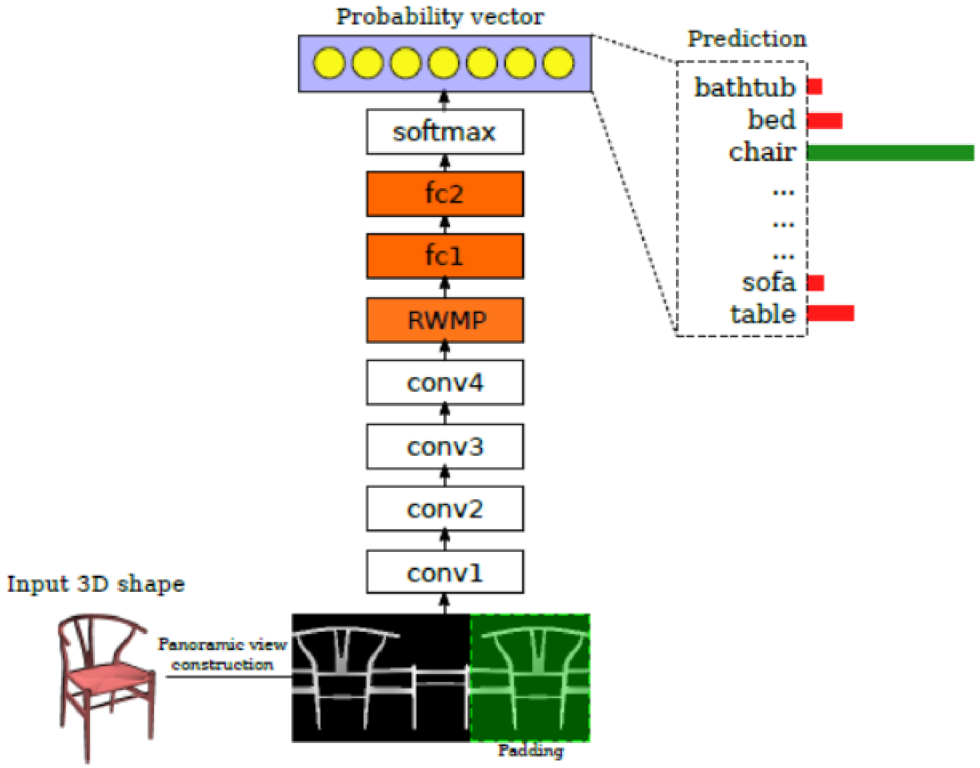
Multi-view representation as 3D input



# DEEP LEARNING ON POINT CLOUD

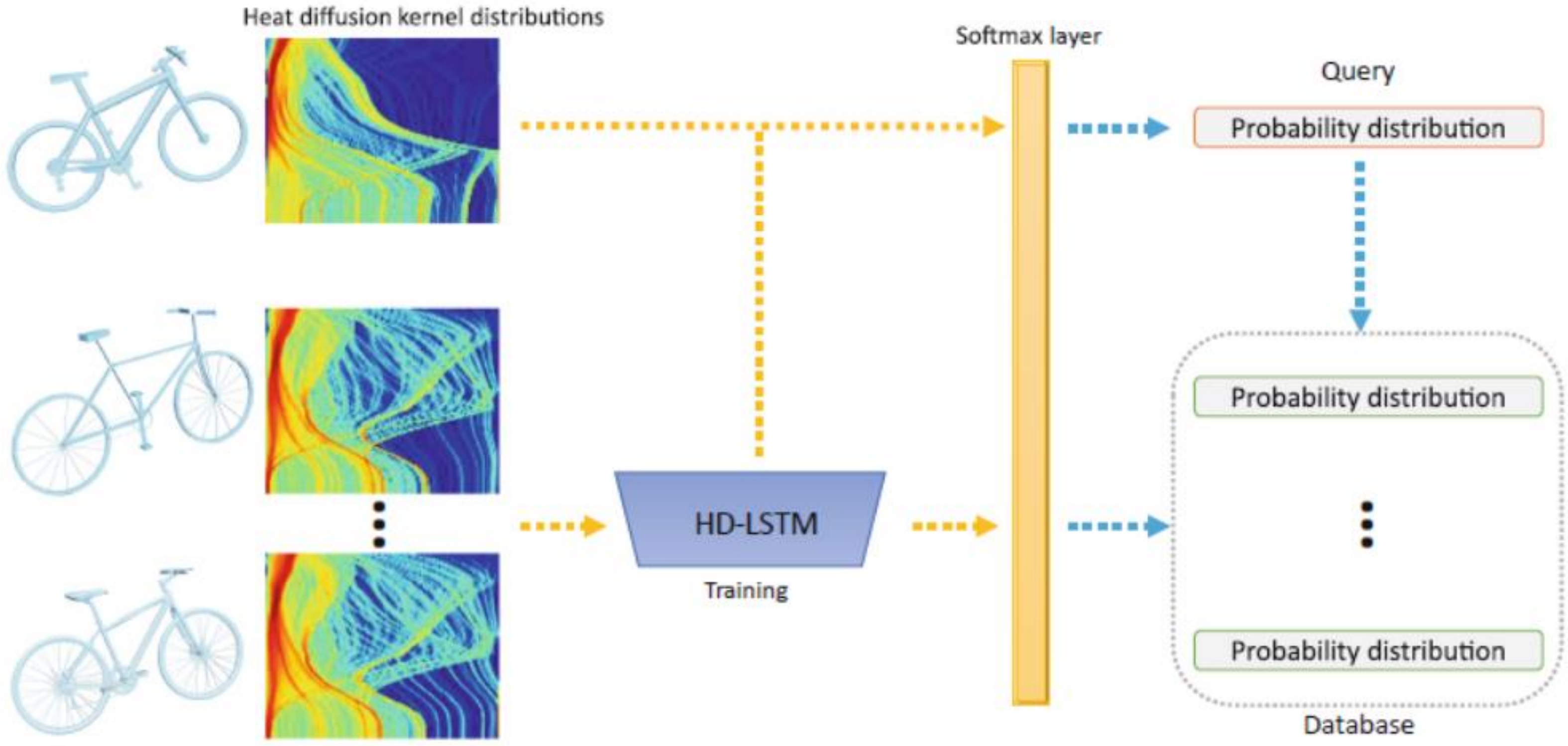
DeepPano



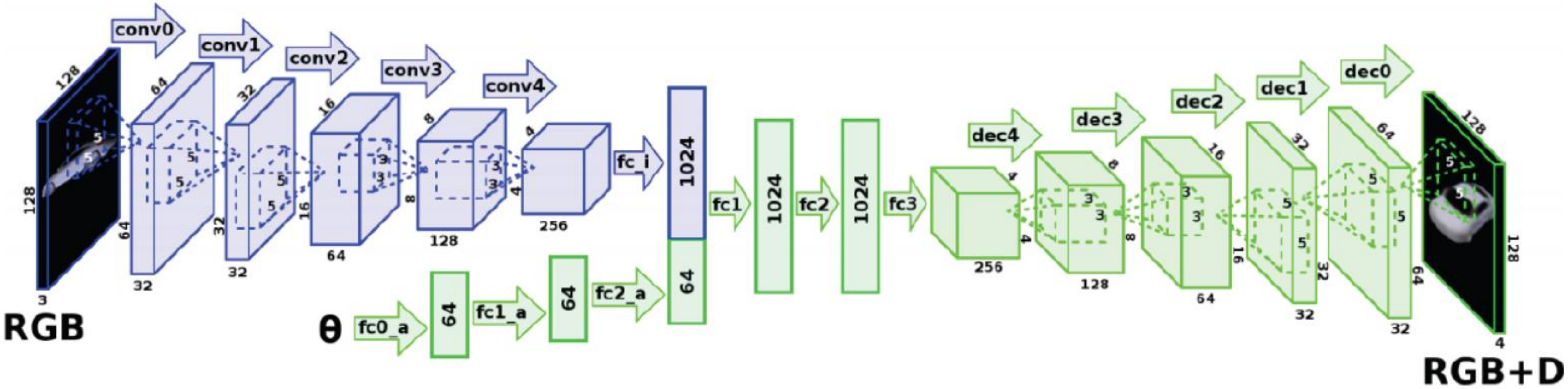
# DEEP LEARNING ON POINT CLOUD

Heat Diffusion LSTM



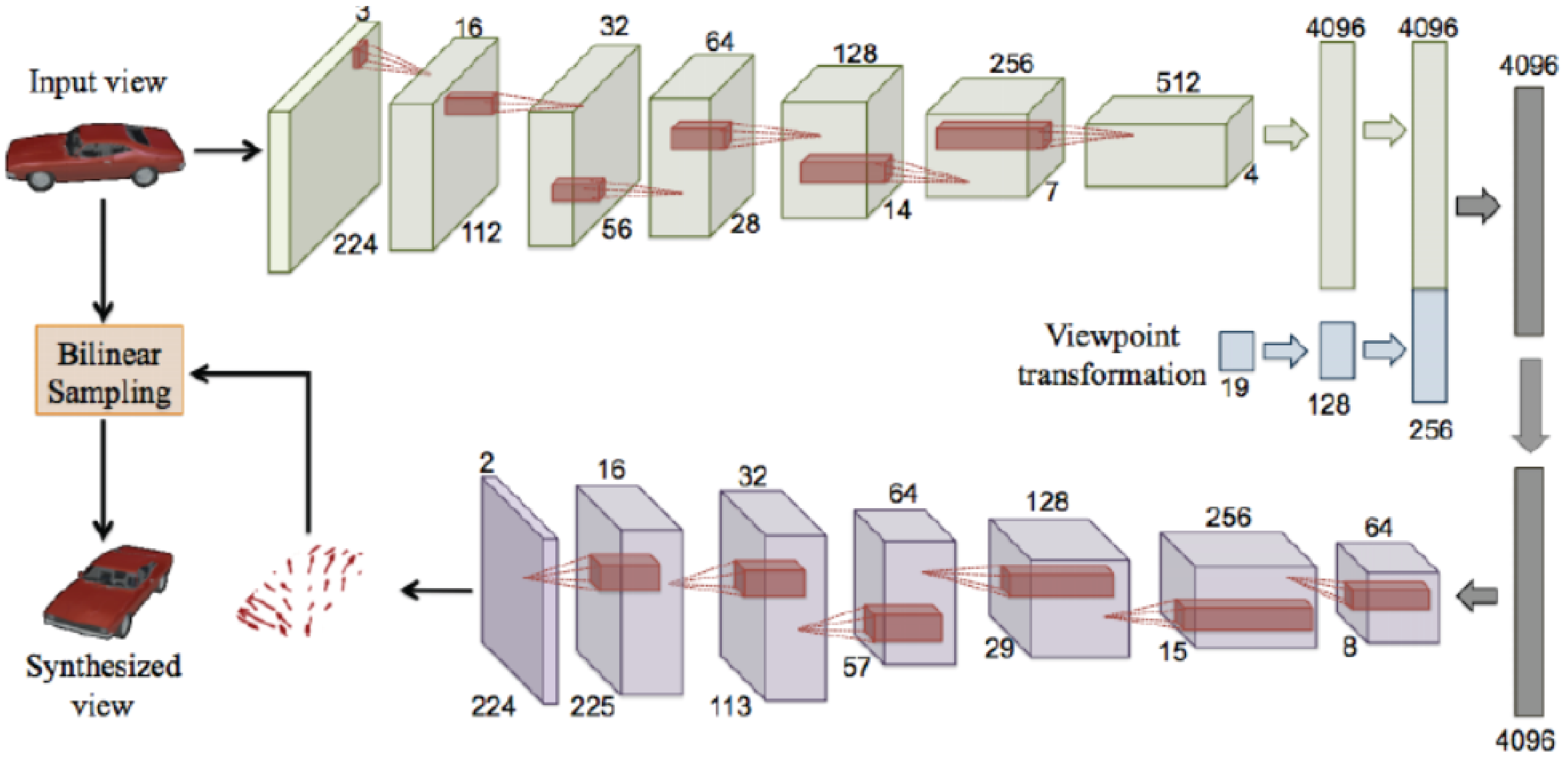
# DEEP LEARNING ON MULTIVIEW REPRESENTATION

Multi-view representation as 3D **output**



# DEEP LEARNING ON MULTIVIEW REPRESENTATION

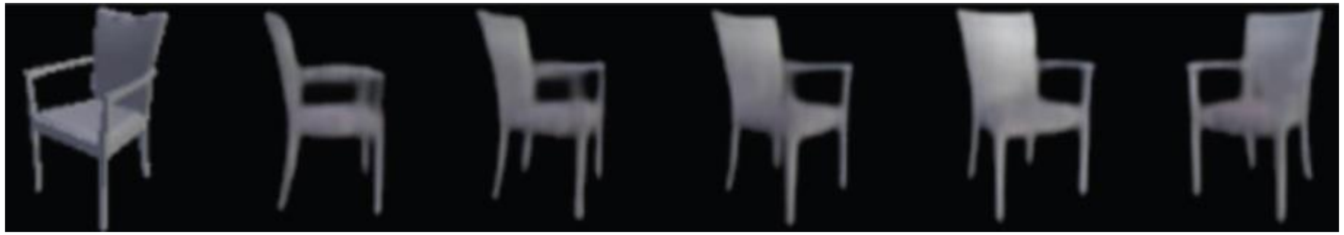
Novel-view RGB(D) image synthesis (flow prediction)

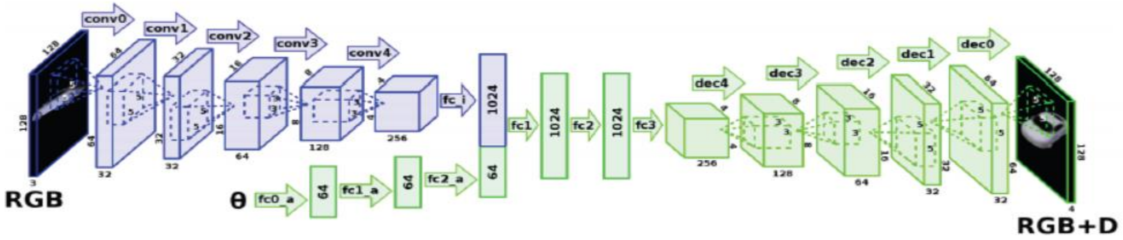


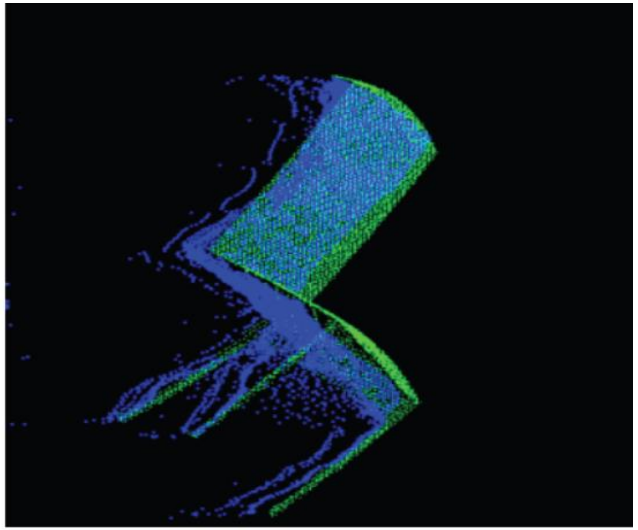
# DEEP LEARNING ON MULTIVIEW REPRESENTATION

Key Challenges

- Each view only contains partial information
- Not trivial to predict across viewpoints
- Cannot see through the surface
- Regular structures in 3D cannot be well captured
  - e.g., symmetry, straightness, roundish

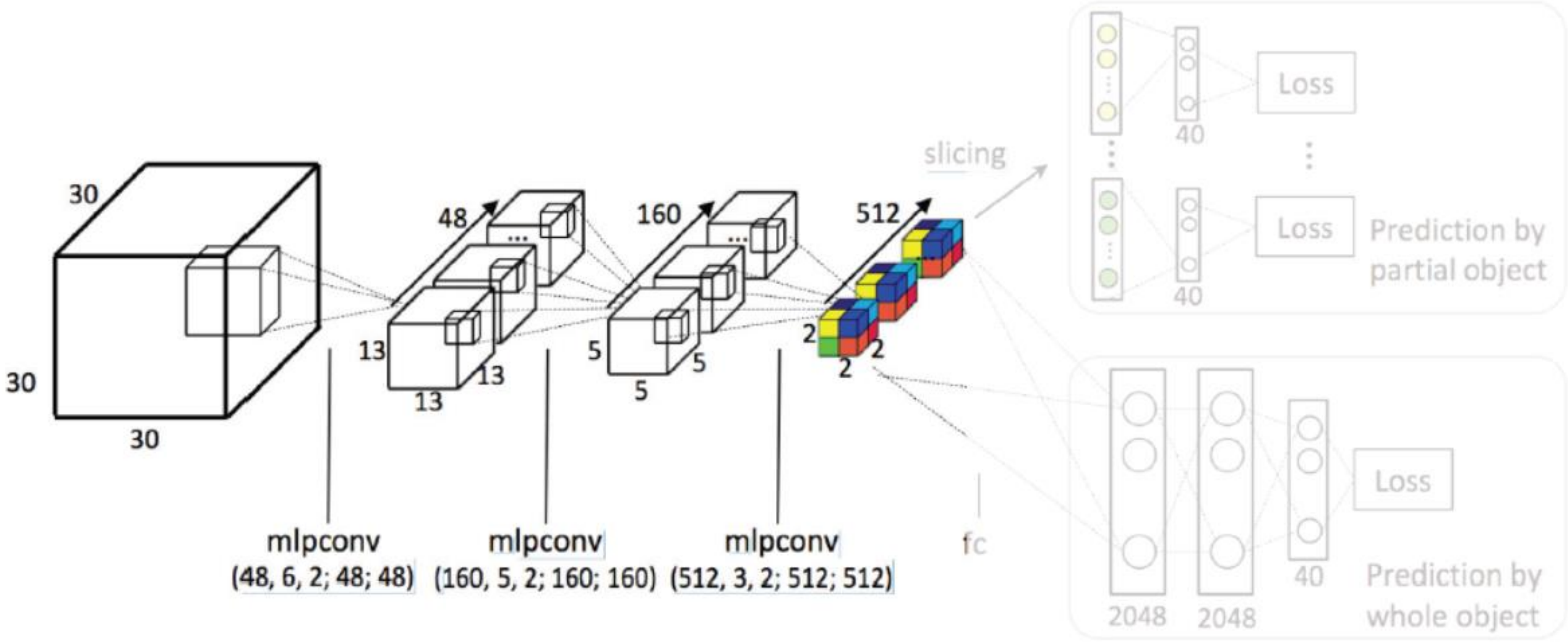




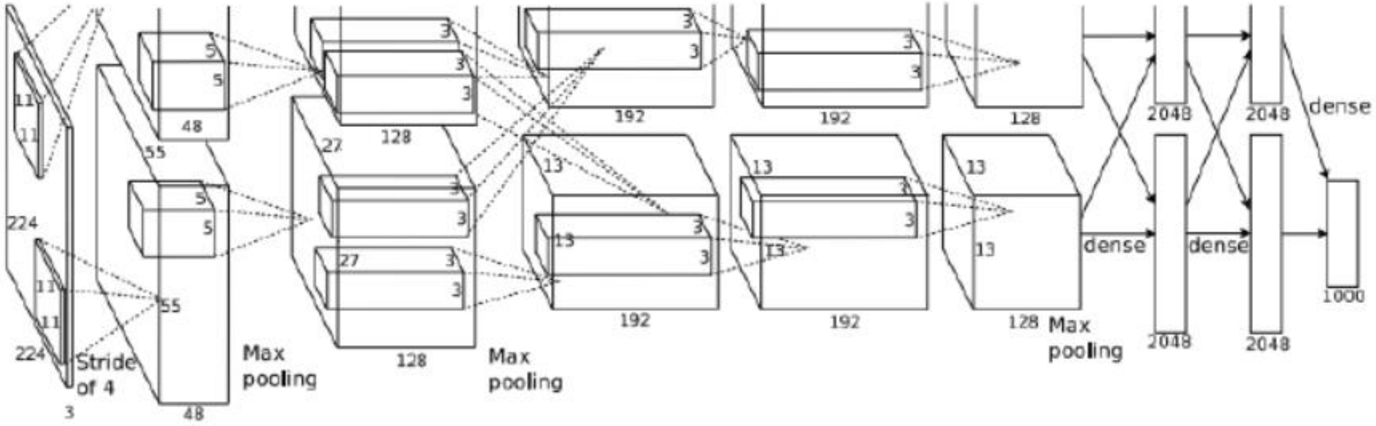
# DEEP LEARNING ON VOLUMETRIC EPRESENTATION

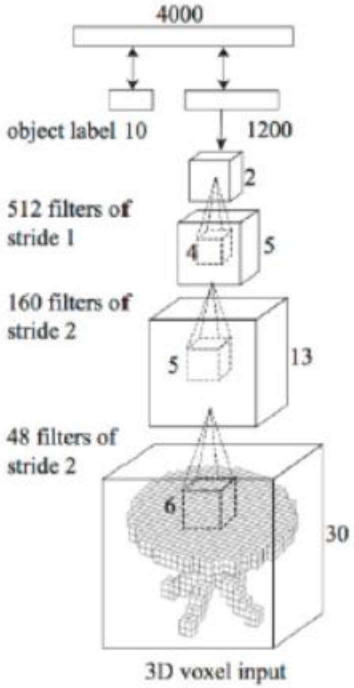
3D convolution uses 4D kernels




High space/time complexity $O(N^3)$

# DEEP LEARNING ON VOLUMETRIC EPRESENTATION




AlexNet, **2012**

3DShapeNets, **2015**

Input resolution: 224x224

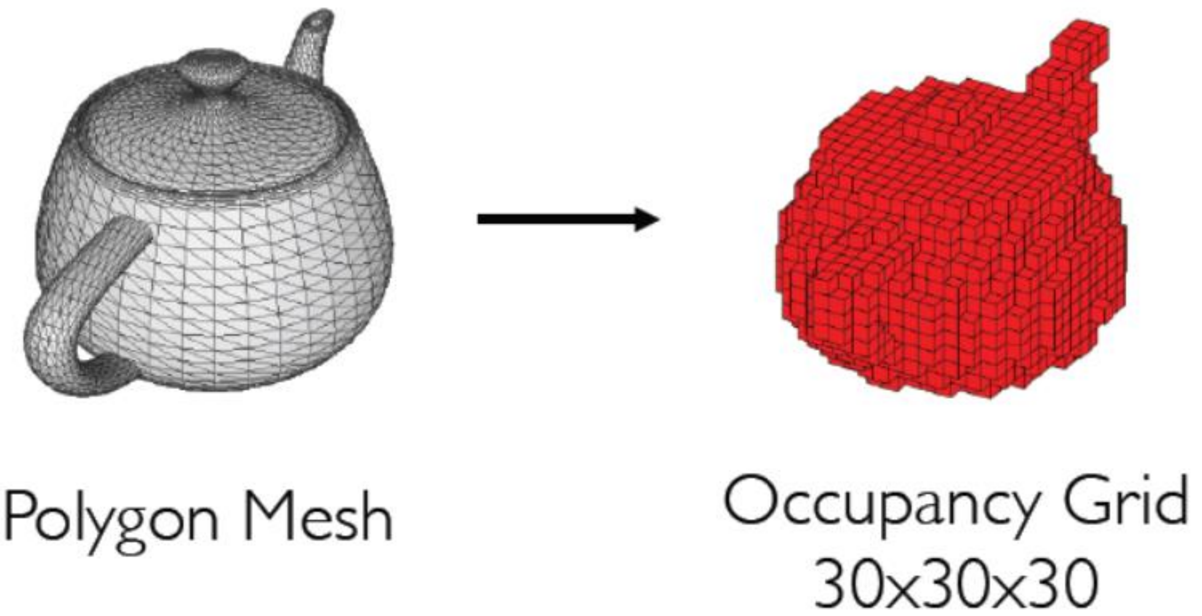
224x224=50176

Input resolution: 30x30x30

224x224=27000

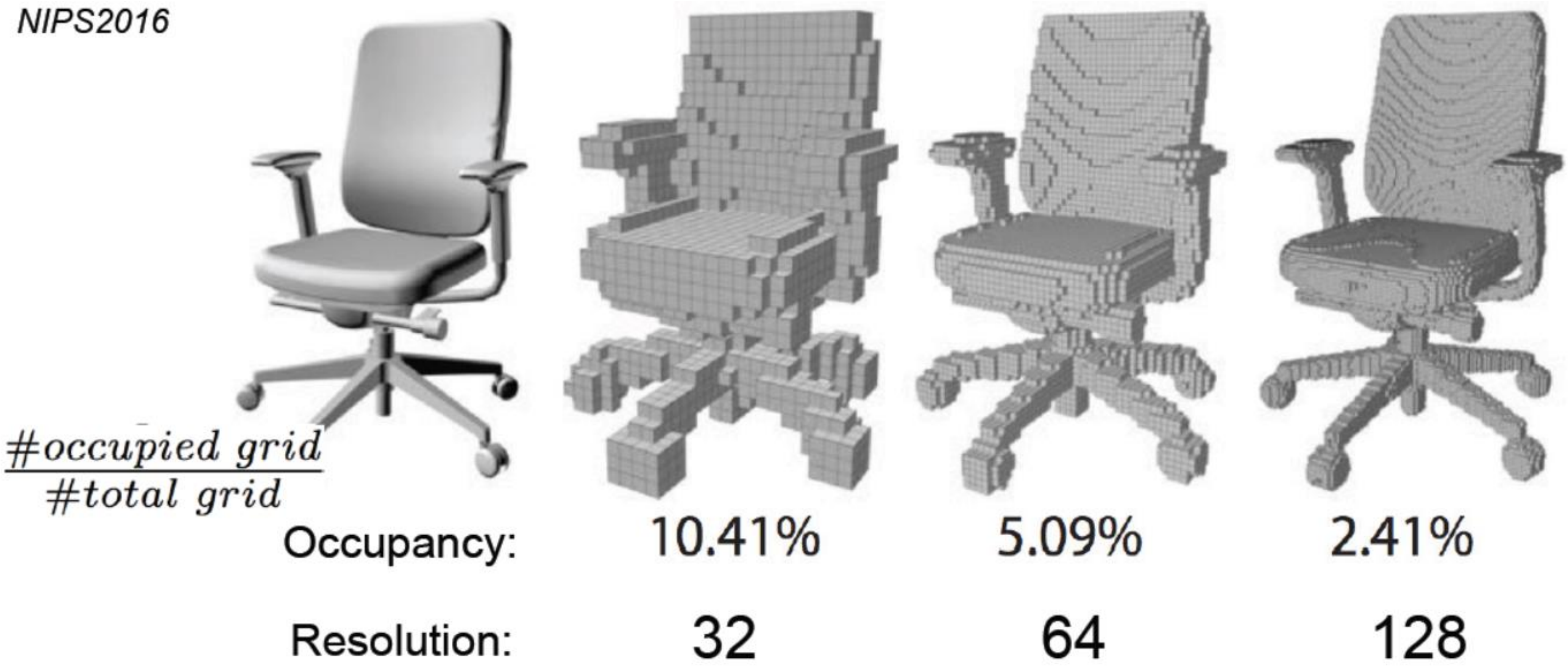
# DEEP LEARNING ON VOLUMETRIC EPRESENTATION

Information loss in voxelization



# DEEP LEARNING ON VOLUMETRIC EPRESENTATION

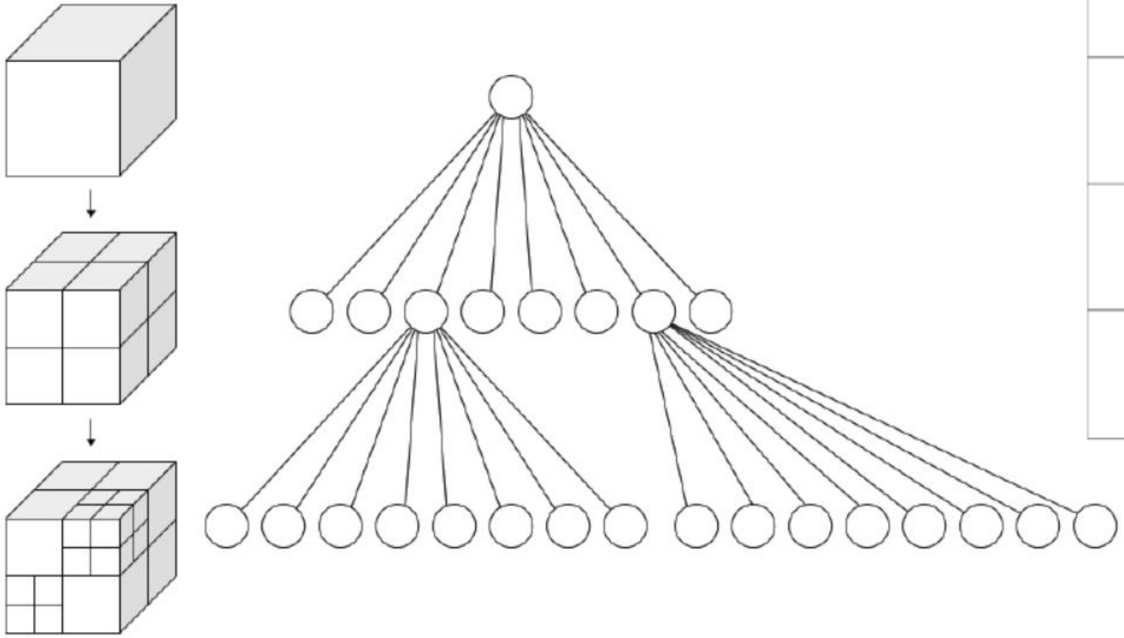
Sparsity characteristic of 3D data

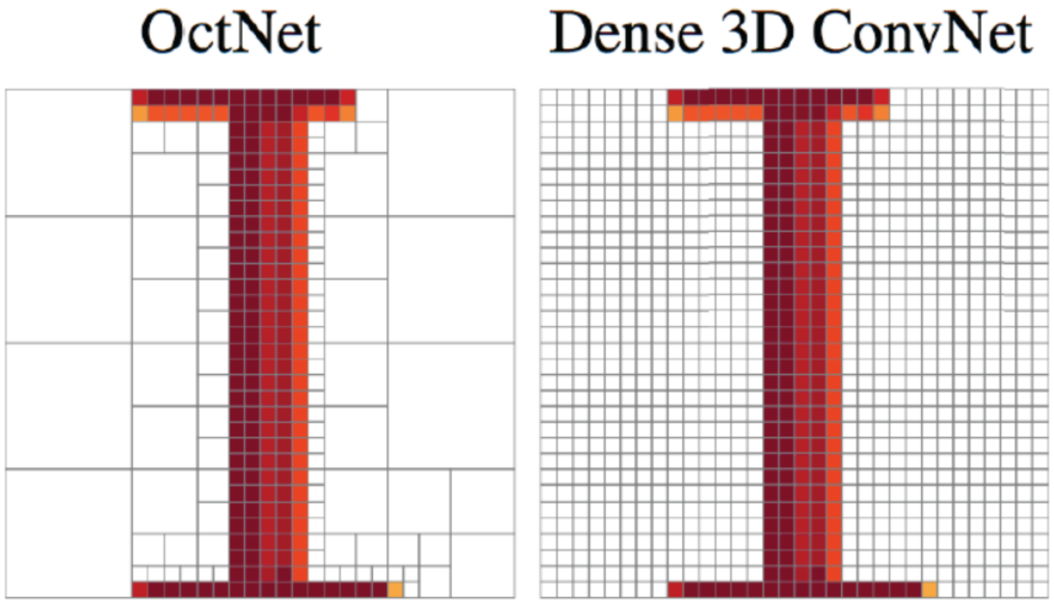


# DEEP LEARNING ON VOLUMETRIC EPRESENTATION

Octree: recursively partition the space

Each internal node has exactly eight children



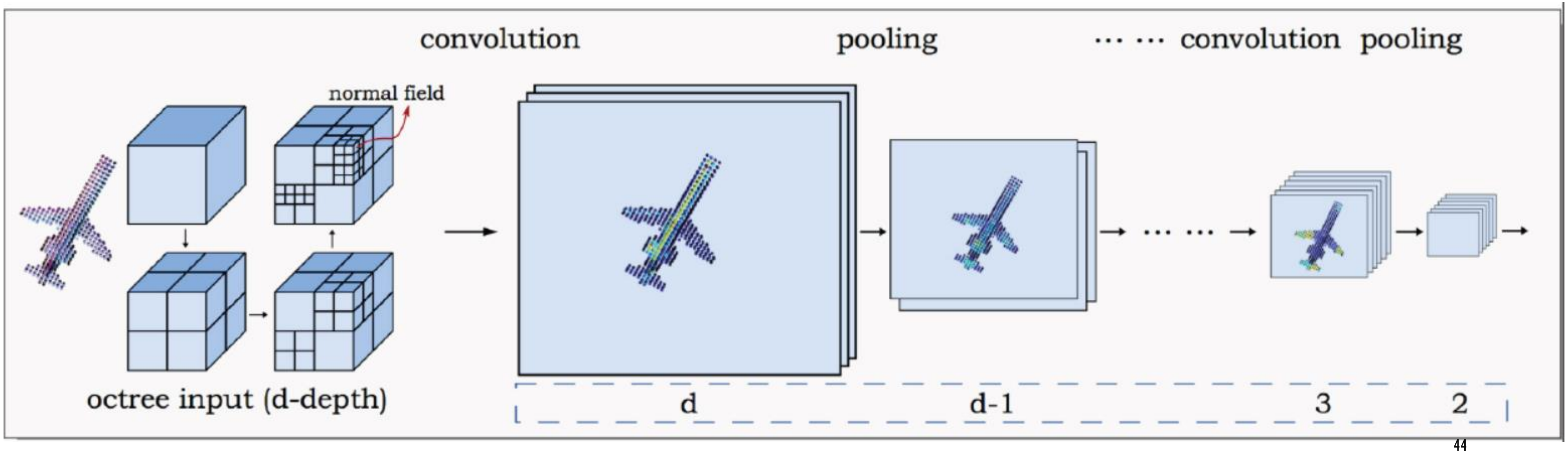


# DEEP LEARNING ON VOLUMETRIC EPRESENTATION

Define convolution and pooling along the octree

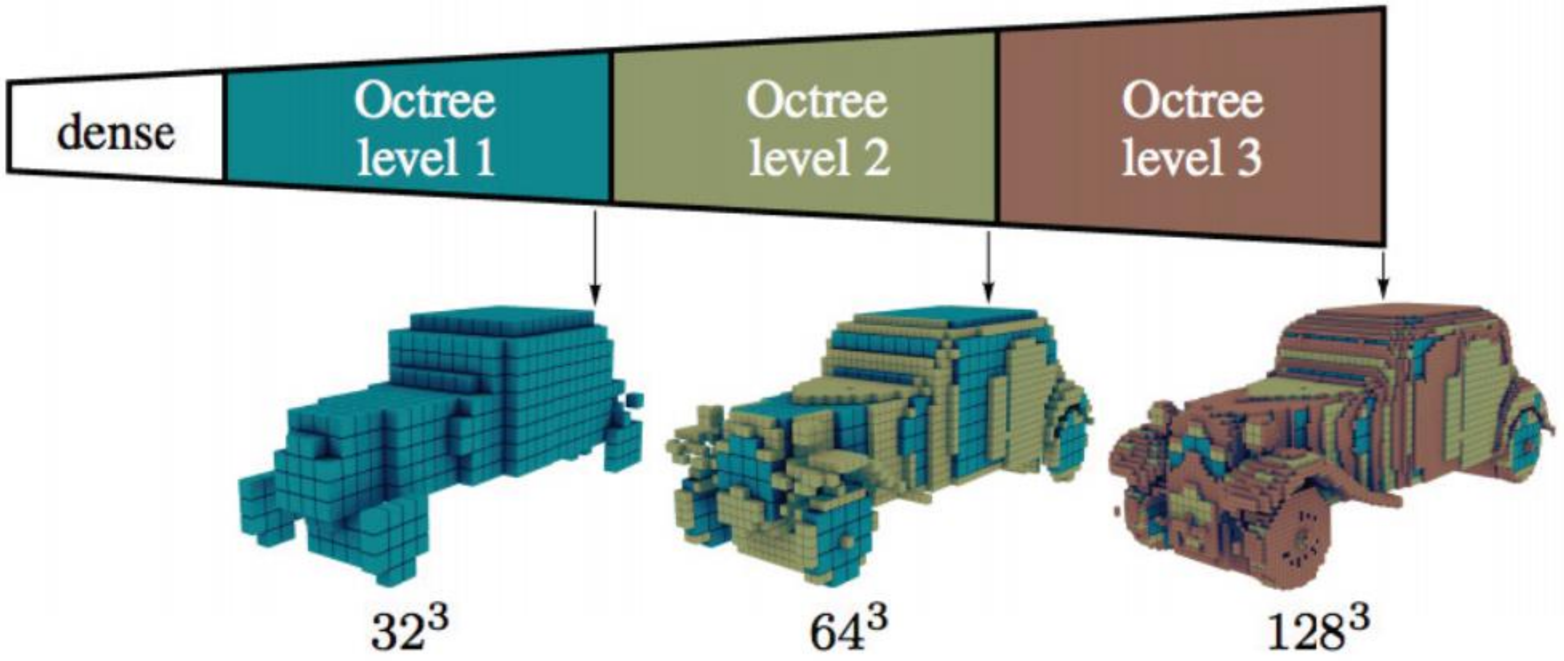
Challenge: how to implement efficiently — build a hash table to index the neighborhood

Restrict the convolution stride to be 2



# DEEP LEARNING ON VOLUMETRIC EPRESENTATION

Towards higher spatial resolution



# DEEP LEARNING ON MESH

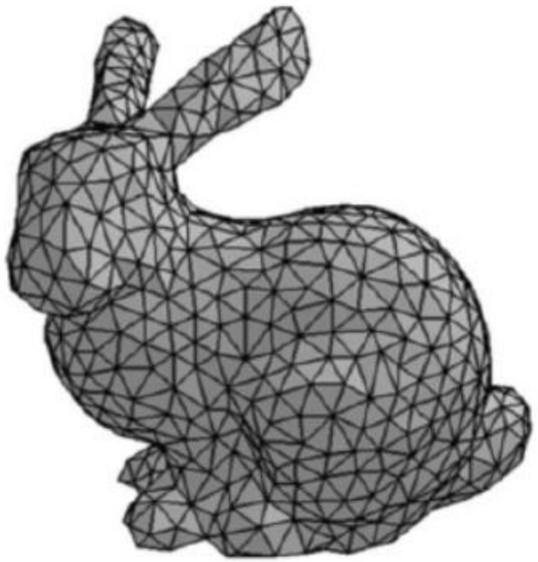
**Directly conduct convolution on graphs**

- Spatial construction (Geodesic CNN)
- Spectral construction (Spectral CNN)

**Conduct convolution on 2D parameterization of 3D surfaces**

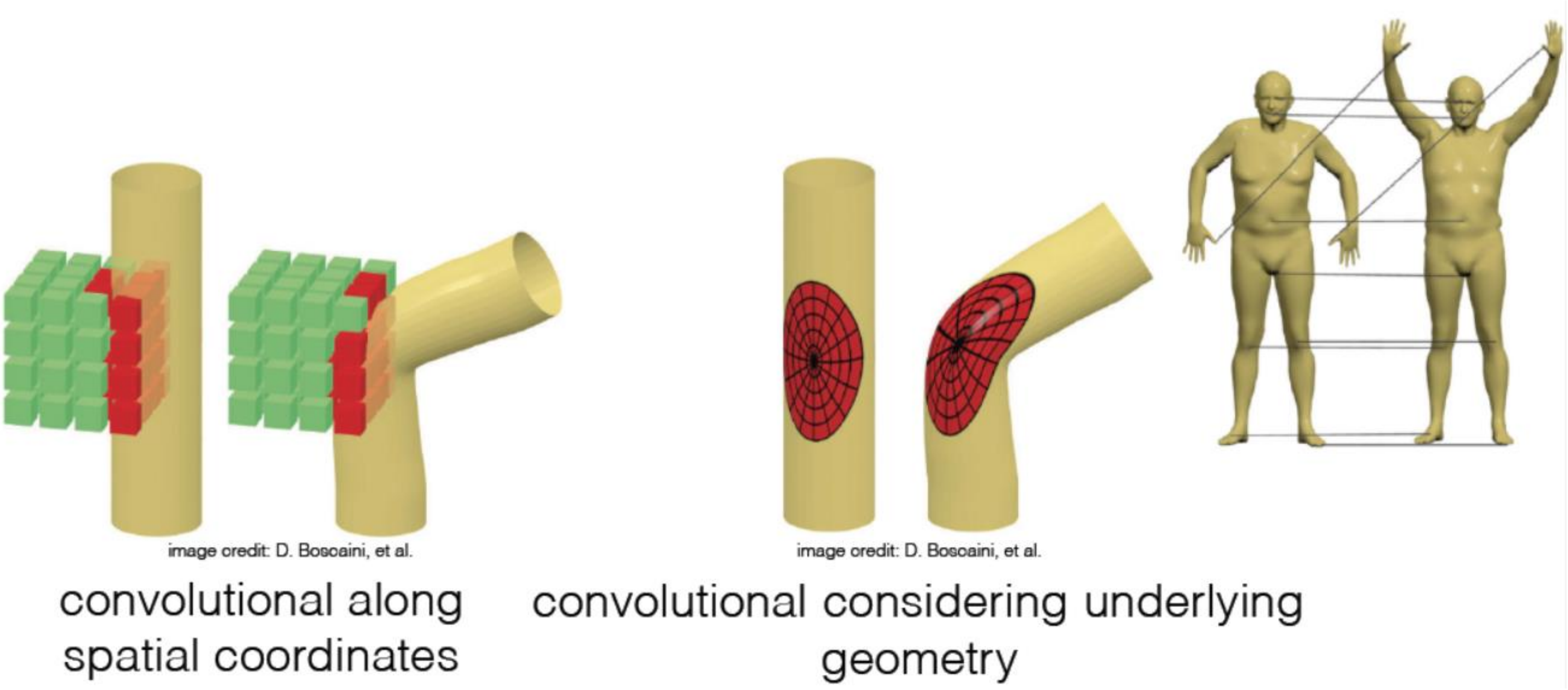
**Desired properties:**

- Locally supported (w.r.t graph metric)
- Allowing weight sharing across different coordinates



3D shape graph

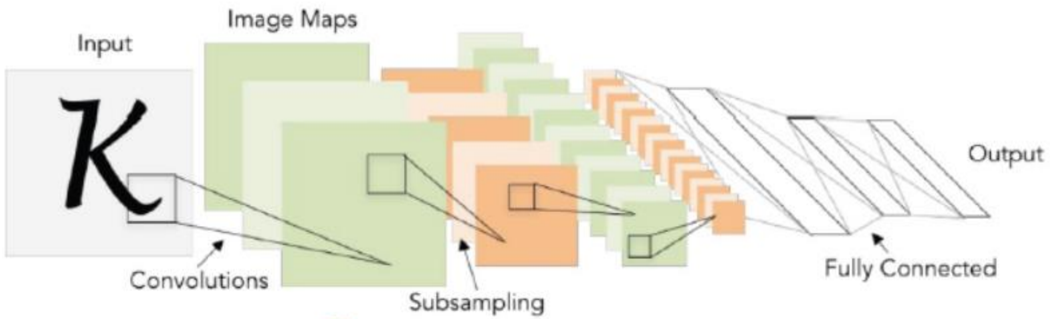
# DEEP LEARNING ON MESH



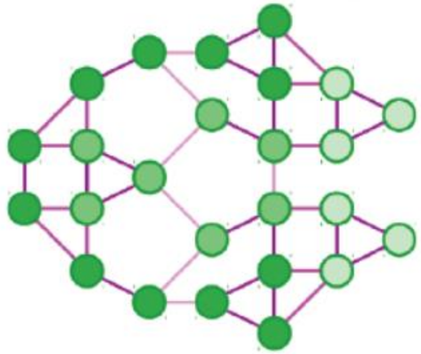
# DEEP LEARNING ON MESH

How to allow multi-scale analysis?

grid structure




graph structure



hierarchical graph coarsening?

from Michaël Defferrard et al. 2016

# DEEP LEARNING ON MESH

## Spatial construction: Geodesic CNN

## Constructing convolution kernels:

- Local system of geodesic polar coordinate
- Extract a small patch at each point x

## Issues

- The local charting method relies on a fast marching-like procedurerequiring a triangular mesh.
- The radius of the geodesic patches must be sufficiently small to acquire a topological disk.
- No effective pooling, purely relying on convolutions to increase receptive field.

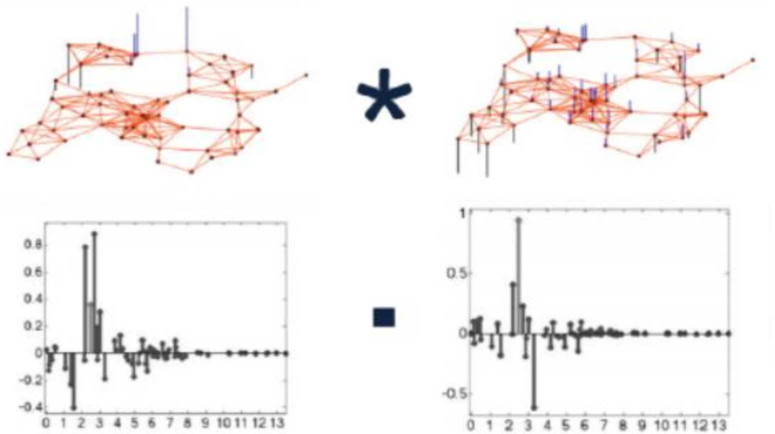
# DEEP LEARNING ON MESH

## Fourier analysis

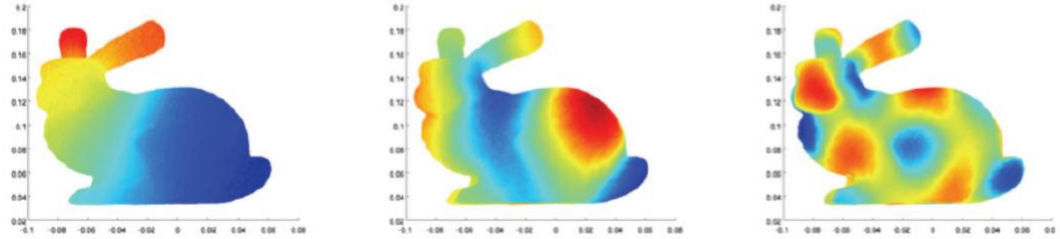
- Convert convolution to multiplication in spectral domain

Generalized convolution of $f, g \in L^2(X)$ can be defined by analogy

$$(f \star g)(x) = \underbrace{\sum_{k \geq 1} \underbrace{\langle f, \phi_k \rangle_{L^2(X)} \langle g, \phi_k \rangle_{L^2(X)}}_{\text{product in the Fourier domain}} \phi_k(x)}_{\text{inverse Fourier transform}}$$

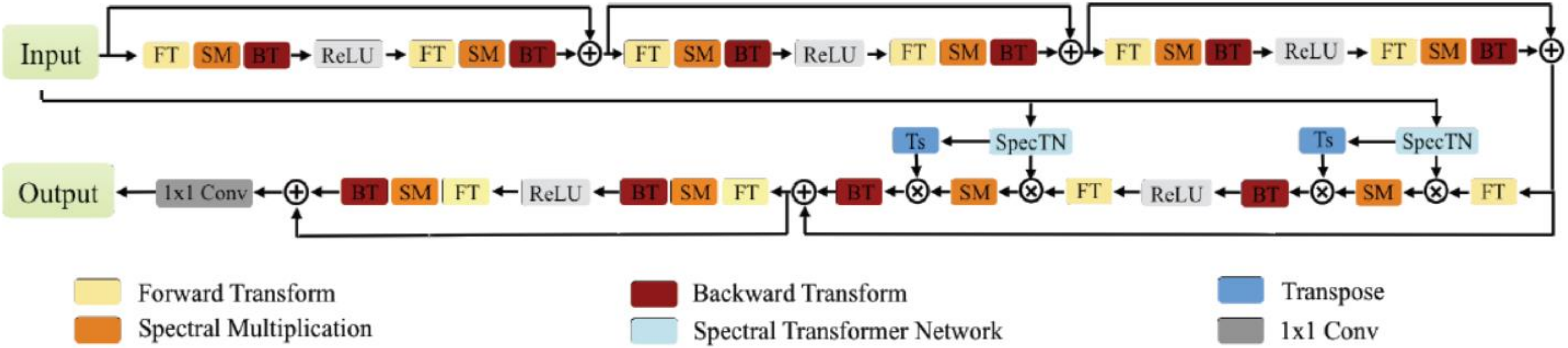


"Fourier basis" of the graph: $V$ : Eigenvectors of $\Delta$

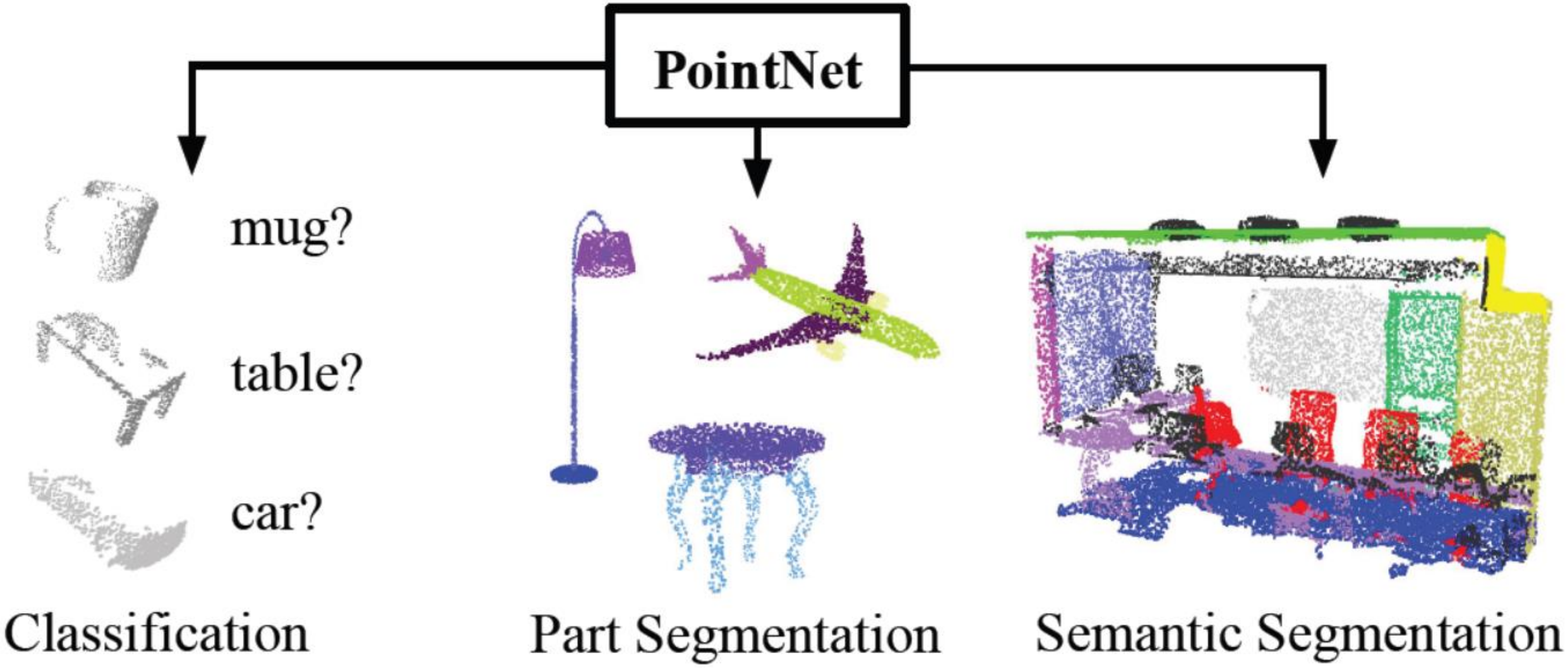


# DEEP LEARNING ON MESH

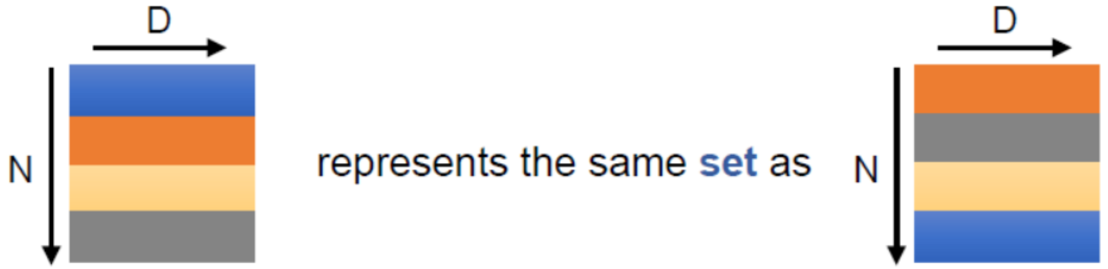
Fourier analysis

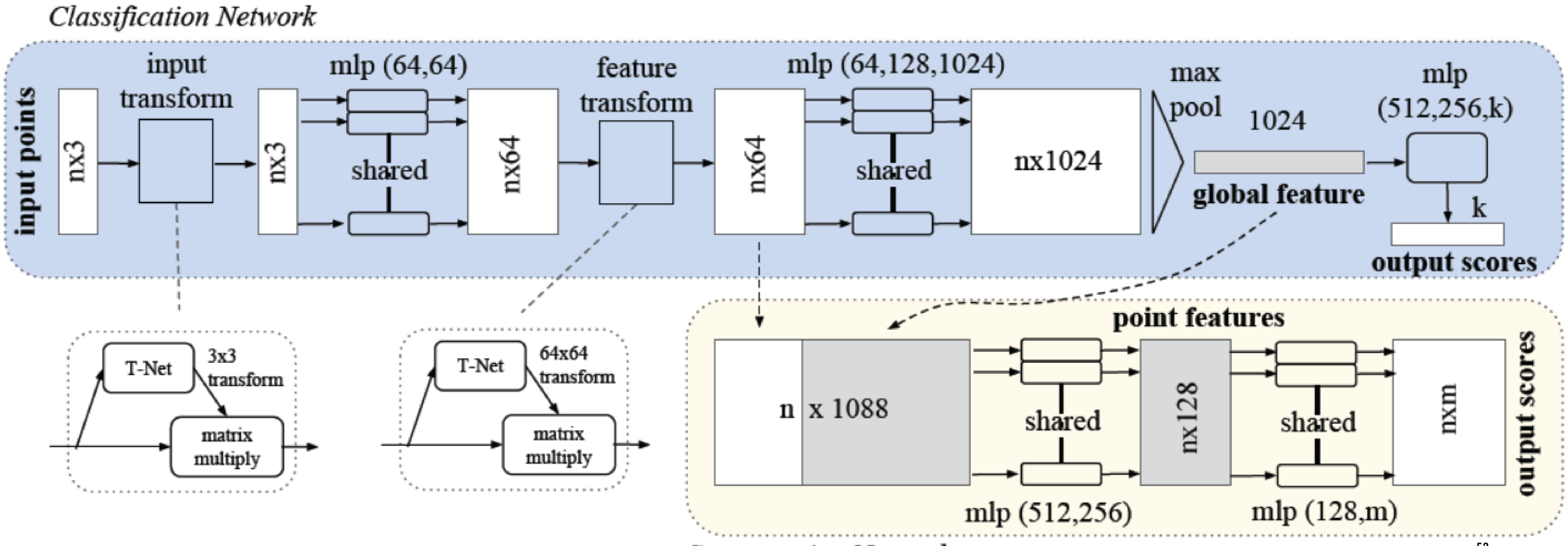


# DEEP LEARNING ON POINT CLOUD



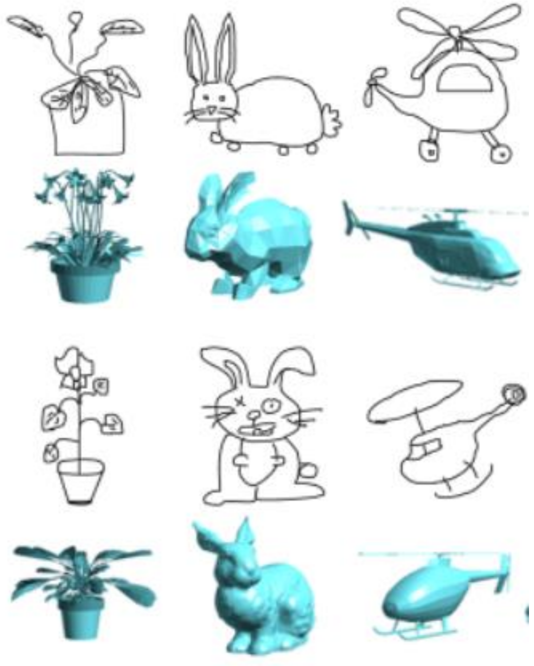
# DEEP LEARNING ON POINT CLOUD



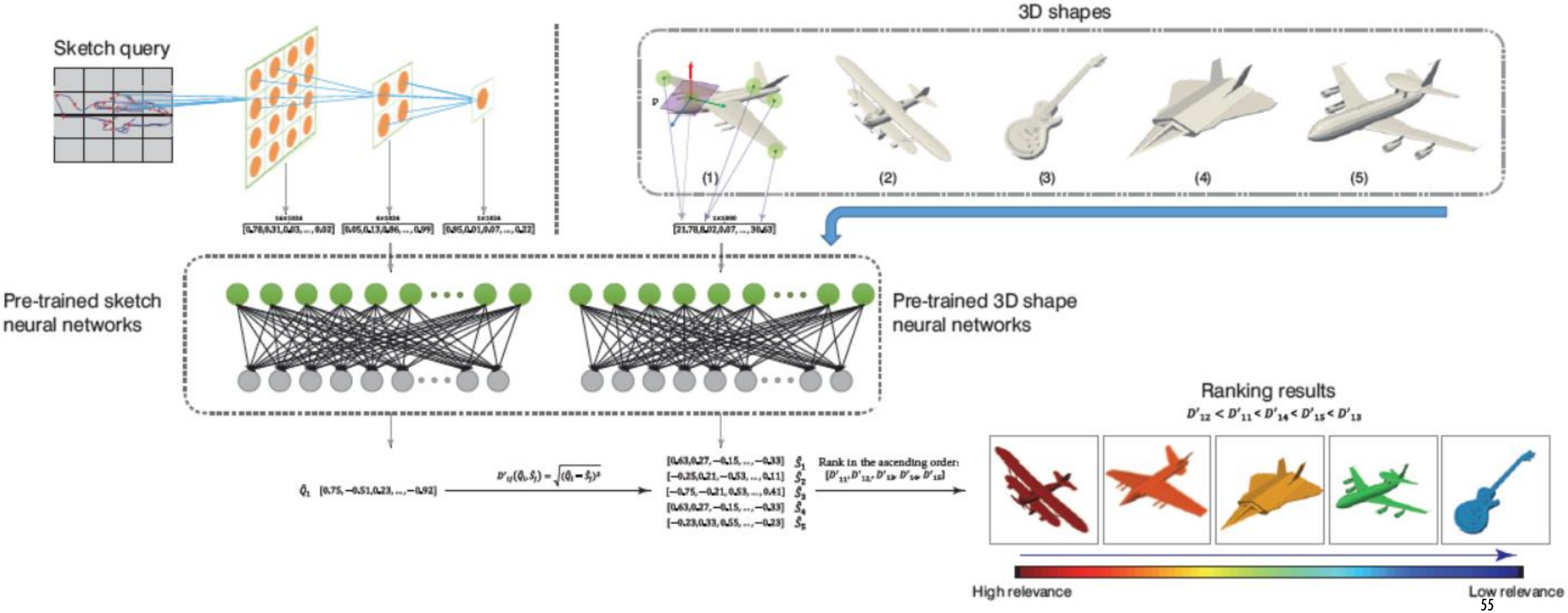
# DEEP LEARNING ON SKETCH

3D objects to 2D sketch



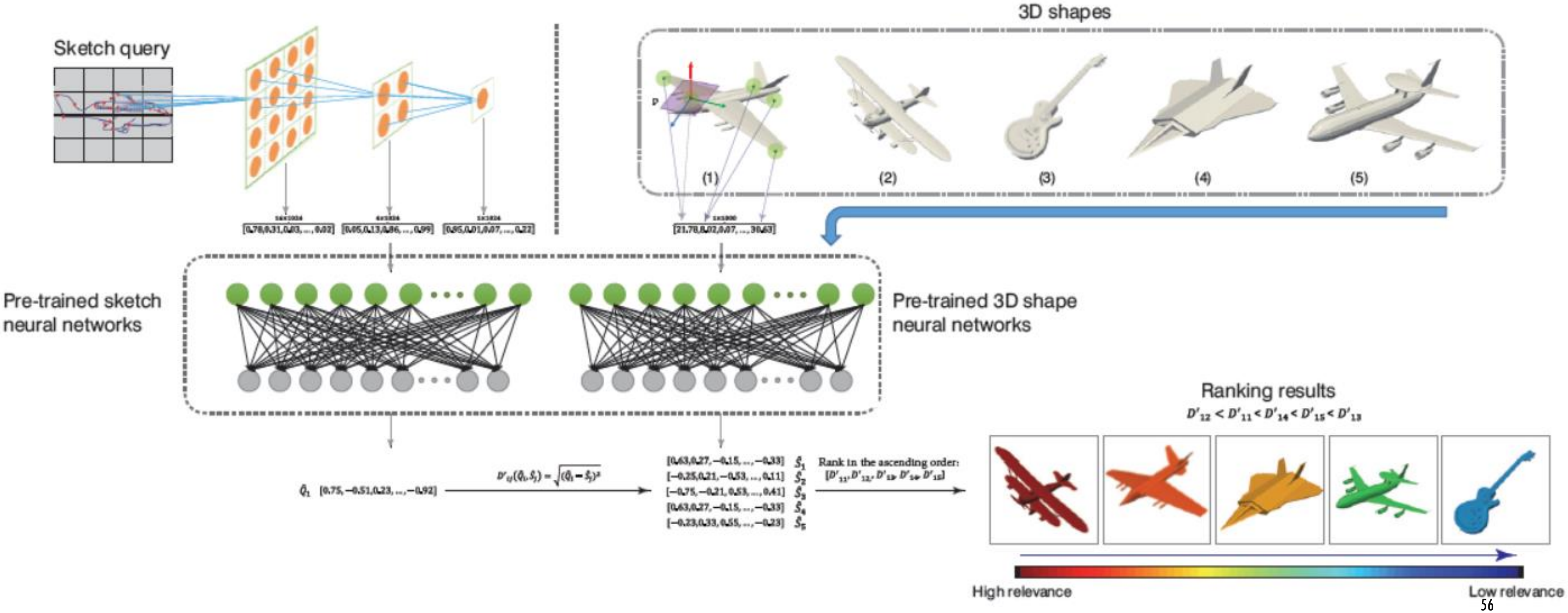
# DEEP LEARNING ON SKETCH

Cross-scenario



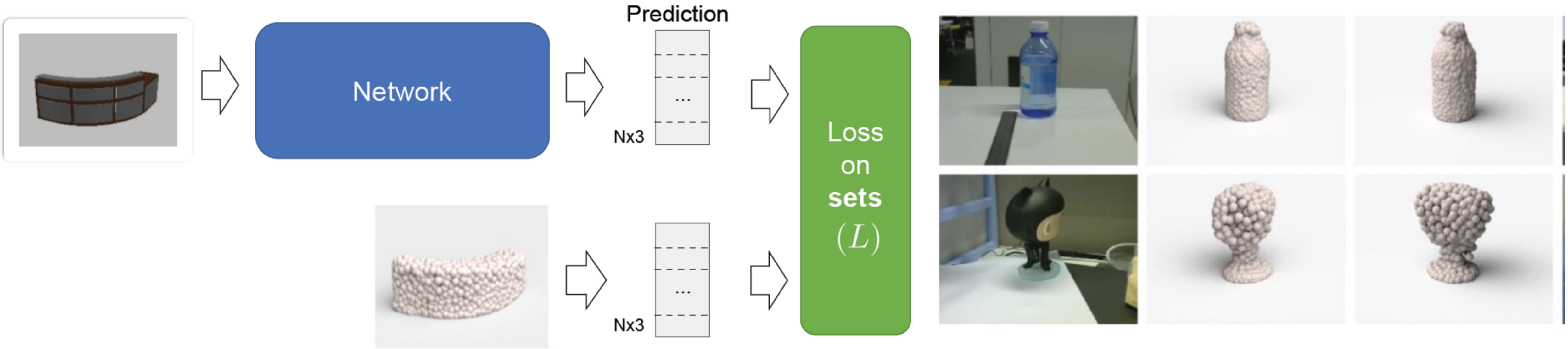
# DEEP LEARNING ON SKETCH

Cross-scenario



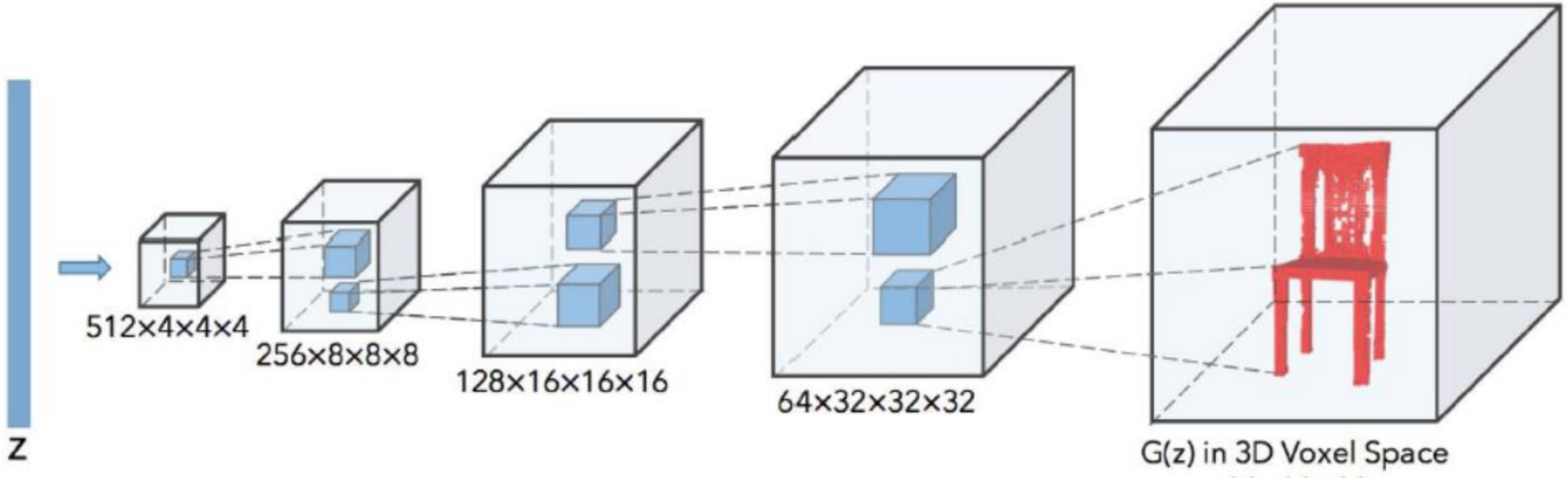
# DEEP LEARNING ON GENERATION

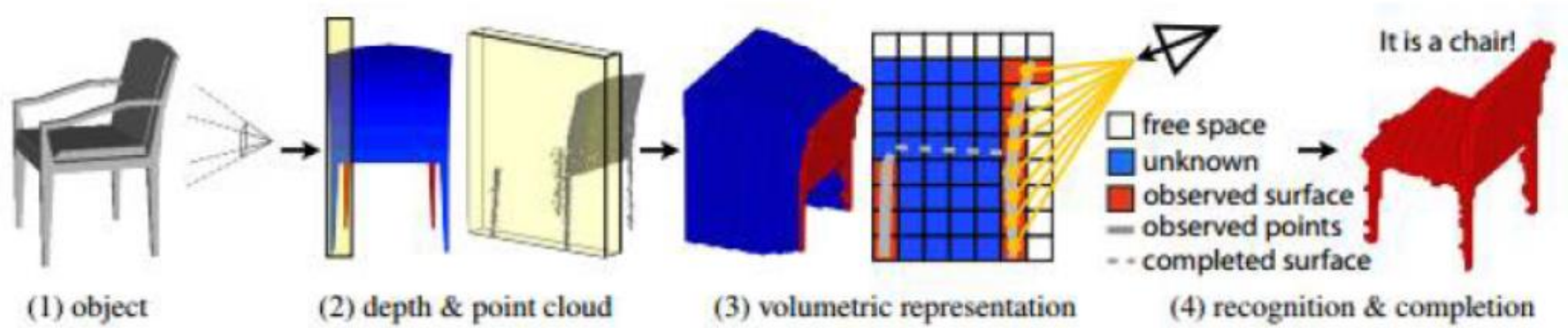
Supervised generation



# DEEP LEARNING ON GENERATION

3D GAN



# CONTENTS/OUTLINE

3D Shape Representation&Applications

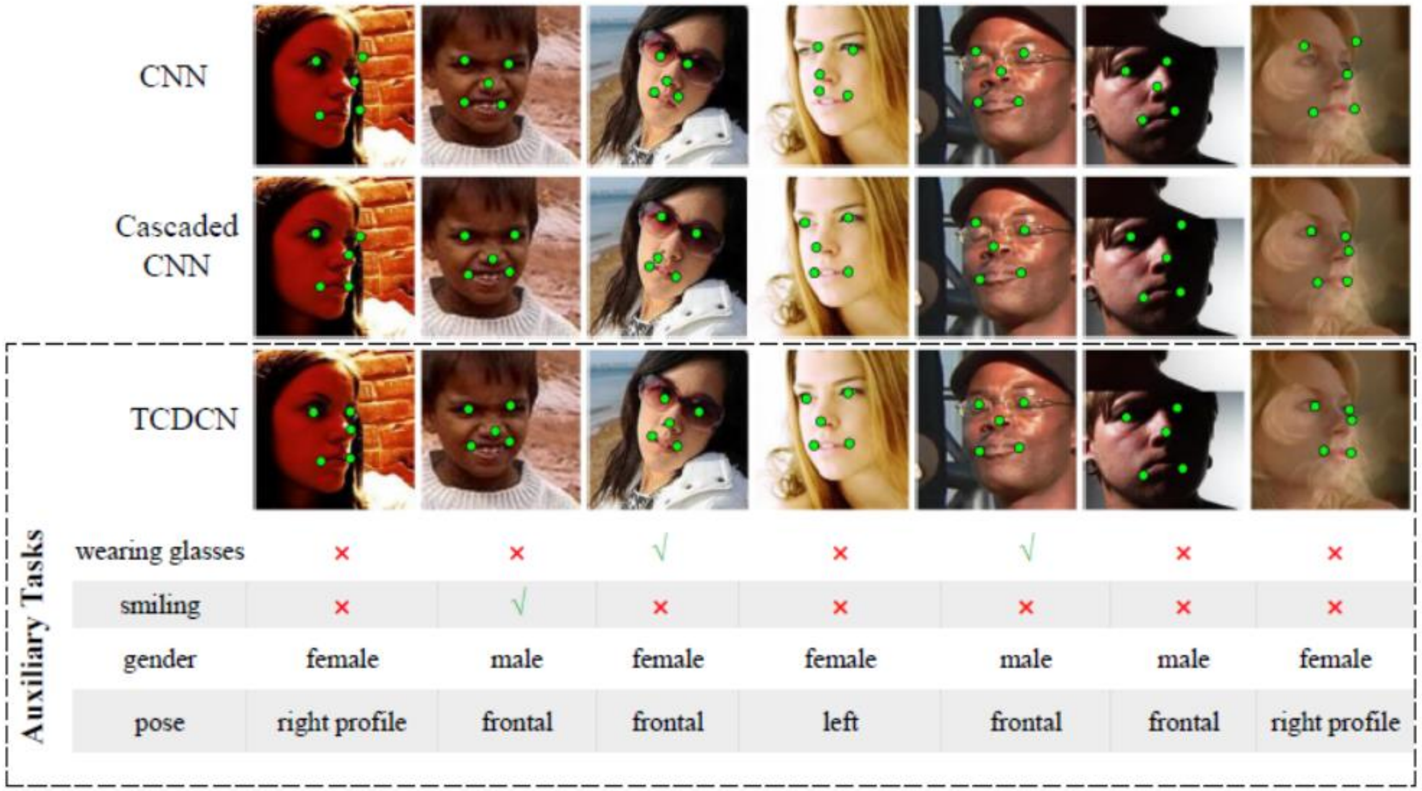
- 1 Introduction
- 2 Traditional Shape Representation
- 3 3D Deep Learning

Others (Brief introduction)

- 4 Multi-Task Learning
- 5 Image Privacy Protection

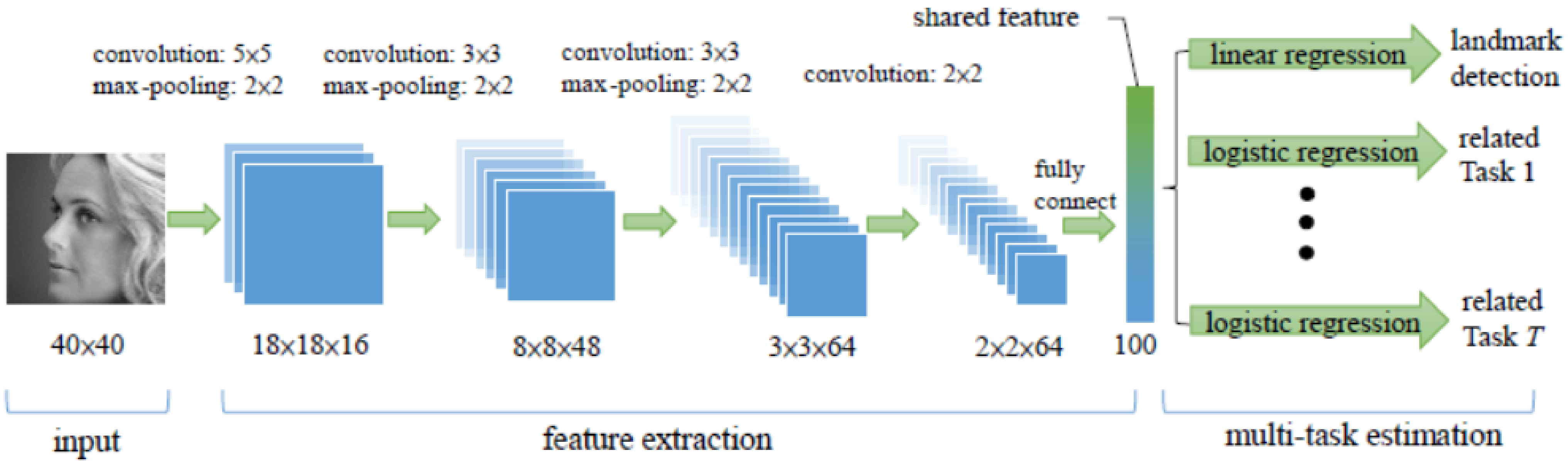
# 4 MULTI-TASK LEARNING FOR IMAGE CLASSIFICATION

- One main task together with several additional tasks

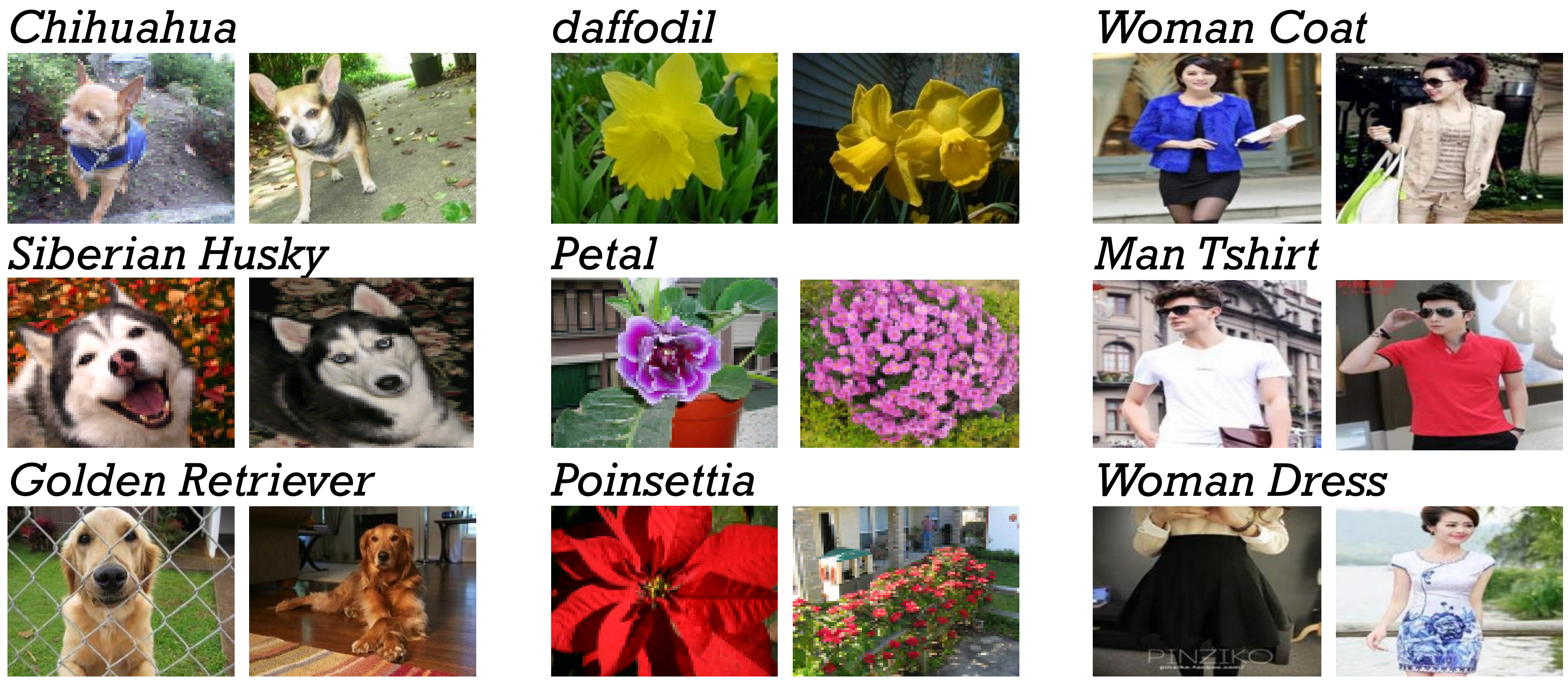


# 4 DEEP MULTI-TASK LEARNING FOR IMAGE CLASSIFICATION

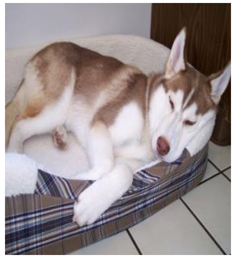
One main task together with several additional tasks



# 4 DEEP MULTI-TASK LEARNING FOR IMAGE CLASSIFICATION




Test



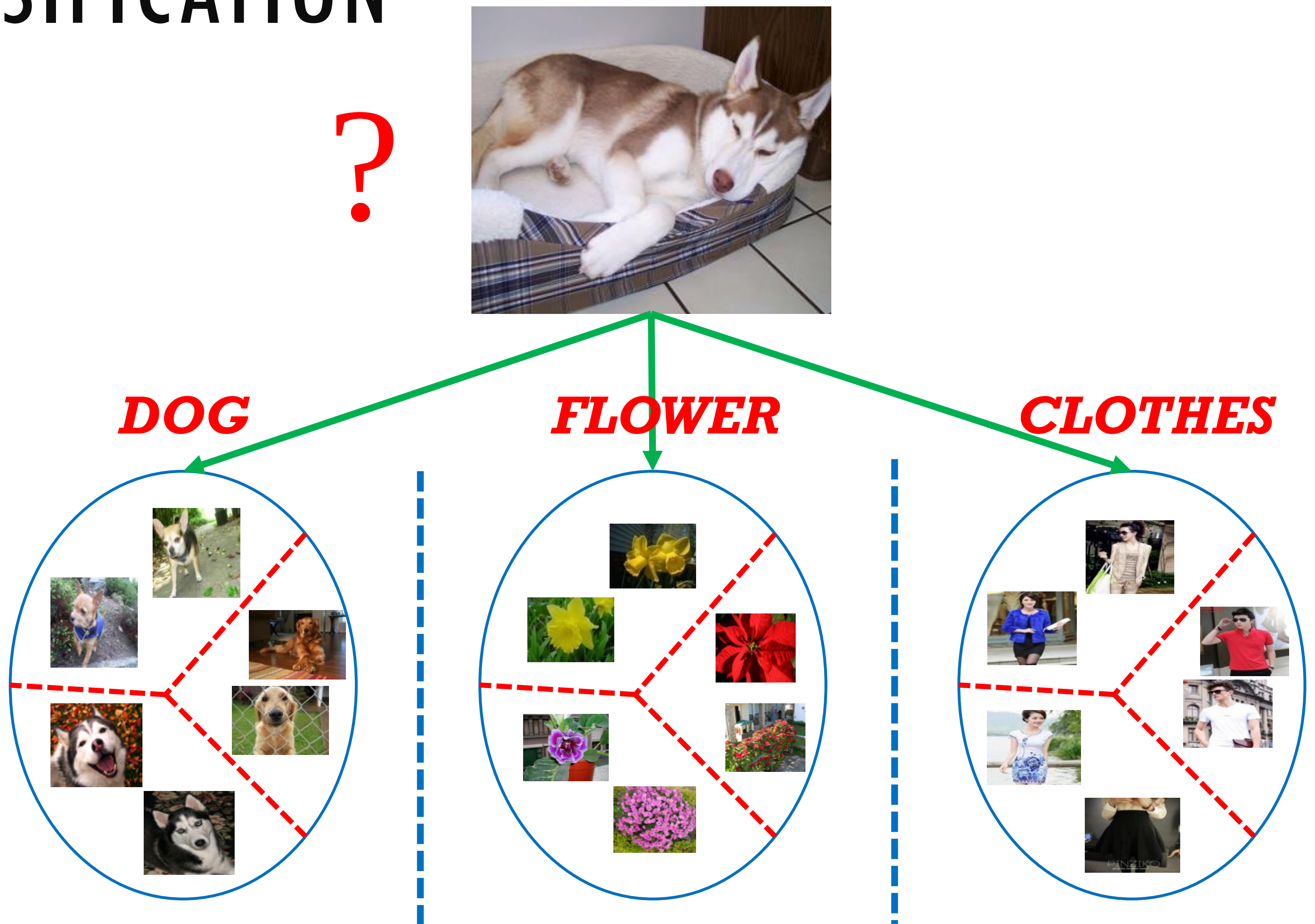
?

**Siberian Husky**

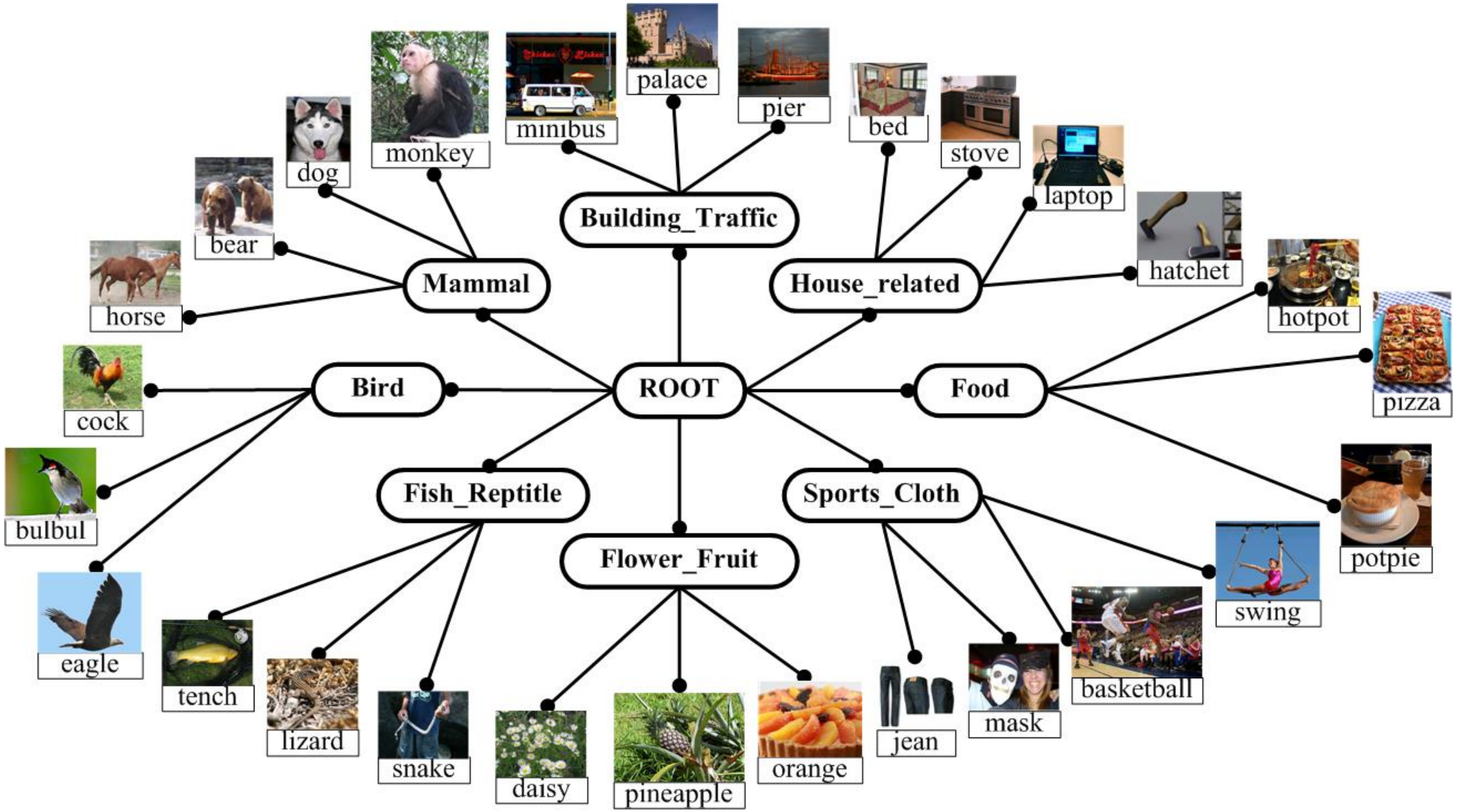
?

**Woman Dress**

# 4 DEEP MULTI-TASK LEARNING FOR IMAGE CLASSIFICATION

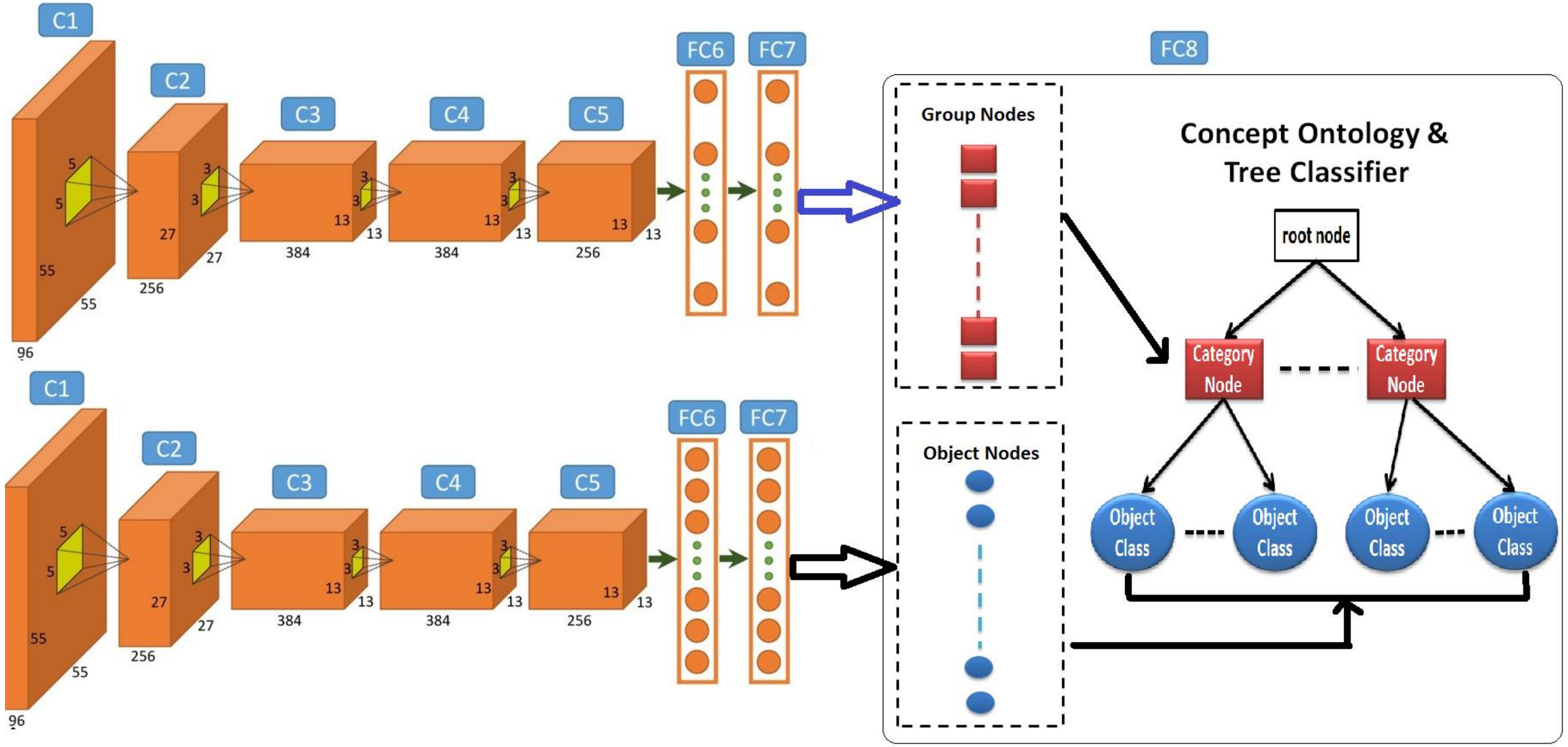


# 4 MULTI-TASK LEARNING FOR IMAGE CLASSIFICATION



# 4 DEEP MULTI-TASK LEARNING FOR IMAGE CLASSIFICATION

- Hierarchical deep multi-task learning



# 4 MULTI-TASK LEARNING FOR IMAGE CLASSIFICATION

Classification Performance

✓Improvement at least 1%

Feature Representation

✓Group specified feature

✓Significant lower storage costs

# 5 IMAGE PRIVACY PROTECTION

We take **lots of photos** everyday! **Share** on social networks!

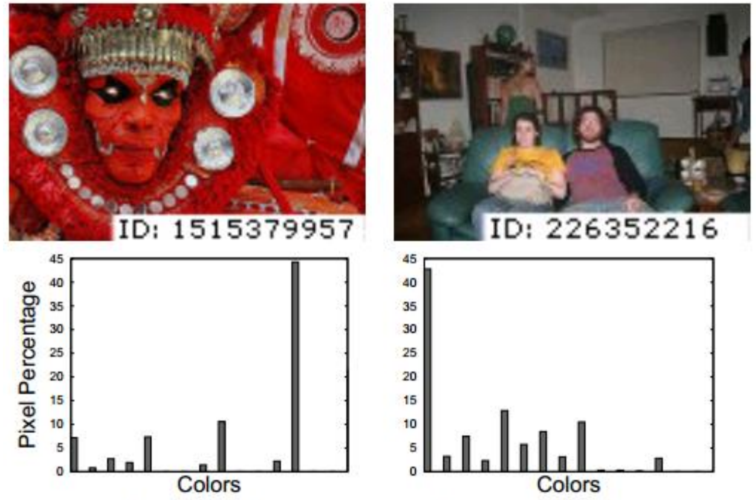
Have you ever considered about the image **safety problem**?

Have you ever encountered any **troubles**?

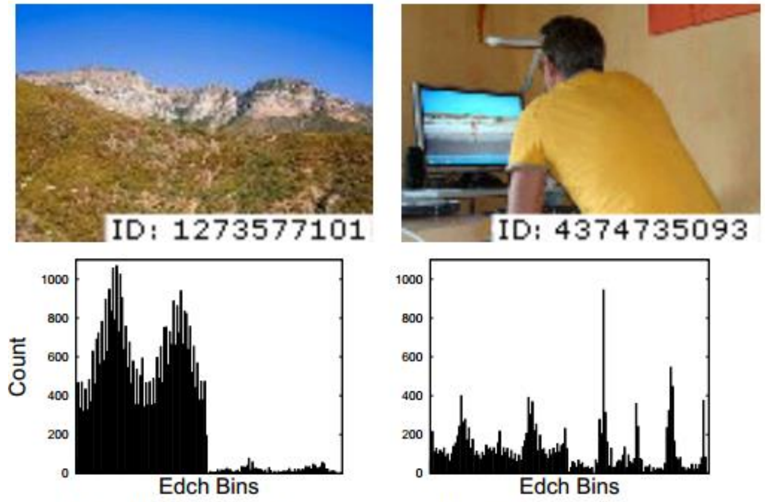
**Image privacy problem**

# 5 IMAGE PRIVACY PROTECTION

- Tag-based approach
  - **Social tags based approach** [Danezis, 2009]
    - User provided tags noisy
    - e.g. missing&spam
- Visual-based approach
  - Visual similar images share similar privacy settings
  - **Handcraft visual features**
    - Color & Edge & SIFT & GIST
    - Non-efficient: e.g. design bias



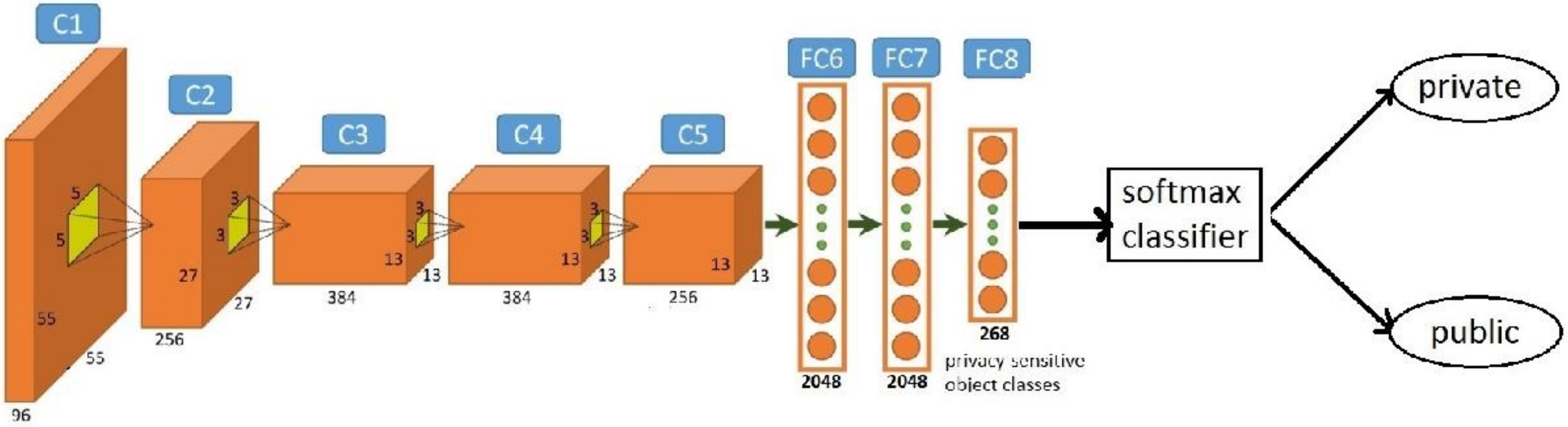
(a) Public photo (b) Private photo



(a) Public photo (b) Private photo

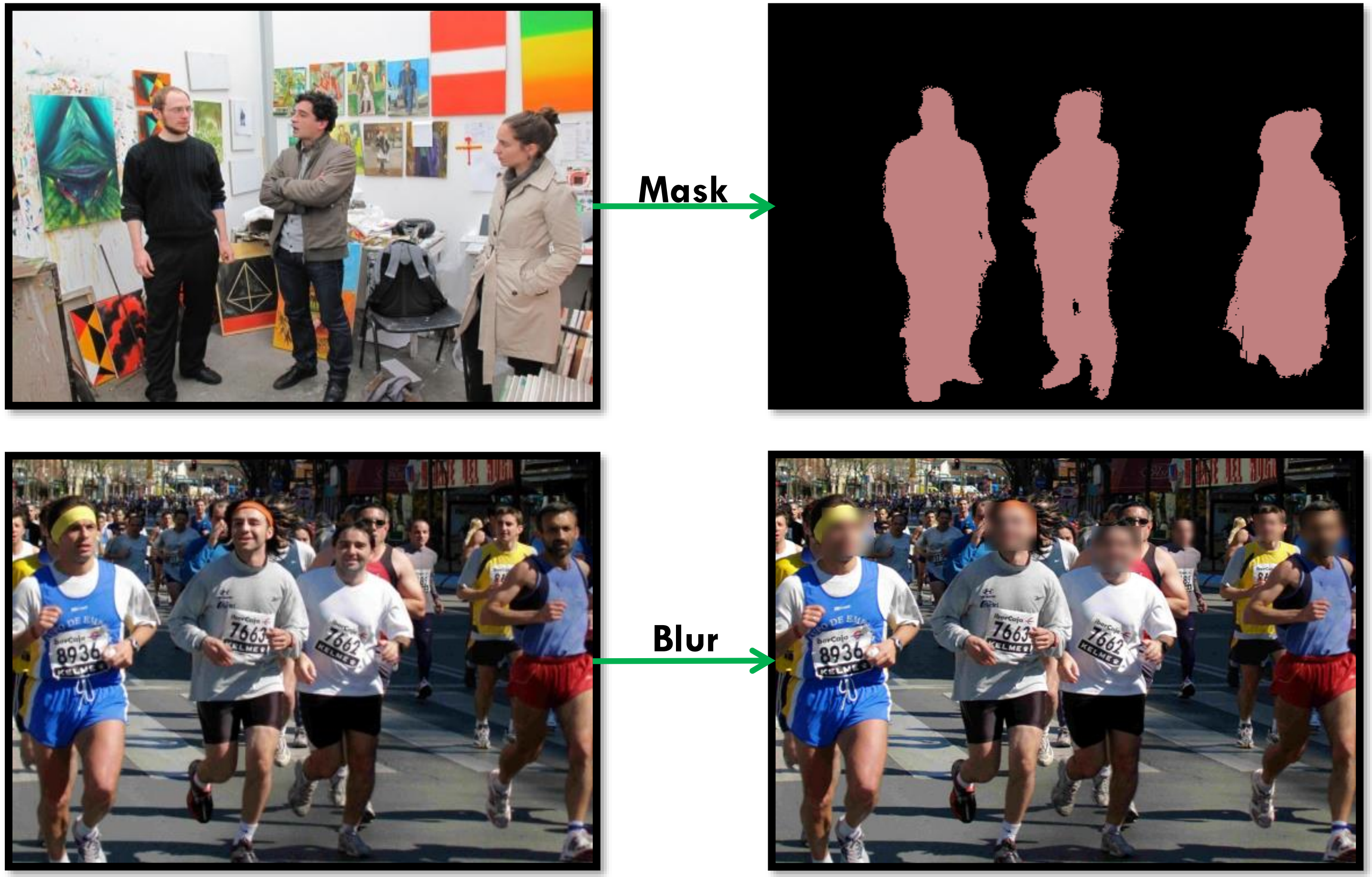
# 5 IMAGE PRIVACY PROTECTION

➢Deep Image based privacy



# 5 IMAGE PRIVACY PROTECTION

- Deep Image based privacy



# Q&A