

Введение в 3D ML

Каратаева Екатерина

БПМИ182
27.04.21

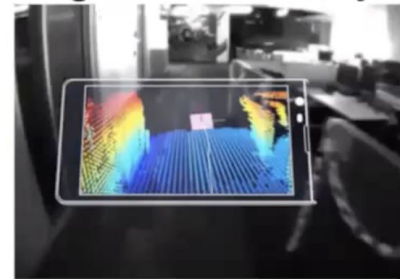
Области применения:

- ❖ Робототехника
- ❖ Компьютерное зрение
- ❖ Виртуальная и дополненная реальность
- ❖ Медицина
- ❖ Промышленность
- ❖ ...

Robot Perception



Augmented Reality

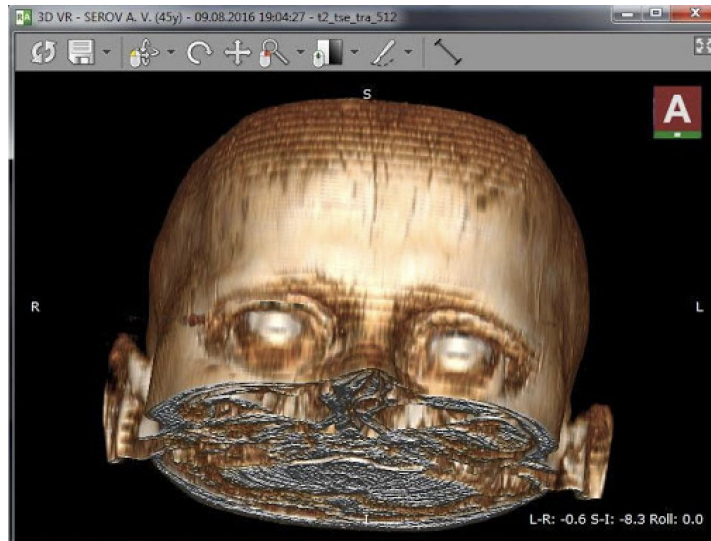
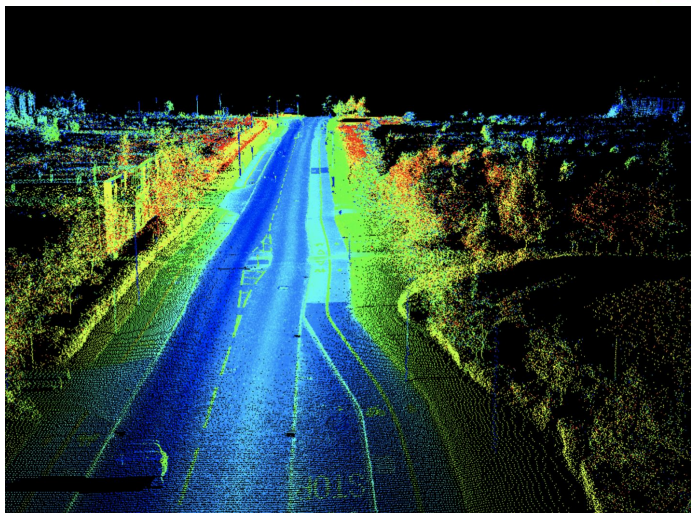
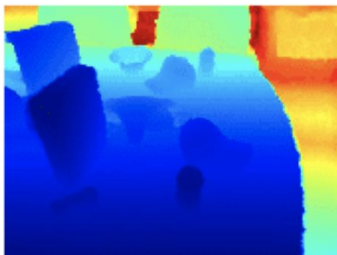
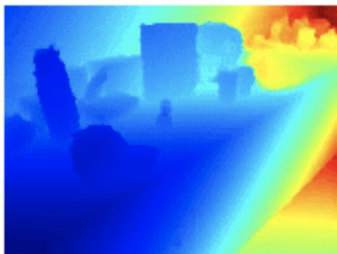


Shape Design



Robotic Bin Picking

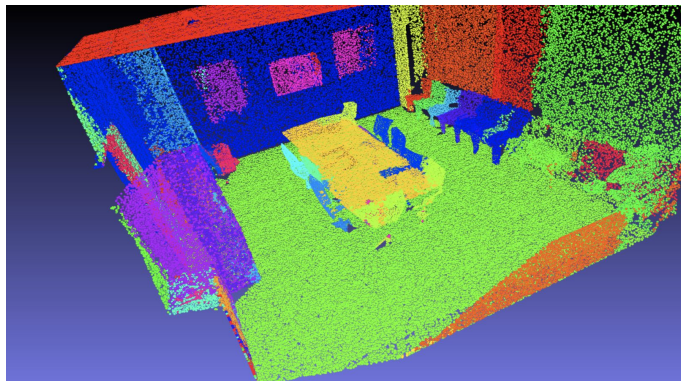




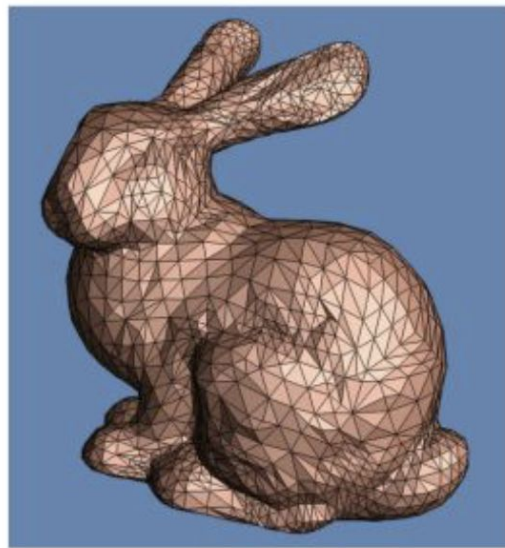
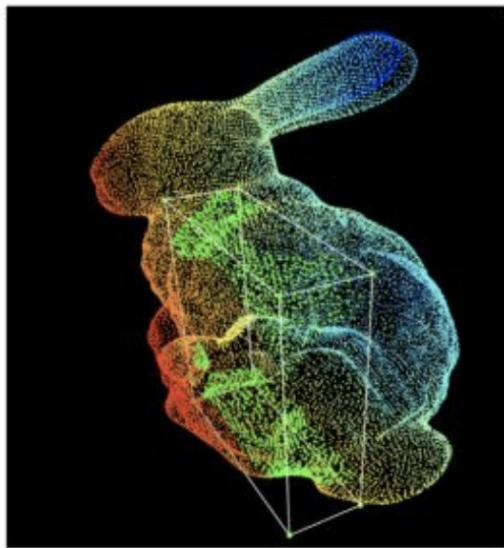


Примеры задач:

- ❖ Классификация и кластеризация 3D данных
- ❖ Генерация 3D контента
- ❖ 2D-to-3D
- ❖ Сегментация 3D сцен или объектов
- ❖ Поиск 3D объектов



Формы представления 3D-данных



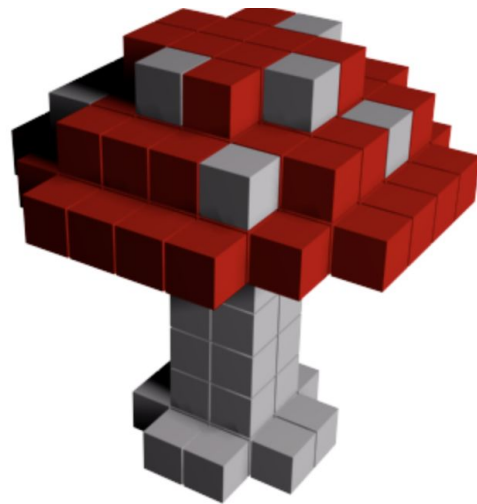
Воксели (voxel = volumetric + pixel)

Преимущества:

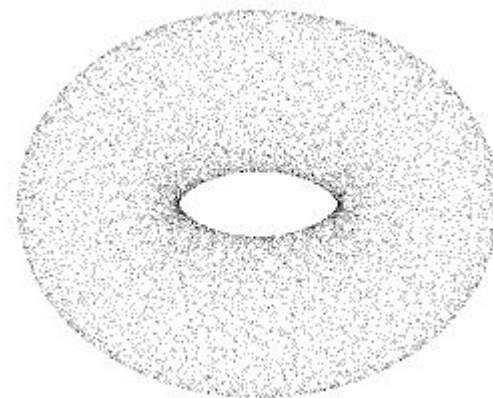
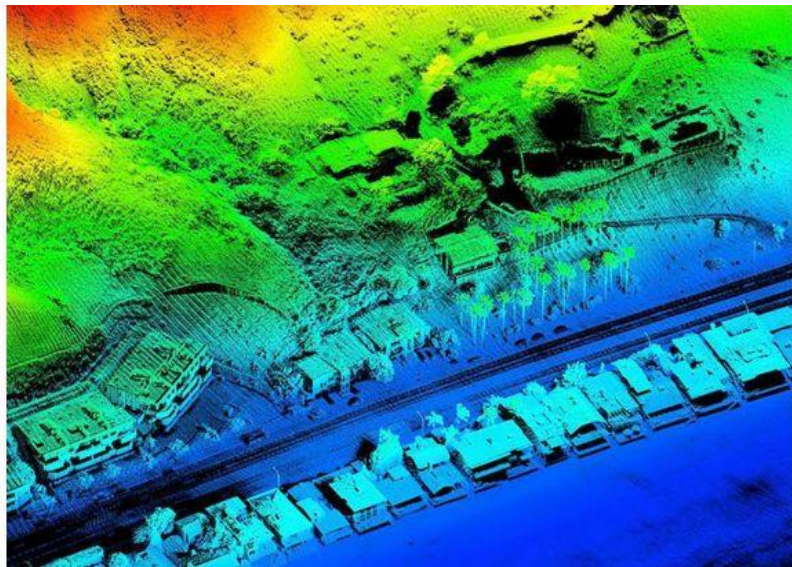
- ❖ Легко получить из другого типа моделей
- ❖ Связь с физическими свойствами объектов
- ❖ Можно представить в виде обычных векторов

Недостатки:

- ❖ “Грубость” аппроксимации формы при малом разрешении
- ❖ Данный подход ограничен низким разрешением пространственной сетки.



Облака точек (Point clouds)



Облака точек (Point clouds)

Преимущества:

- ❖ Естественный формат данных в задачах пространственного сканирования
- ❖ Можно легко получены из других моделей

Недостатки:

- ❖ Неупорядоченность данных
- ❖ Отсутствует информация о связях между точками

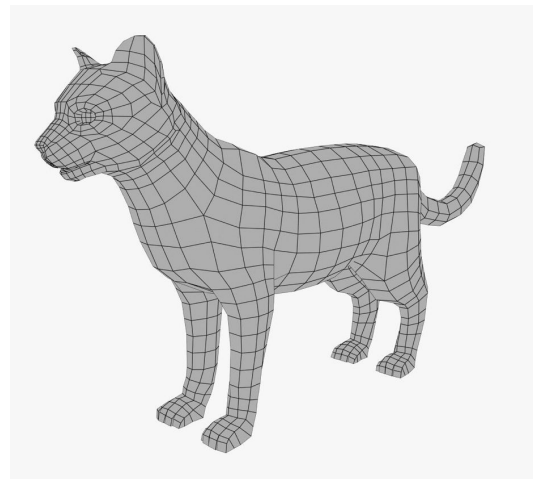
Полигональные модели (Mesh, polygonal models)

Преимущества:

- ❖ Естественный формат для использования в компьютерной графике, игровых движках
- ❖ Лучше описывают пространственные особенности объектов

Недостатки:

- ❖ Чувствительность формата к выбросам в данных.



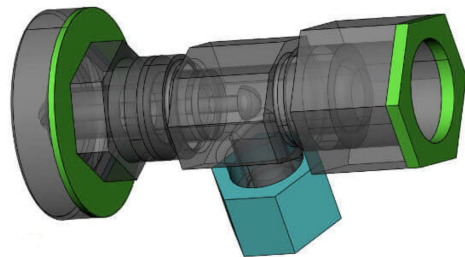
Функциональные модели (CAD models)

Преимущества:

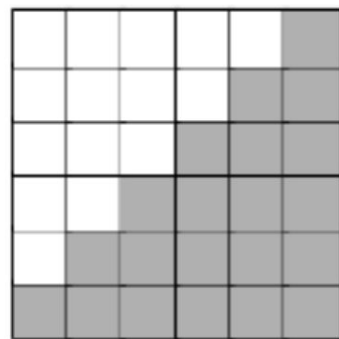
- ❖ Компактность, масштабируемость, физически корректная модель
- ❖ Можно получить модель любого другого формата

Недостатки:

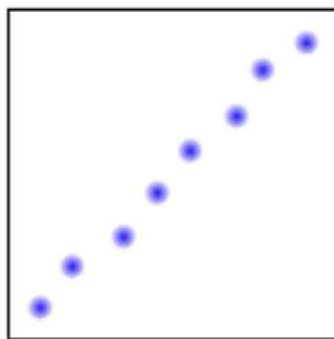
- ❖ Мало обучающих выборок
- ❖ Другие форматы проблематично приводить к функциональному



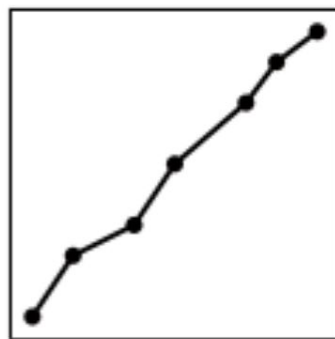
$$SDF(\mathbf{x}) = s : \mathbf{x} \in \mathbb{R}^3, s \in \mathbb{R}$$



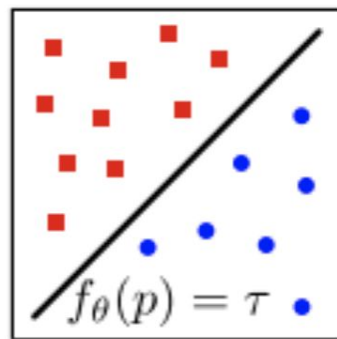
(a) Voxel



(b) Point



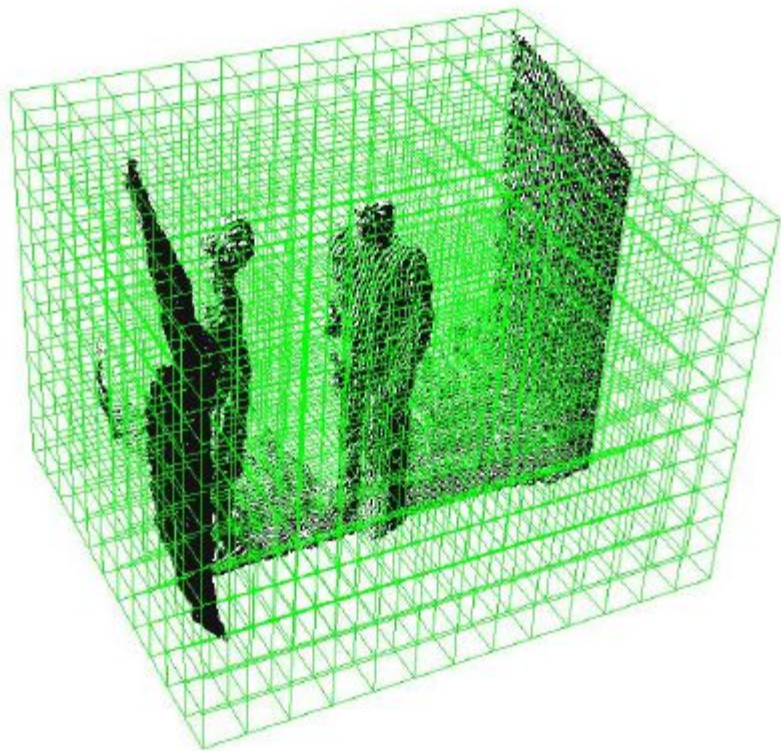
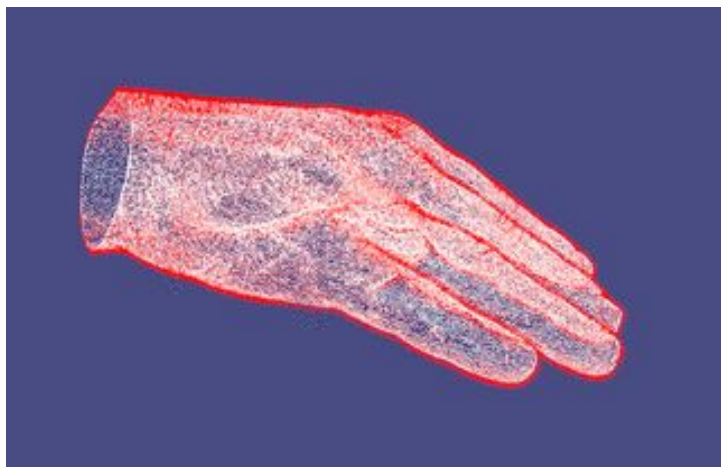
(c) Mesh

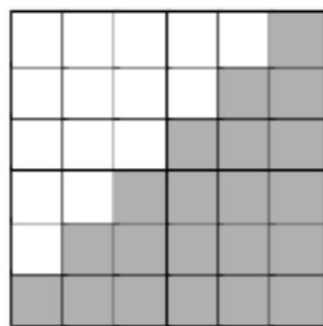


(d) Functions

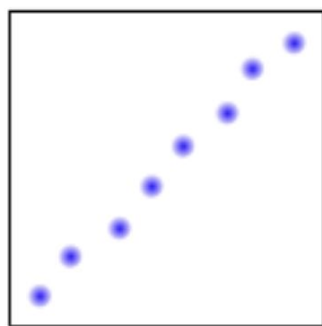


Information

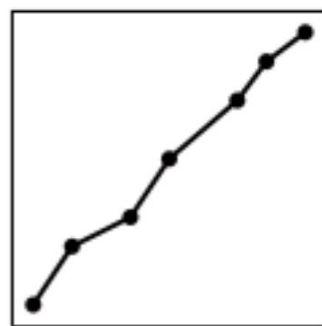




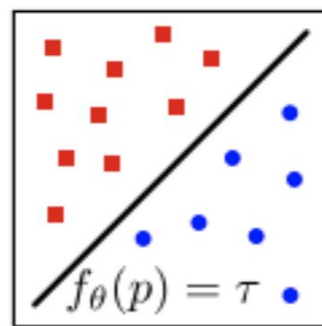
(a) Voxel



(b) Point



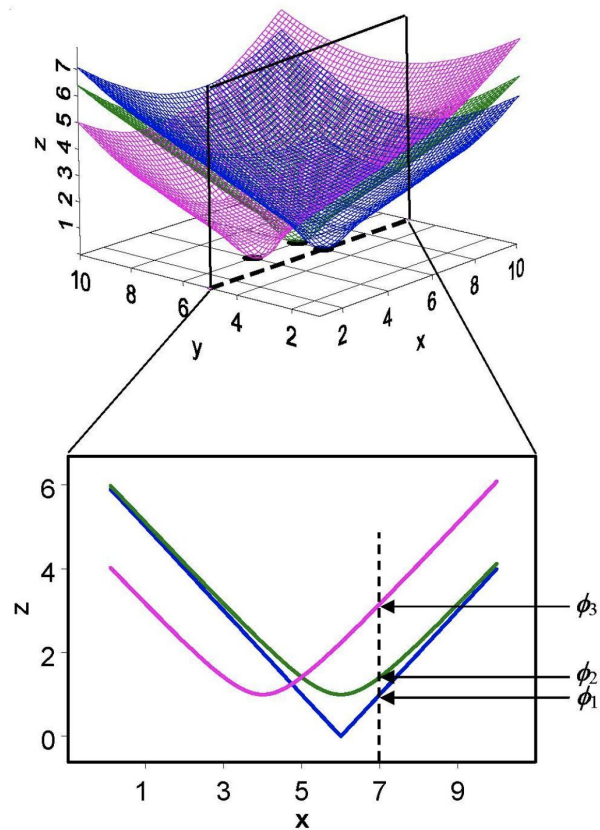
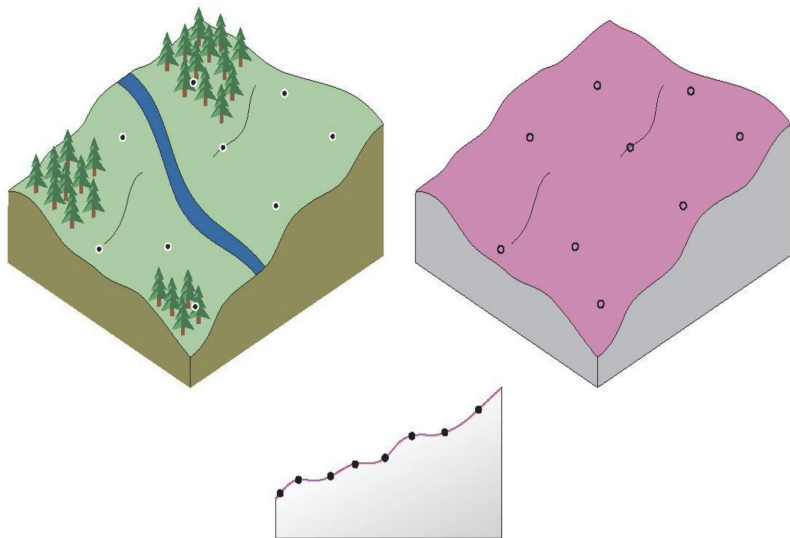
(c) Mesh



(d) Functions

Information

Метод радиальных базисных функций



Occupancy Net



Input

ONet [27]

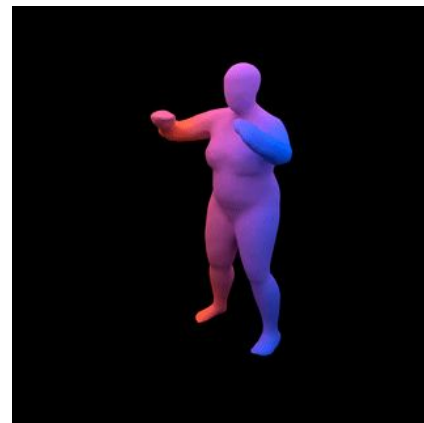
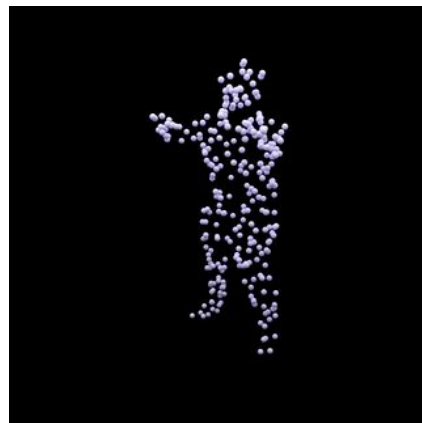
PointConv

Ours-2D



Ours-2D

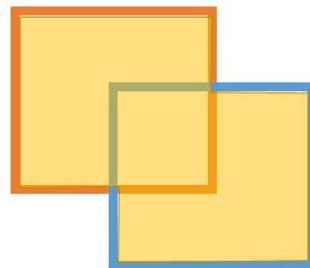
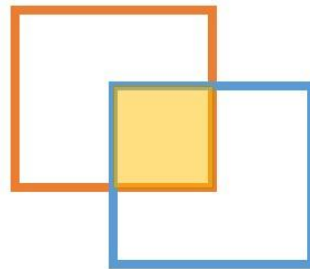
Ours-3D

GT mesh

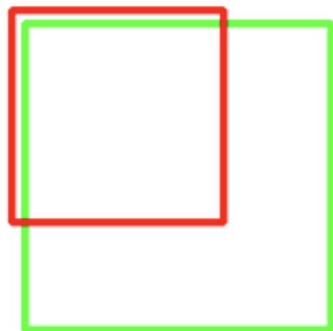


$$\textit{Intersection over Union (IoU)} = \frac{\textit{Area of Overlap}}{\textit{Area of Union}}$$

 Prediction
 Ground-truth

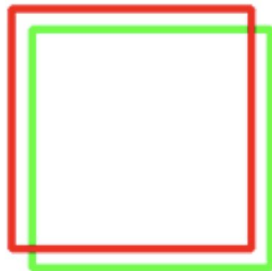


IoU: 0.4034



Poor

IoU: 0.7330



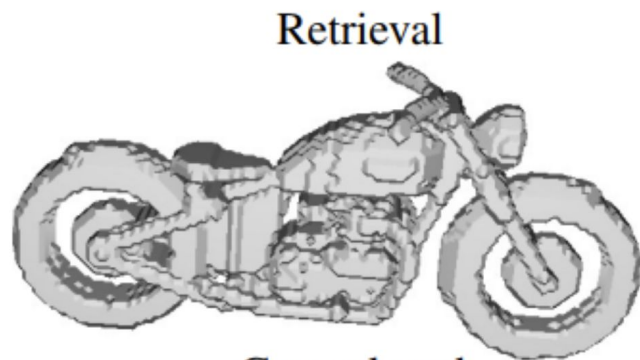
Good

IoU: 0.9264

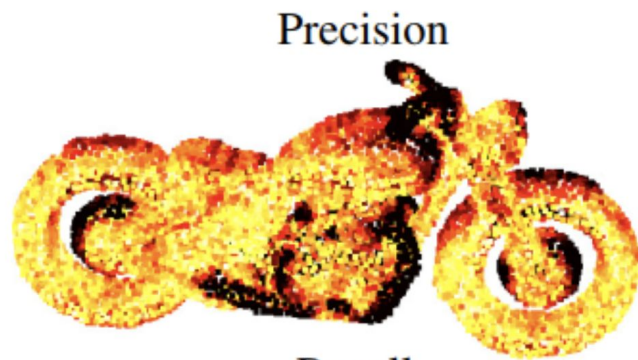
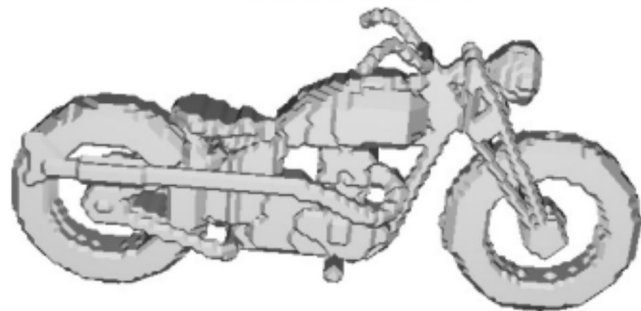


Excellent

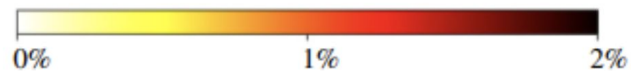
F1 score



Ground truth



Recall



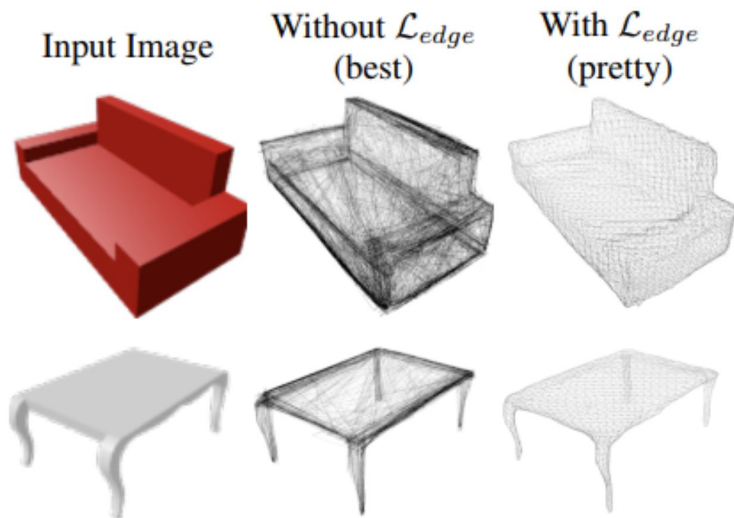
Chamfer loss

$$\Lambda_{P,Q} = \{(p, \arg \min_q \|p - q\|) : p \in P\}$$

$$\mathcal{L}_{\text{cham}}(P, Q) = |P|^{-1} \sum_{(p,q) \in \Lambda_{P,Q}} \|p - q\|^2 + |Q|^{-1} \sum_{(q,p) \in \Lambda_{Q,P}} \|q - p\|^2$$

Edge regularizer (loss)

$$\mathcal{L}_{\text{edge}}(V, E) = \frac{1}{|E|} \sum_{(v, v') \in E} \|v - v'\|^2, \quad E \subseteq V \times V$$

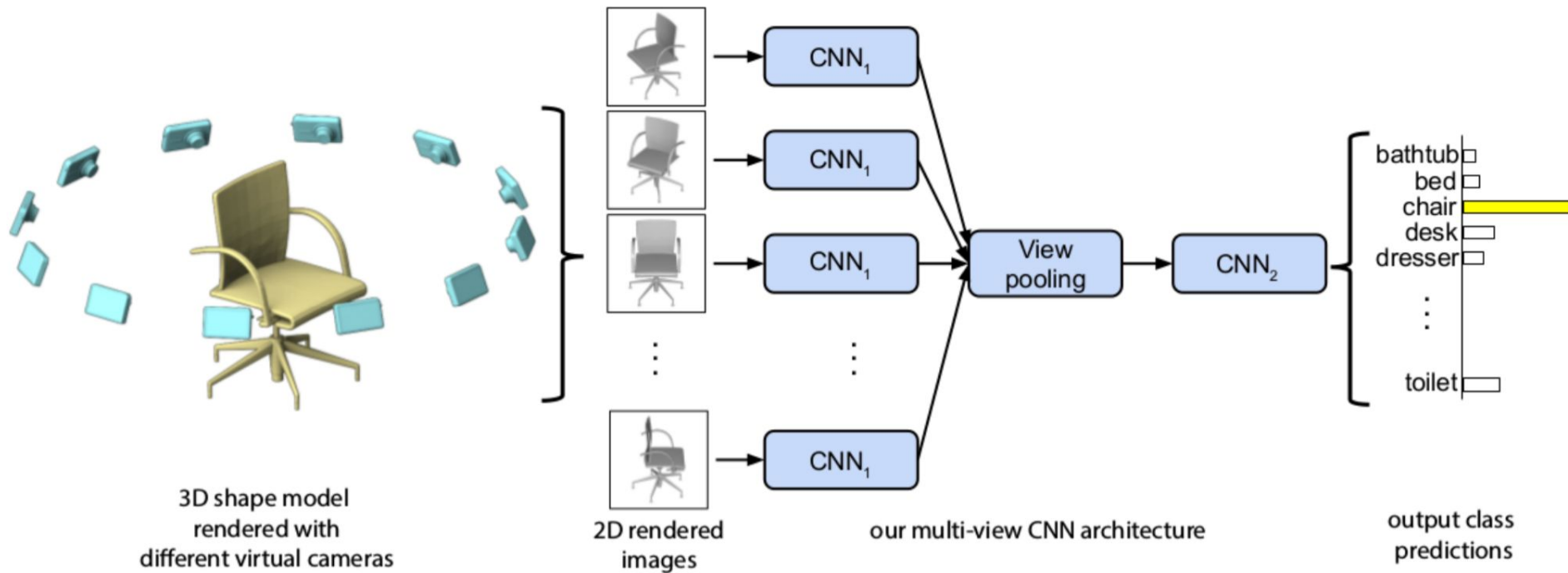


Smooth regularizer (loss)

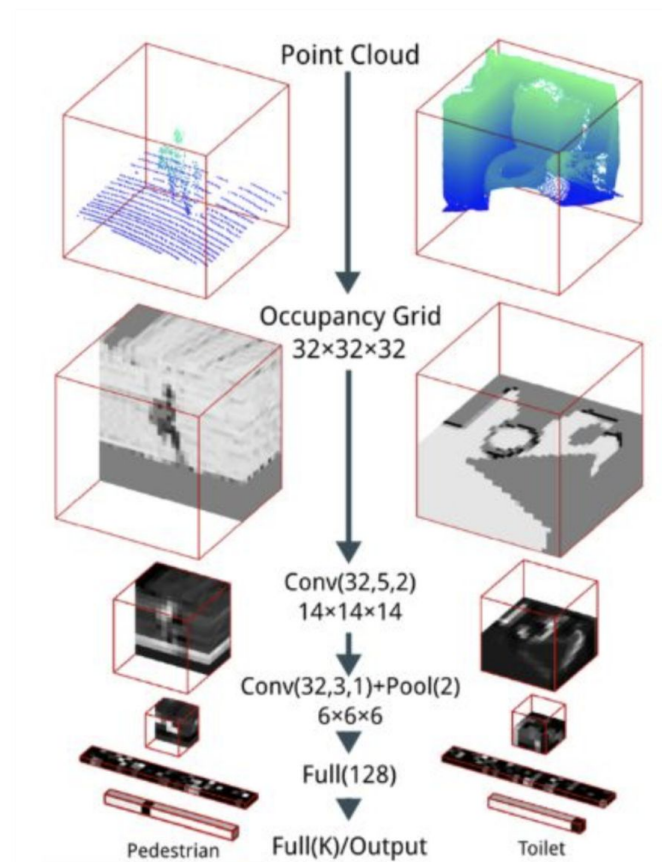
$$\mathcal{L}_{\text{sm}}(x) = \sum_{\theta_i \in \mathcal{E}} (\cos \theta_i + 1)^2$$



Multi-view Convolutional Neural Networks for 3D Shape Recognition (2015)

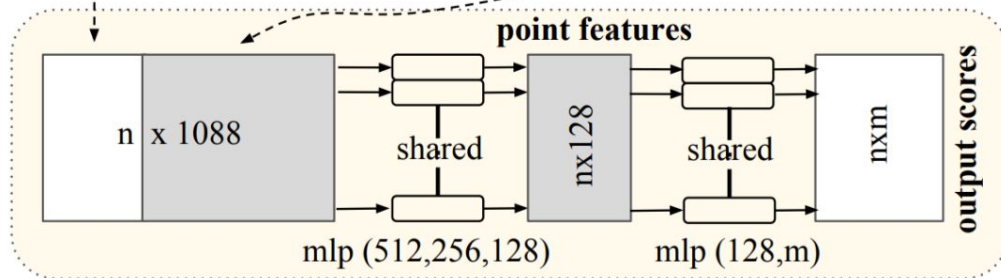
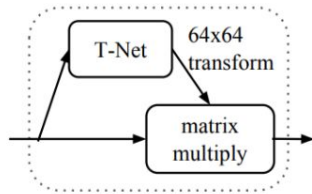
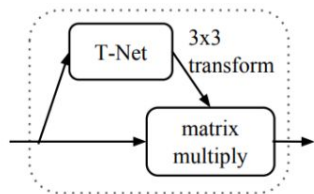
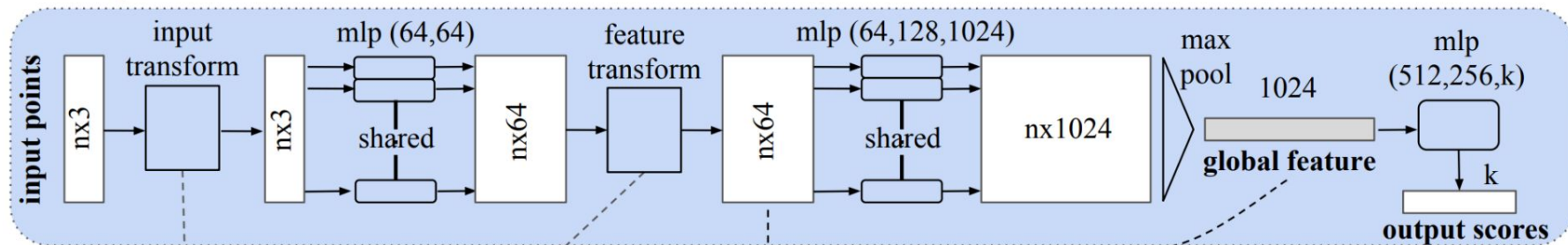


VoxNet (2015)



PointNet (2016)

Classification Network



Segmentation Network

	input	#views	accuracy avg. class	accuracy overall
SPH [11]	mesh	-	68.2	-
3DShapeNets [28]	volume	1	77.3	84.7
VoxNet [17]	volume	12	83.0	85.9
Subvolume [18]	volume	20	86.0	89.2
LFD [28]	image	10	75.5	-
MVCNN [23]	image	80	90.1	-
Ours baseline	point	-	72.6	77.4
Ours PointNet	point	1	86.2	89.2

Table 1. **Classification results on ModelNet40.** Our net achieves state-of-the-art among deep nets on 3D input.

	#params	FLOPs/sample
PointNet (vanilla)	0.8M	148M
PointNet	3.5M	440M
Subvolume [18]	16.6M	3633M
MVCNN [23]	60.0M	62057M

Источники:

<https://github.com/timzhang642/3D-Machine-Learning>

<https://habr.com/ru/company/itmai/news/t/548734/>

<https://studfile.net/preview/6226531/page:10/>

<https://medium.com/@nabil.madali/introduction-to-3d-deep-learning-740c199b100c>

<https://medium.com/phygitalism/3d-neural-networks-review-6a5908439d54>

<https://medium.com/phygitalism/3d-ml-metrics-loss-functions-9708ff0476e2>