MAI4CAREU Master programmes in Artificial Intelligence 4 Careers in Europe

University of Cyprus

University of Cyprus

MAI645 - Machine Learning for Graphics and Computer Vision

Marios Loizou, PhD candidate

Spring Semester 2023

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D Vision

These notes are mainly based on the following works:

- Fei-Fei Li, Jiajun Wu, Ruohan Gao, **CS231n - Deep Learning for Computer Vision**, Stanford University
- Hao Su, Jiayuan Gu, Minghua Liu, **Tutorial on 3D Deep Learning**, University of California San Diego
- Evangelos Kalogerakis, **Deep learning architectures for 3D shape analysis and synthesis**, University of Massachusetts Amherst

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D Vision

Notes have been prepared by **Mr. Marios Loizou**

Research Associate at Visual Computing Group at CYENS Centre of Excellence

Visual Computing Group

CYENS CENTRE OF EXCELLENCE

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today's Agenda

- Who are we?
- What is 3D Vision
- 3D shape representations
- 3D shape datasets
- 3D Deep Learning architectures
- What we do

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Today’s Agenda

- Who are we?
- What is 3D Vision
- 3D shape representations
- 3D shape datasets
- 3D Deep Learning architectures
- What we do

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Who are we?

Visual Computing Group at CYENS Centre of Excellence

Melinos Averkiou

MRG Leader

Yiangos Georgiou

Research Associate

Marios Loizou

Research Associate

Yeshwanth Kumar Adimoolam

Research Associate

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today's Agenda

- Who are we?
- What is 3D Vision
- 3D shape representations
- 3D shape datasets
- 3D Deep Learning architectures
- What we do

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

What is 3D Vision: *Overview*

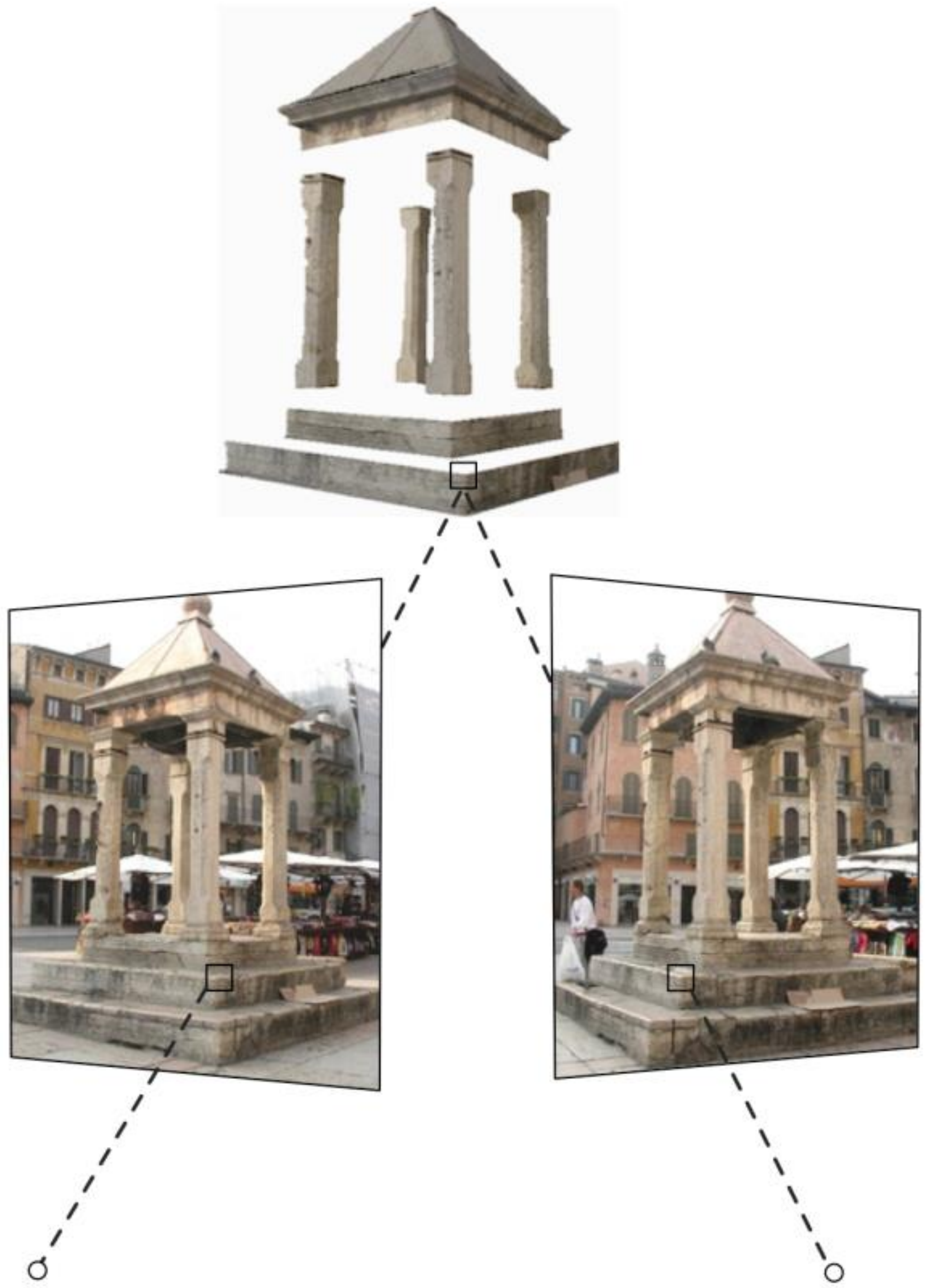
- **Teaching** the computer (**learning**) to **understand the 3D world** around it
- In 3D Vision the input data lie in the **3D space**, rather the 2D domain as in the case of images (2D Vision)
- Deep Learning algorithms and architectures are specifically designed to process this type of data

Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

What is 3D Vision: *Overview*

Traditional 3D Vision

- Multi-view Geometry: Structure from Motion (SfM)



Hao Su et al.

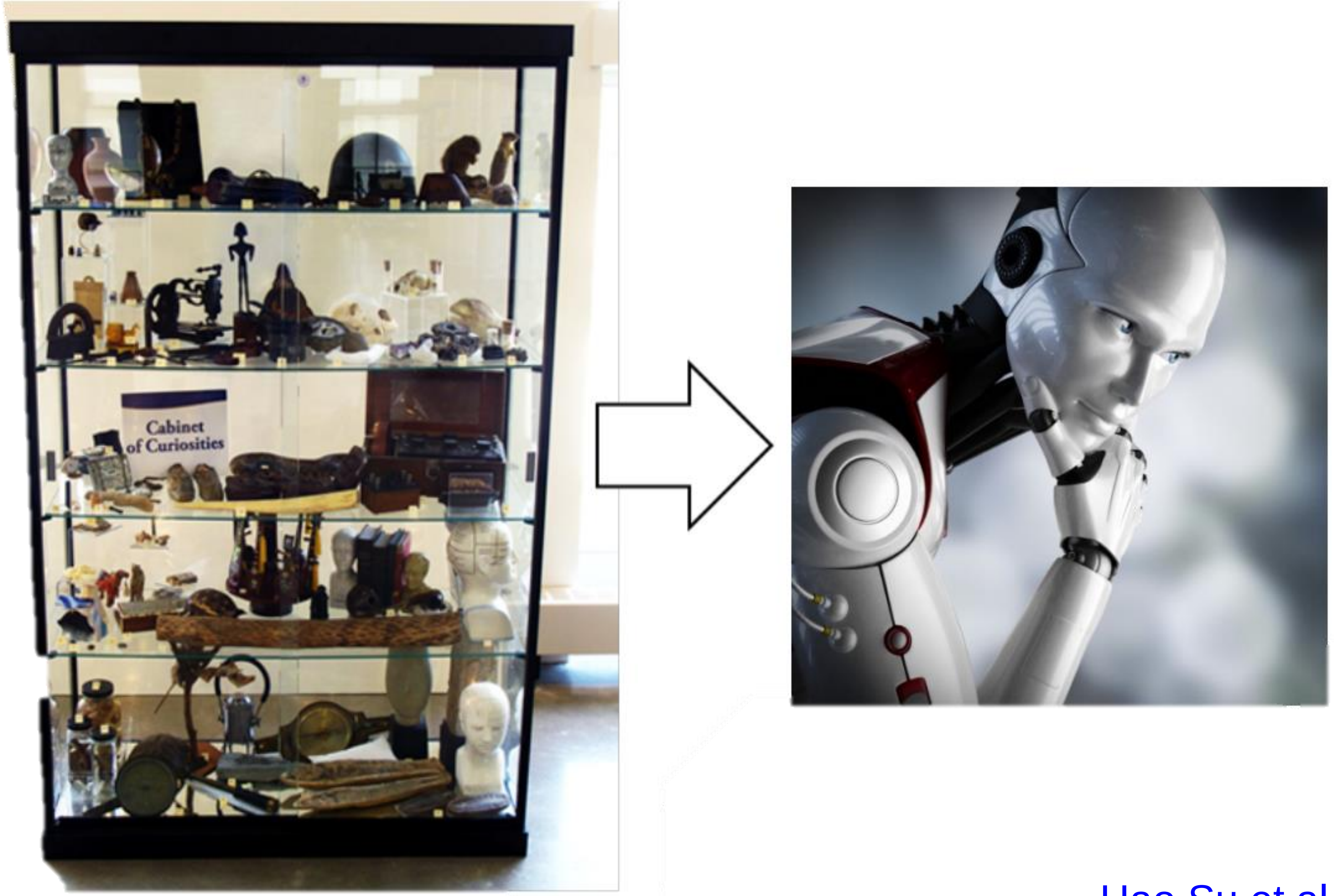
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

What is 3D Vision: *Overview*

Now

- Acquire knowledge of the 3D world by **Learning**



Hao Su et al.

What is 3D Vision: *Tasks (a very small subset)*

Object Classification

ShapeNetCore, Chang et al.

What is 3D Vision: *Tasks (a very small subset)*

Indoor Scene Segmentation

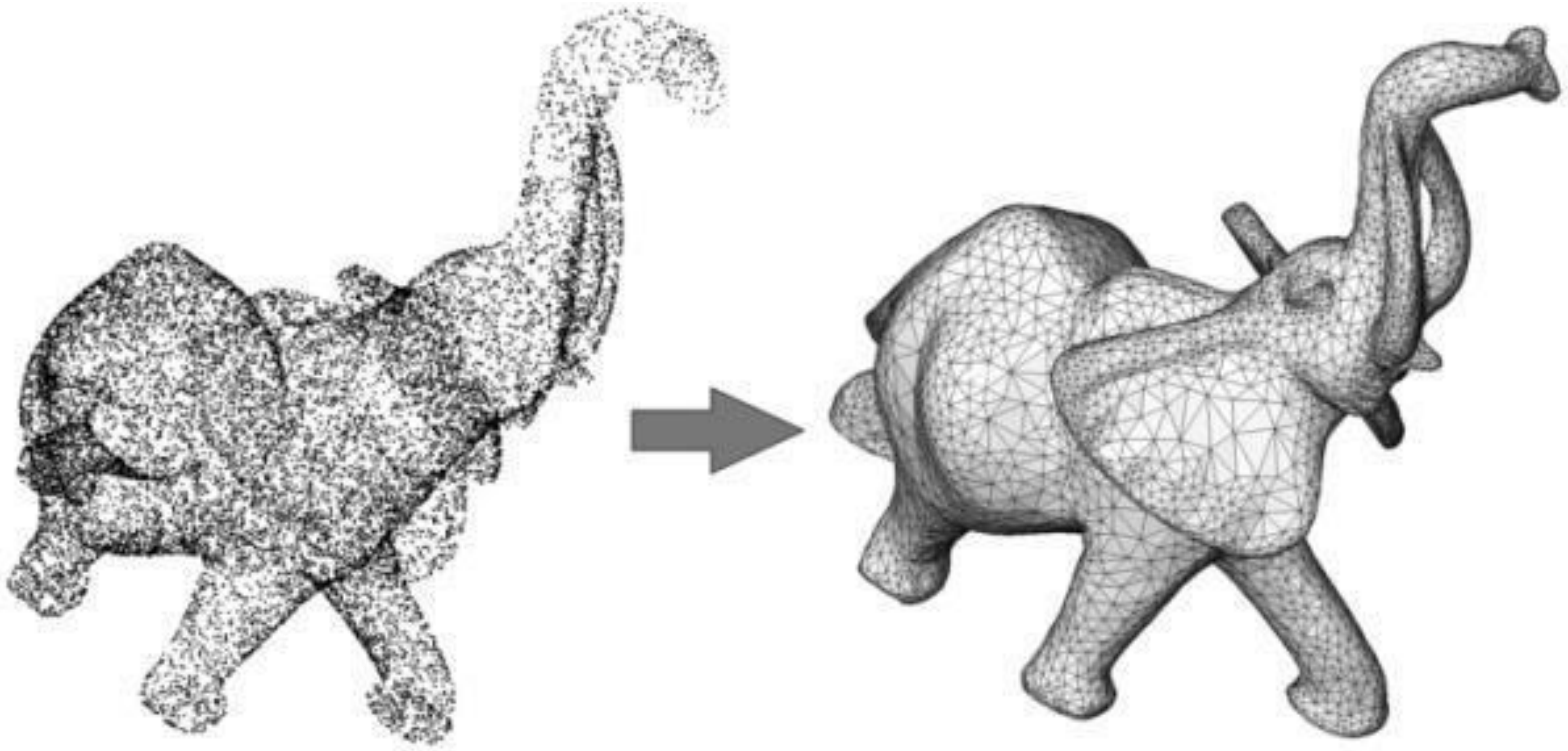
ScanNet,
Angela Dai et al.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

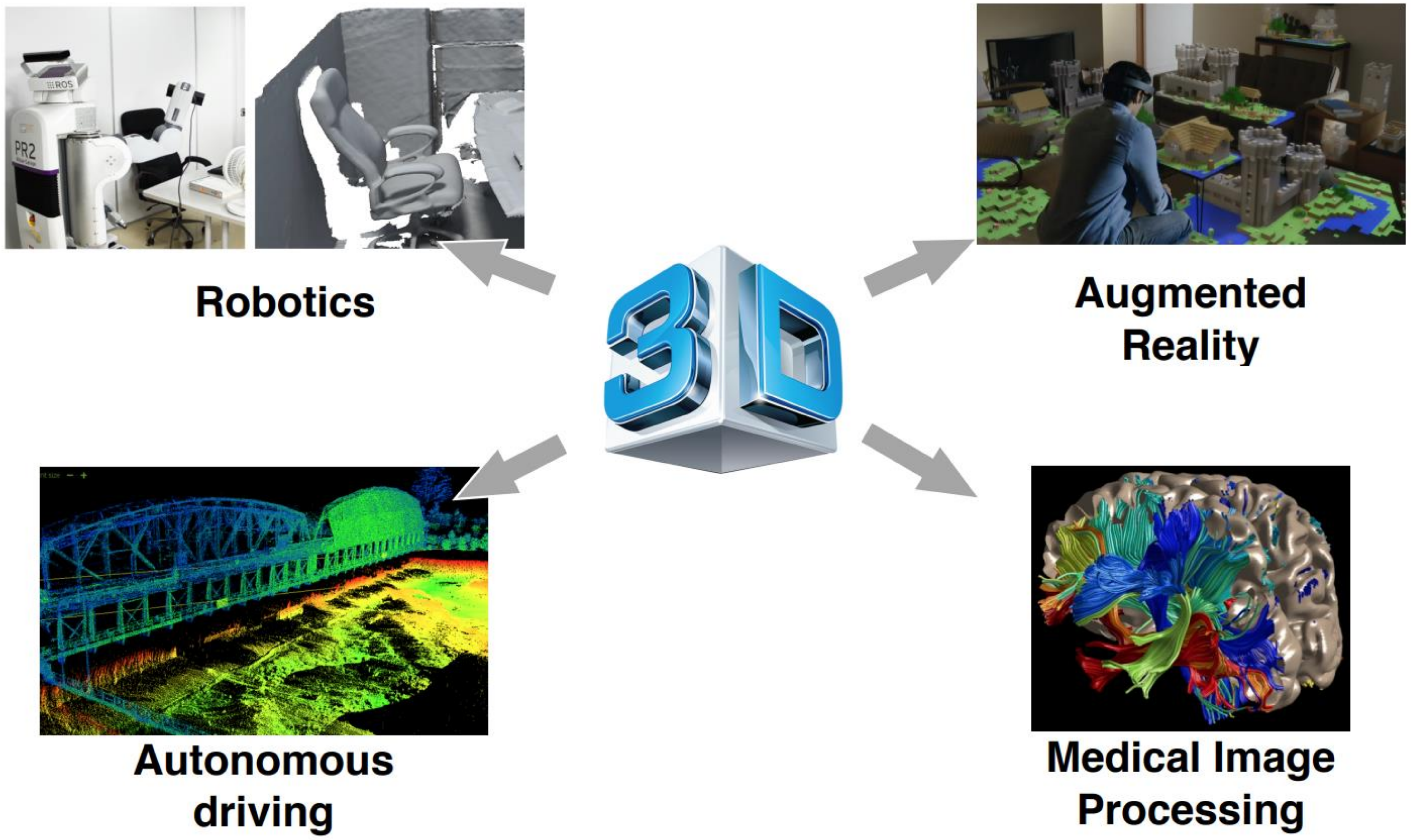
What is 3D Vision: *Tasks (a very small subset)*

Surface reconstruction



This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

What is 3D Vision: *Applications*



Hao Su et al.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

Today’s Agenda

- Who are we?
- What is 3D Vision
- 3D shape representations
- 3D shape datasets
- 3D Deep Learning architectures
- What we do

Co-financed by the European Union
Connecting Europe Facility

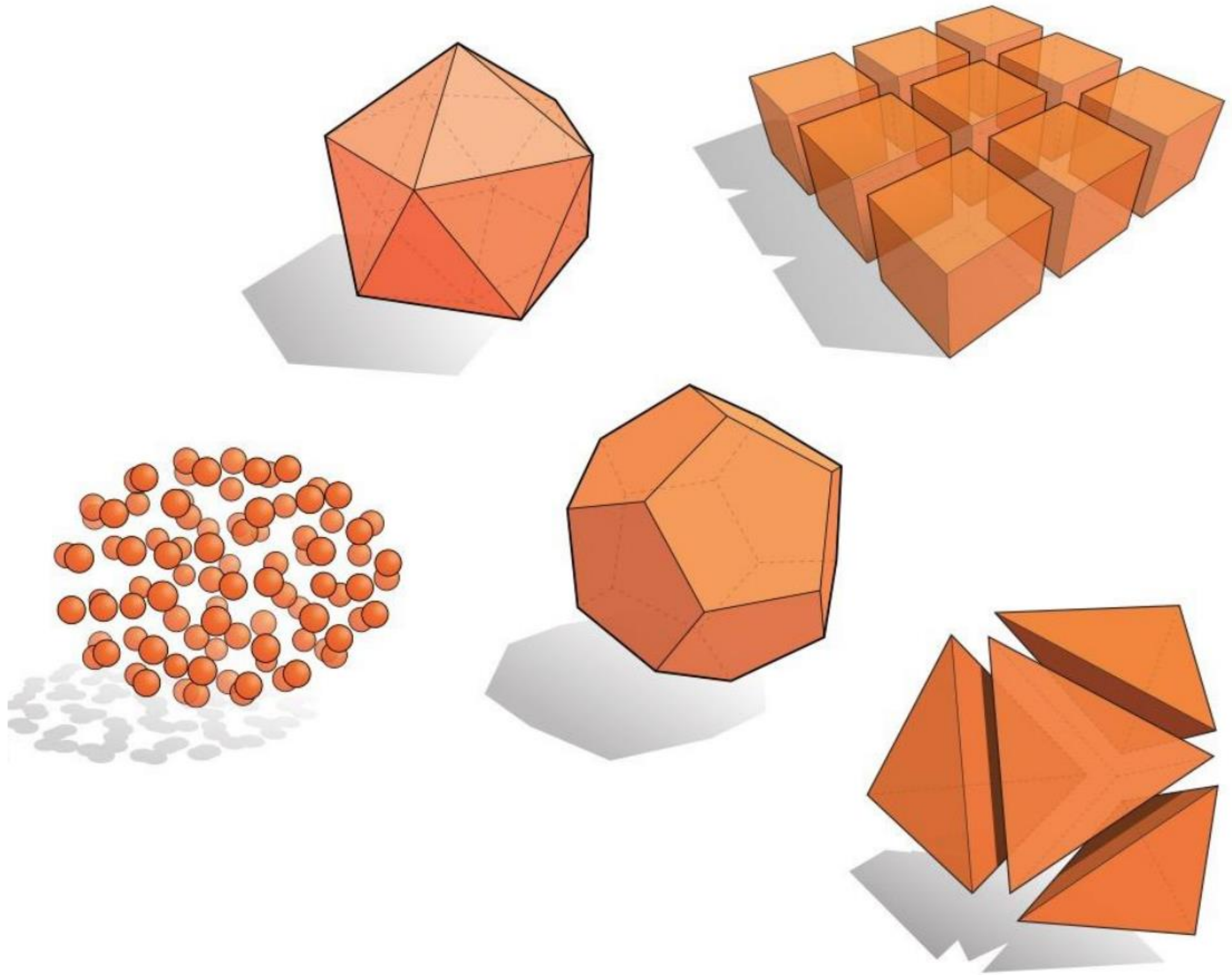
This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Many ways to represent geometry*

Explicit
- point cloud
- polygon mesh
- …

Implicit
- level sets
- distance functions
- …

- Voxels

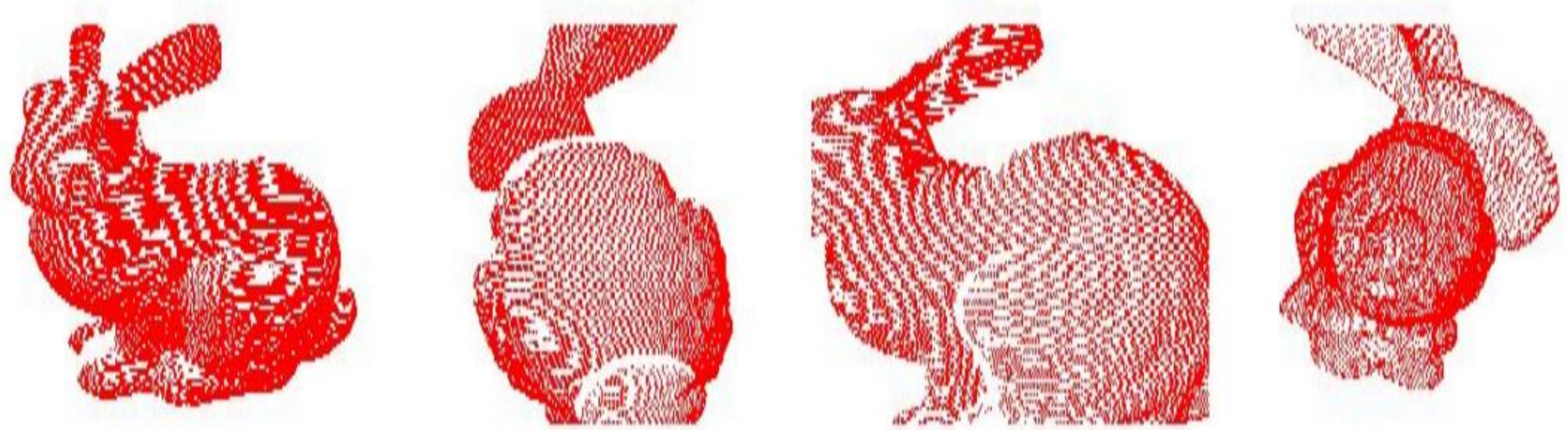


Jiajun Wu

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Point clouds*

- Simplest representation: **only points**, no connectivity



Jiajun Wu

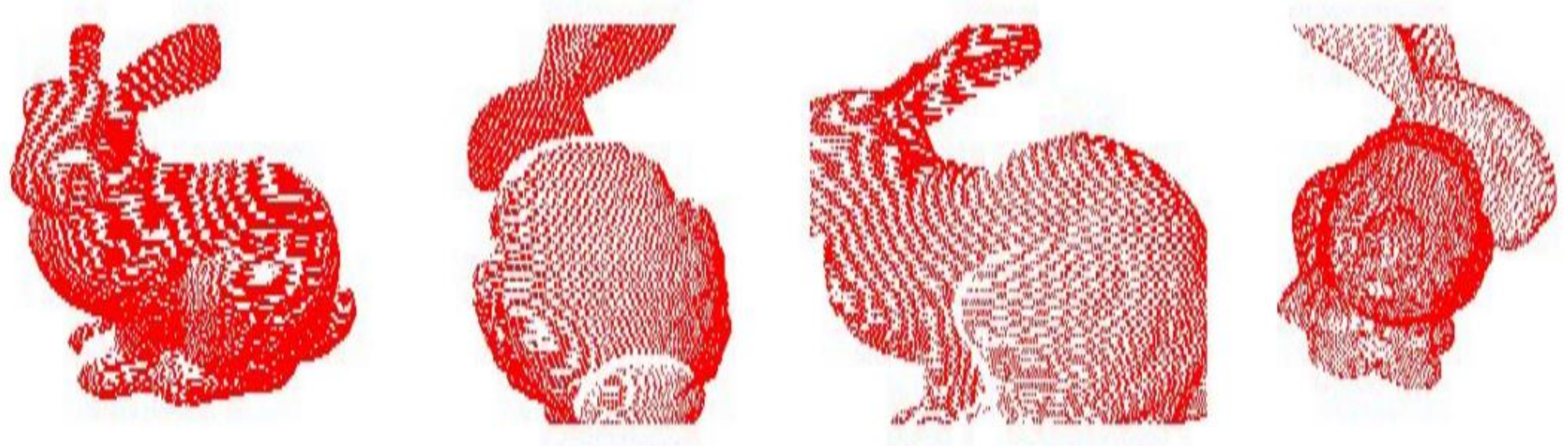
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Point clouds*

- Simplest representation: **only points**, no connectivity

- Collection of (x, y, z) coordinates, possibly with **normal** (perpendicular to the underlying surface)



List of N **unordered** points

$$\left\{\begin{array}{l|l} p_1 & x_1, y_1, z_1, n_{x_1}, n_{y_1}, n_{z_1} \\ p_2 & x_2, y_2, z_2, n_{x_2}, n_{y_2}, n_{z_2} \\ p_3 & x_3, y_3, z_3, n_{x_3}, n_{y_3}, n_{z_3} \\ \vdots & \vdots \\ p_N & x_N, y_N, z_N, n_{x_N}, n_{y_N}, n_{z_N} \end{array}\right.$$

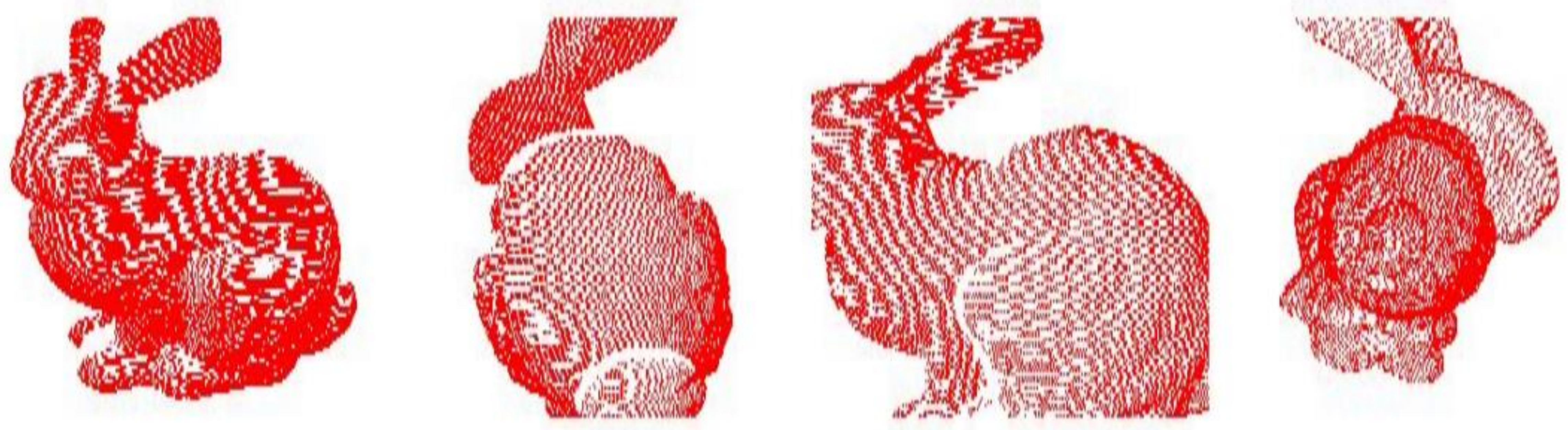
Jiajun Wu

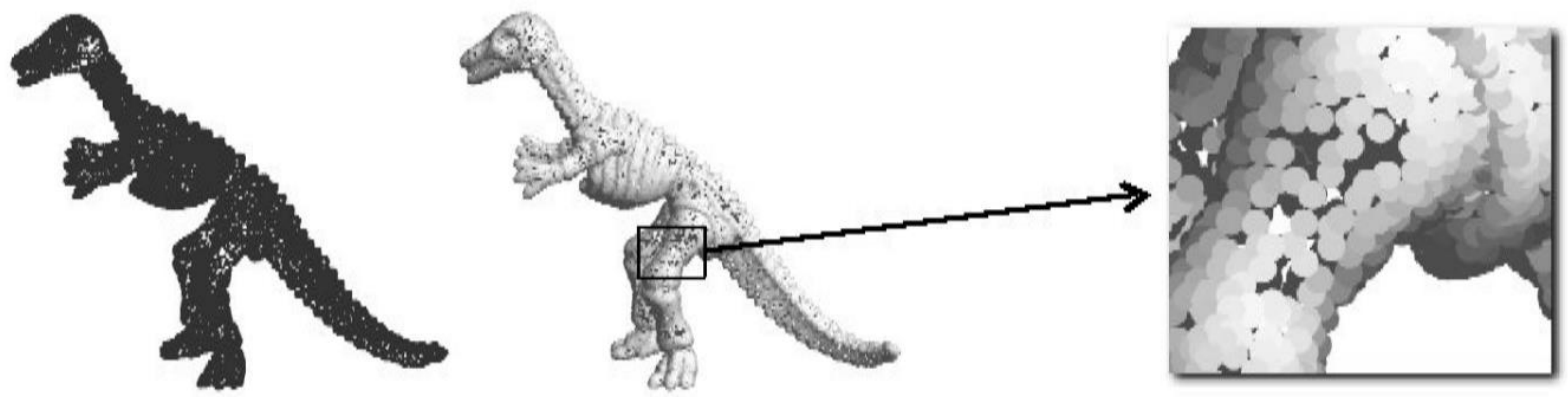
3D shape representations: *Point clouds*

- Simplest representation: **only points**, no connectivity

- Collection of (x, y, z) coordinates, possibly with **normal** (perpendicular to the underlying surface)

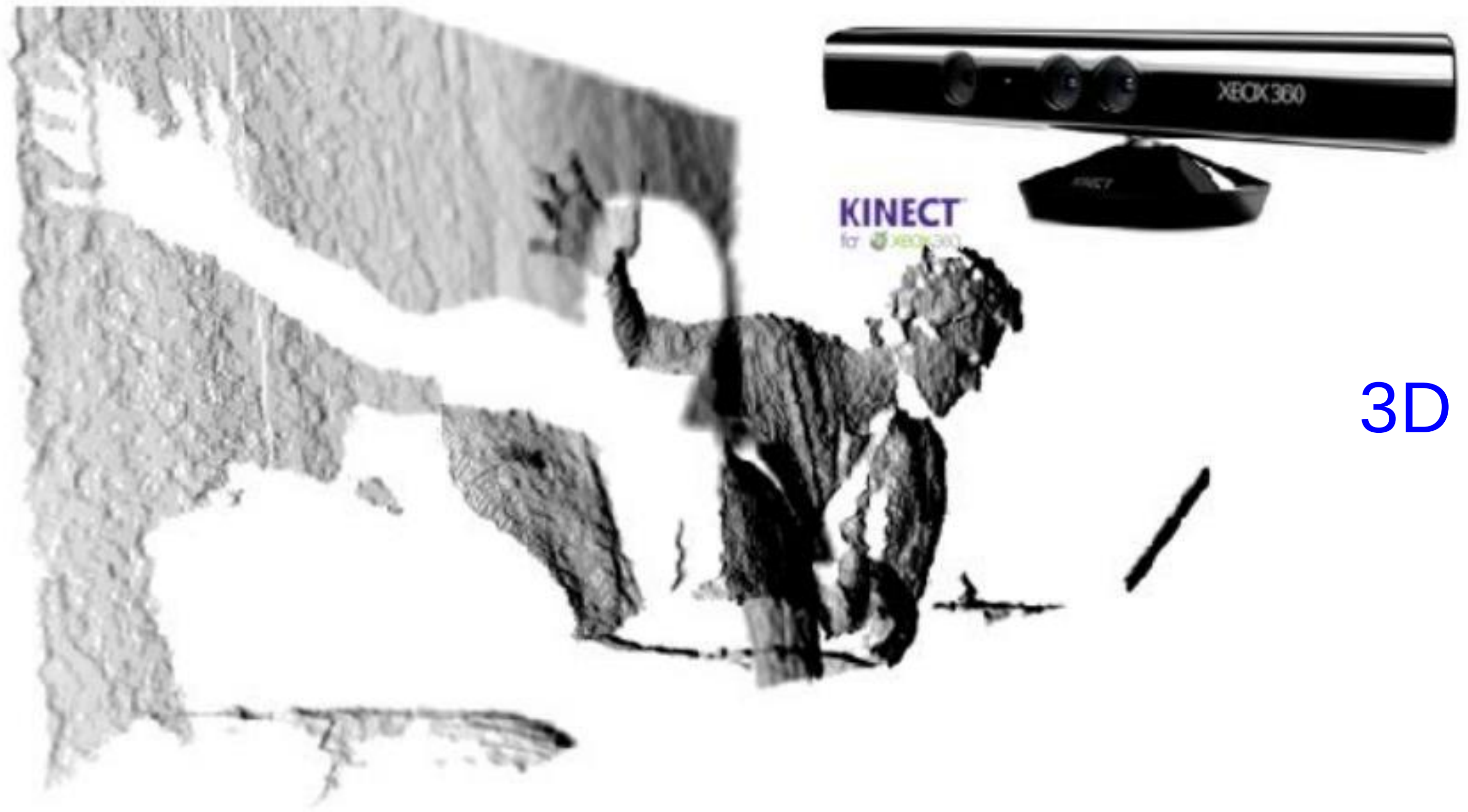
- Points with orientation (normal) are called **surfels**





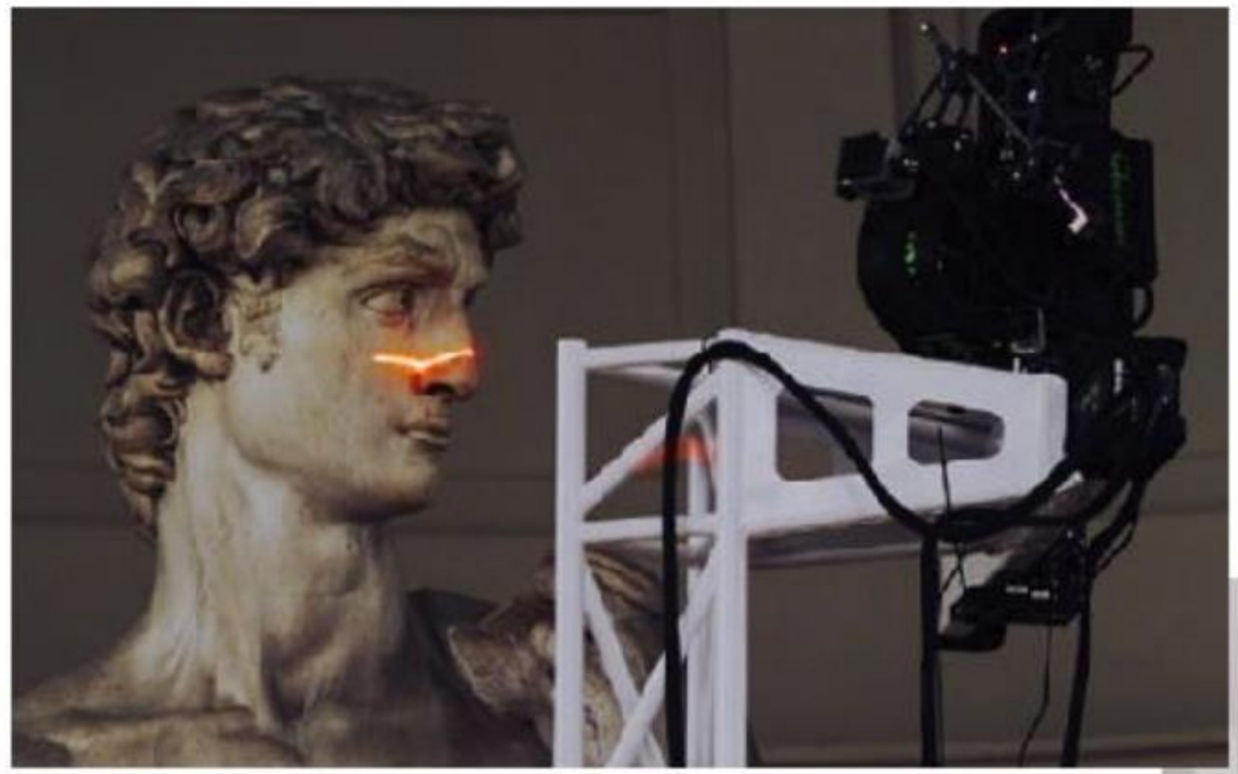
Jiajun Wu

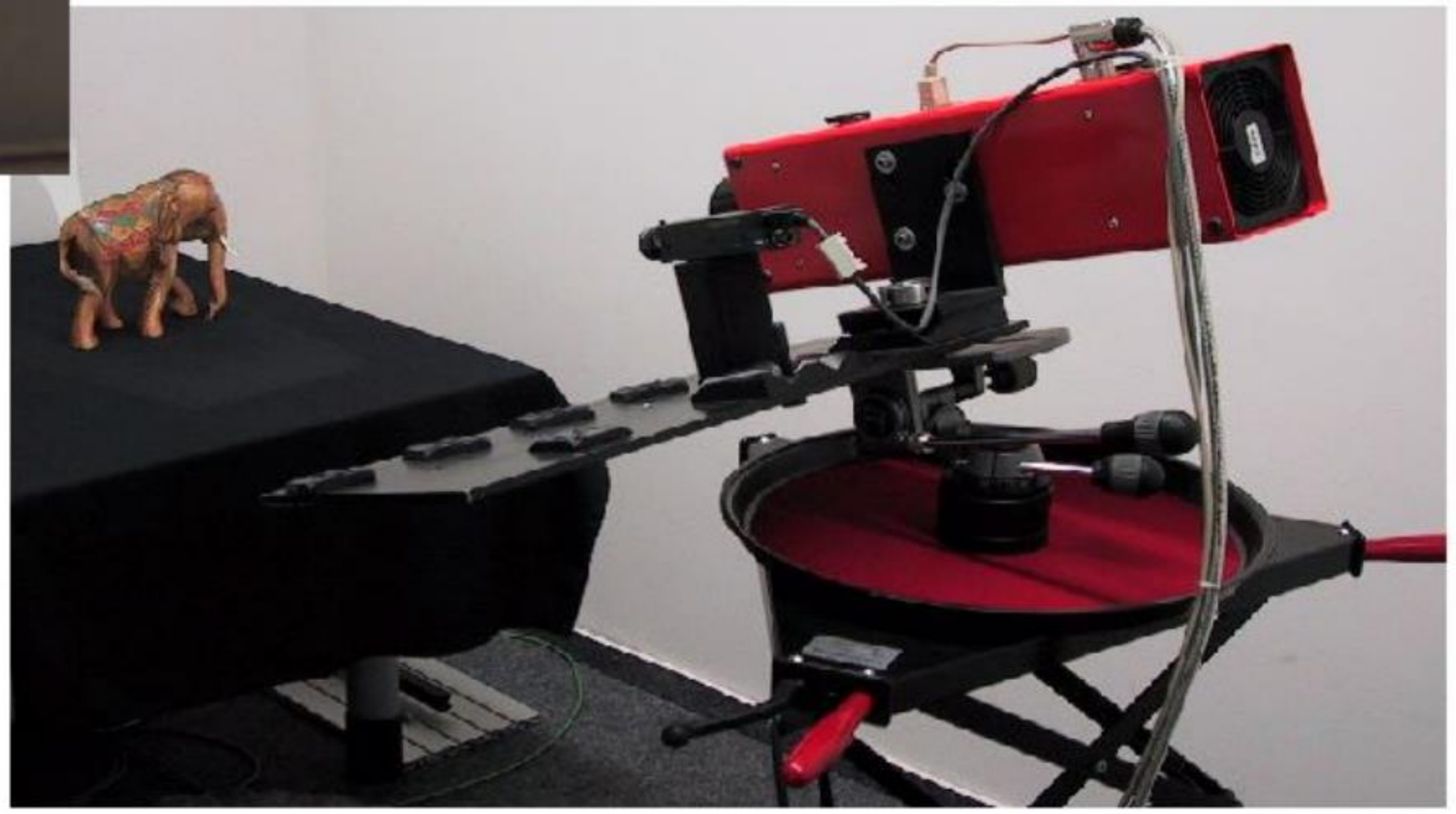
3D shape representations: *Point clouds acquisition*



3D Depth sensor

Laser triangulation rangefinder





3D Laser scanner

Jiajun Wu

Co-financed by the European Union
Connecting Europe Facility

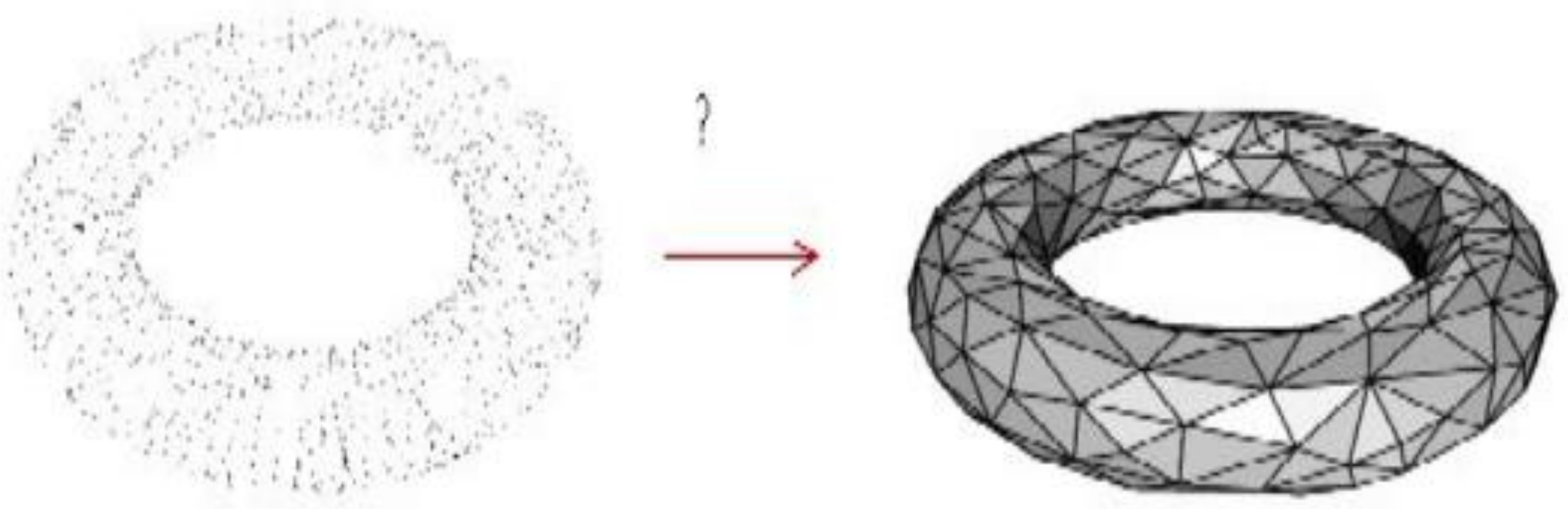
This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Point clouds pros & cons*

- Pros:
  - ✓ Easily represents any kind of geometry
  - ✓ Useful for large datasets

- Cons
  - x Incomplete/noisy point clouds
  - x No topological information

Incomplete scans



No topology

Jiajun Wu

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Polygonal Meshes*

- A 3D polygonal mesh is the **structural build** of a **3D model** consisting of **polygons**
- **Boundary representation** of objects

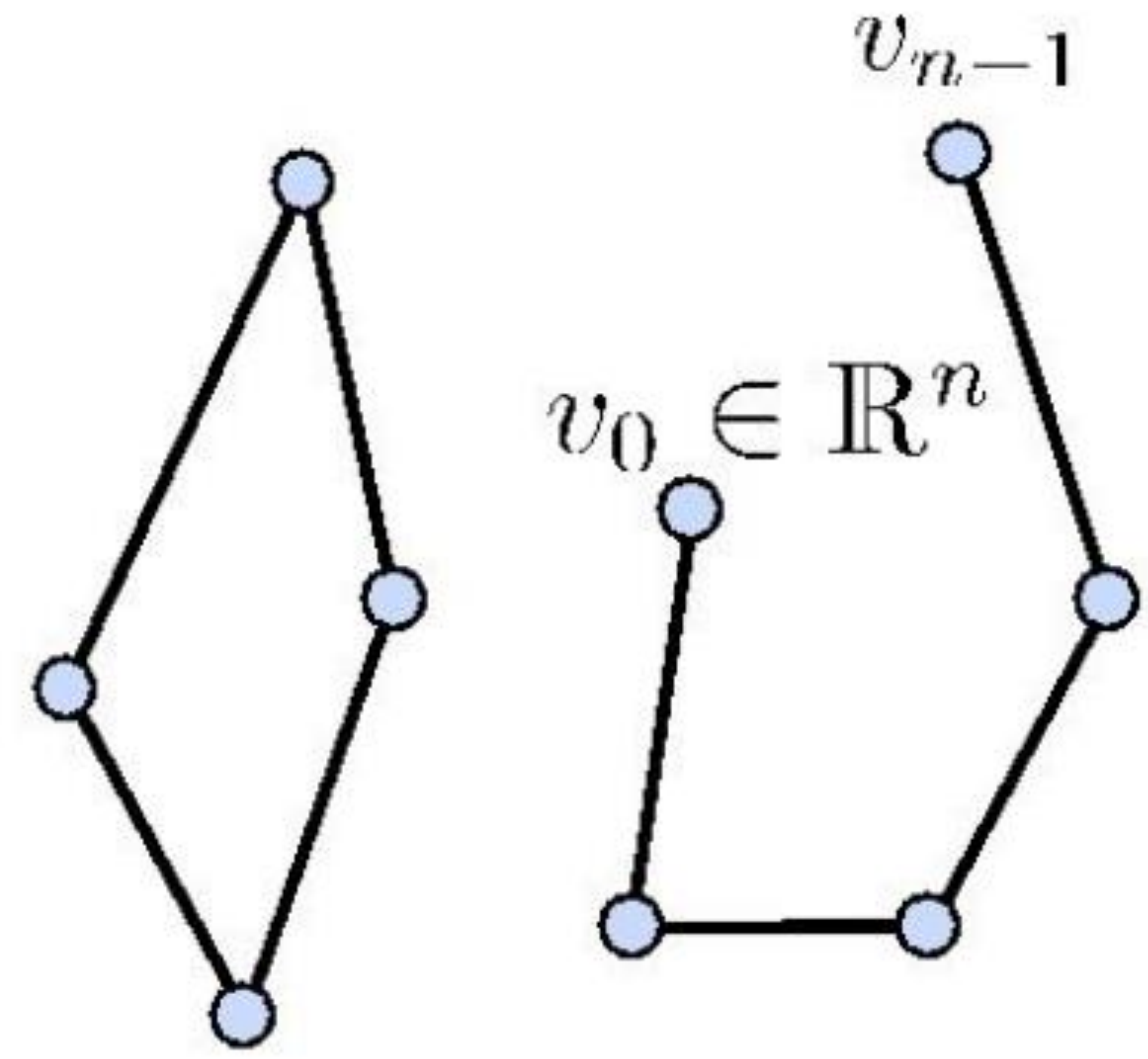
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Polygonal Meshes*

- Polygon:
  - **Vertices**: $v_0, v_1, \cdots, v_{n-1}$
  - **Edges**: $\{(v_0, v_1), \cdots, (v_{n-2}, v_{n-1})\}$

- Types of polygons:
  - **Closed**: $v_0 = v_{n-1}$
  - **Planar**: all vertices on a plane
  - **Simple**: not self-intersecting



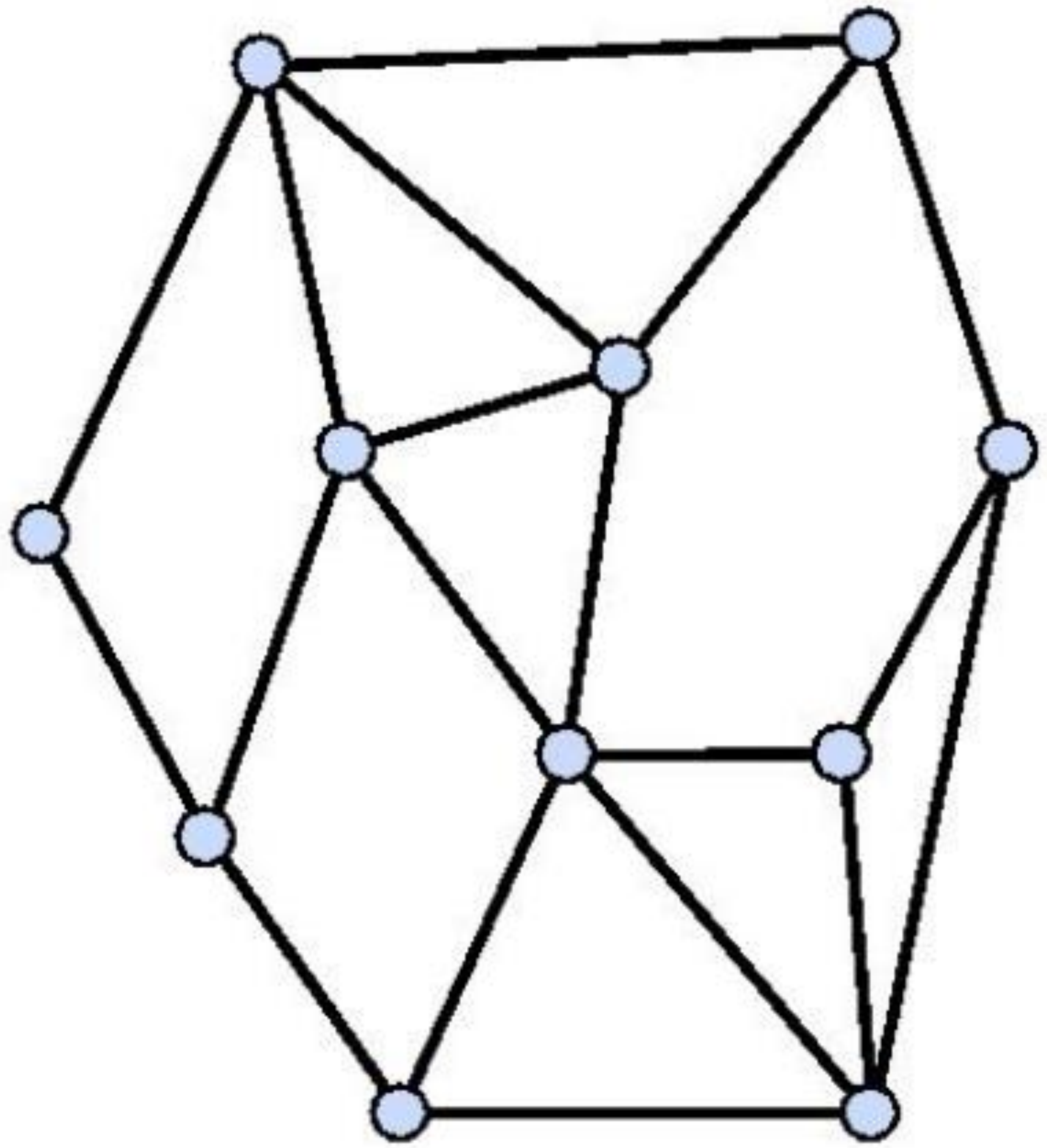
Jiajun Wu

3D shape representations: *Polygonal Meshes*

- Polygonal Mesh:
  - A finite set M of **closed, simple** polygons Q_i

$$\boldsymbol{M} = \langle \boldsymbol{V}, \boldsymbol{E}, \boldsymbol{F} \rangle$$

$\boldsymbol{V}$ = set of vertices
$\boldsymbol{E}$ = set of edges
$\boldsymbol{F}$ = set of faces



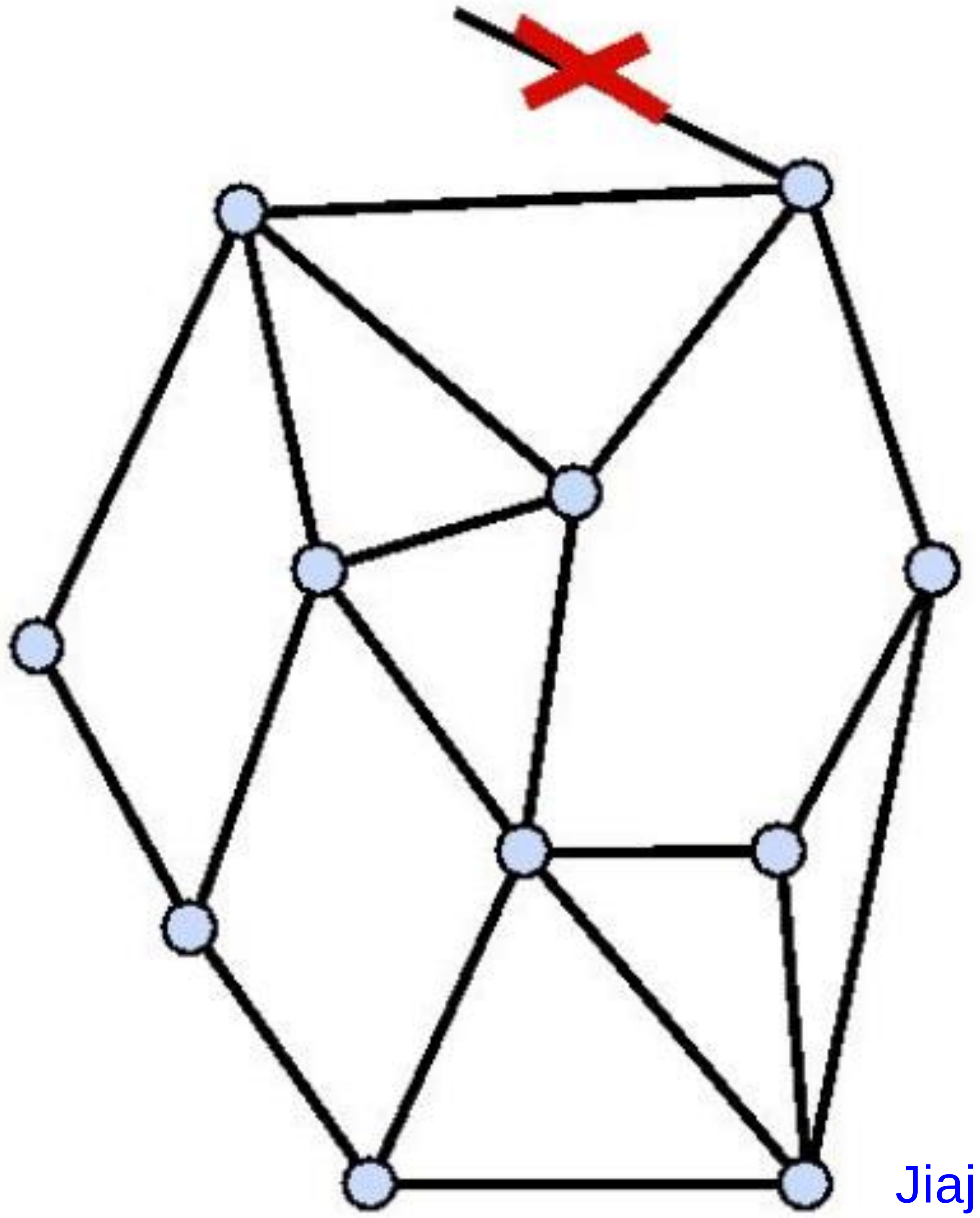
Jiajun Wu

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Polygonal Meshes*

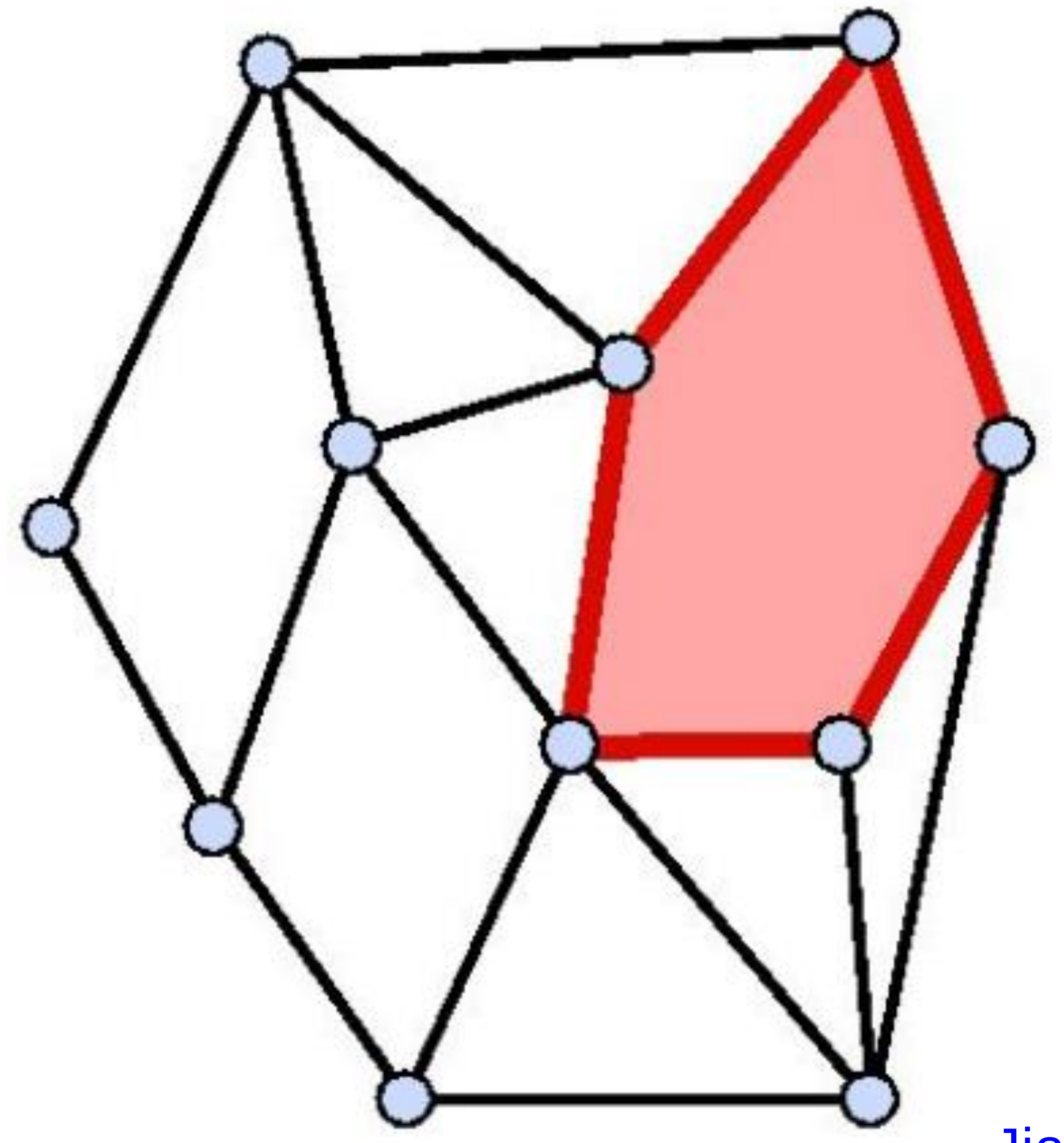
- Polygonal Mesh:
  - A finite set M of **closed, simple** polygons Q_i
  - Every **edge** belongs to **at least one polygon**



Jiajun Wu

3D shape representations: *Polygonal Meshes*

- Polygonal Mesh:
  - A finite set M of **closed, simple** polygons Q_i
  - Every **edge** belongs to **at least one polygon**
  - Each Q_i defines a **face** of the polygonal mesh



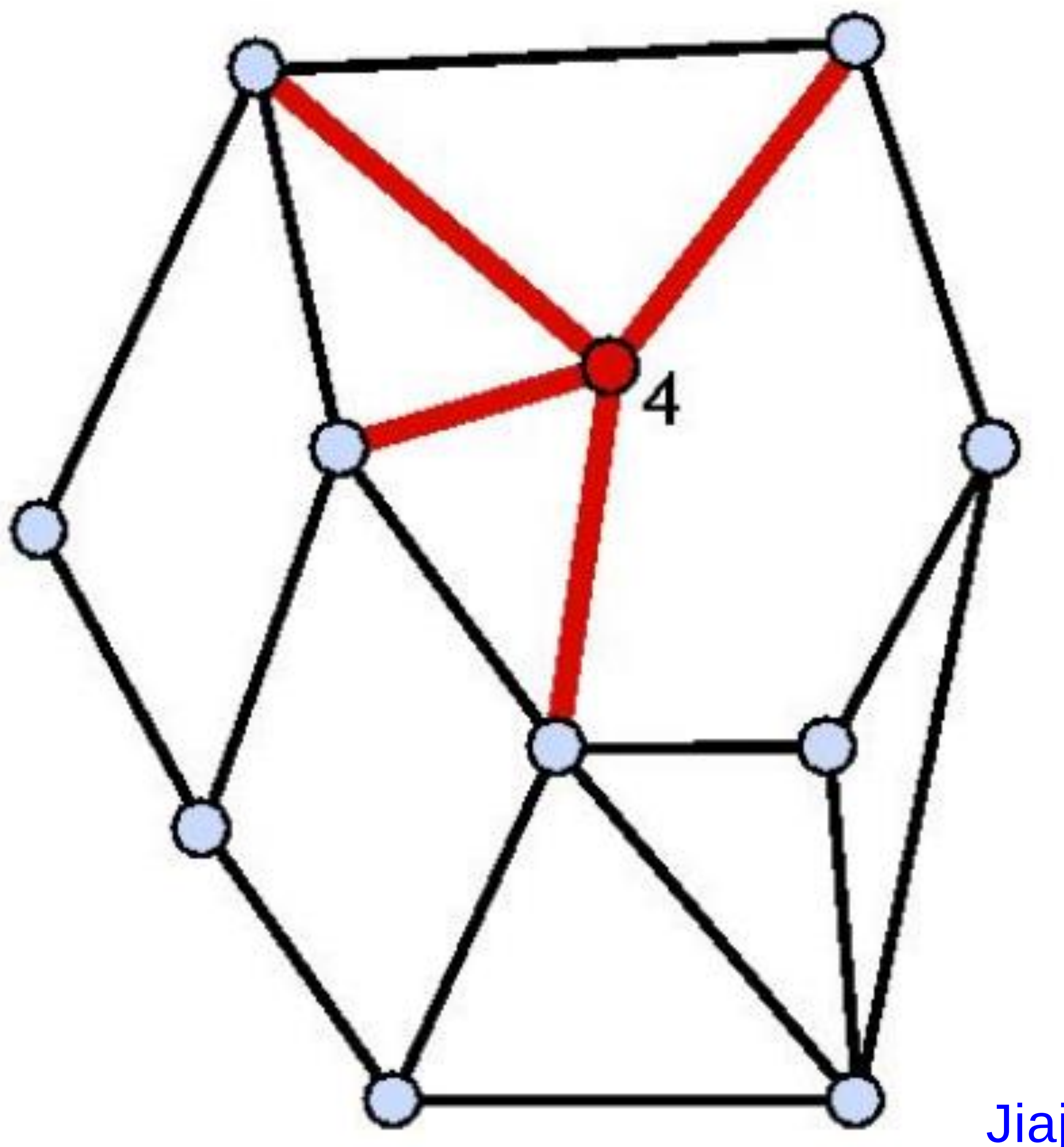
Jiajun Wu

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Polygonal Meshes*

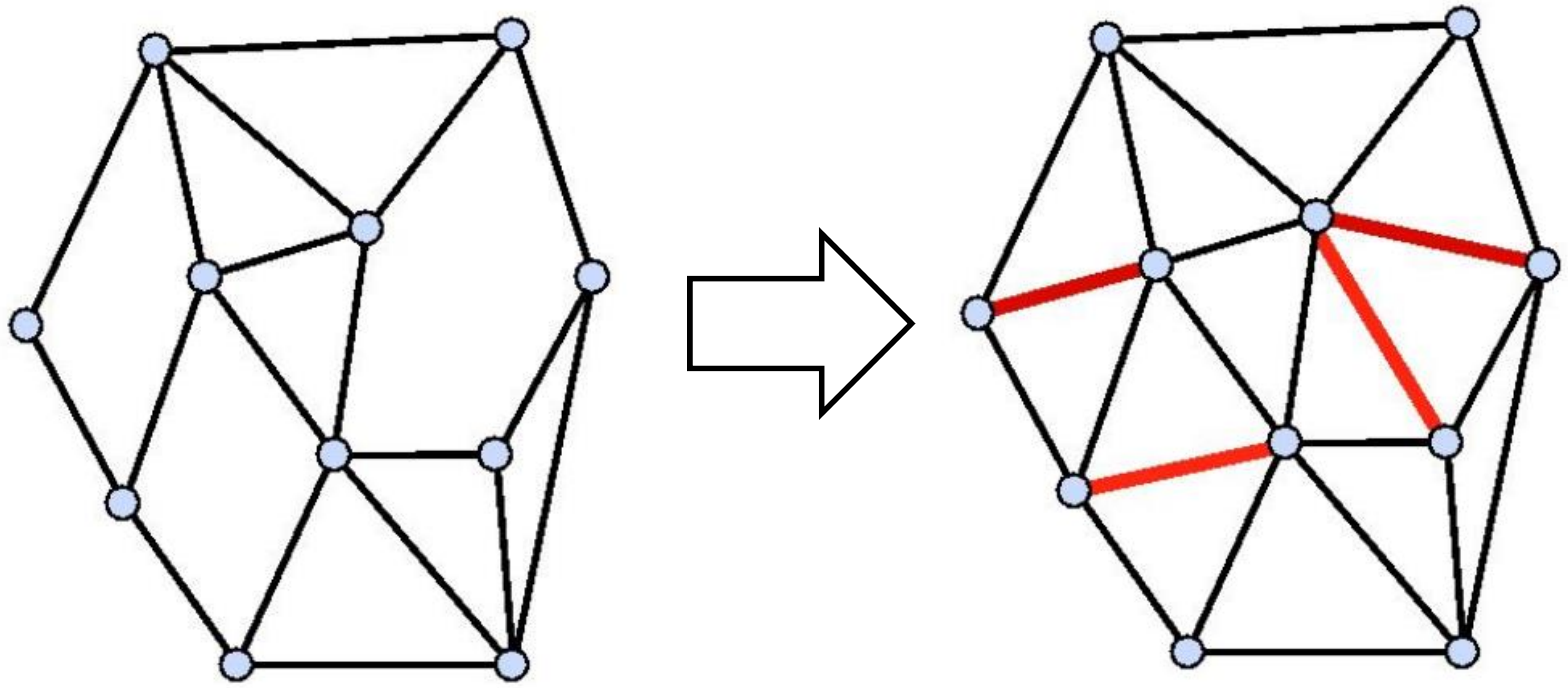
- Polygonal Mesh:
  - A finite set M of **closed, simple** polygons Q_i
  - Every **edge** belongs to **at least one polygon**
  - Each Q_i defines a **face** of the polygonal mesh
  - Vertex **degree** or **valance** = number of incident edges



Jiajun Wu

3D shape representations: *Polygonal Meshes*

- Polygonal Mesh **Triangulation**:
  - Polygonal mesh where every face is a triangle → **triangular mesh**
  - Simplifies data structures
  - Simplifies rendering
  - Simplifies algorithms
  - Each face is planar and convex
  - Any polygon can be triangulated

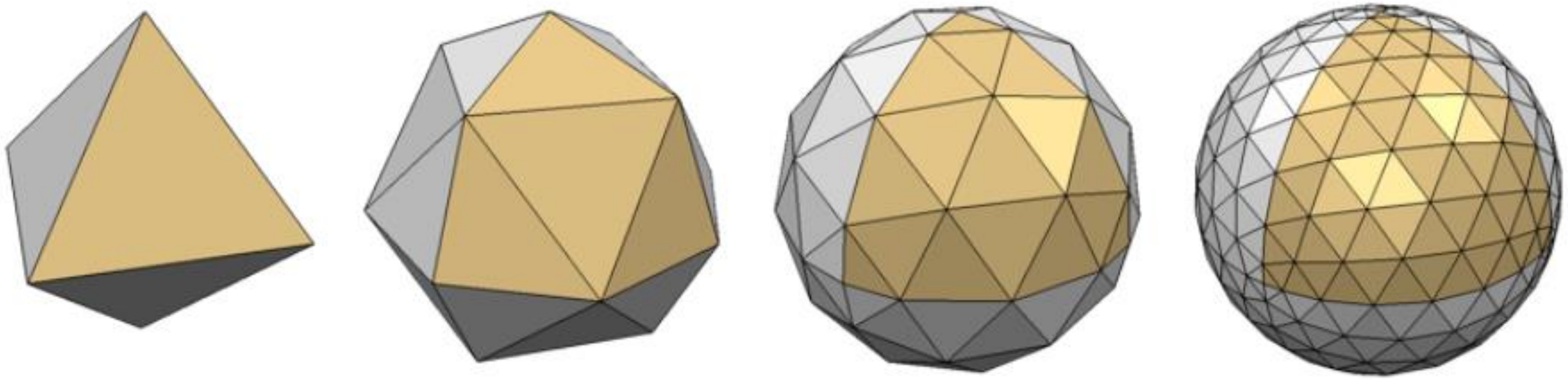


Jiajun Wu

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Polygonal Meshes*



- A polygonal mesh consists of **three kinds** of **mesh elements**: **vertices**, **edges** and **faces**
- **Mesh connectivity** or **topology**: describes the incidence relation amongst mesh elements
- **Mesh geometry**: specifies the position and other geometric characteristics of each vertex

Co-financed by the European Union
Connecting Europe Facility
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Polygonal Meshes*

- **Data Structures**:
  - What should be stored?
    - **Geometry**: 3D coordinates
    - **Connectivity**: Adjacency relationships
    - **Attributes**:
      - Normal, color, texture coordinates
      - Per vertex, face, edge

Jiajun Wu

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Polygonal Meshes*

- **Indexed Face Set**
  - Used in formats like OBJ and OFF
  - Storage
    - Vertex: position
    - Face: vertex indices
  - No explicit neighborhood info

Vertices			
v0	x0	y0	z0
v1	x1	x1	z1
v2	x2	y2	z2
v3	x3	y3	z3
v4	x4	y4	z4
v5	x5	y5	z5
v6	x6	y6	z6
...	...	...	...

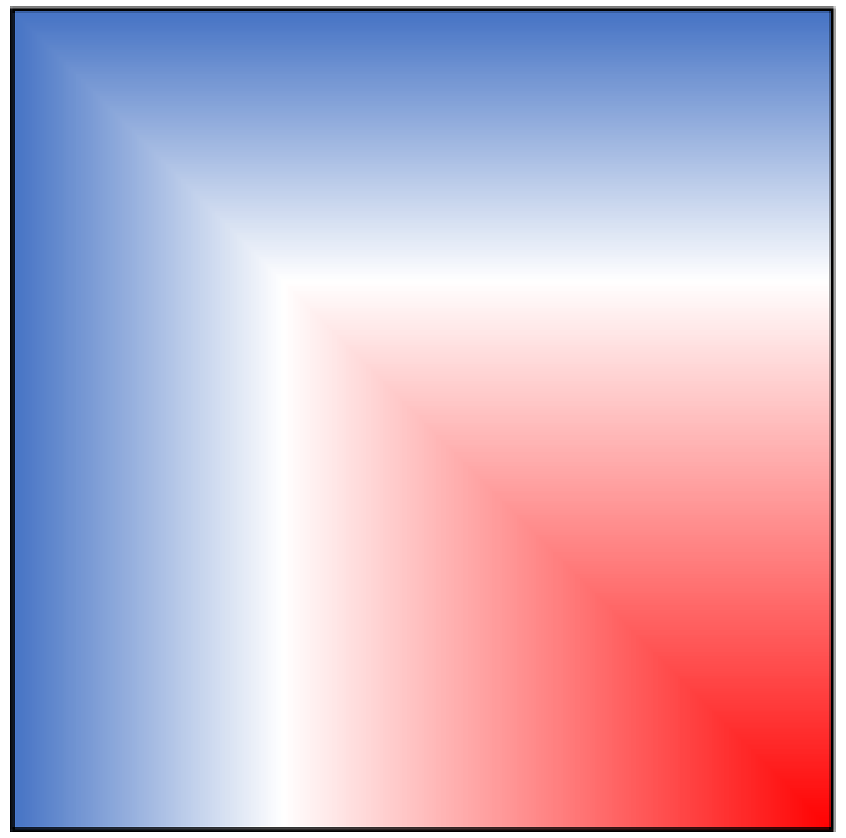
Triangles			
t0	v0	v1	v2
t1	v0	v1	v3
t2	v2	v4	v3
t3	v5	v2	v6
...	...	...	...

Jiajun Wu

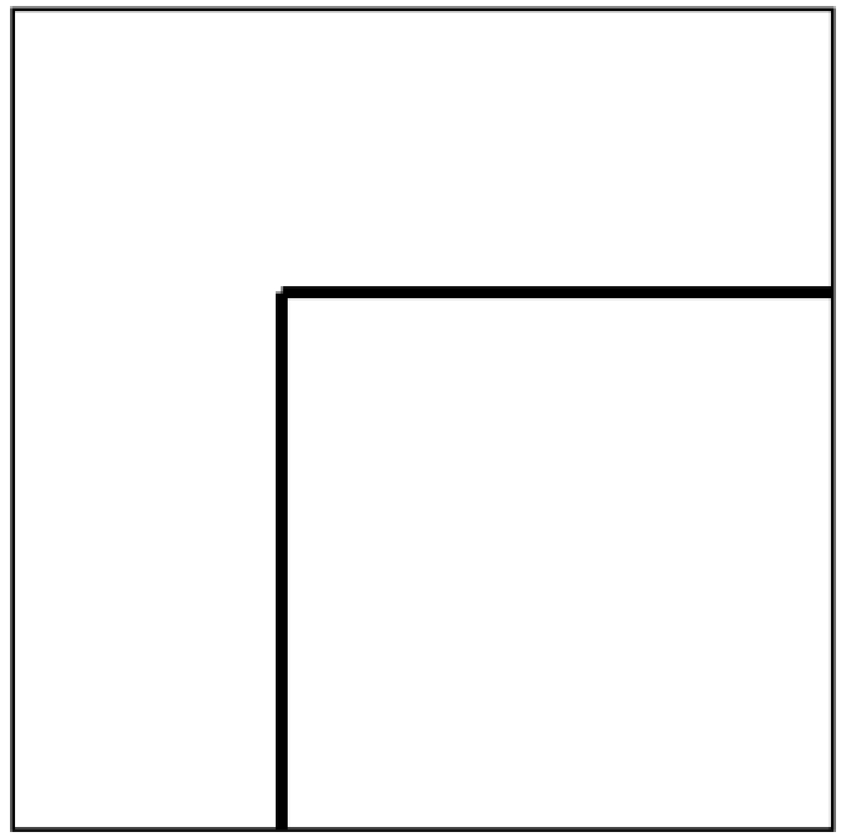
This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape representations: *Implicit Functions*

- **Implicit function**
  - Classifies arbitrary $3D$ points as inside / outside the shape



Implicit function



Explicit Shape

Justin Solomon

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

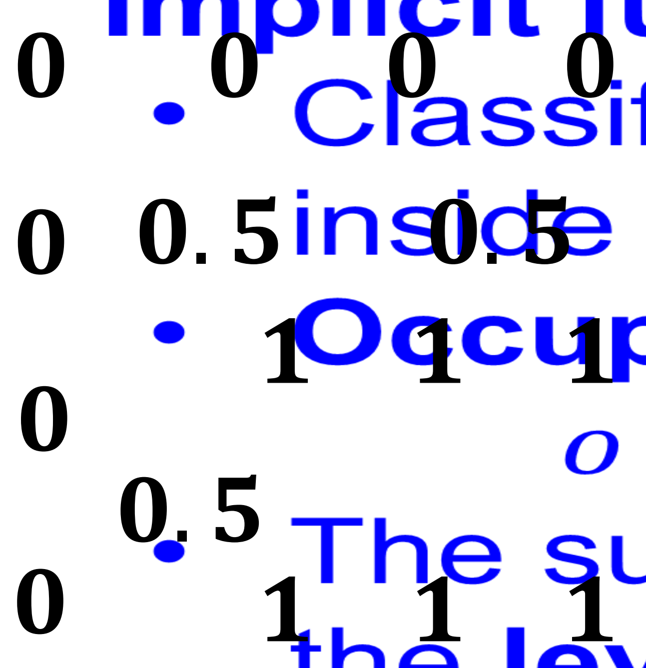
3D shape representations: *Implicit Functions*

- **Implicit function**
  - Classifies arbitrary $3D$ points as inside / outside the shape
  - **Occupancy function**:

    $$o : \mathbb{R}^3 \rightarrow \{0,1\}$$

  - The surface of the 3D object is the **level set**:

    $$\{x : \boldsymbol{o}(x) = \frac{1}{2}\}$$



- **Implicit function**
  - Classifies arbitrary $3D$ points as inside / outside the shape
  - **Occupancy function**:

    $$o : \mathbb{R}^3 \rightarrow \{0,1\}$$

  - The surface of the 3D object is the **level set**:

    $$\{x : \boldsymbol{o}(x) = \frac{1}{2}\}$$

Justin Solomon

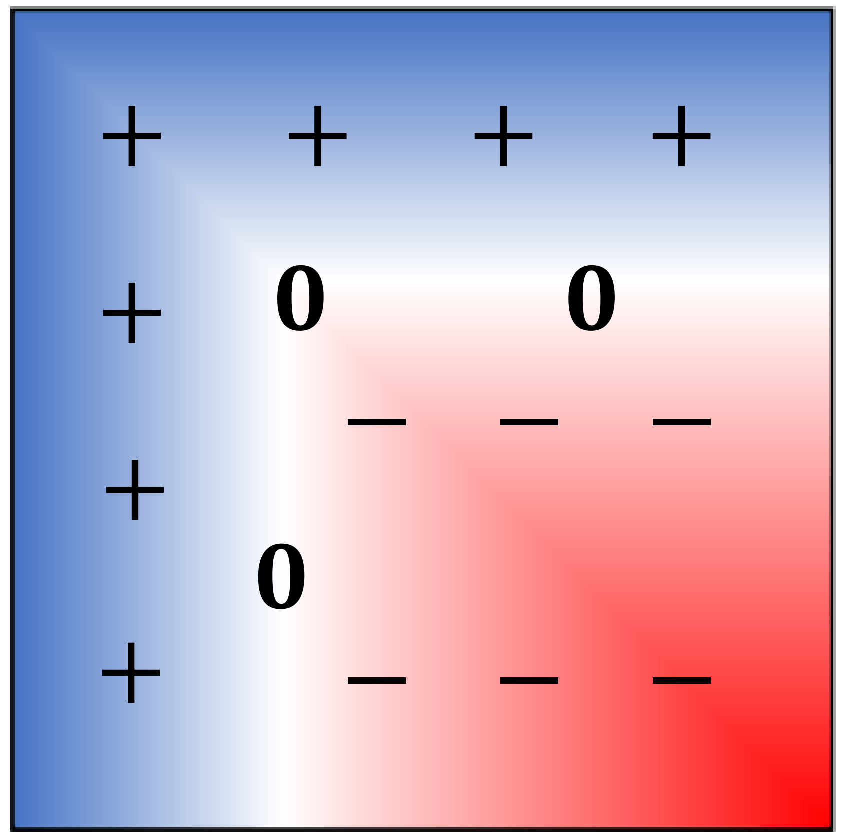
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

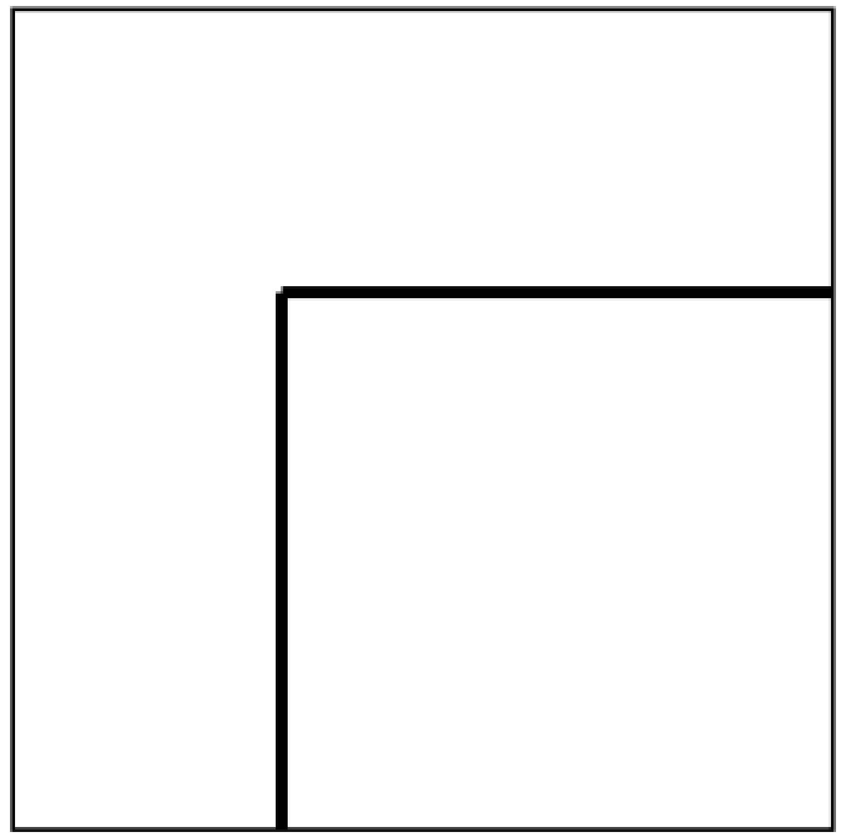
3D shape representations: *Implicit Functions*

- **Implicit function**
  - Classifies arbitrary $3D$ points as inside / outside the shape
  - **Signed Distance Function:** Euclidean distance to the surface of shape; sign gives inside / outside
  - The surface of the 3D object is the **level set**:

$$\{x: \boldsymbol{SDF}(x) = 0\}$$



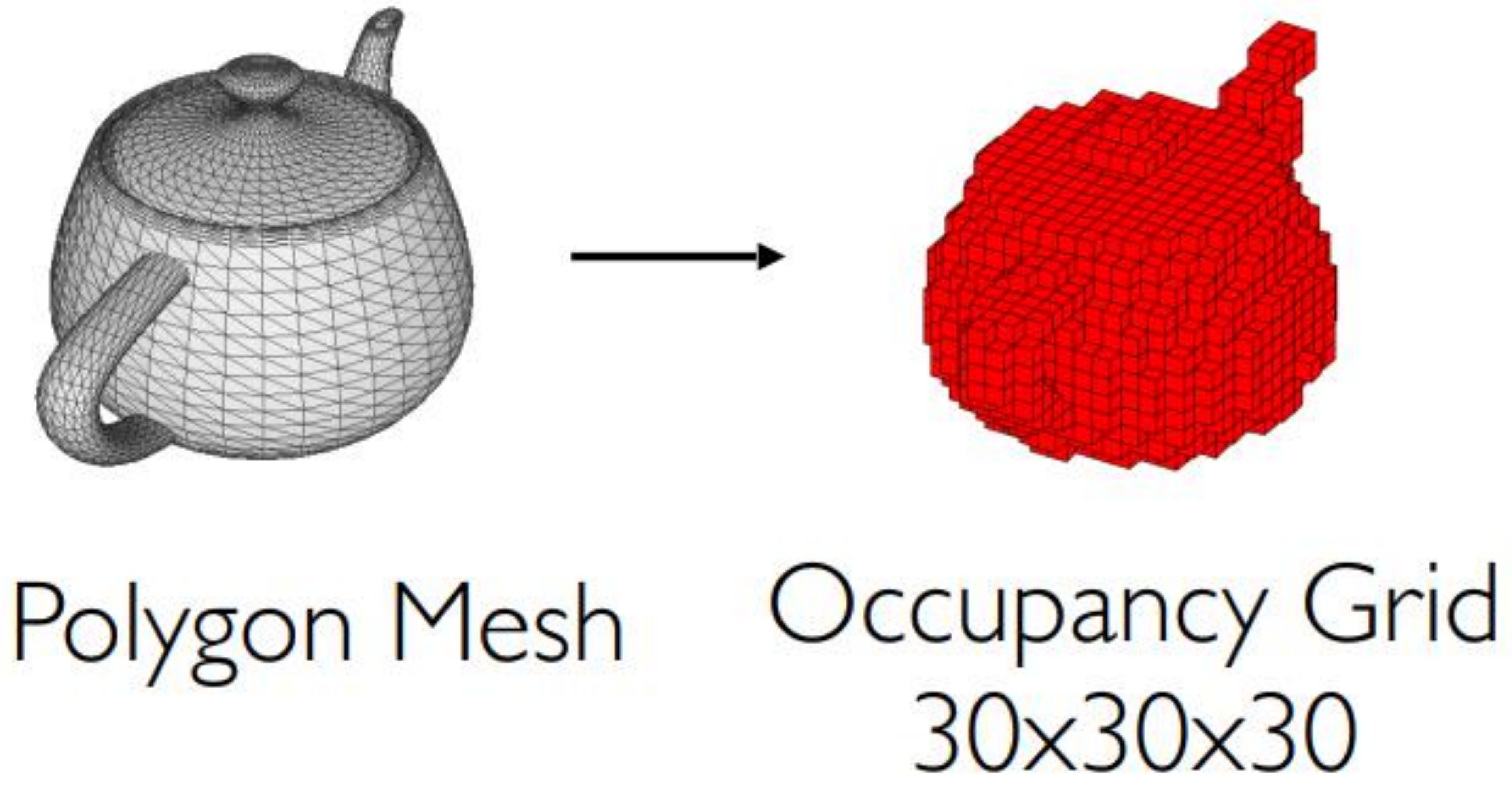
Implicit function


Explicit Shape

Justin Solomon

3D shape representations: *Volumetric Grid*

- **Volumetric Grid**
  - Represent a shape with a $V \times V \times V$ grid of occupancies or SDFs
  - Conceptually simple → just a 3D **regular Euclidean** grid
  - Like an image
    - Pixels –> Voxels
  - Straightforward to apply 3D convolutions



Hao Su et al.

Today’s Agenda

- Who are we?
- What is 3D Vision
- 3D shape representations
- 3D shape datasets
- 3D Deep Learning architectures
- What we do

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape datasets: *Datasets for 3D Objects*

- **Princeton Shape Benchmark**
  - # Models: 1,814
  - # Categories: 182

Shilane et al., 2004

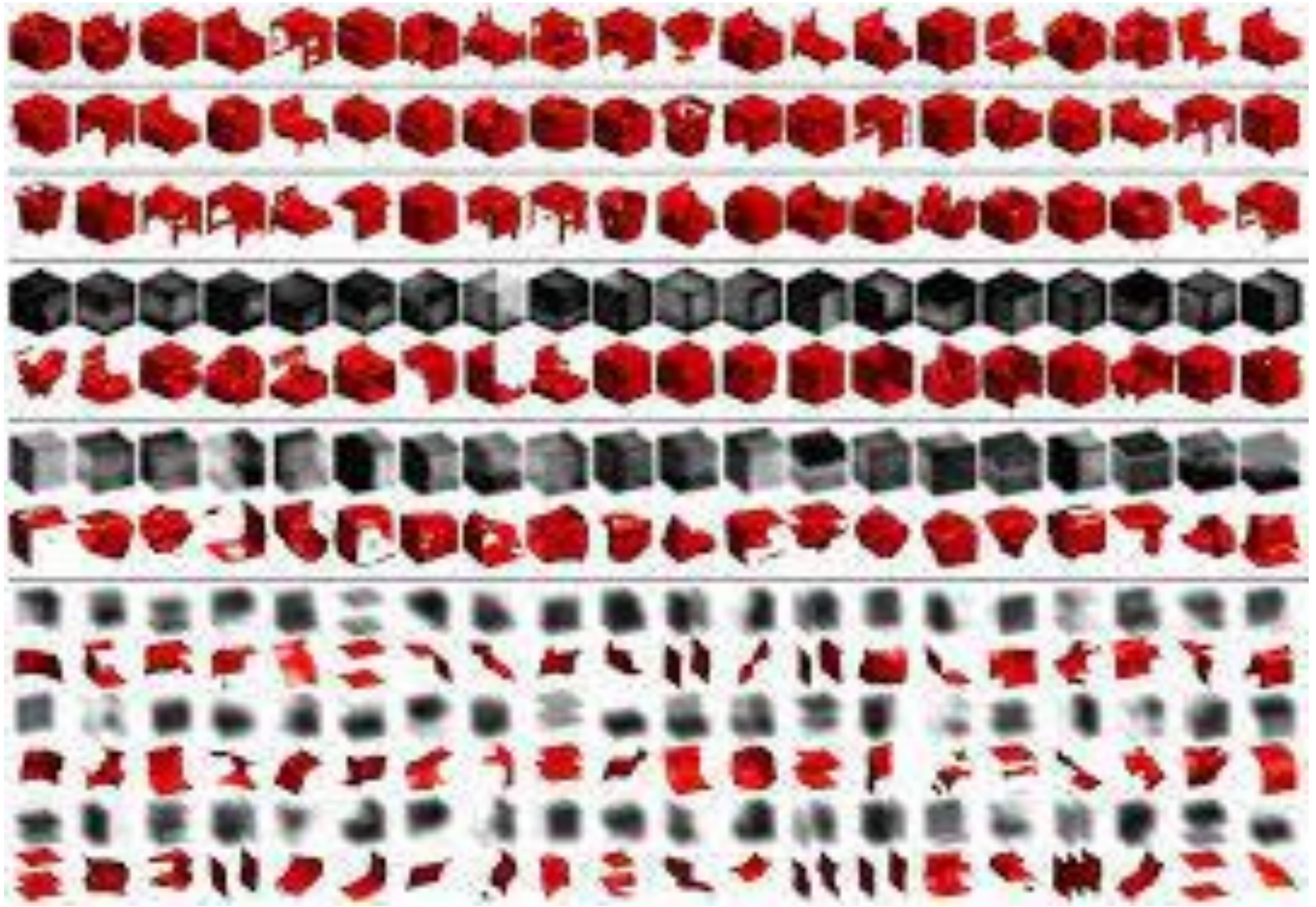
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape datasets: *Datasets for 3D Objects*

- **ModelNet40 and ModelNet40**
  - # Models: 12,311
  - # Categories: 40

- **ModelNet10** (subset of ModelNet10)
  - # Models: 4,899
  - # Categories: 10



Z. Wu et al., 2015

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape datasets: *Datasets for 3D Objects*

- **ShapeNet**
  - # Models: 3M (not publicly available)
- **ShapeNetCore** (subset of ShapeNet)
  - # Models: 51,300
  - # Categories: 55

Change et al., 2015

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

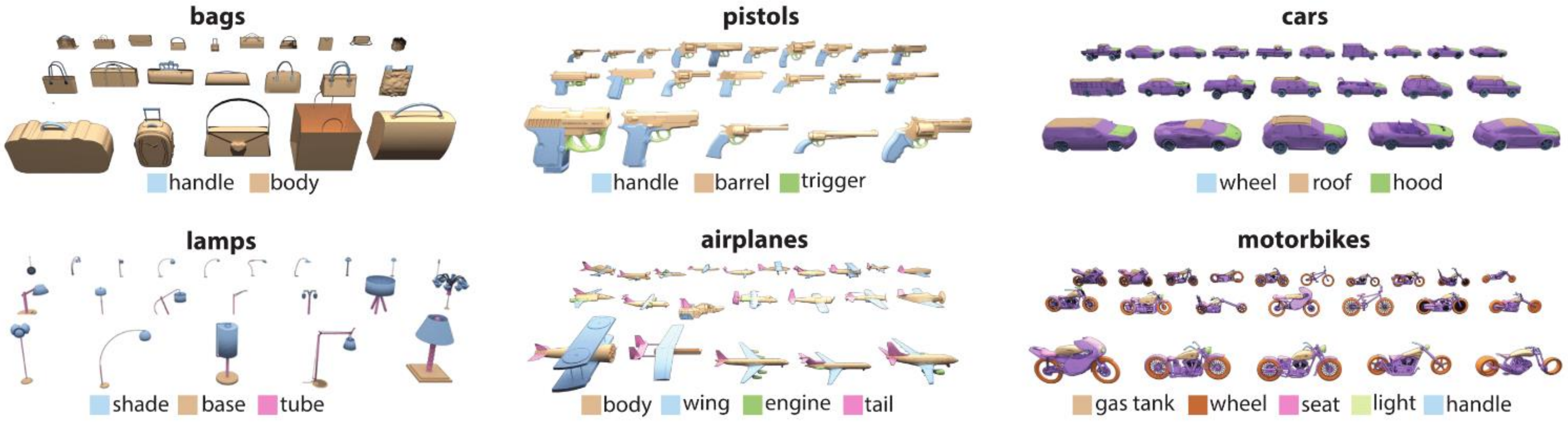
3D shape datasets: *Datasets for 3D Objects*

- **Pix3D**
  - # Images: 10,069
  - # Models: 395 (2D-3D aligned)

Sun et al., 2018

3D shape datasets: *Datasets for 3D Objects Parts*

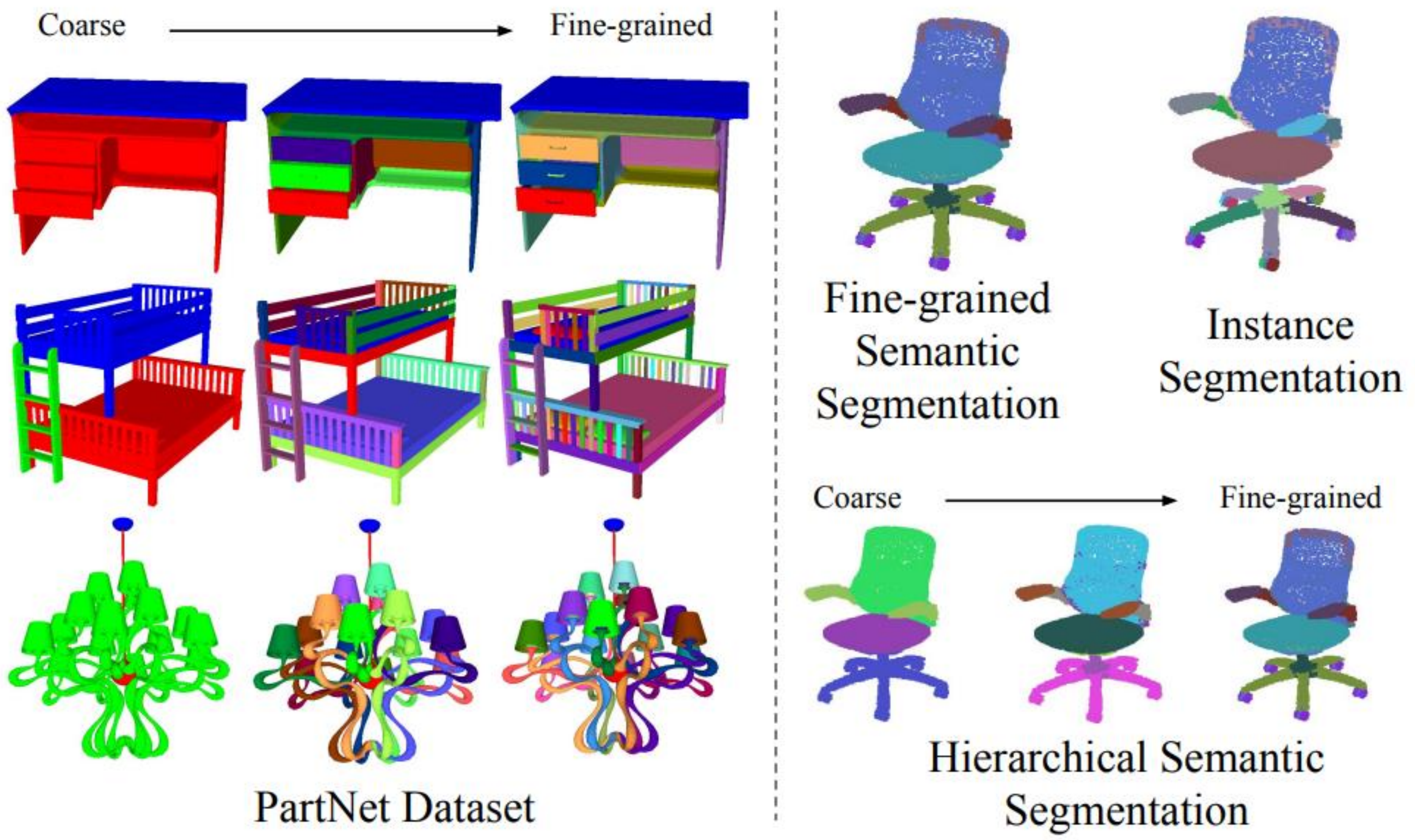
- **ShapeNet-Part** (subset of ShapeNet)
  - # Models: 16,881
  - # Categories: 16
  - # Semantic parts: 50



Yi et al., 2016

3D shape datasets: *Datasets for 3D Objects Parts*

- **PartNet** (subset of ShapeNet)
    - # Models: 26,671
    - # Categories: 24
    - # Part instances: 573,585
    - # Semantic parts: 480
    - Fine-grained
    - Hierarchical



Mo et al., 2019

3D shape datasets: *Datasets for Indoor 3D Scenes*

- **Large-scale Scanned Real Scenes: ScanNet**
  - # Views: 2.5M
  - # RGBD scans: 1,500
  - 3D camera poses
  - Surface reconstruction
  - Instance-level semantic segmentations

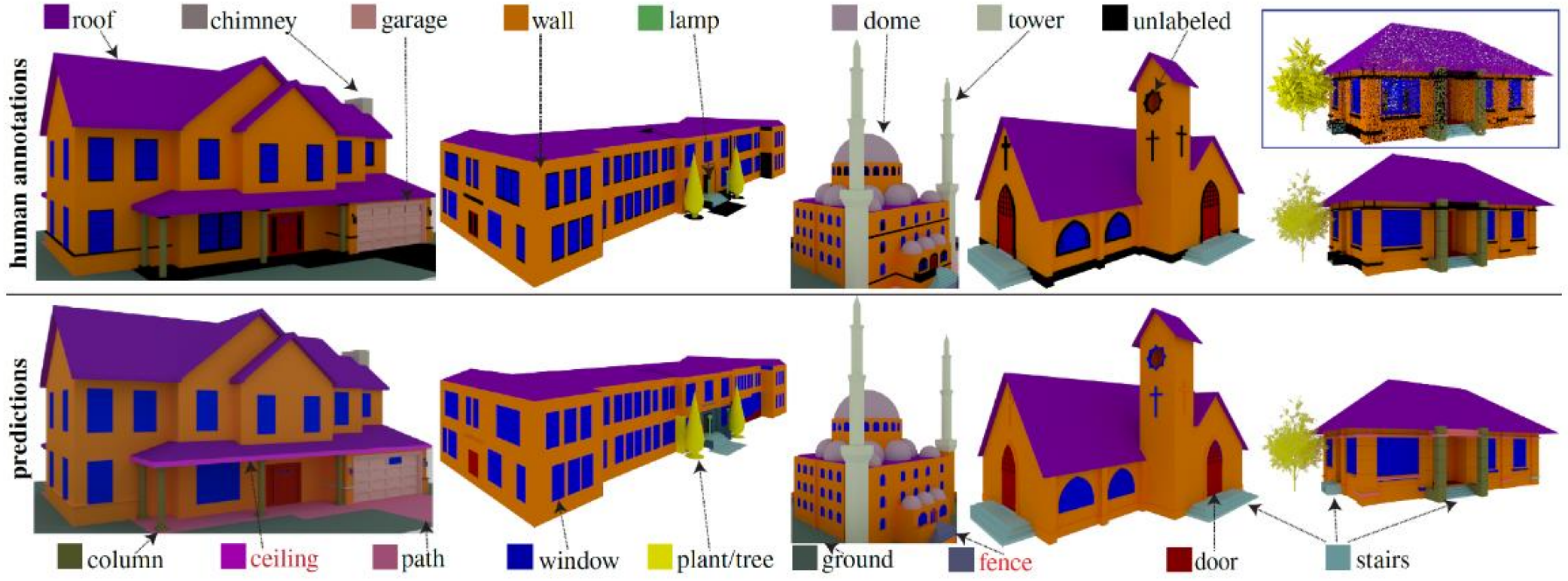
Dai et al., 2017

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D shape datasets: *Datasets for 3D Buildings*

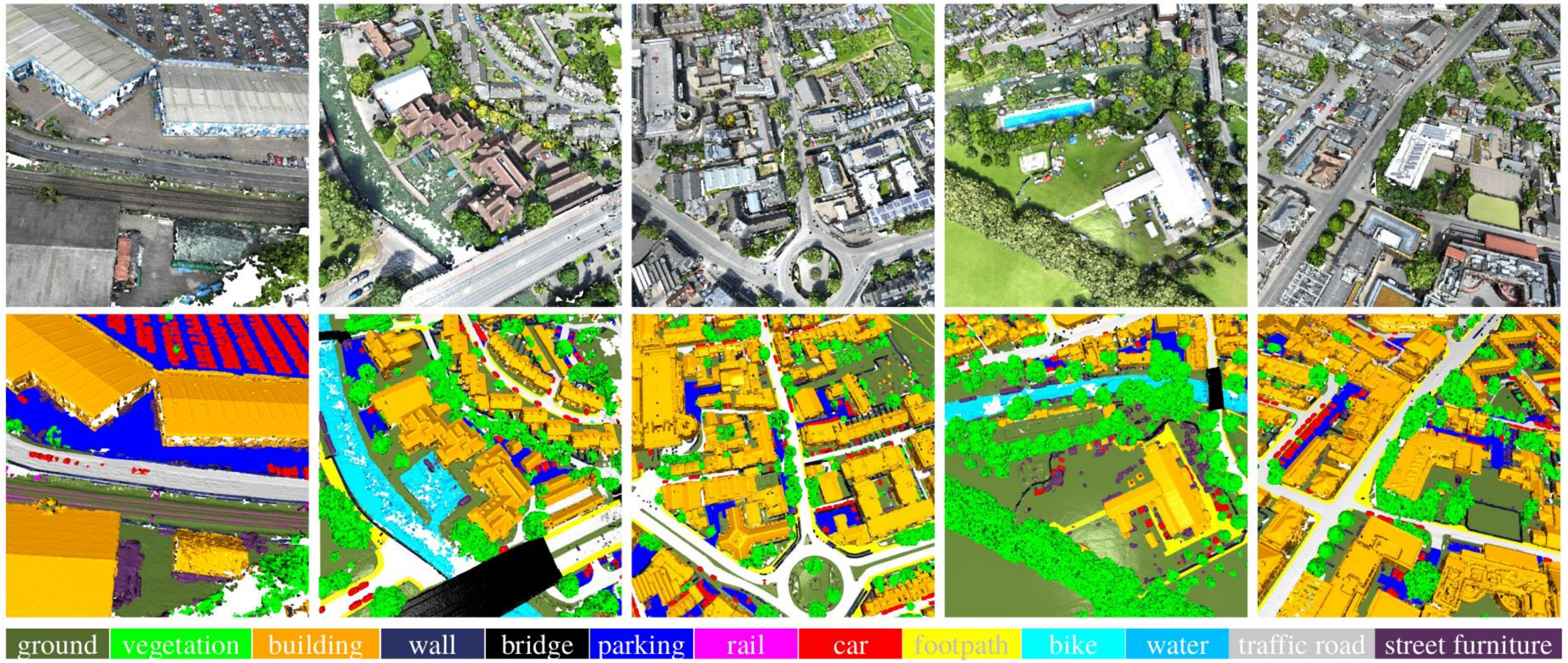
- **BuildingNet**
  - # Models: 2,000
  - # Semantic Components: 292K
  - # Semantic Parts: 31
  - Semantic segmentation
  - Surface reconstruction



Selvaraju et al., 2021

3D shape datasets: *Datasets for Urban Areas*

- **SensatUrban**
  - # Points: 3B
  - # Semantic Classes: 13



Qingyong et al., 2022

Figure 3: Examples of our SensatUrban dataset. Different semantic classes are labeled by different colors.

Today’s Agenda

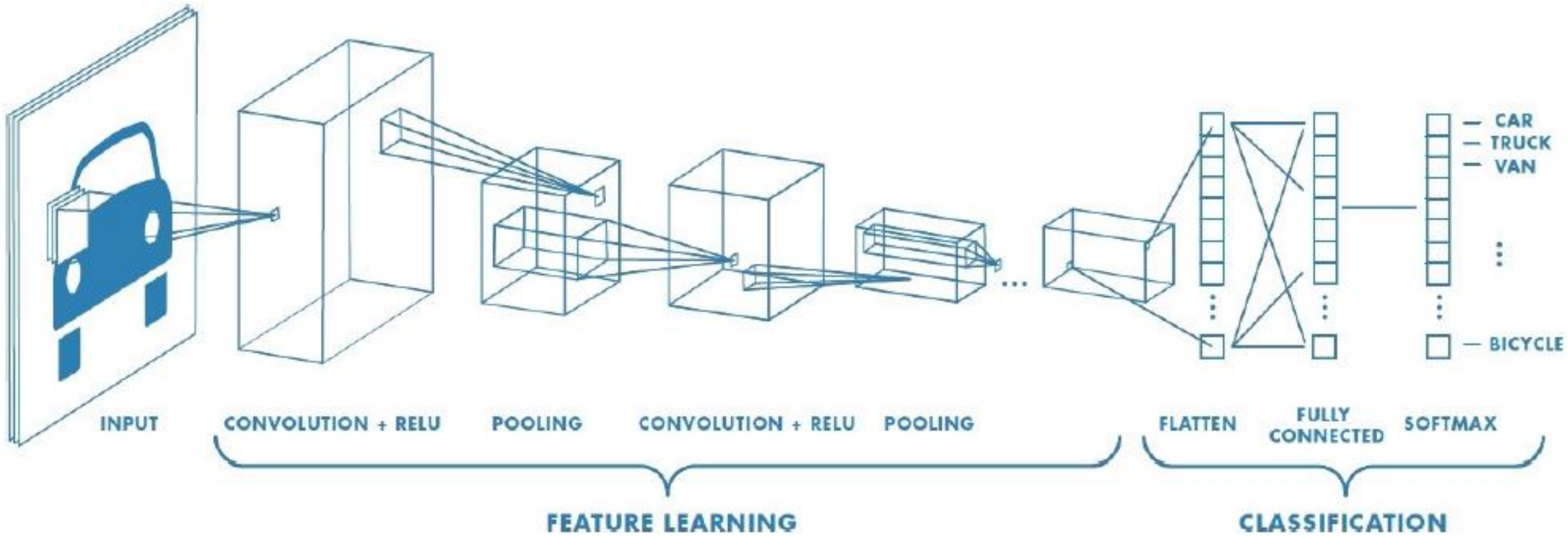
- Who are we?
- What is 3D Vision
- 3D shape representations
- 3D shape datasets
- 3D Deep Learning architectures
- What we do

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *2D architectures “success story”*

Layers of **convolutional filters** trained to extract descriptors + **learned functions** that map descriptors to high-level concepts



Kalogerakis E.

3D DL architectures: *2D architectures "success story"*

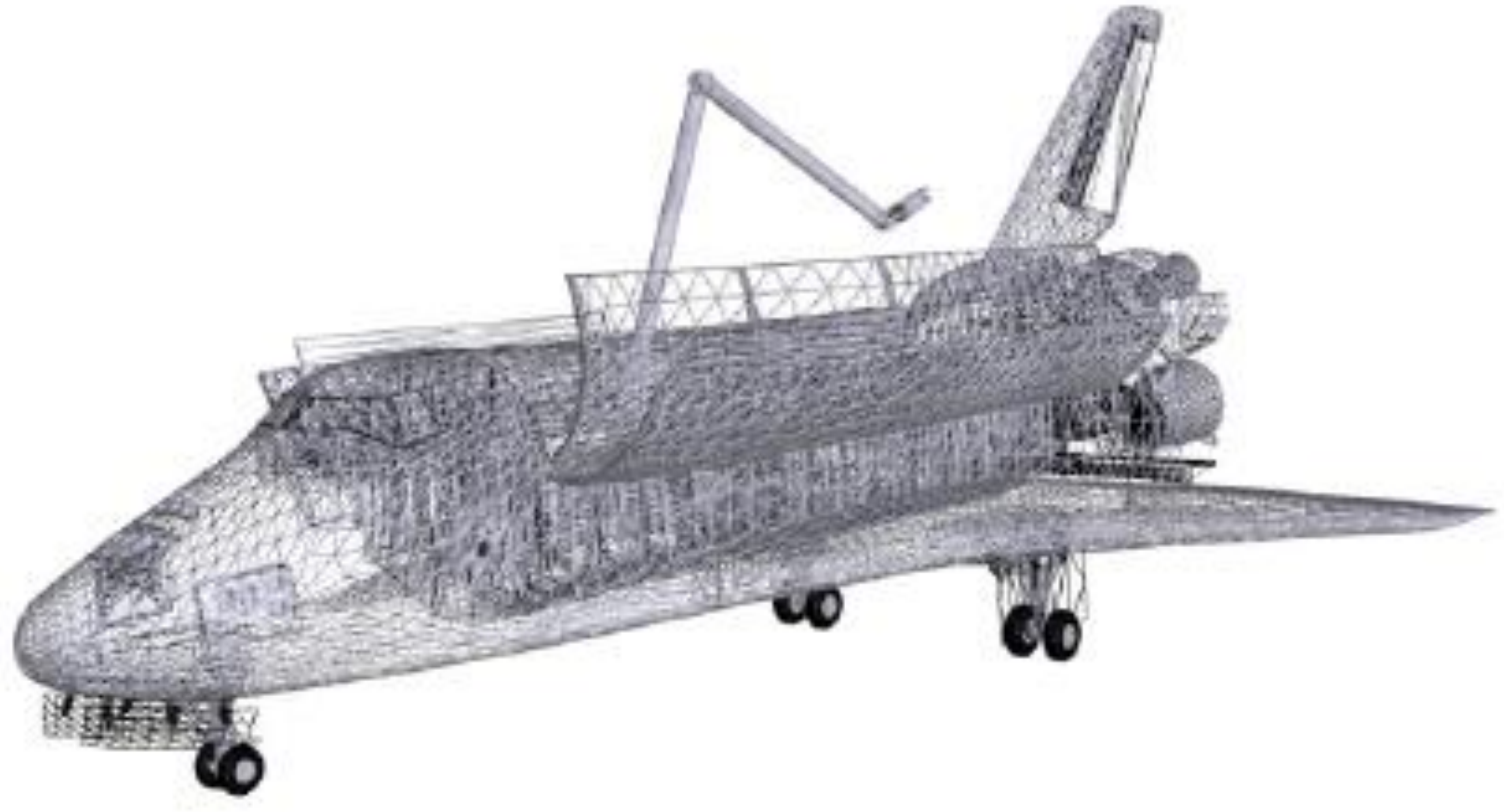
Can capture various **low-level** and **high-level** features through hierarchical representation learning. **Very good performance** in 2D vision tasks (class., seg., obj. det….)

Kalogerakis E.

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Challenges – How do we apply convnets in 3D shapes*

Geometric representations are **irregular** and **unordered**: arbitrary point order, different #points, different #neighbor per point etc.



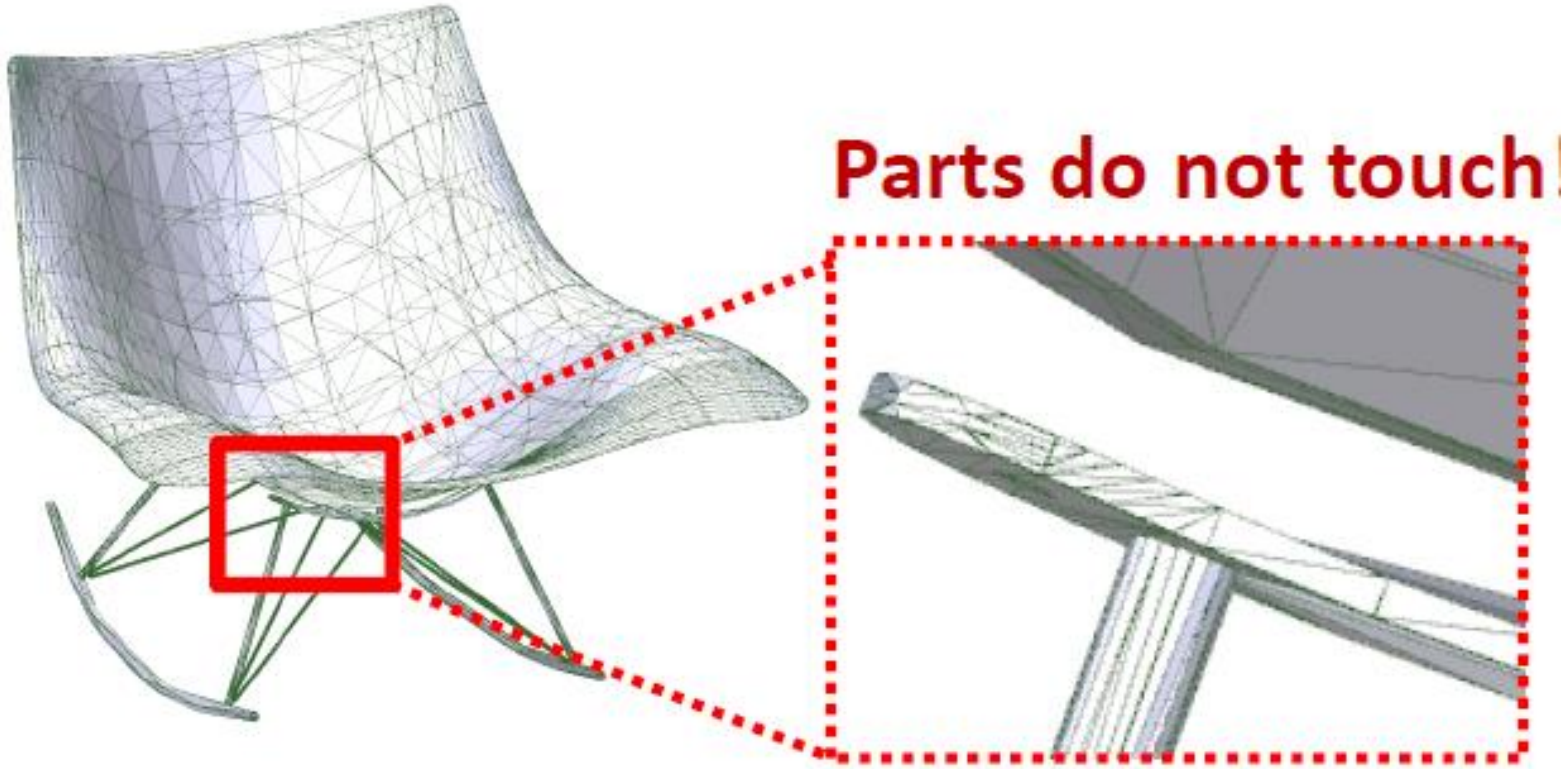
Polygon mesh

Point clouds

Kalogerakis E.

3D DL architectures: *Challenges – Artifacts*

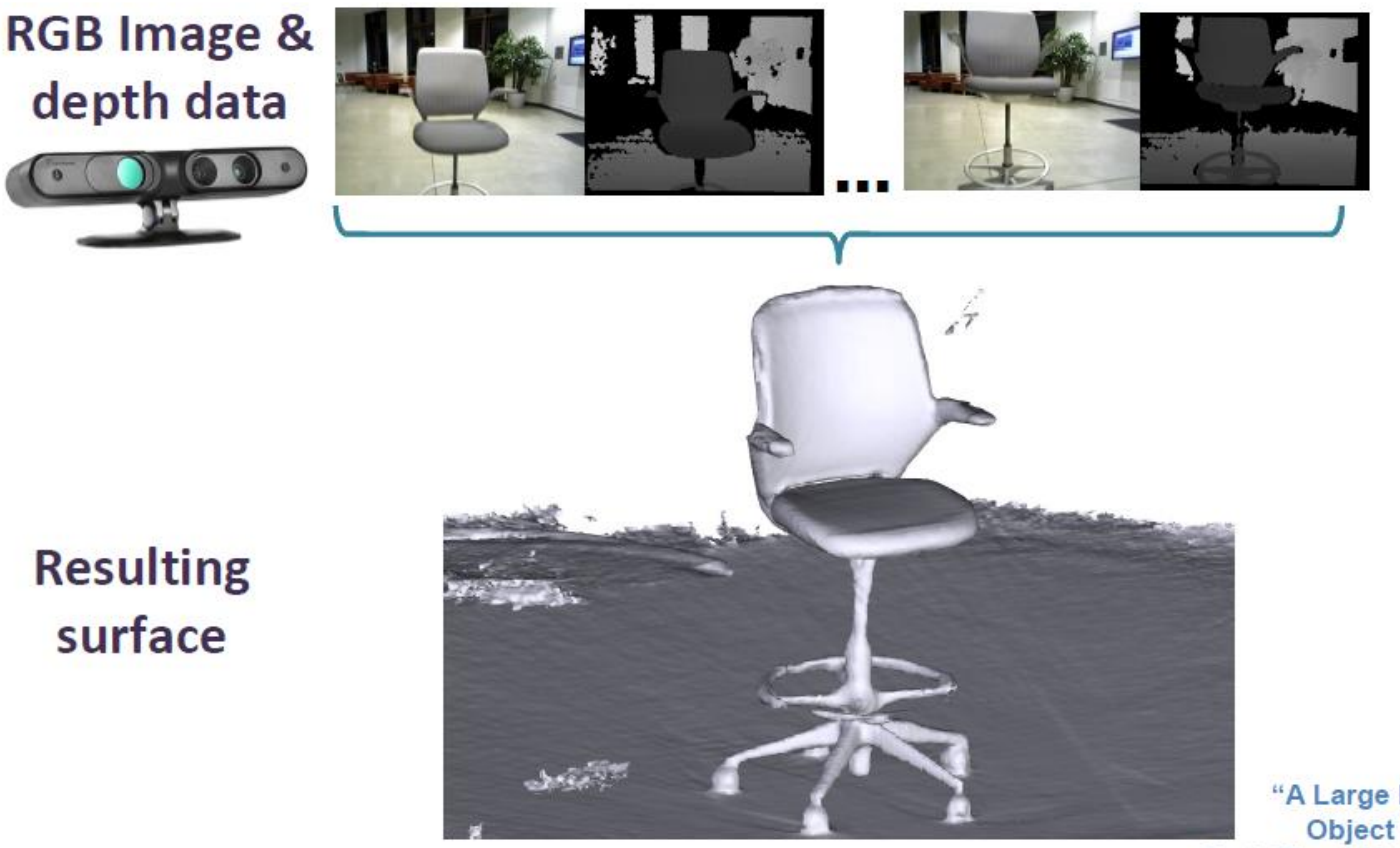
3D models can have several **artifacts**



Kalogerakis E.

3D DL architectures: *Challenges – Noise*

Scanned surfaces have **noisy** and **missing parts**



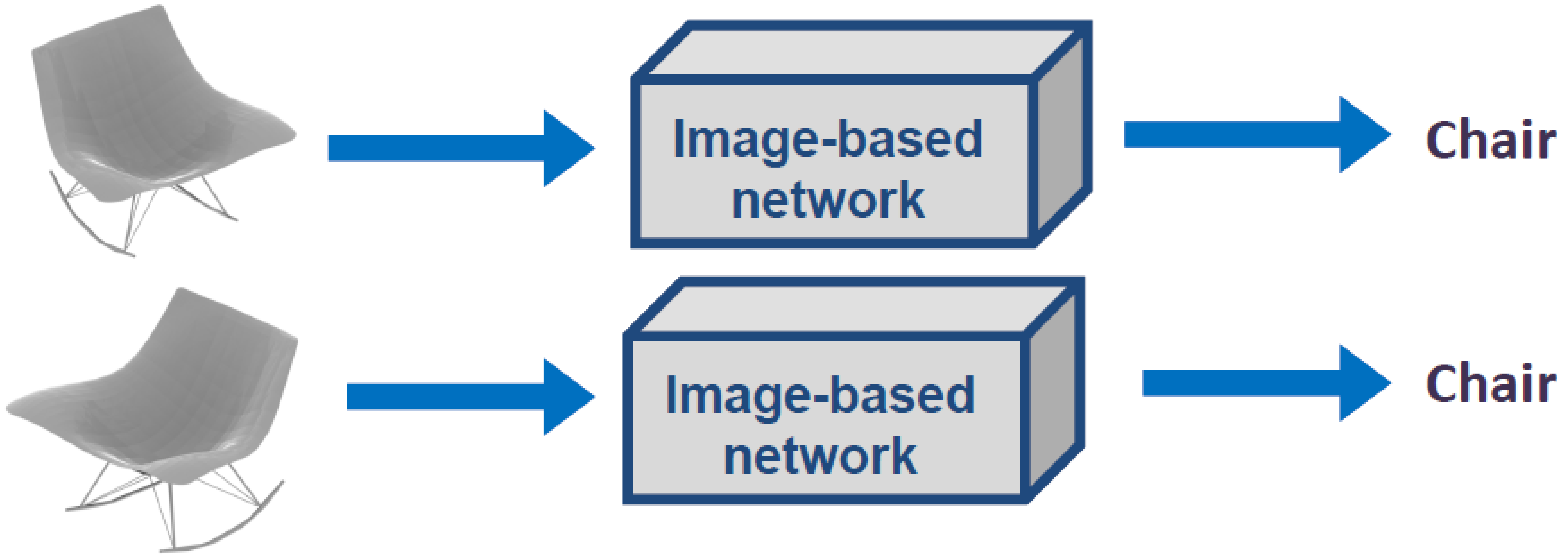
"A Large Dataset of Object Scans"
Choi, Zhou, Miller, Koltun 2016

Kalogerakis E.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Multi-view approach*

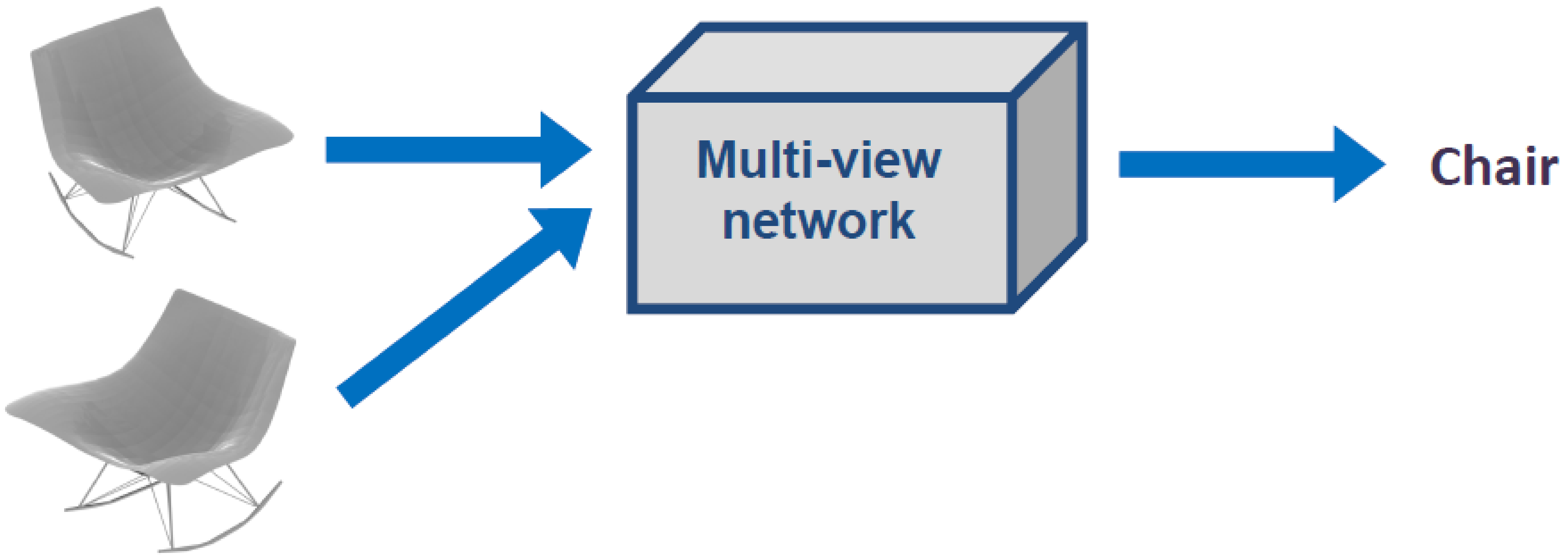
Image-based networks can process individual shape renderings



Kalogerakis E.

3D DL architectures: *Multi-view approach*

Image-based networks can process individual shape renderings



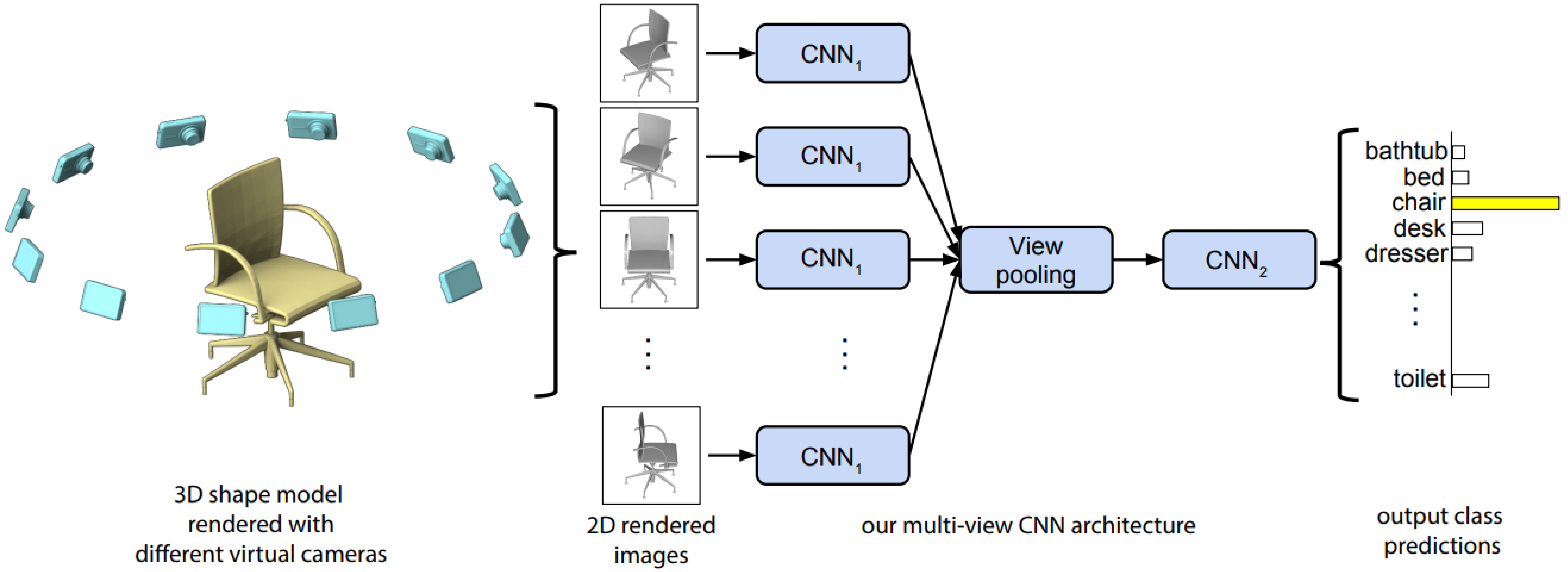
Kalogerakis E.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Multi-view approach*

Multi-view Convolutional Neural Networks for 3D Shape Recognition



Hang Su et al.
ICCV 2015

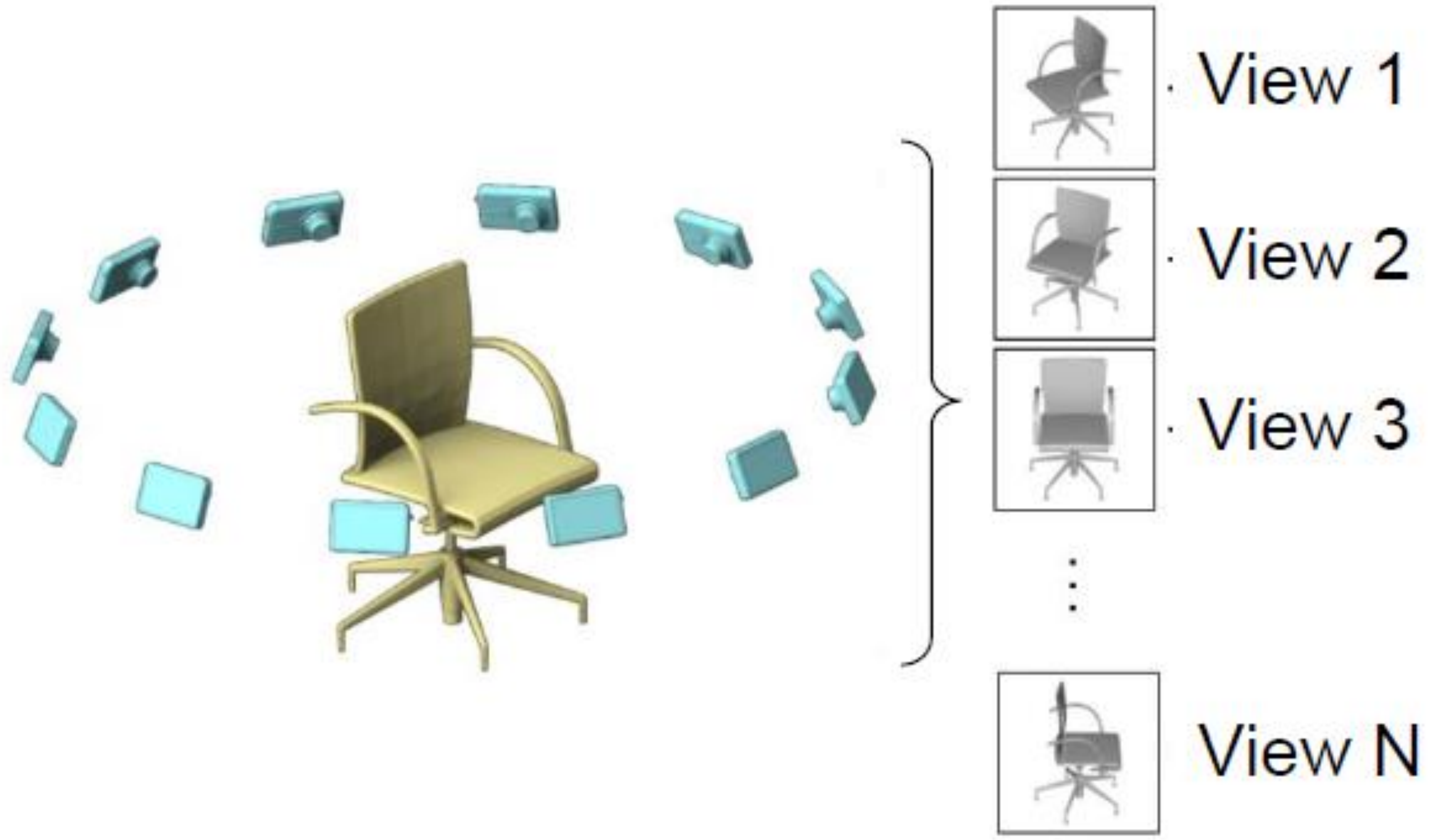
3D DL architectures: *Multi-view approach*

Multi-view Convolutional Neural Networks for 3D Shape Recognition

Hang Su et al.
ICCV 2015

3D DL architectures: *Multi-view approach*

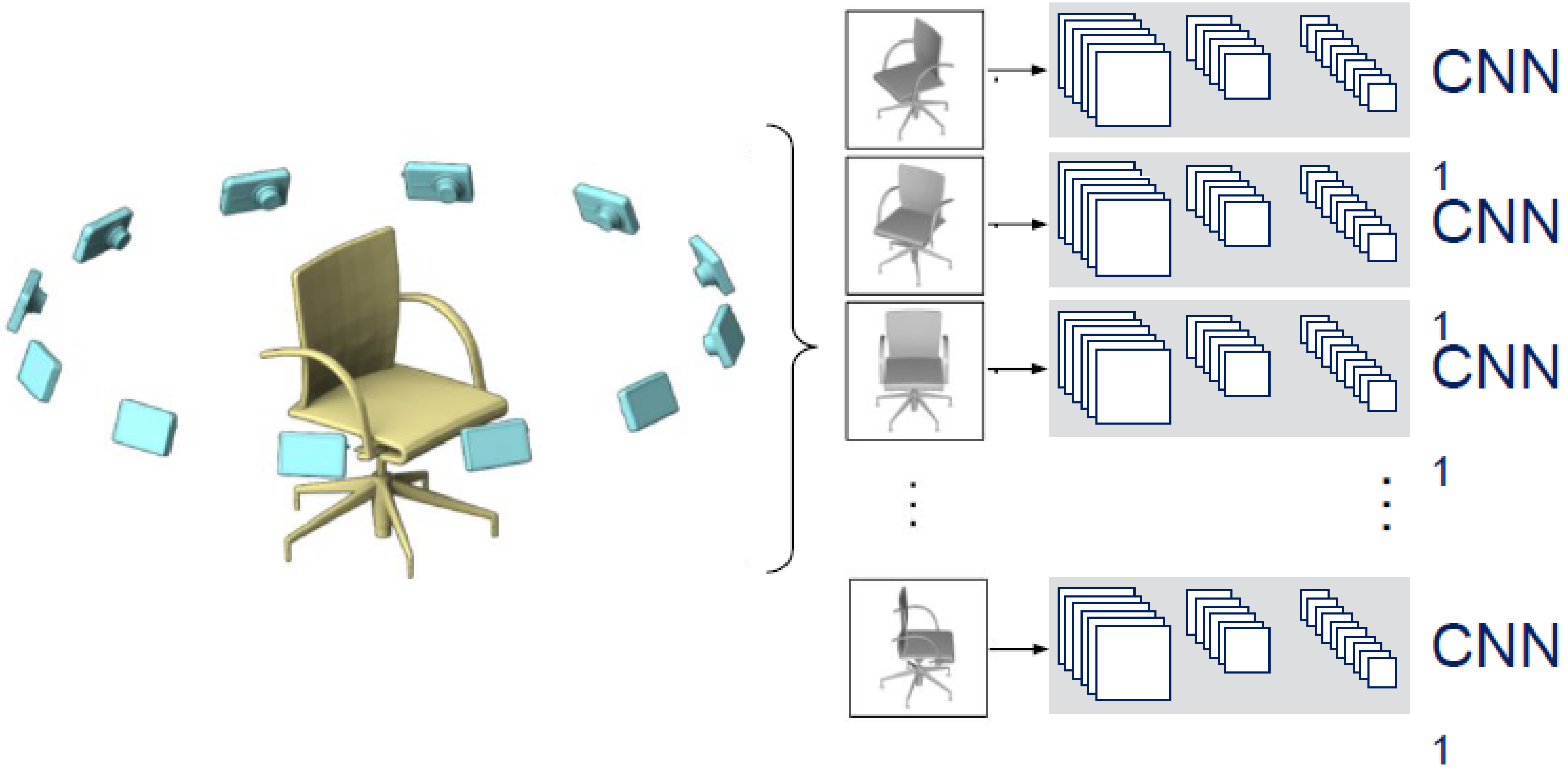
Multi-view Convolutional Neural Networks for 3D Shape Recognition



Hang Su et al.
ICCV 2015

3D DL architectures: *Multi-view approach*

Multi-view Convolutional Neural Networks for 3D Shape Recognition

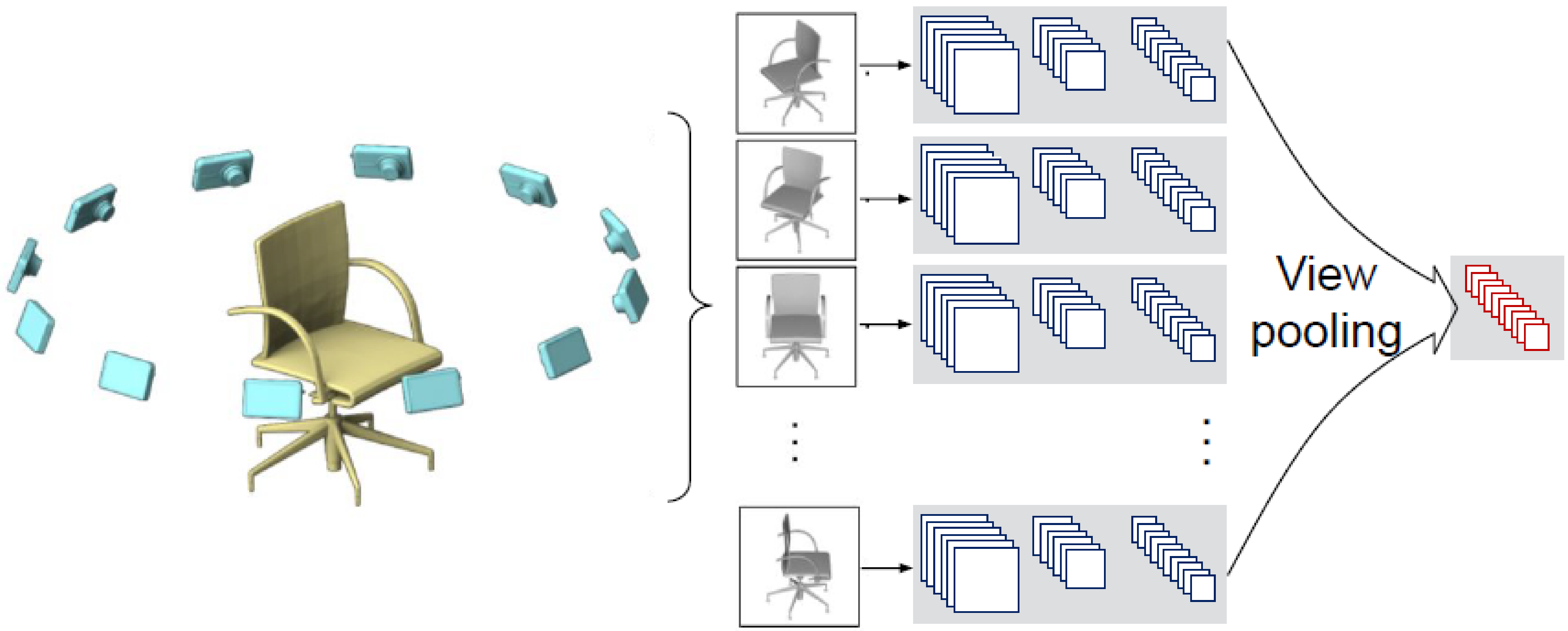


CNN_1: a ConvNet extracting image features

Hang Su et al.
ICCV 2015

3D DL architectures: *Multi-view approach*

Multi-view Convolutional Neural Networks for 3D Shape Recognition

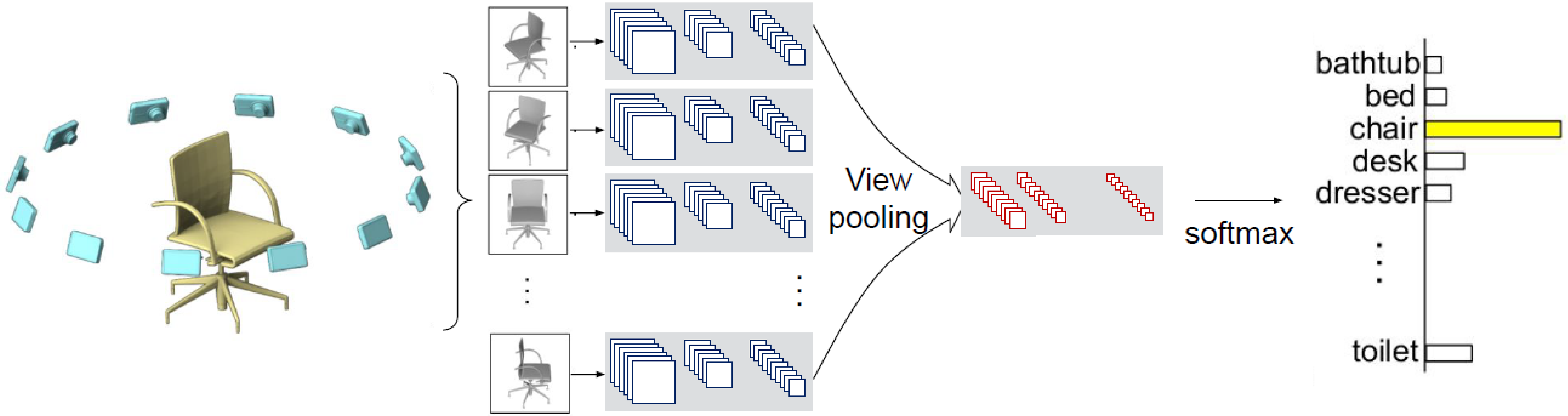


View pooling: element-wise
max-pooling across all views

Hang Su et al.
ICCV 2015

3D DL architectures: *Multi-view approach*

Multi-view Convolutional Neural Networks for 3D Shape Recognition



CNN$_2$: a second ConvNet producing shape descriptors

Hang Su et al. ICCV 2015

3D DL architectures: *Multi-view approach*

Multi-view Convolutional Neural Networks for 3D Shape Recognition

ModelNet40: Classification & Retrieval

Method	Classification (Accuracy)
Spherical Harmonics [Kazhdan et al.]	68.2%
LightField [Chen et al.]	75.5%
Volumetric Net [Wu et al.]	77.3%
ImageNet-trained CNN (VGG-M, 1 view)	83.0%
Multi-view convnet (MVCNN)	**90.1%**

Hang Su et al.
ICCV 2015

3D DL architectures: *Multi-view approach*

Multi-view Networks

- Pros:
  - ✓ Good performance
  - ✓ Can leverage vast literature of image classification
  - ✓ Can use pretrained features
- Cons
  - x Need projection
  - x Issue with noisy and/or incomplete input, e.g., point cloud

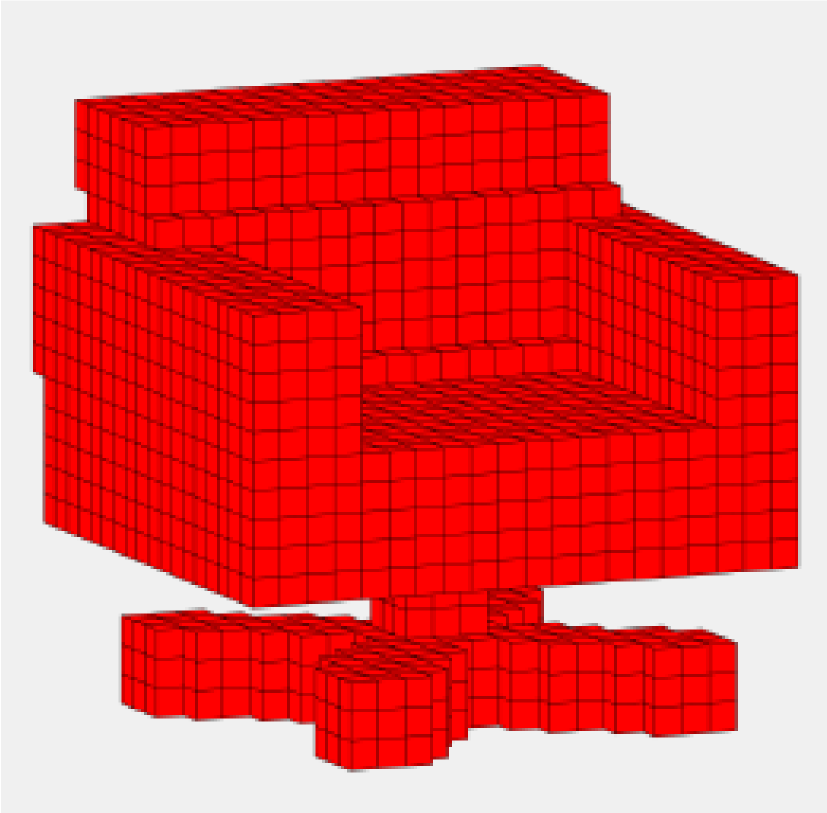
Jiajun Wu

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Volumetric approach*

Voxelization: Convert shape to 3D regular volumetric grid

3D polygon mesh



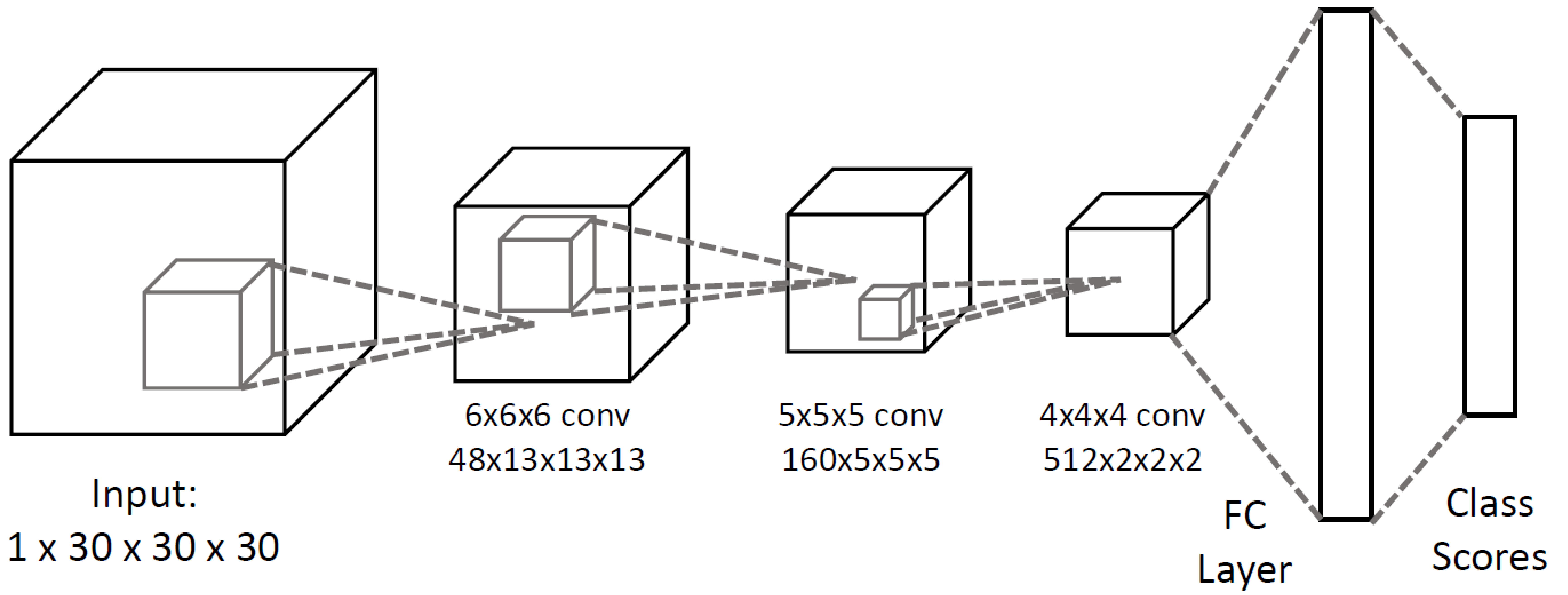
Voxels

Kalogerakis E.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Volumetric approach*

Processing Voxel Inputs –> **3D Convolution**

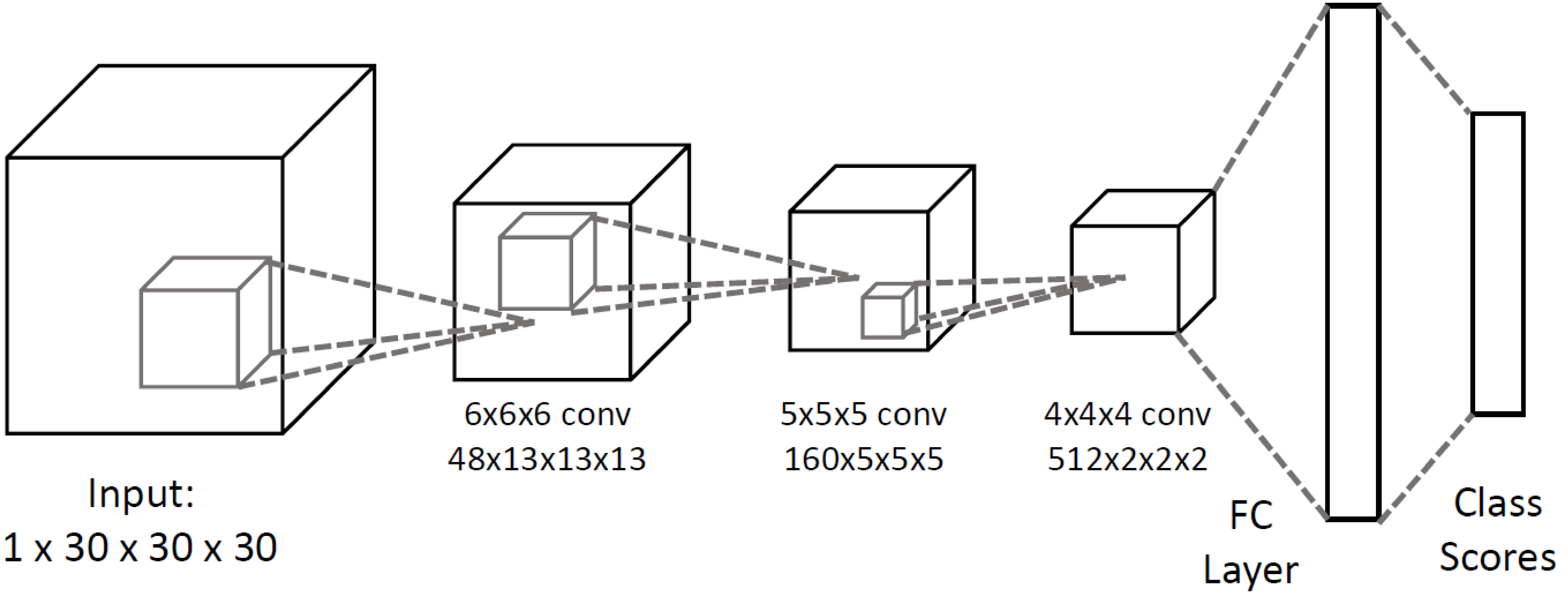


"3D ShapeNet",
Wu et al., CVPR 2015

3D DL architectures: *Volumetric approach*

Processing Voxel Inputs –> **3D Convolution**

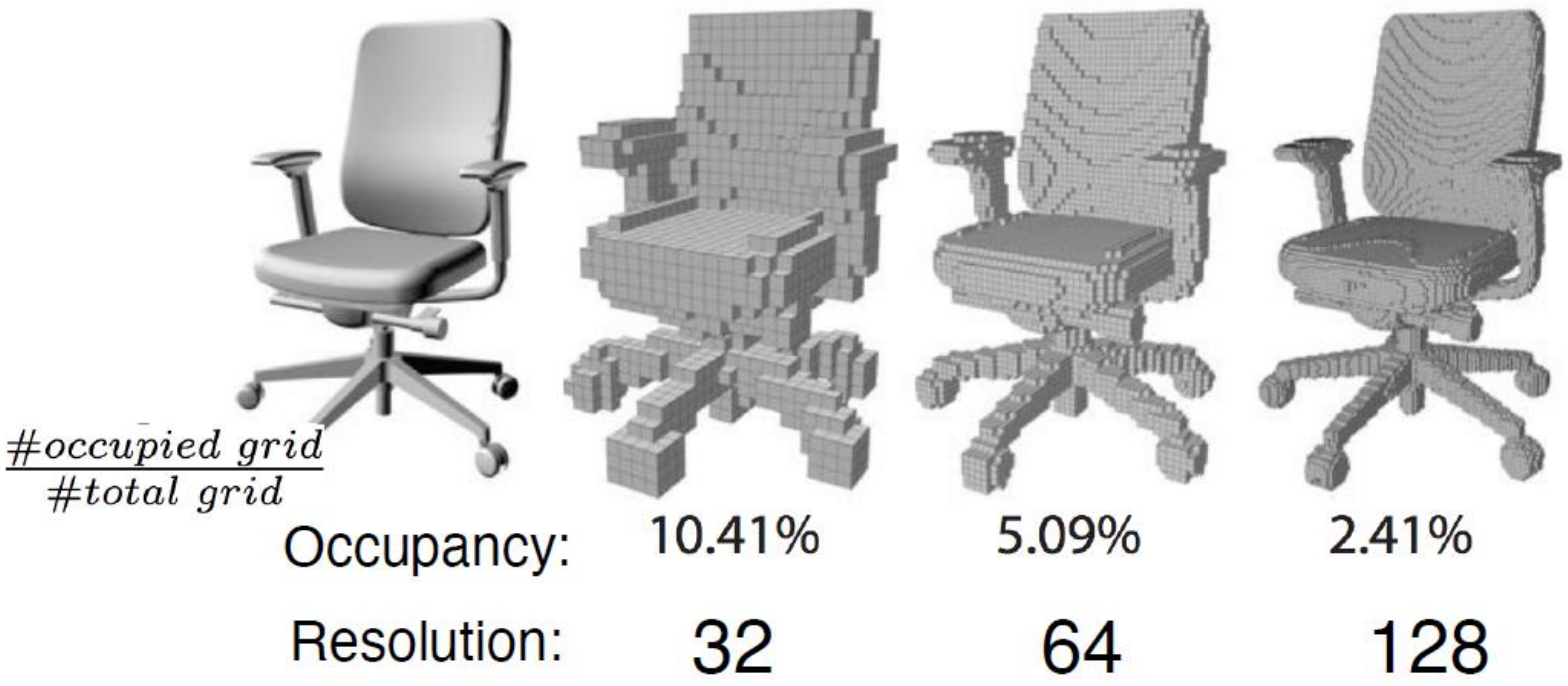
Computationally and memory expensive! Requires low-res input



"3D ShapeNet", Wu et al., CVPR 2015

3D DL architectures: *Volumetric approach*

Sparsity of 3D data

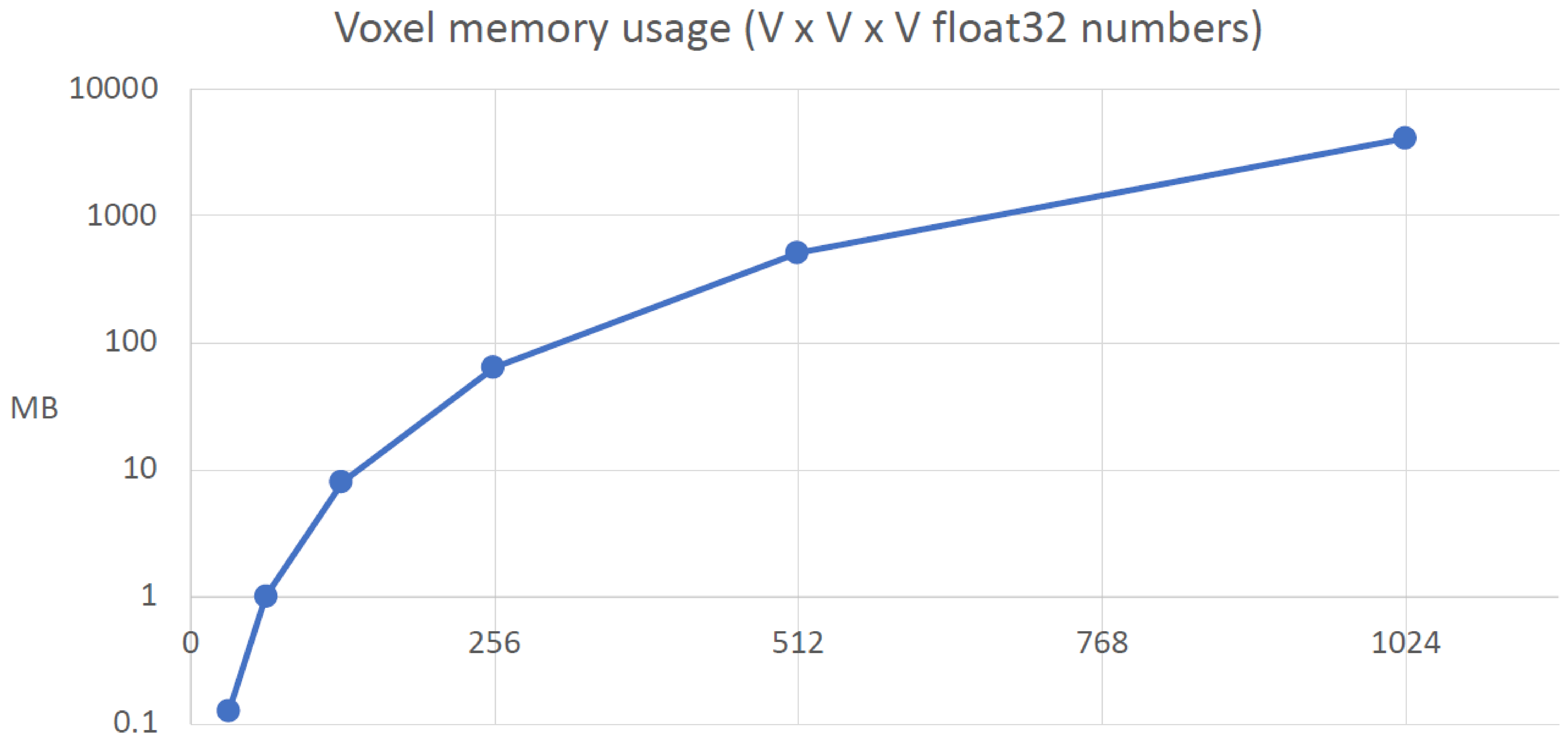


Running convolution on so much empty space is wasteful!

Hao Su et al.

3D DL architectures: *Volumetric approach*

Memory usage



Storing 1024^3 voxel grid takes 4GB of memory!

Justin Solomon

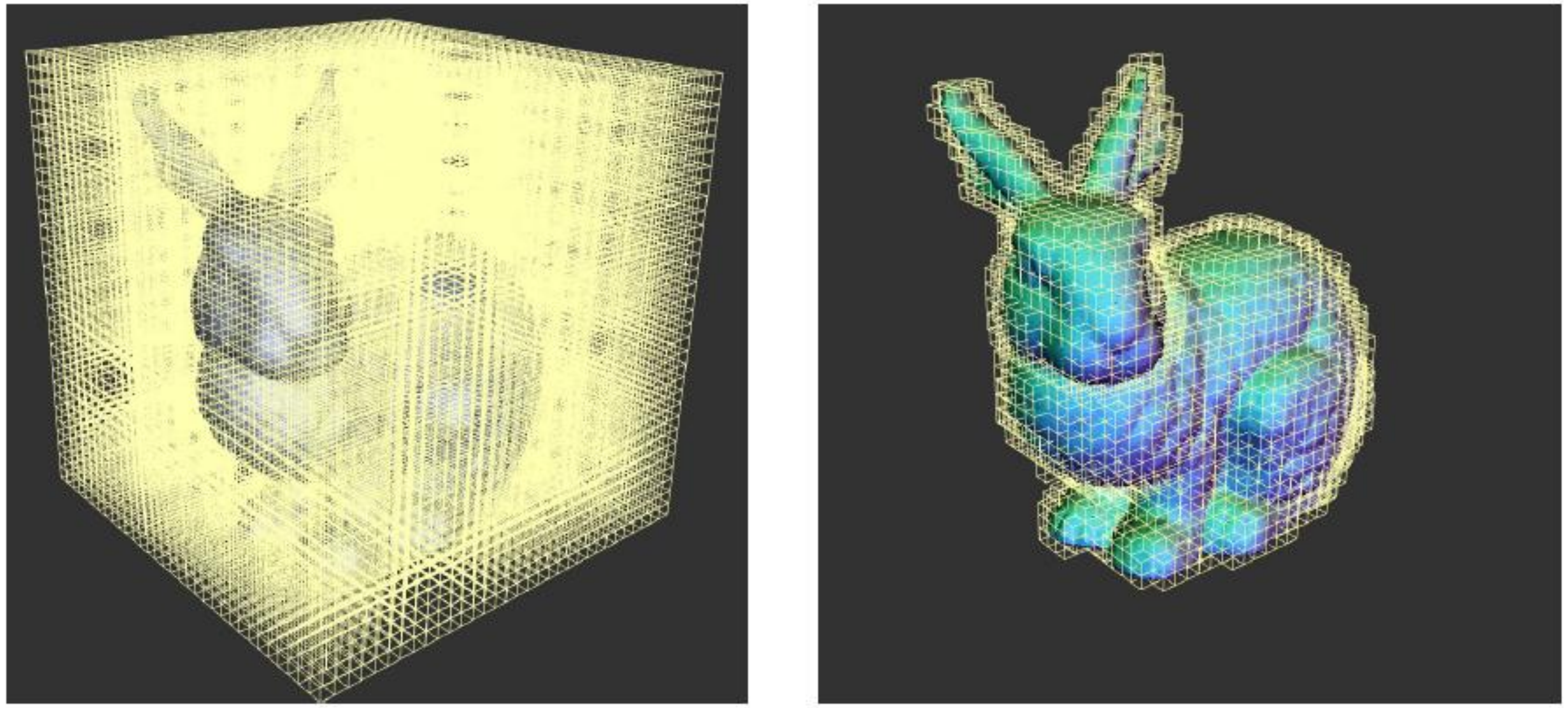
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Volumetric approach*

Solution –> **Octave Tree Representations**

- Store the sparse surface signals
- Constrain the computation near the surface

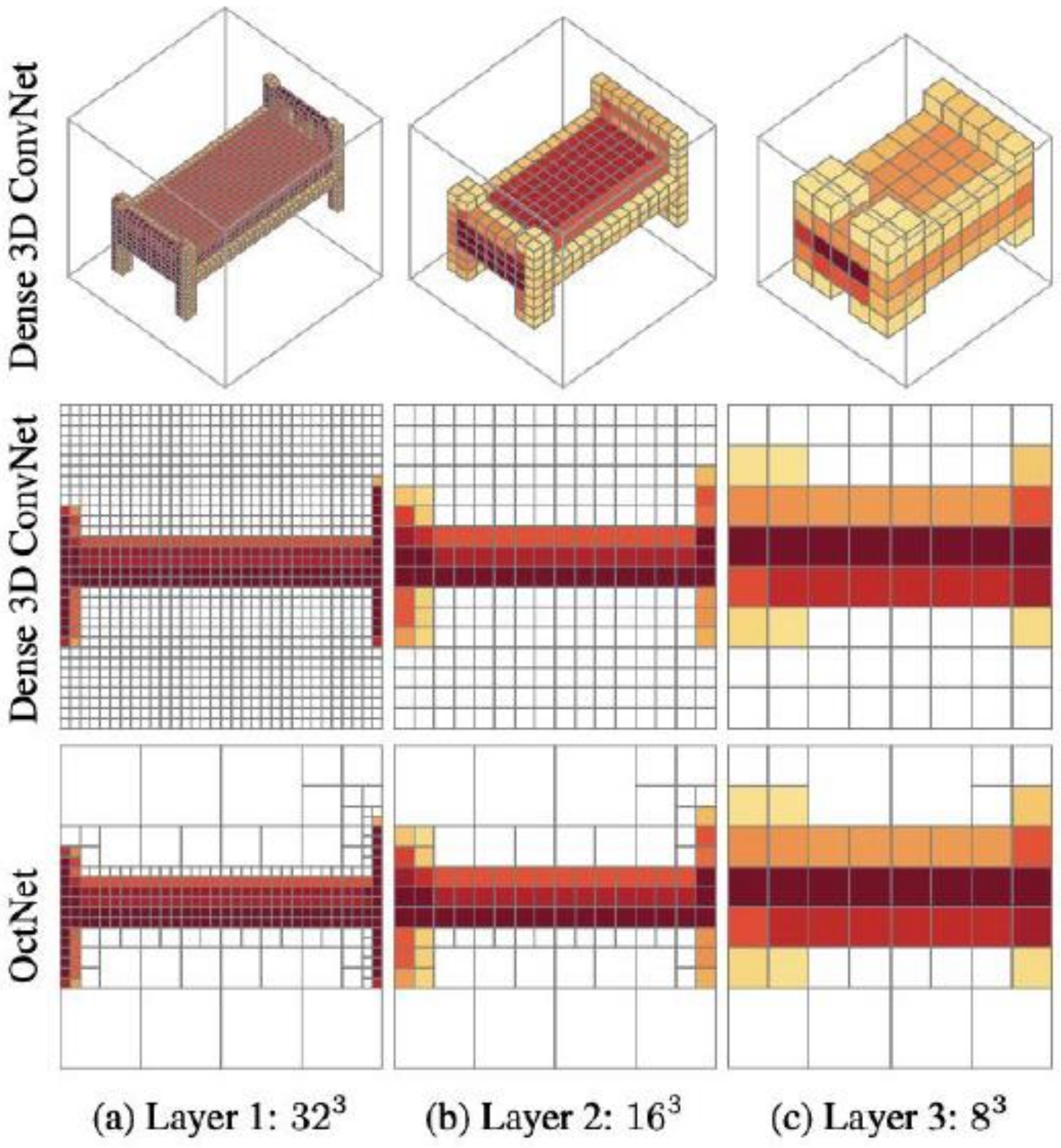


Hao Su et al.

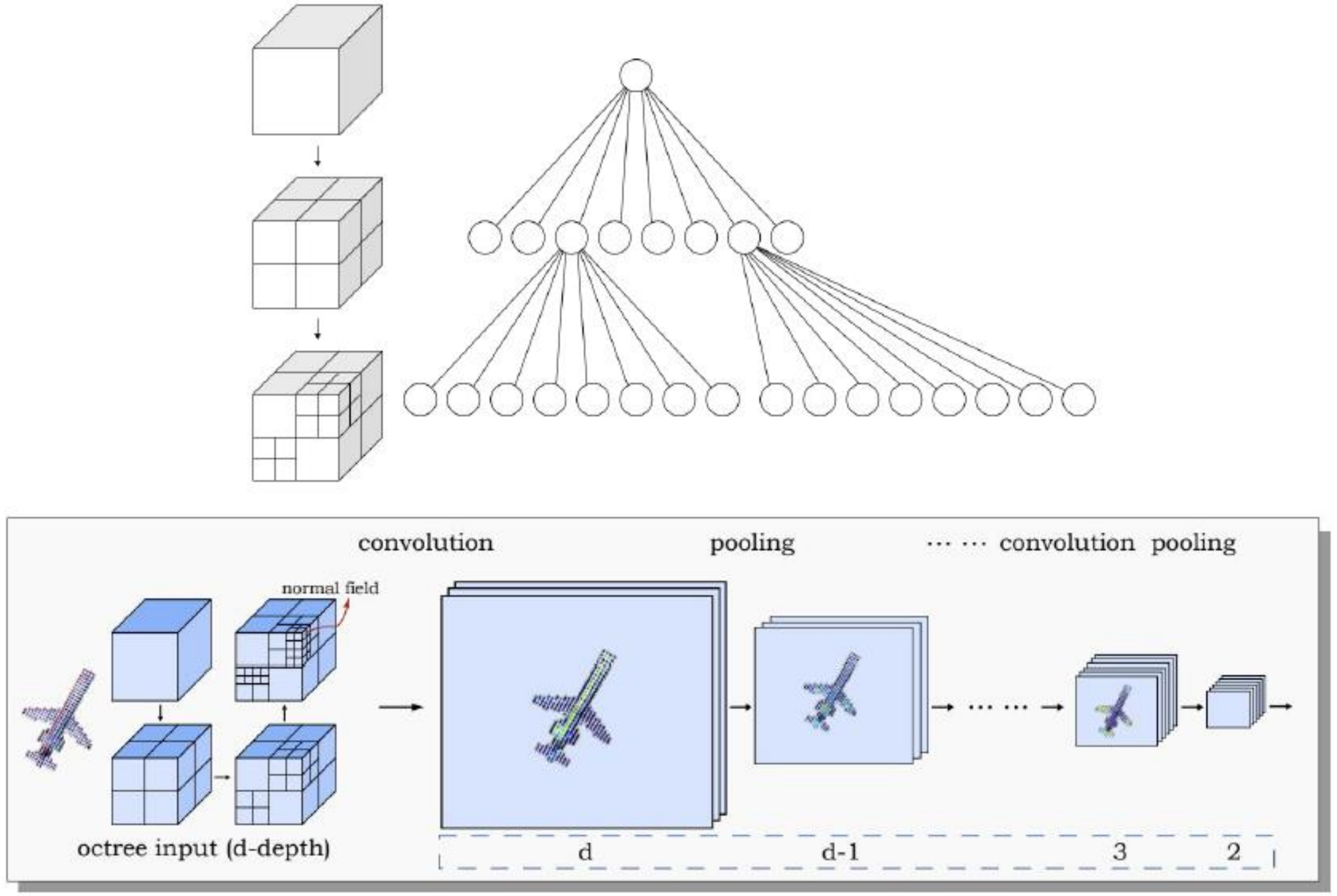
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Volumetric approach*

Octree: Recursively Partition the Space



Riegler et al. OctNet. CVPR 2017

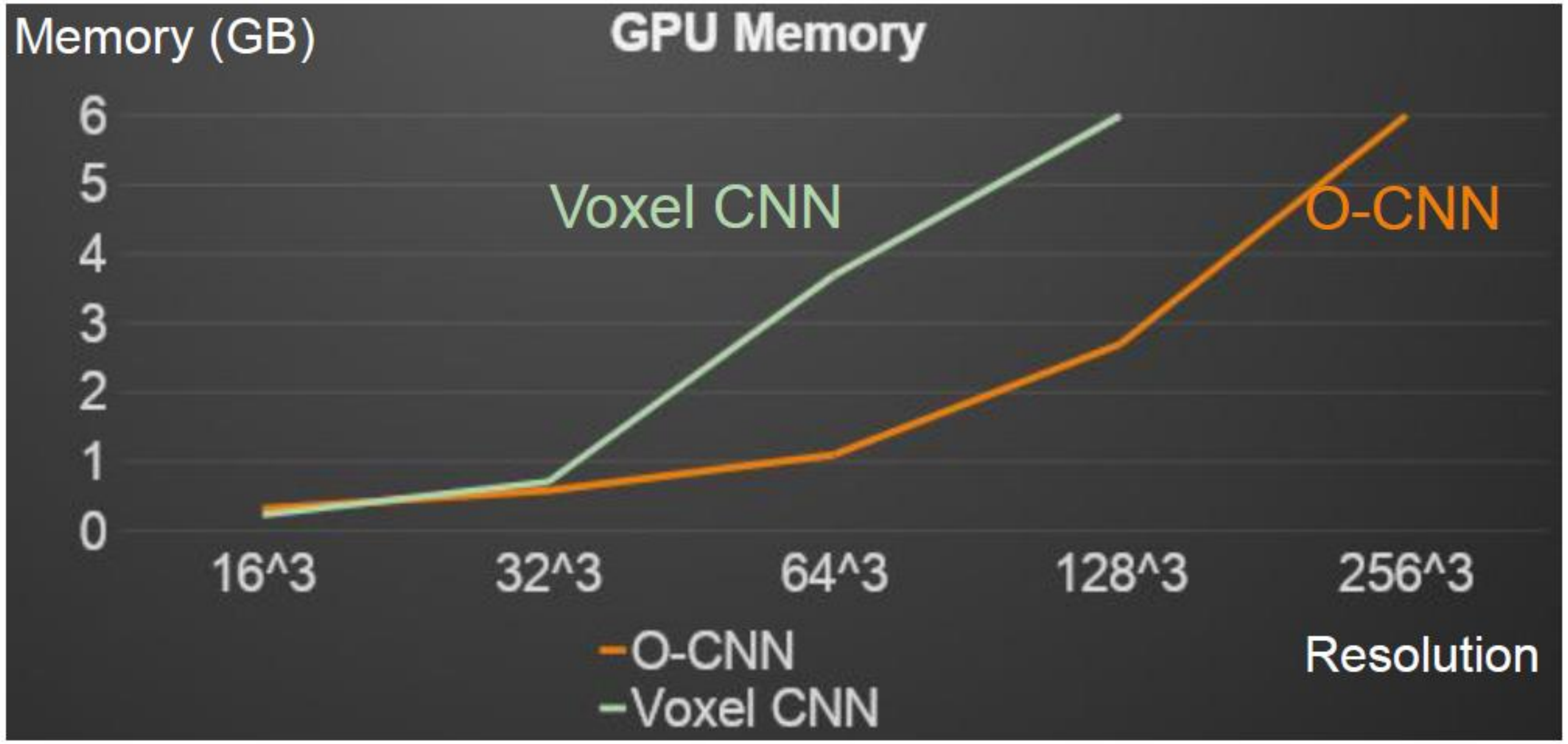


Wang et al. O-CNN. SIGGRAPH 2017

Hao Su et al.

3D DL architectures: *Volumetric approach*

Memory Efficiency

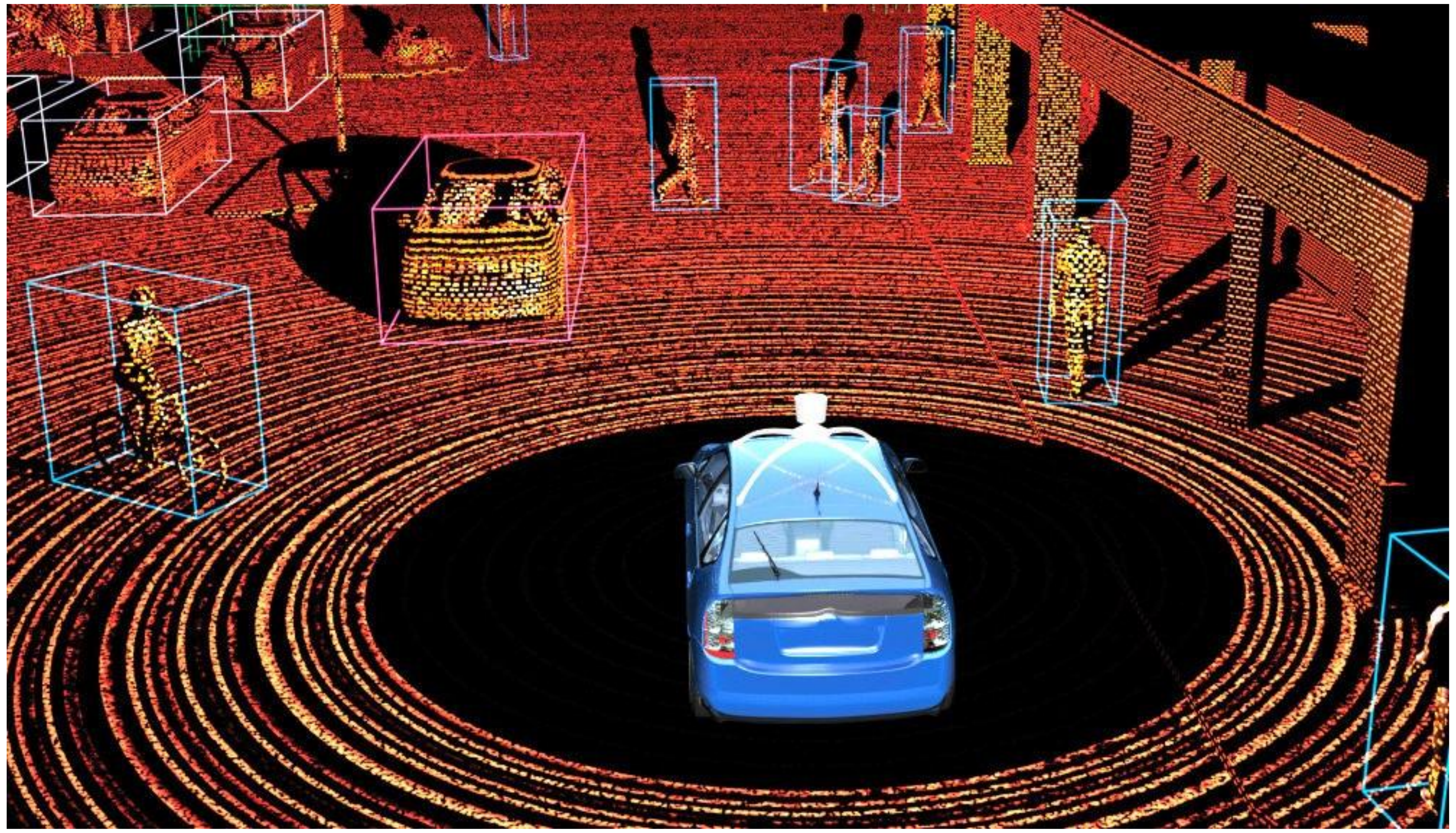


Hao Su et al.

3D DL architectures: *Point-based approach*

Motivation:

- Lots of **scanned data** are raw 3D point clouds
- **Process raw input**, i.e., point cloud, without any preprocessing



Kalogerakis E.

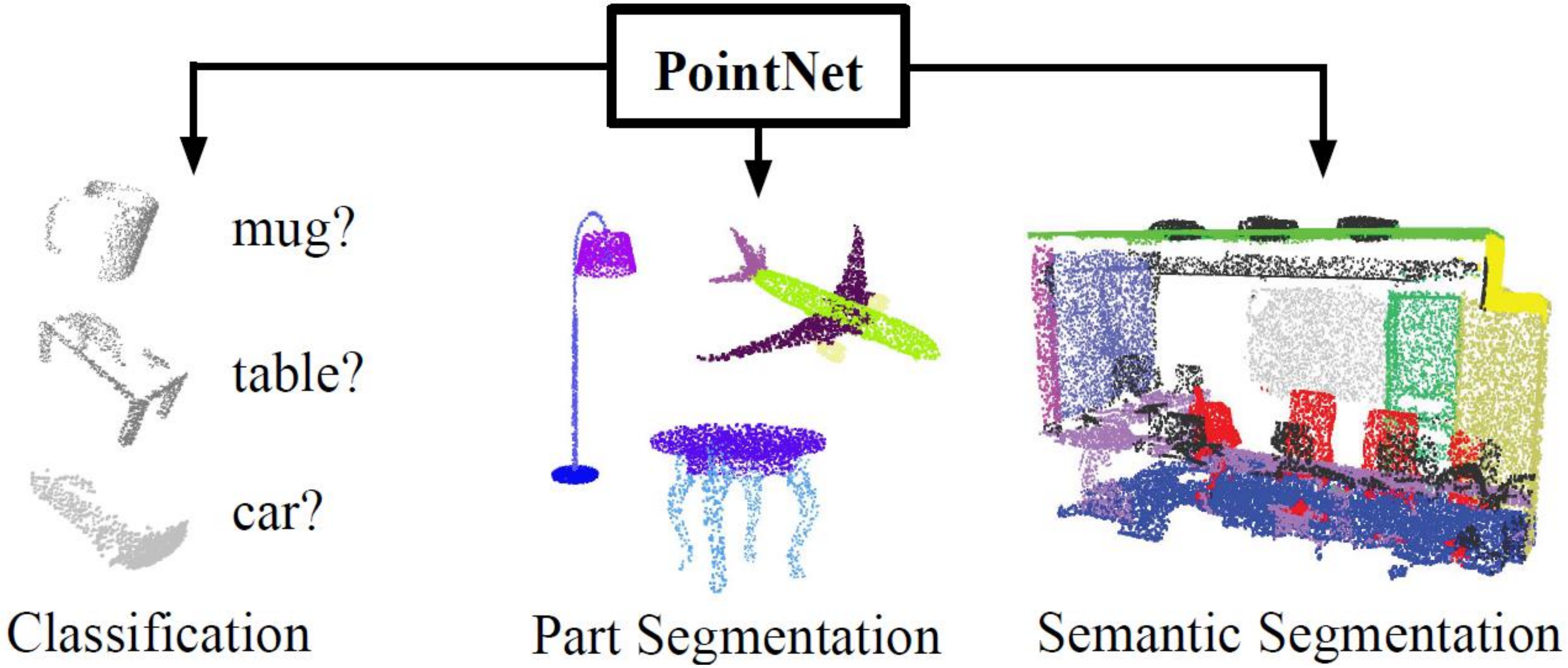
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Point-based approach*

PointNet: (Qi et al., CVPR 2017)

- Processes input point clouds for various tasks

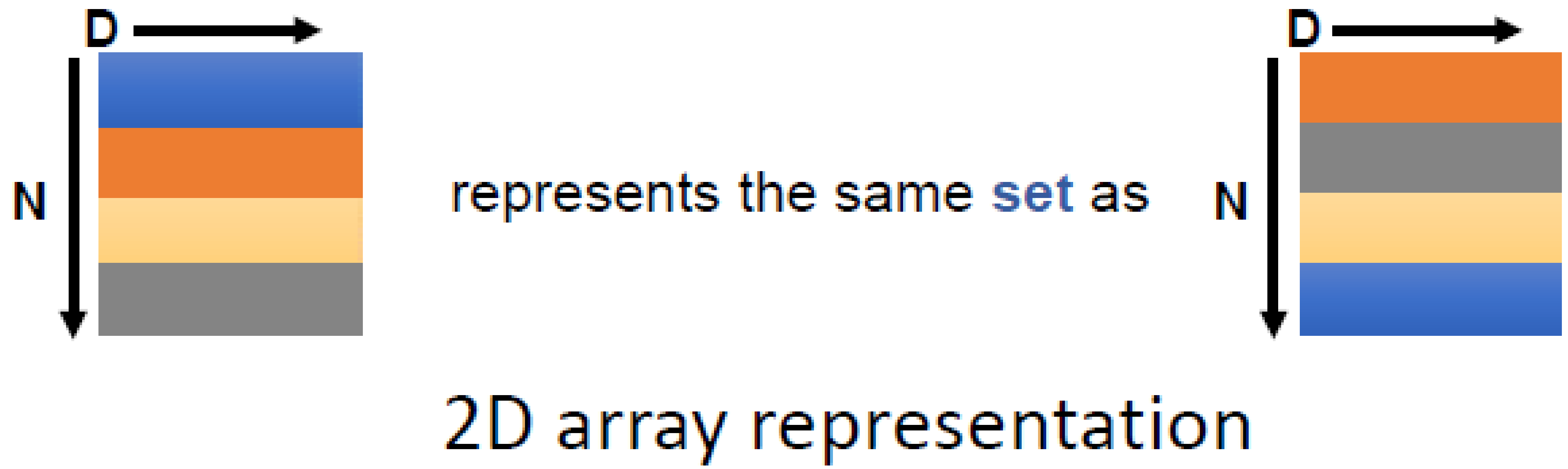


Hao Su

3D DL architectures: *Point-based approach*

Desired Properties of PointNet:

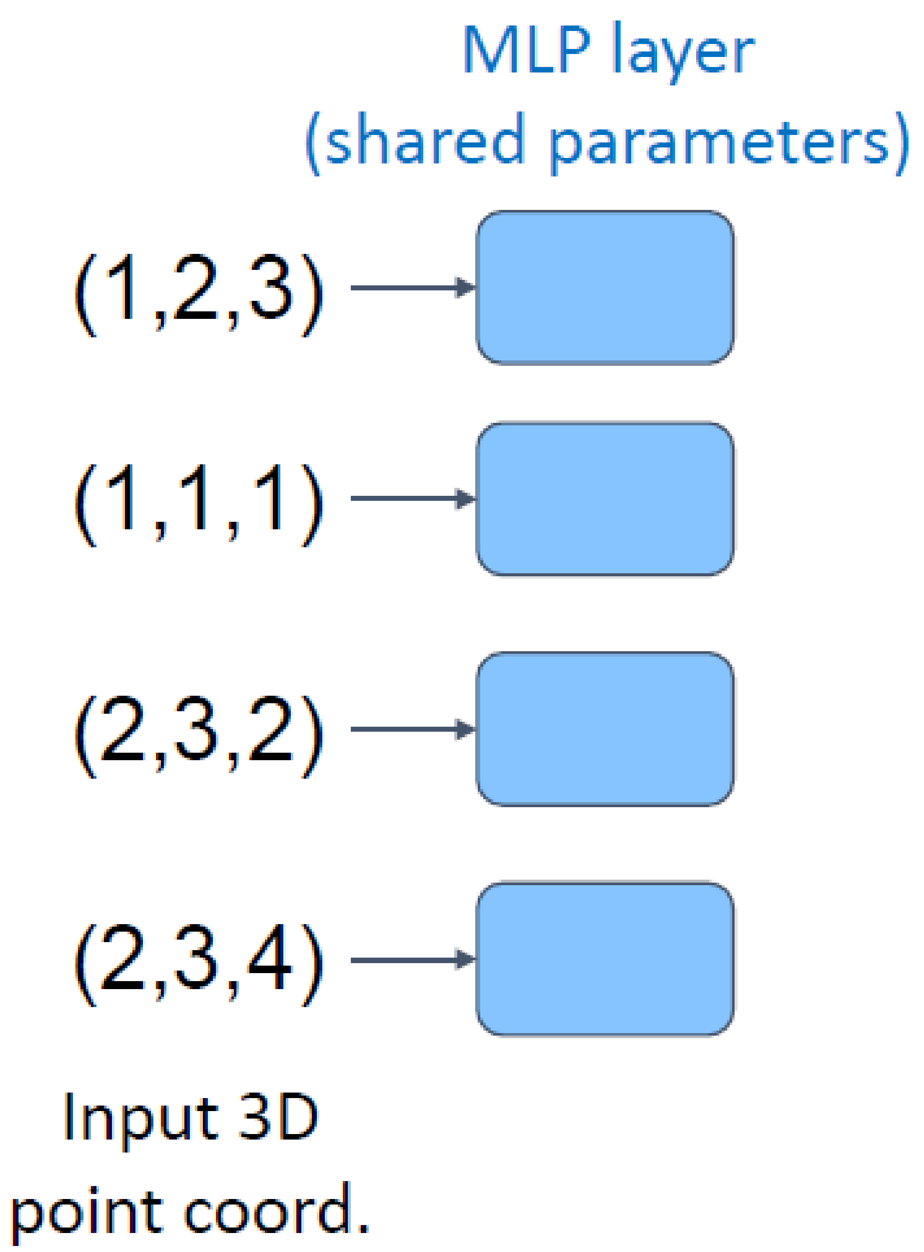
- **Permutation invariance**



Kalogerakis E.

3D DL architectures: *Point-based approach*

PointNet architecture:

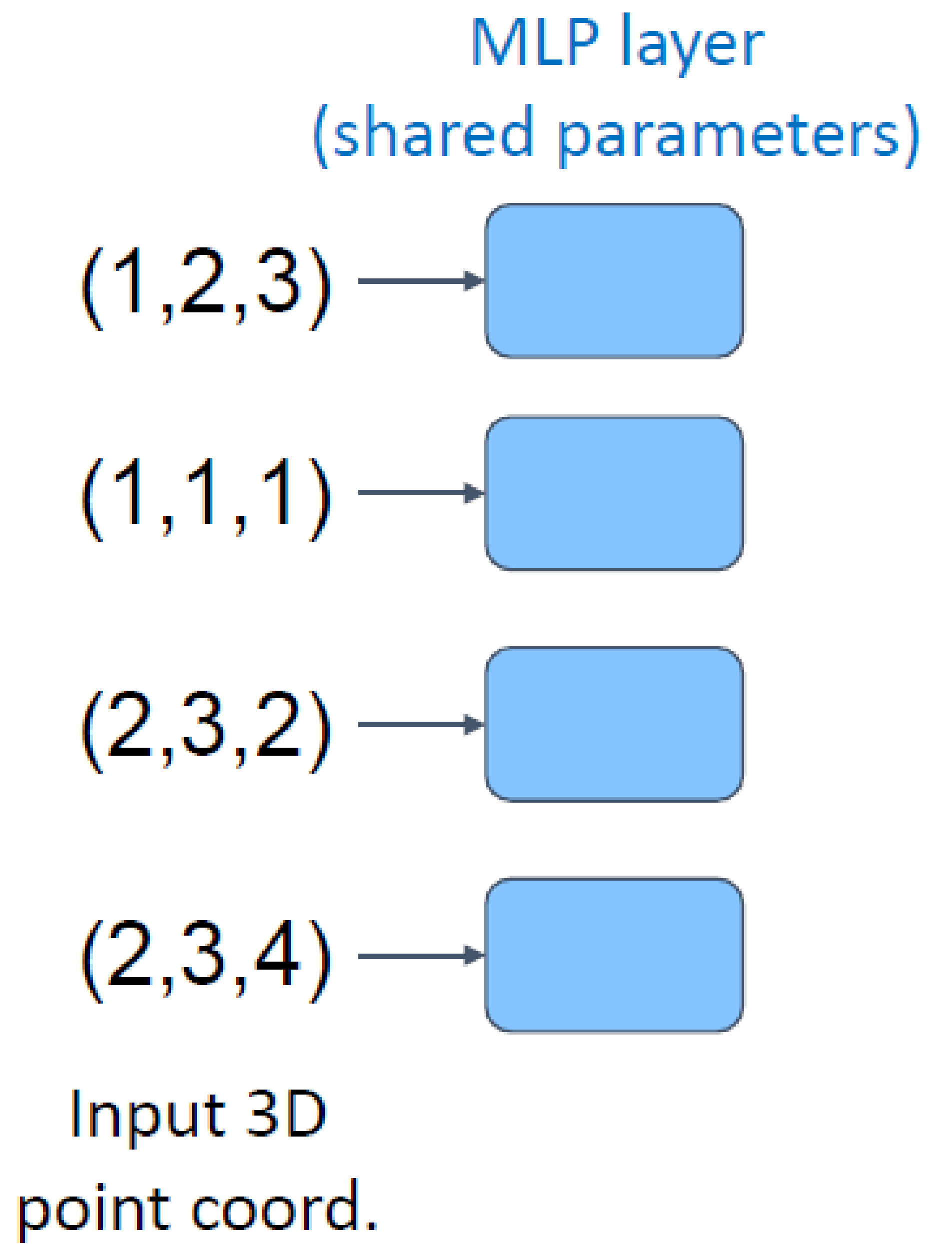


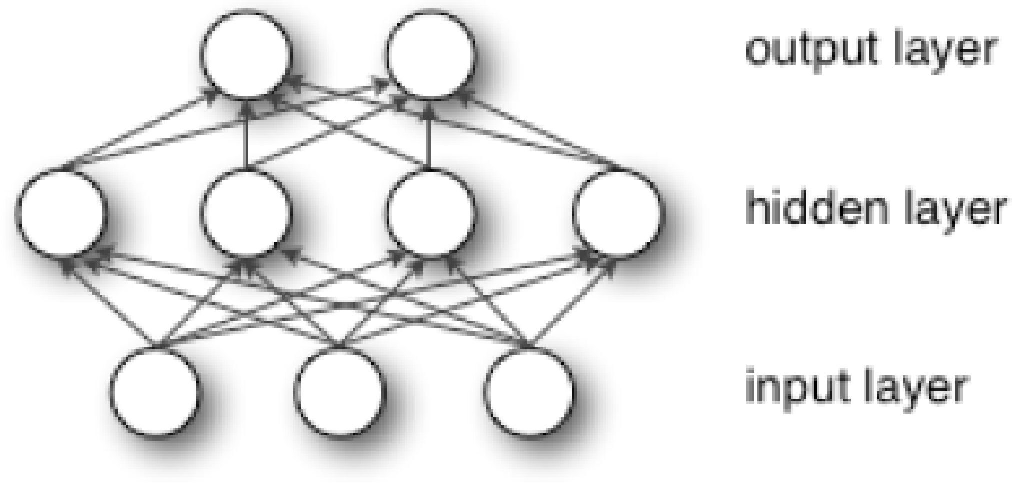
Kalogerakis E.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Point-based approach*

PointNet architecture:





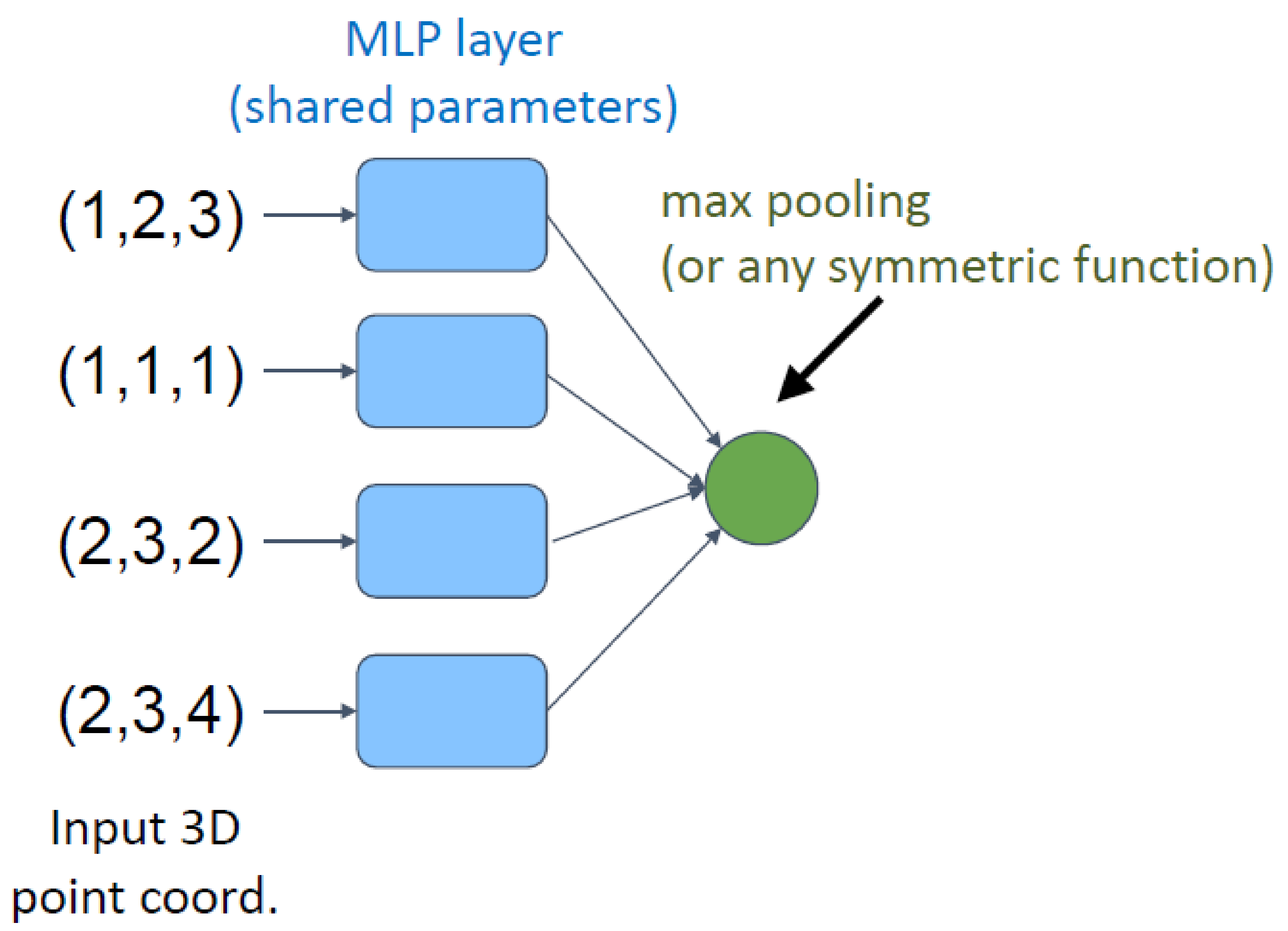
Simply a fully connected NN
with one hidden layer,
3 inputs for 3D points, and T
outputs (T is layer parameter)

Kalogerakis E.

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Point-based approach*

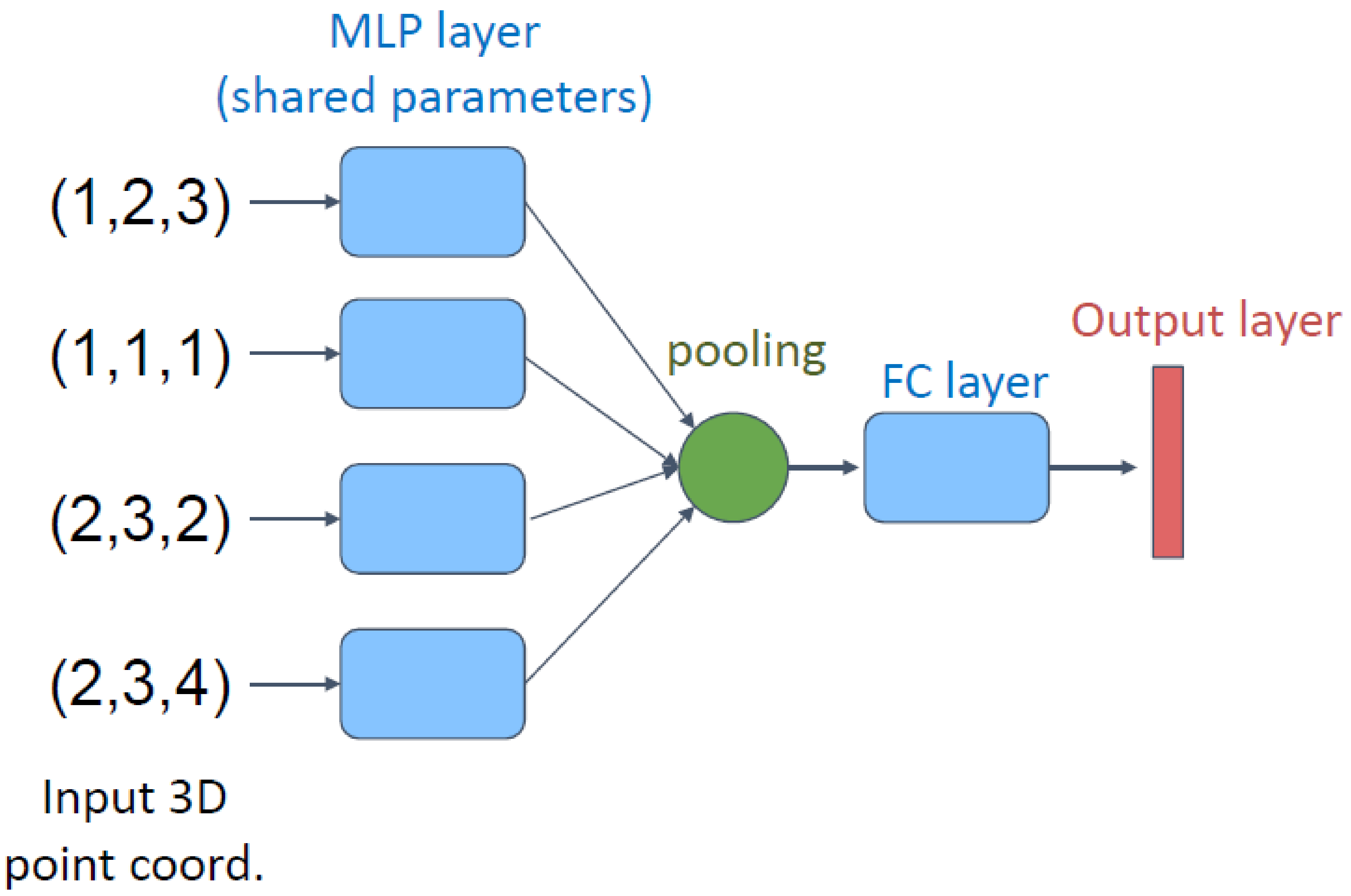
PointNet architecture:



Kalogerakis E.

3D DL architectures: *Point-based approach*

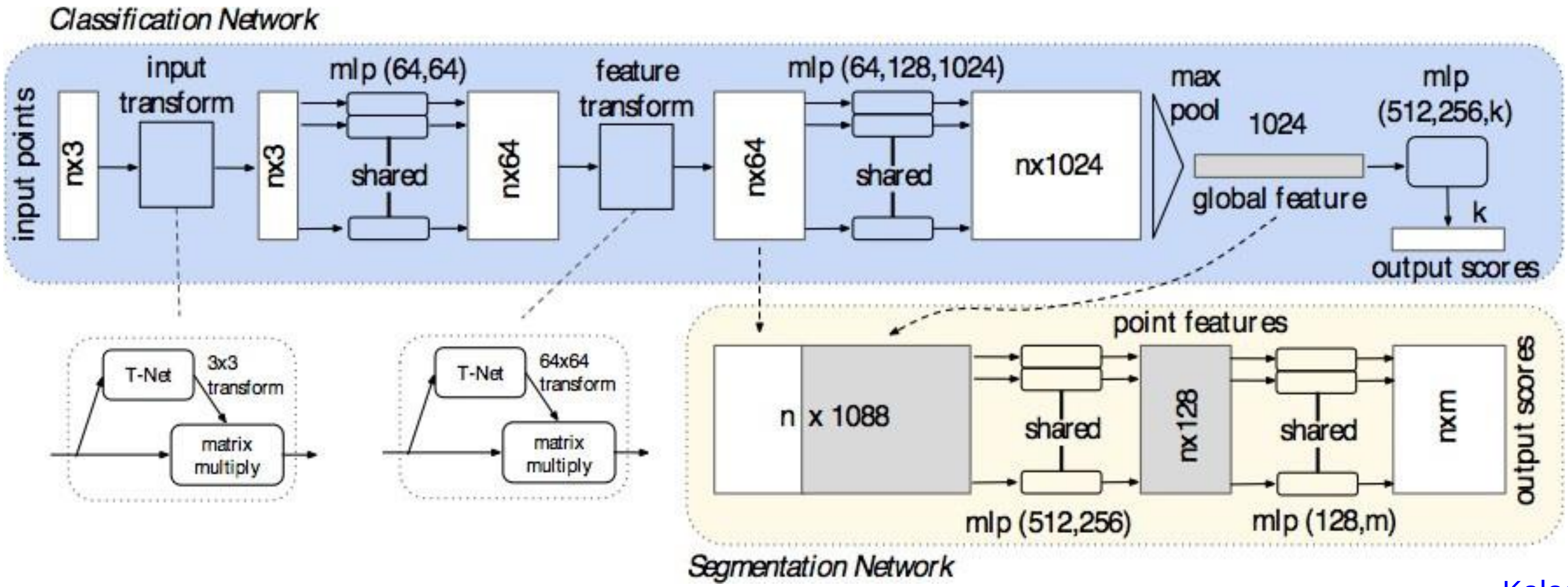
PointNet architecture:



Kalogerakis E.

3D DL architectures: *Point-based approach*

PointNet architecture:



Kalogerakis E.

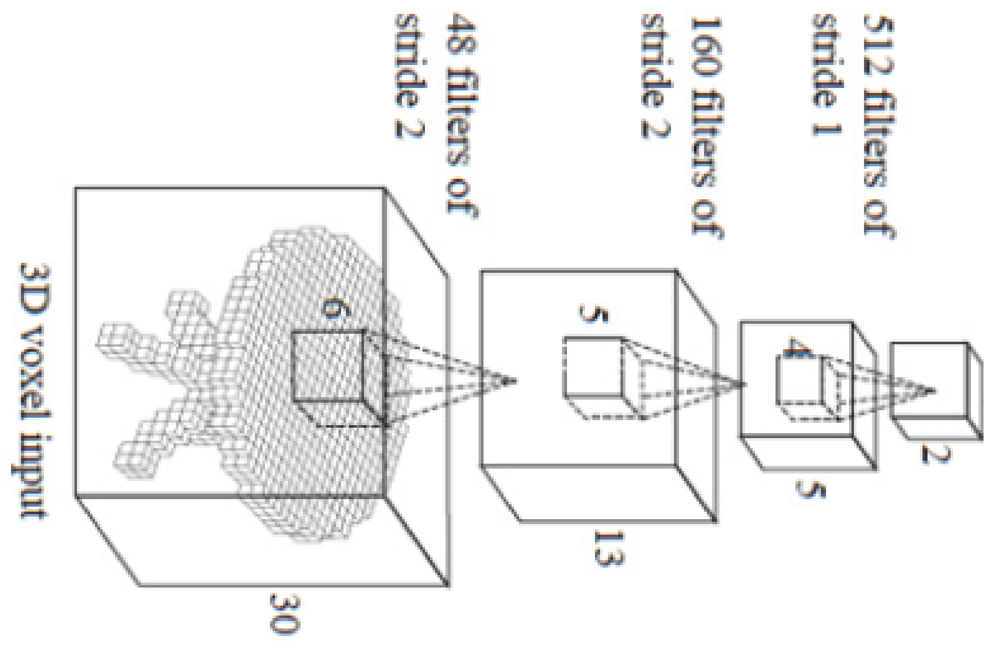
3D DL architectures: *Point-based approach*

Limitations of PointNet

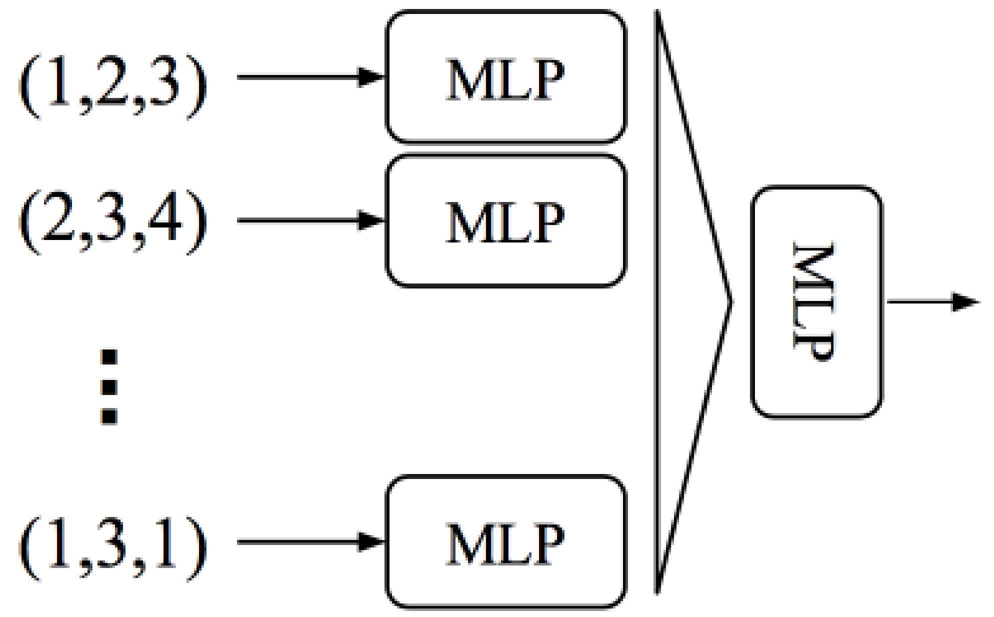
Hierarchical feature learning
Multiple levels of abstraction

v.s.

Global feature learning
Either one point or all points



3D CNN (Wu et al.)



PointNet (vanilla) (Qi et al.)

No local context for each point!

Hao Su et al.

3D DL architectures: *Point-based approach*

Points in Metric Space

- Learn “kernels” in 3D space and conduct convolution
- Kernels have compact spatial support
- For convolution, we need to find neighboring points
- Possible strategies for range query
  - Ball query (results in more stable features)
  - k-NN query (faster)

Hao Su et al.

Co-financed by the European Union
Connecting Europe Facility

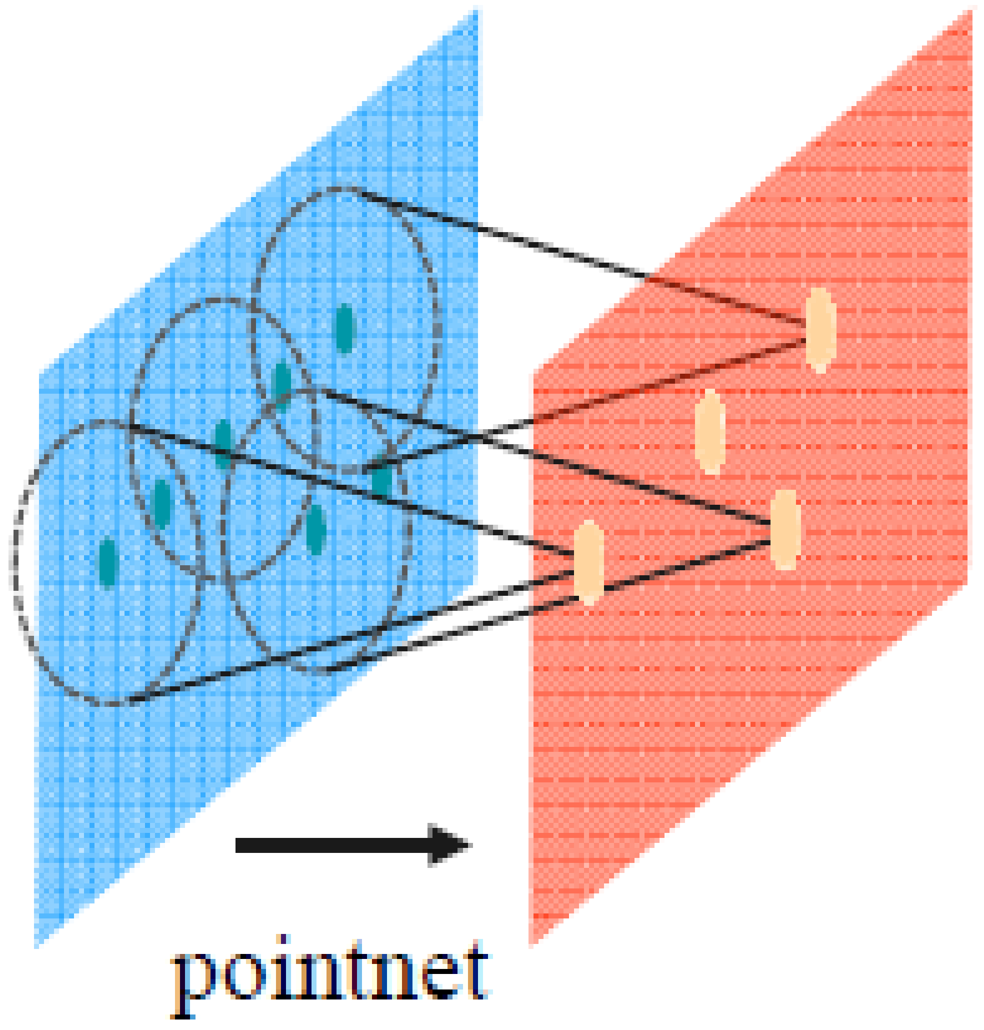
This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Point-based approach*

PointNet++: (Qi et al., NIPS 2017)

- Use PointNet in local regions
- Aggregate local features by PointNet again

–> **Hierarchical feature learning**

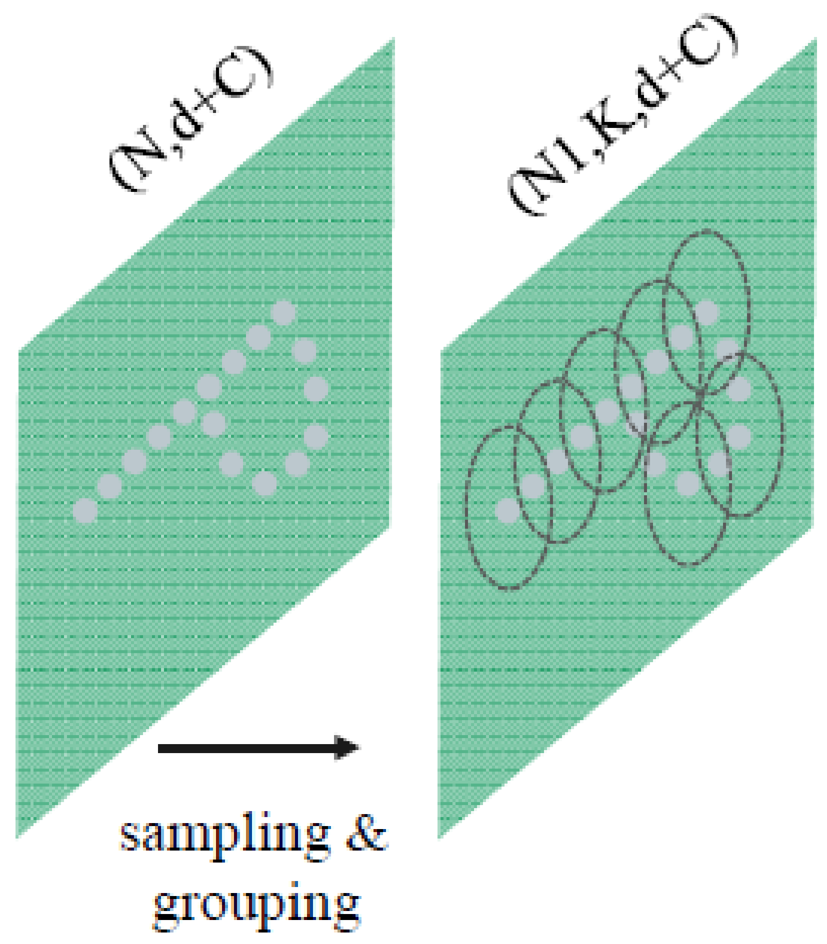


Kalogerakis E.

3D DL architectures: *Point-based approach*

PointNet++:

- **Sampling**: Farthest Point Sampling (FPS)
- **Grouping**: Radius-based ball query

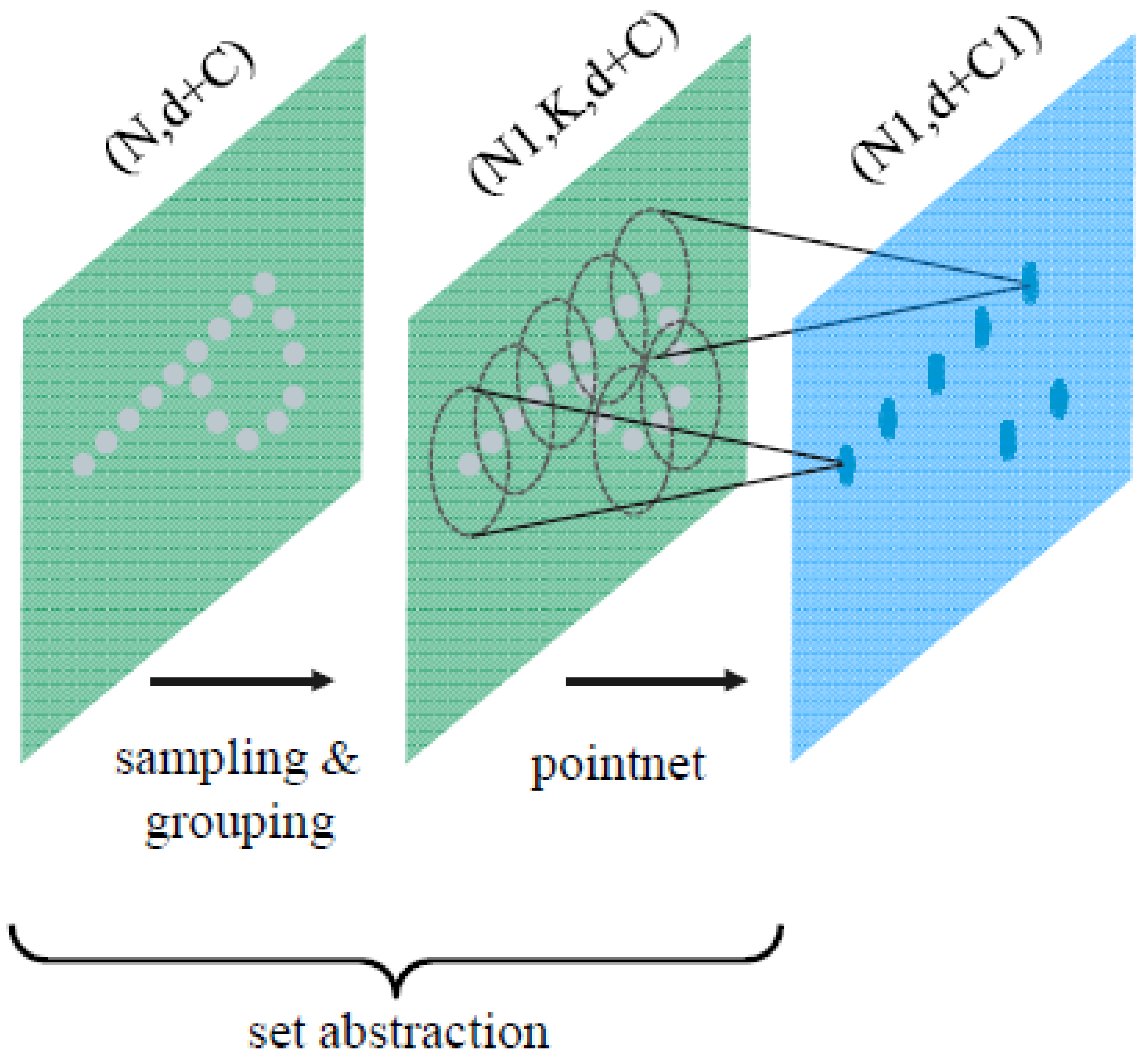


Kalogerakis E.

3D DL architectures: *Point-based approach*

PointNet++:

- Shared PointNet applied in each local region using local coordinates

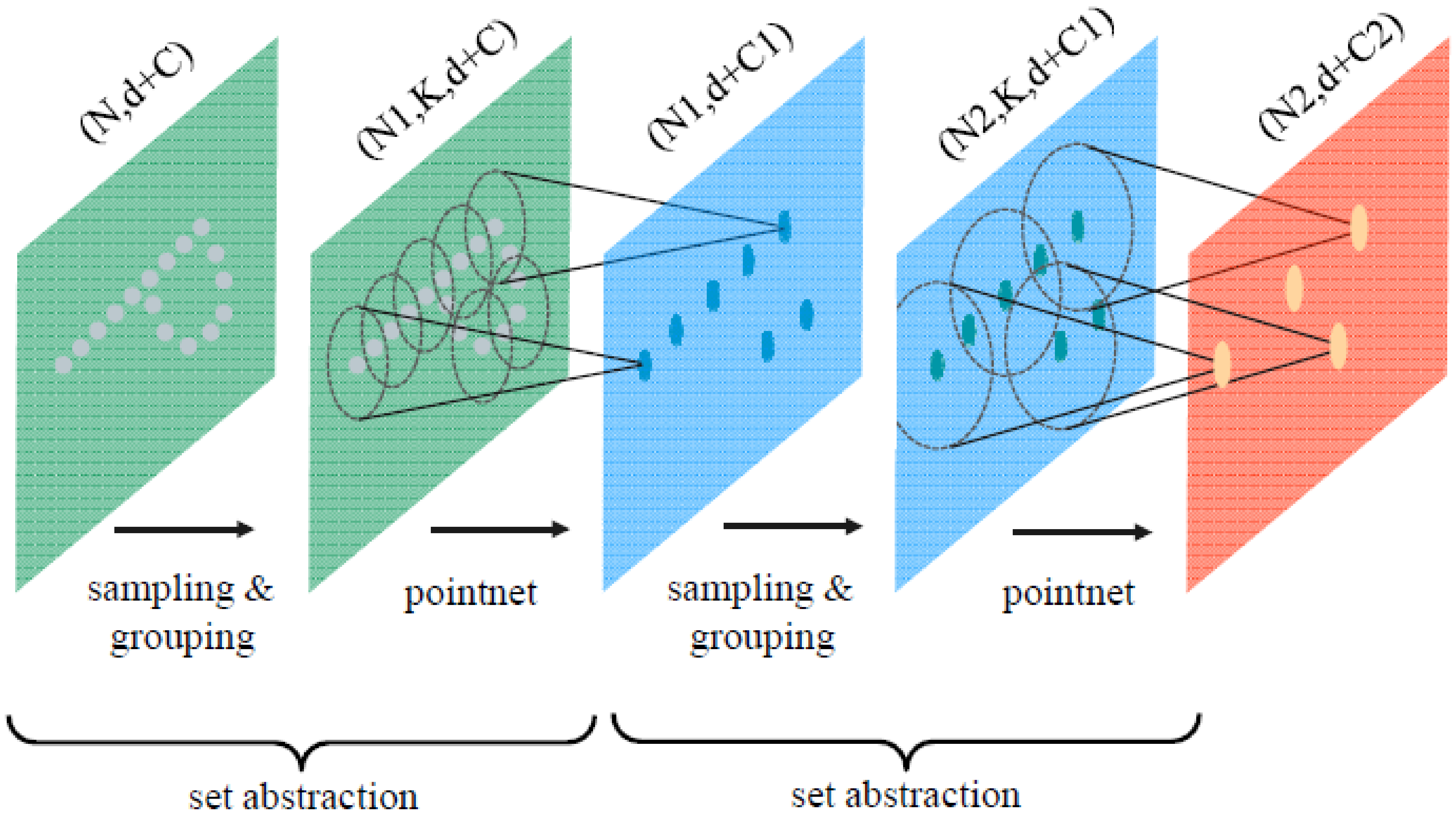


Kalogerakis E.

3D DL architectures: *Point-based approach*

PointNet++:

- Shared PointNet applied in each local region using local coordinates



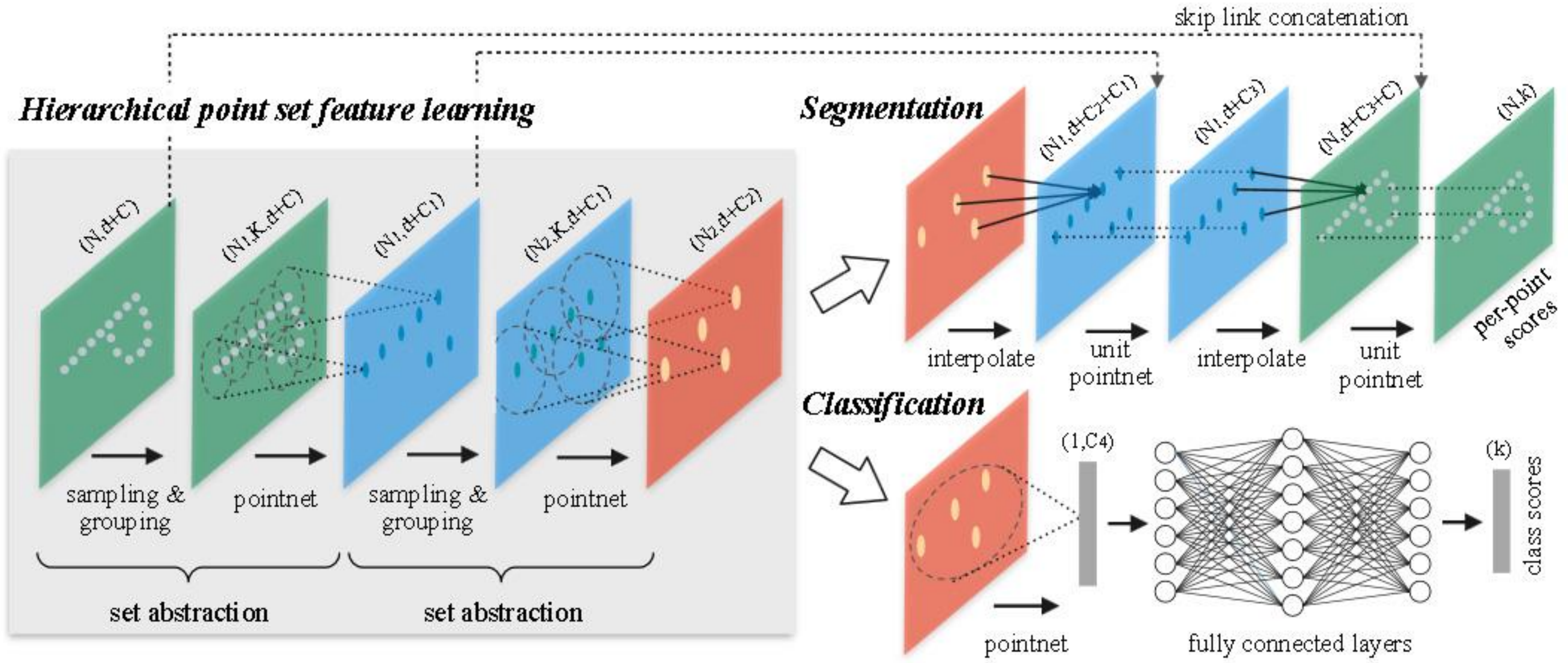
Kalogerakis E.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Point-based approach*

PointNet++ architecture

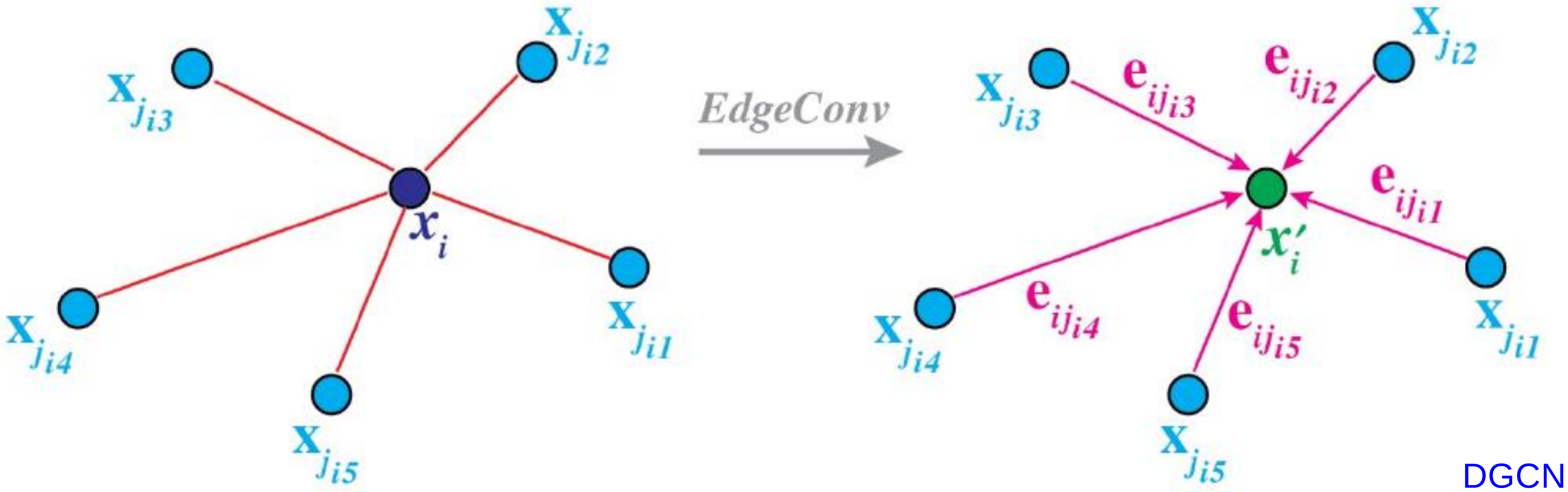


Qi et al., NIPS 2017

3D DL architectures: *Point-based approach*

Point Convolution as Graph Convolution: **Dynamic Graph CNN**

- Points –> Nodes
- Neighborhood –> Edges
- Graph CNN for point cloud processing

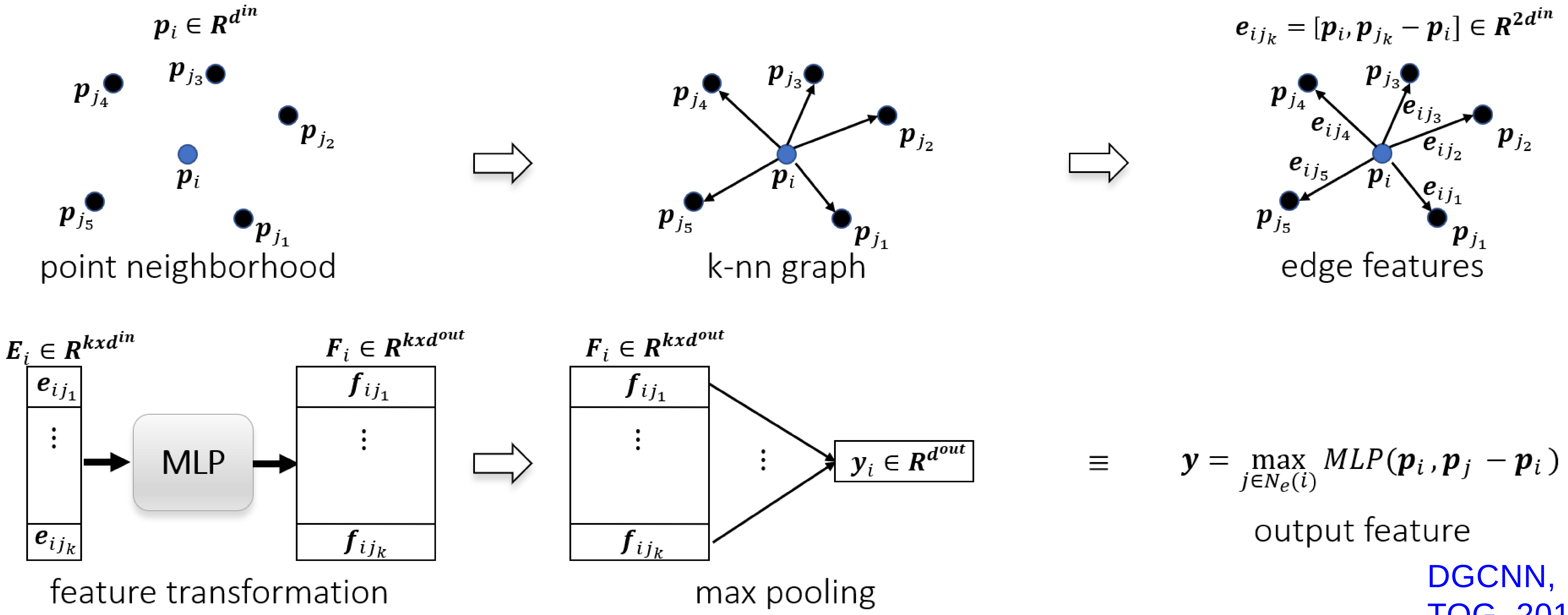


DGCNN, Wang et al., TOG, 2019

3D DL architectures: *Point-based approach*

Dynamic Graph CNN

- **EdgeConv Layer**

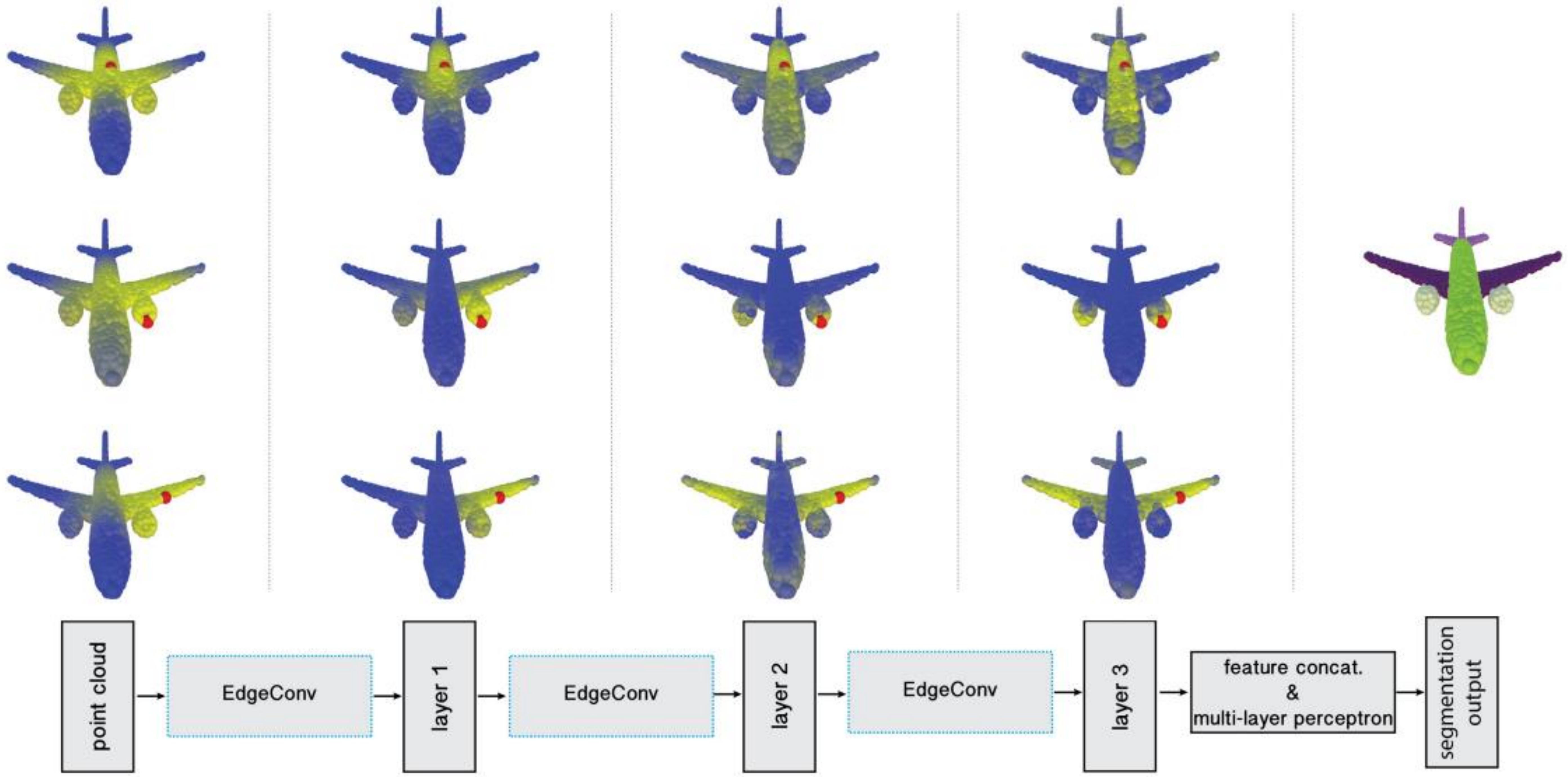


DGCNN, Wang et al., TOG, 2019

3D DL architectures: *Point-based approach*

Dynamic Graph CNN:

- At each layer, each local graph is rebuilt upon the feature space of the previous EdgeConv layer



DGCNN, Wang et al., TOG, 2019

3D DL architectures: *Point-based approach*

Standard GCNs are not Geometry Aware:

- Note that points are **sampled** from surfaces
- Ideally, features describe the geometry of the underlying surface
- Should be **sample invariant**
- But GCNs lack design to address sample invariance
- **Solution**: Estimate the continuous kernel and point density for continuous convolution

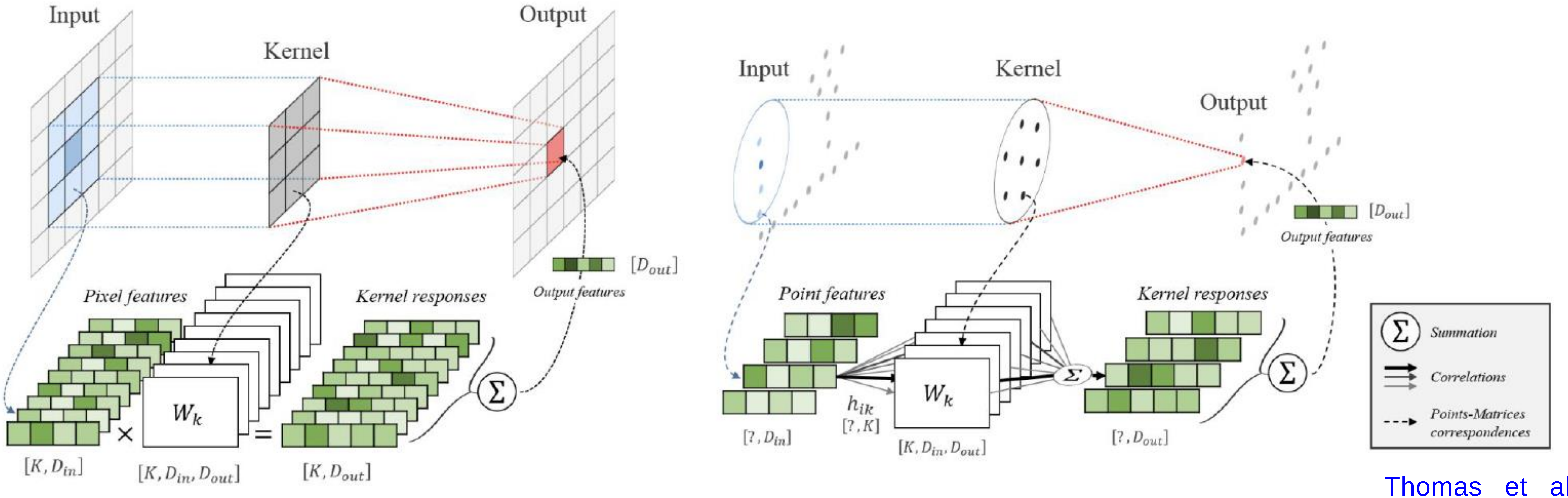
Hao Su et al.

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

3D DL architectures: *Point-based approach*

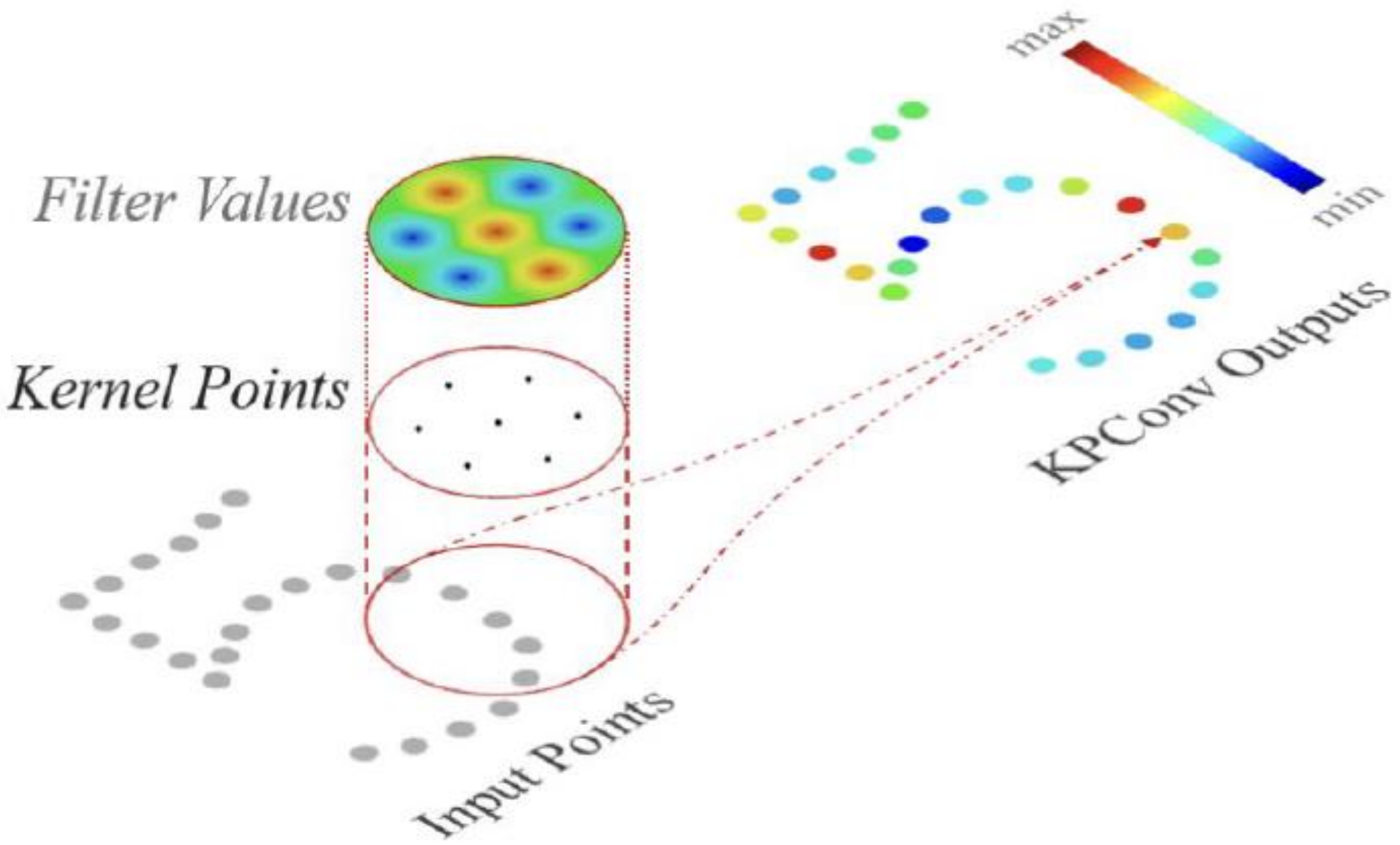
Kernel Point Convolution (KPConv)



Thomas et al., ICCV, 2019

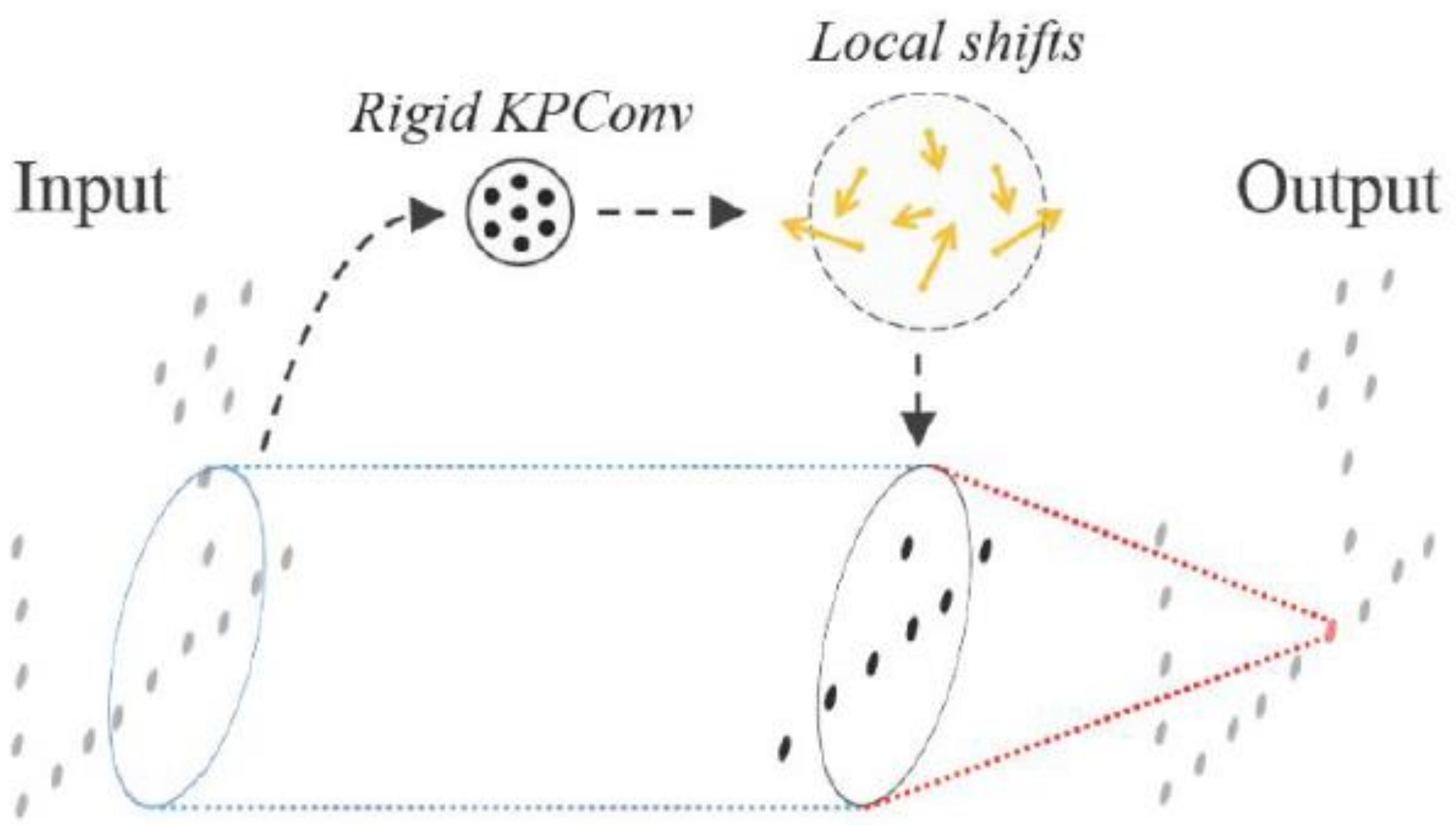
3D DL architectures: *Point-based approach*

Kernel Point Convolution (KPConv)



Deformable point-based kernel

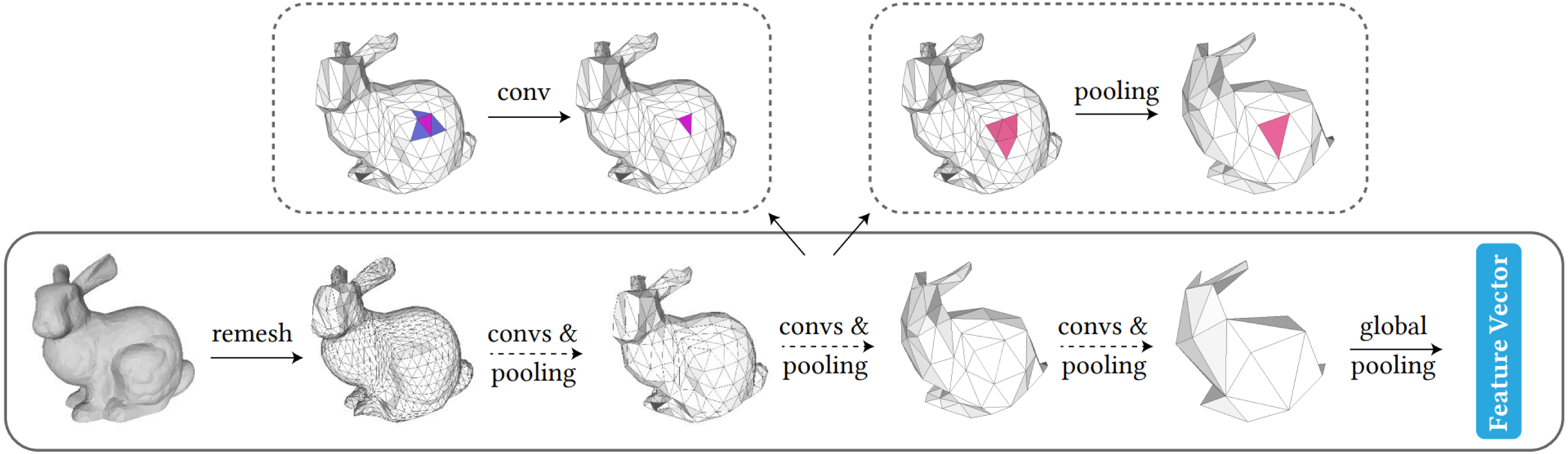
- 3D version of 2D deformable convolution



Hao Su et al.

3D DL architectures: *Mesh-based approach*

Subdivision-Based Mesh Convolution Networks (SubdivNet)

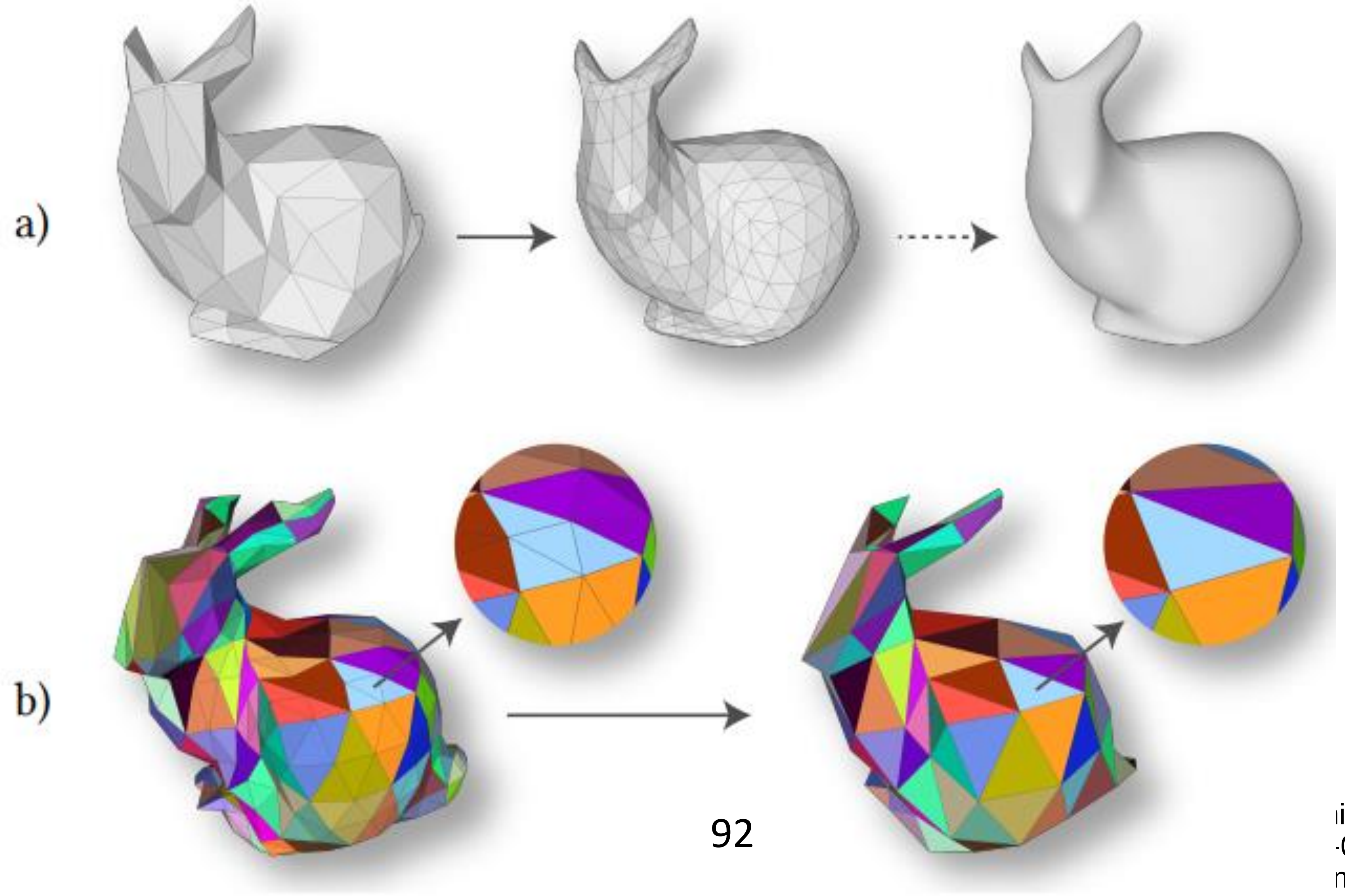


Shi-Min Hu et al., TOG, 2021

3D DL architectures: *Mesh-based approach*

SubdivNet:

- A **subdivision surface** provides a h**ierarchical multi-resolution structure**, in which **each face** in a **closed triangle** mesh is exactly **adjacent to three faces**



Shi-Min Hu et al., TOG, 2021

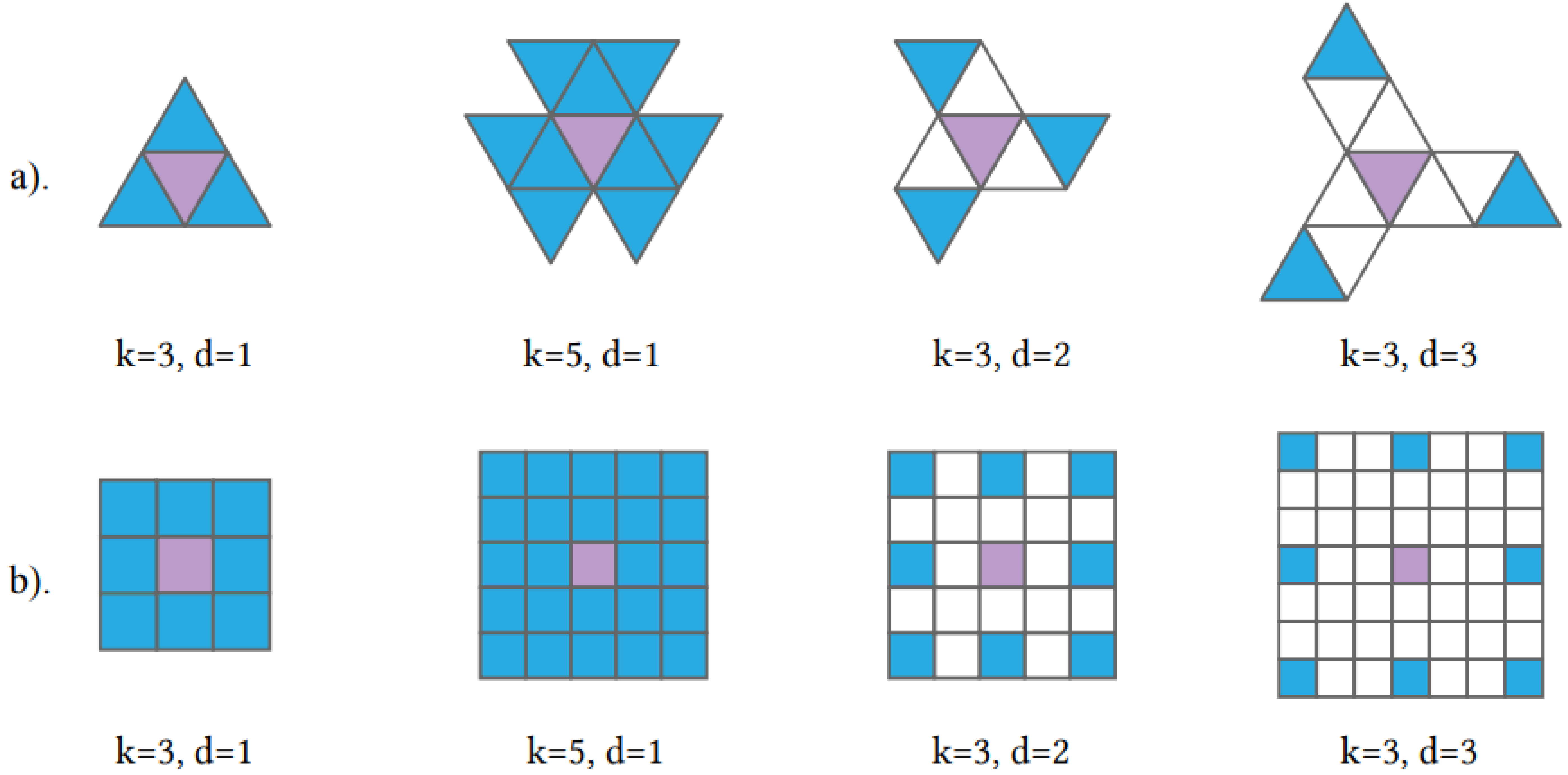
3D DL architectures: *Mesh-based approach*

SubdivNet:

- Can support **mesh convolution**

Mesh conv. kernels

2D conv. kernels

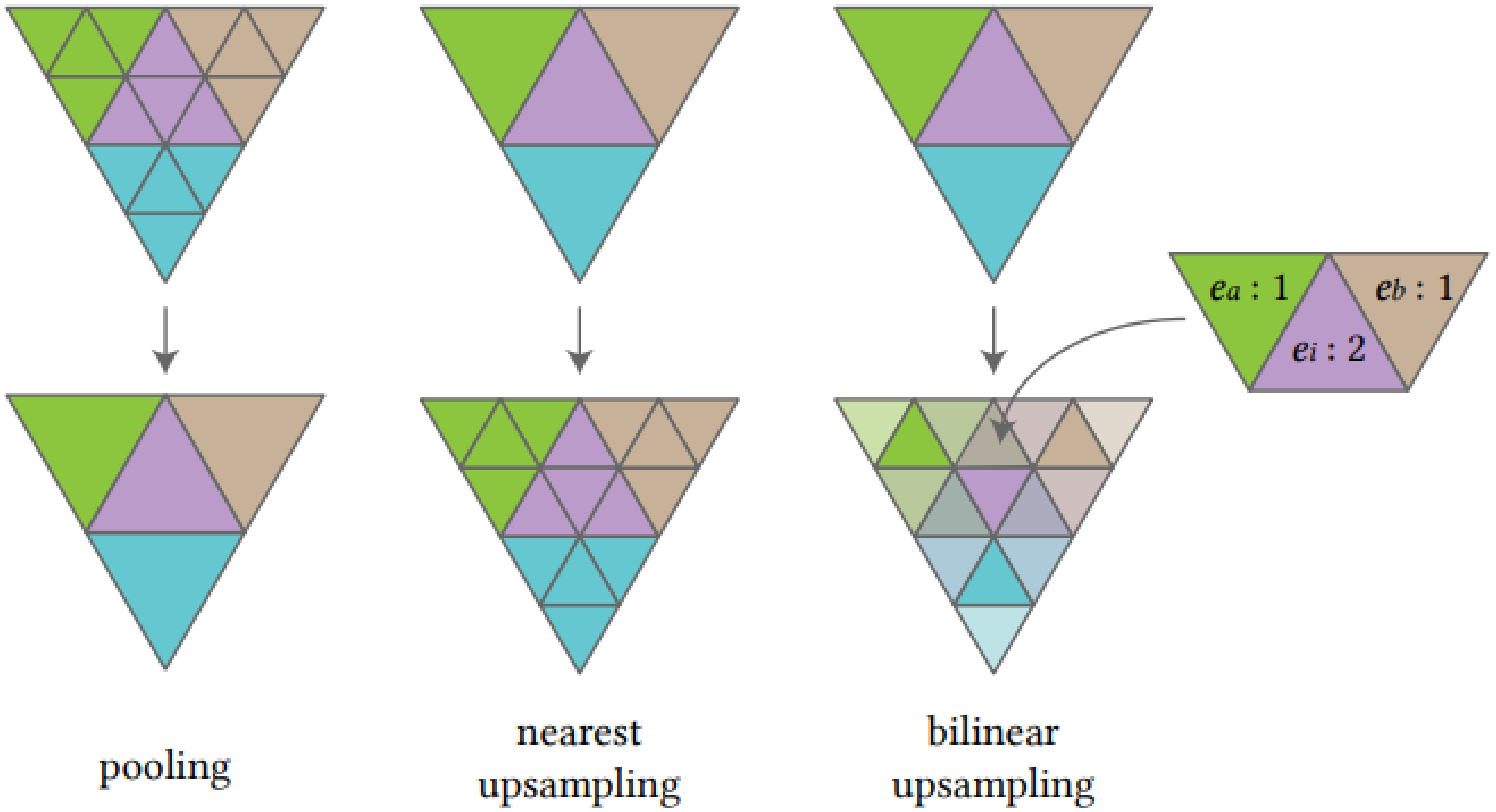


Shi-Min Hu et al., TOG, 2021

3D DL architectures: *Mesh-based approach*

SubdivNet:

- Can support **pooling** and **upsampling**



Shi-Min Hu et al., TOG, 2021

Today's Agenda

- Who are we?
- What is 3D Vision
- 3D shape representations
- 3D shape datasets
- 3D Deep Learning architectures
- What we do

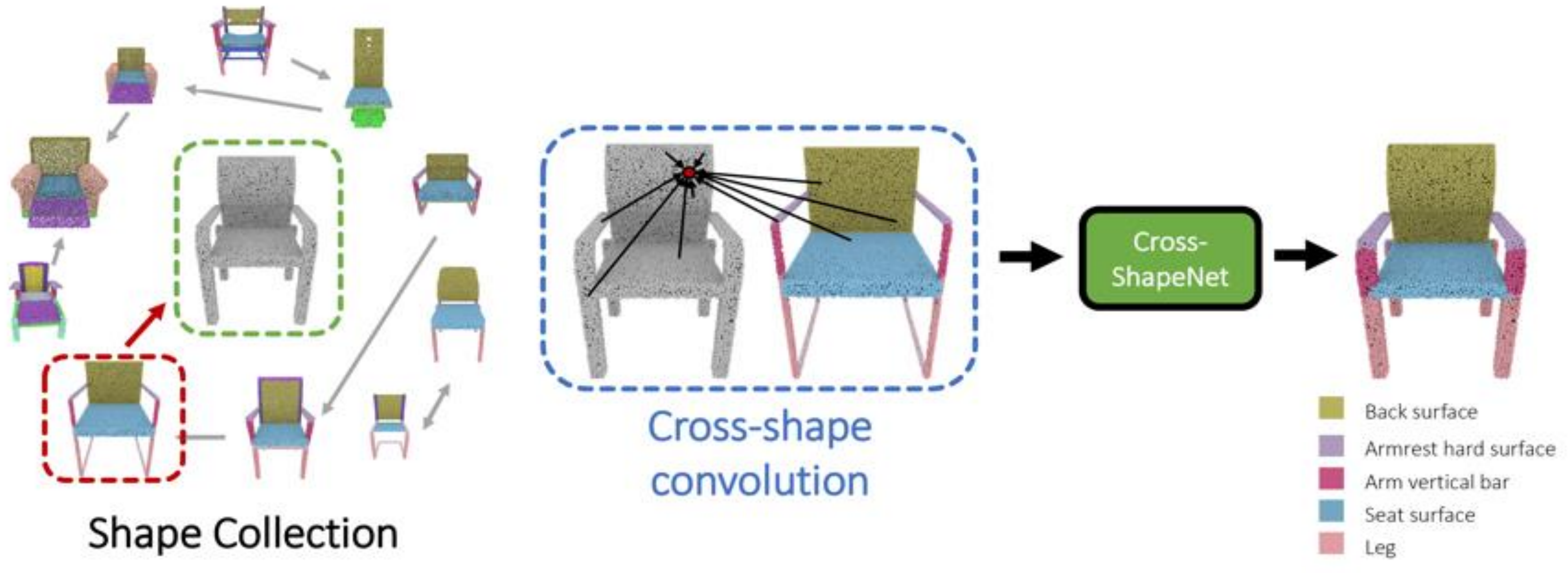
Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423

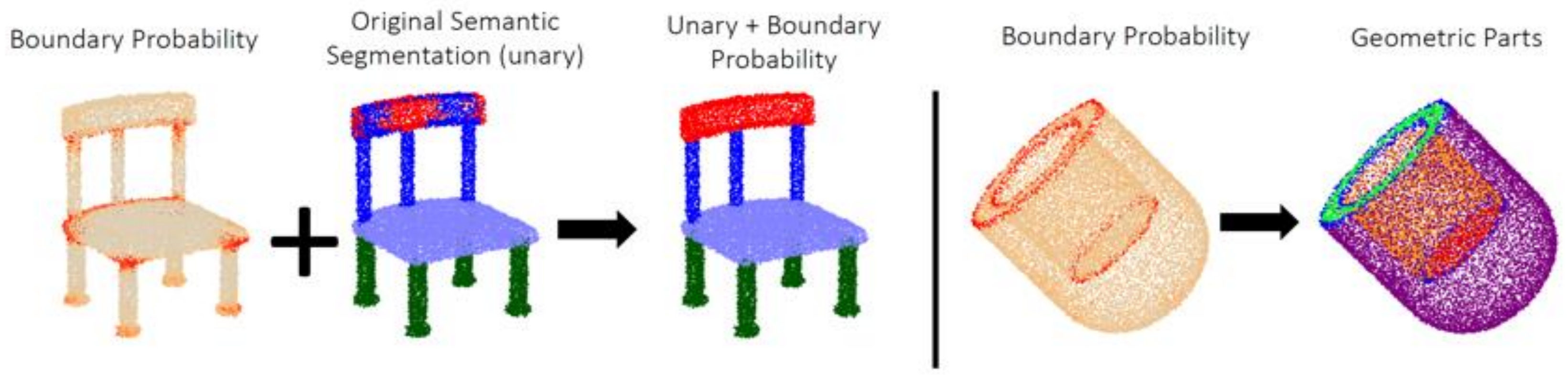
What we do: *3D shape understanding*

3D Building Semantic Understanding

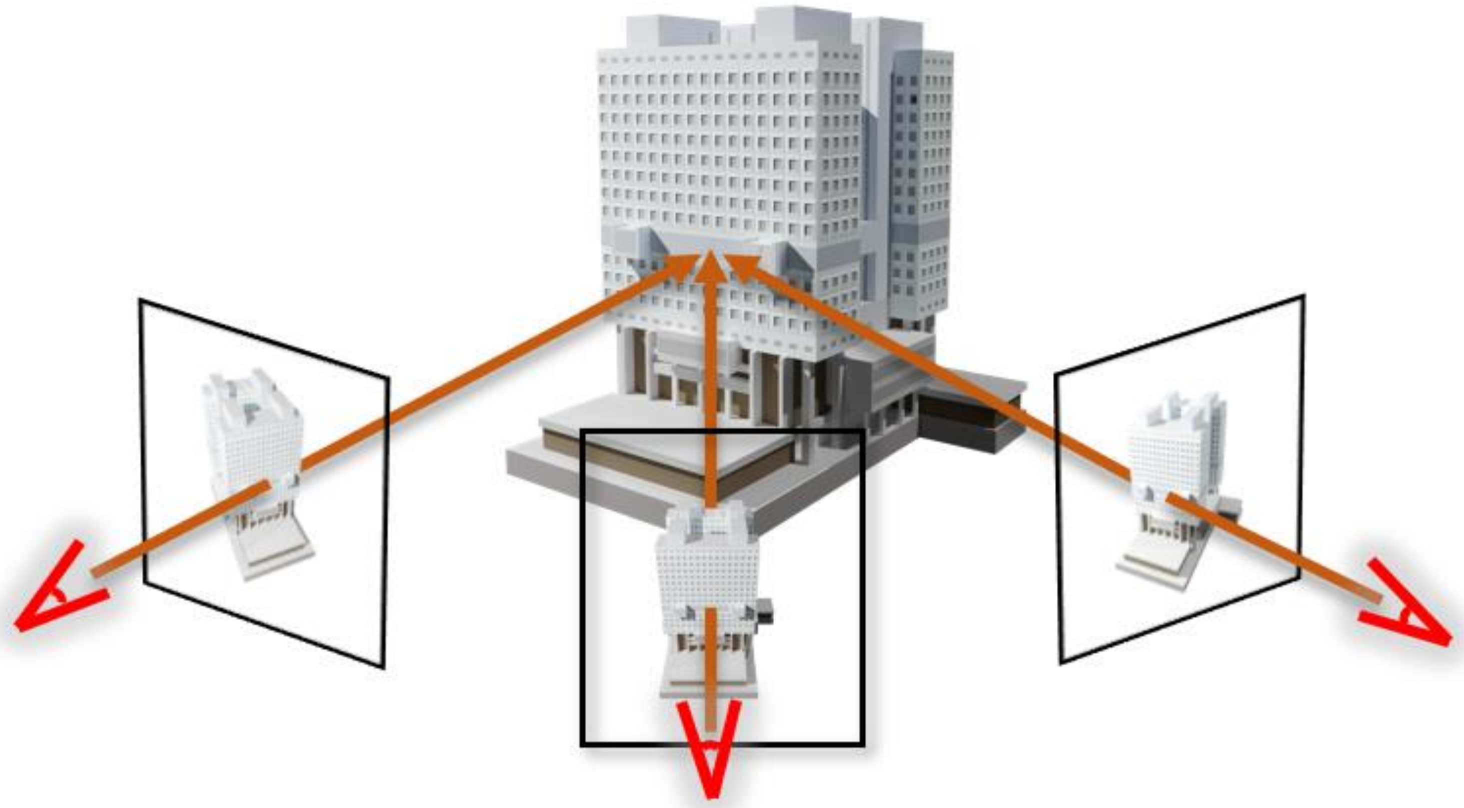
Cross-shape semantic segmentation



Geometric/Semantic Decomposition

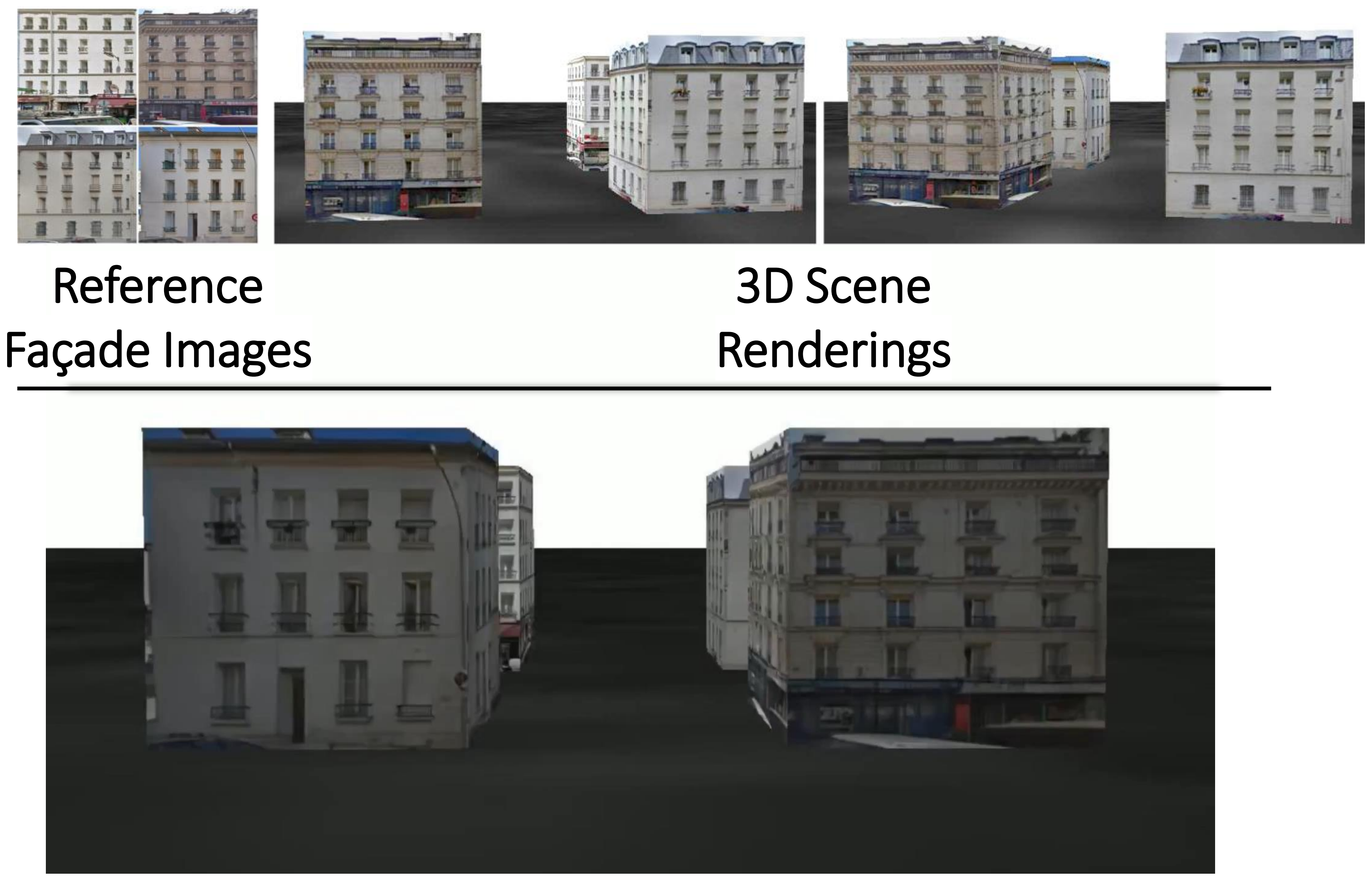


Neural 3D Reconstruction

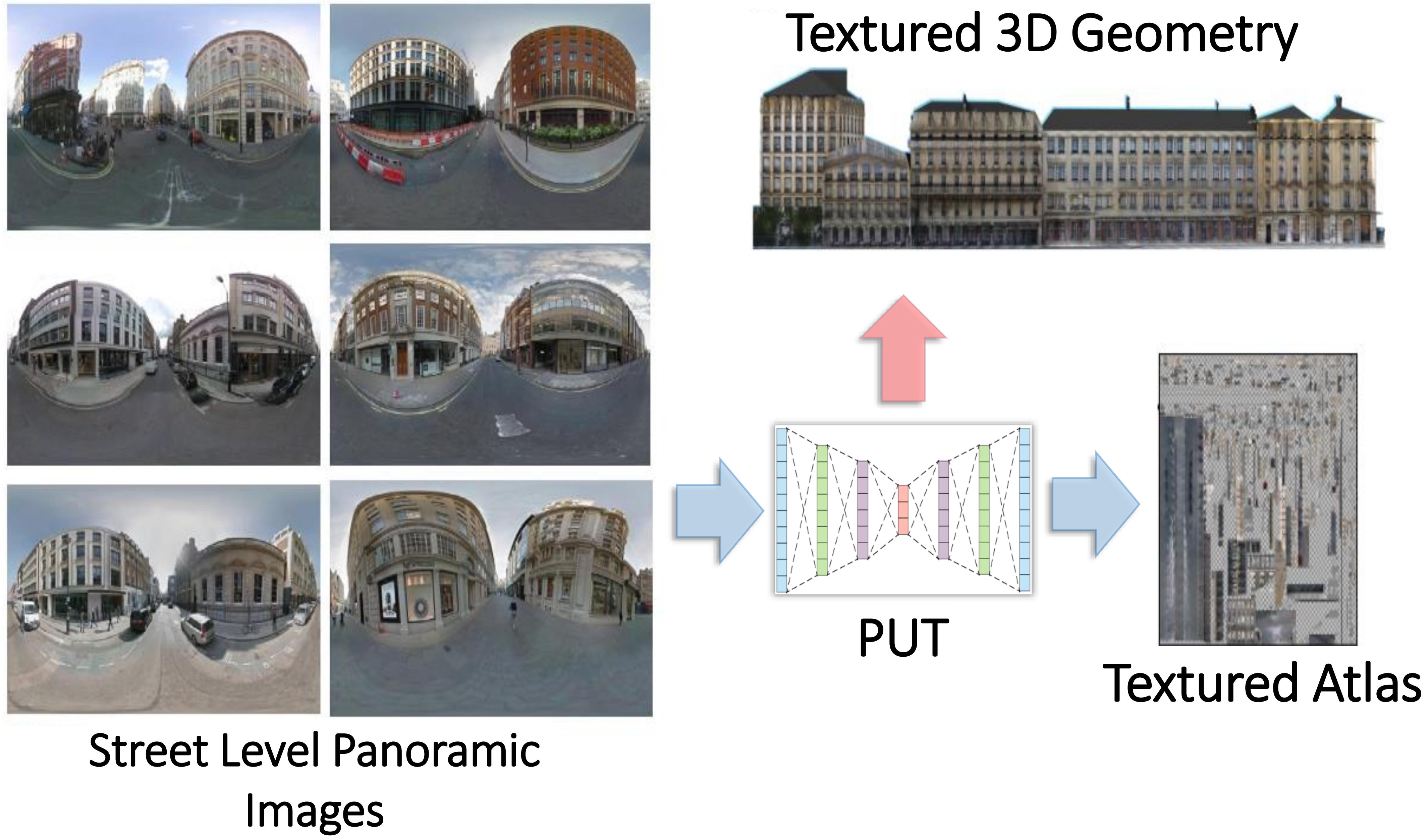


What we do: *Texture Generation for 3D Data*

Single-View Guided Façade Synthesis

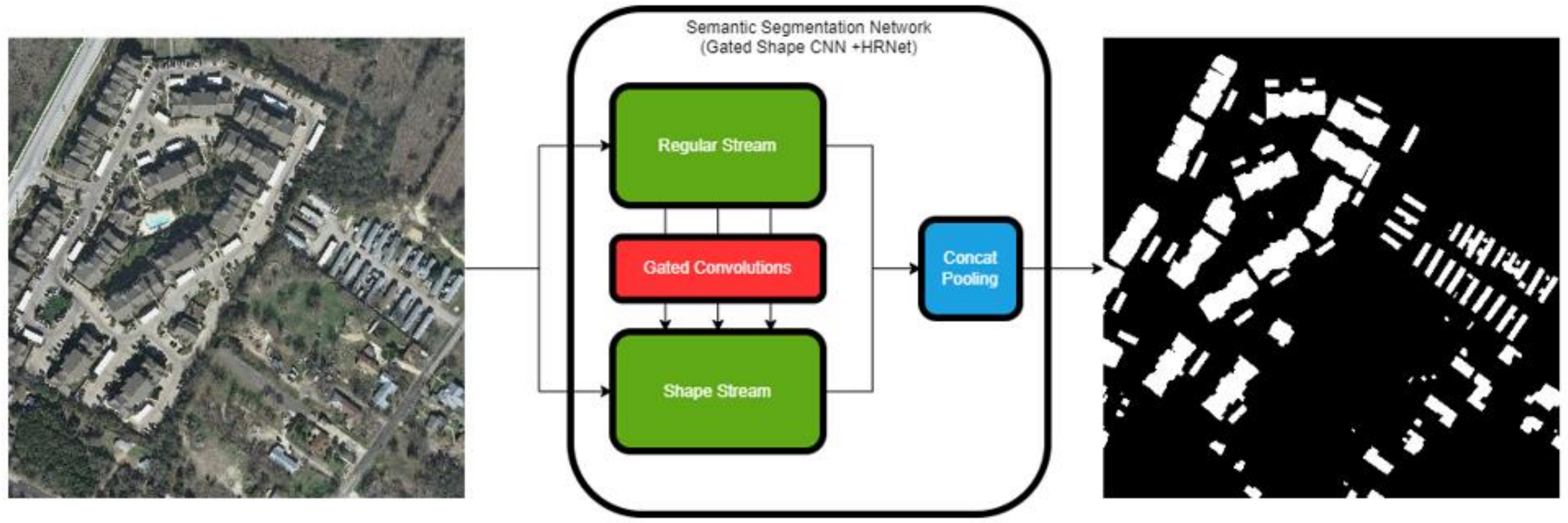


Projective Urban Texturing

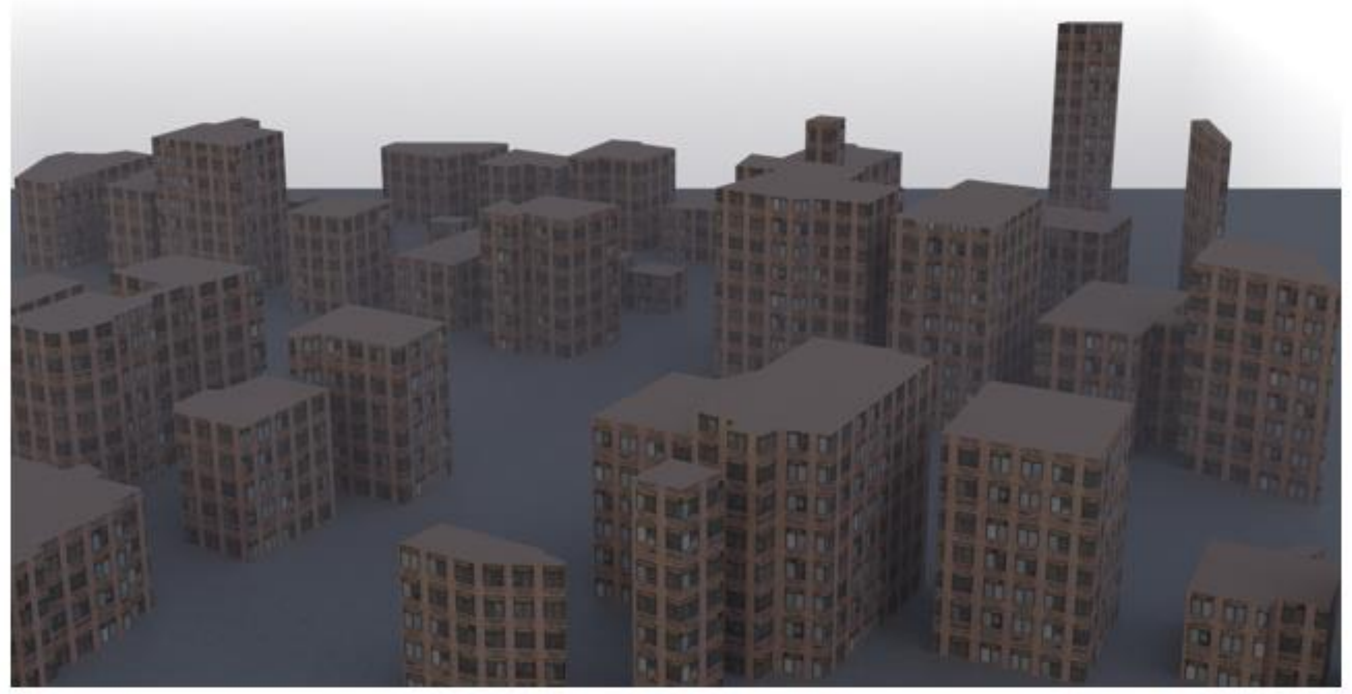


What we do: *Urban Semantic Understanding from Remote Sensing Data Sources*

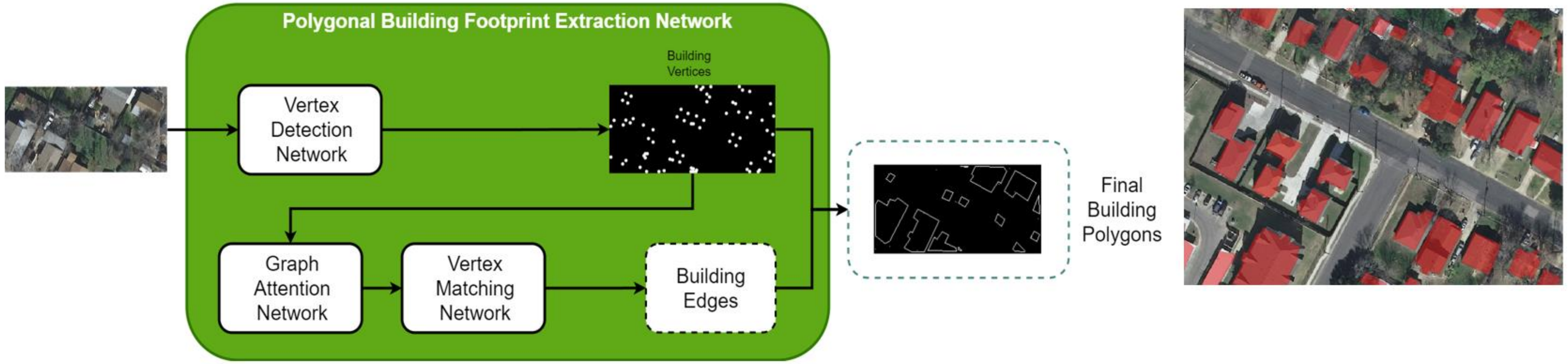
Semantic Segmentation of Buildings



Urban 3D Reconstruction



Building Footprint Extraction



Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action
No 2020-EU-IA-0087, co-financed by the EU CEF Telecom
under GA nr. INEA/CEF/ICT/A2020/2267423

Research in Visual Computing

Melinos Averkiou

Team Leader
Visual Computing Group

email: m.averkiou@cyens.org.cy

VCG

CYENS
CENTRE OF EXCELLENCE

Research Interests:

Geometry processing, acquisition, understanding and modeling of 3D geometry, deep learning for 3D objects, including part segmentation, material identification and style detection.

https://www.cyens.org.cy/en-gb/research/pillars-groups/visual-sciences/deep-camera/people/alessandro-artusi/

https://vcg.cyens.org.cy/

m.averkiou@cyens.org.cy

+357 227 475 75

Dimarchias Square 23 STOA, Nicosia
Nicosia, Nicosia 1016, Cyprus

Thank you!

See you next week

Co-financed by the European Union
Connecting Europe Facility

This Master is run under the context of Action No 2020-EU-IA-0087, co-financed by the EU CEF Telecom under GA nr. INEA/CEF/ICT/A2020/2267423