3D COMPUTER VISION: INTRO

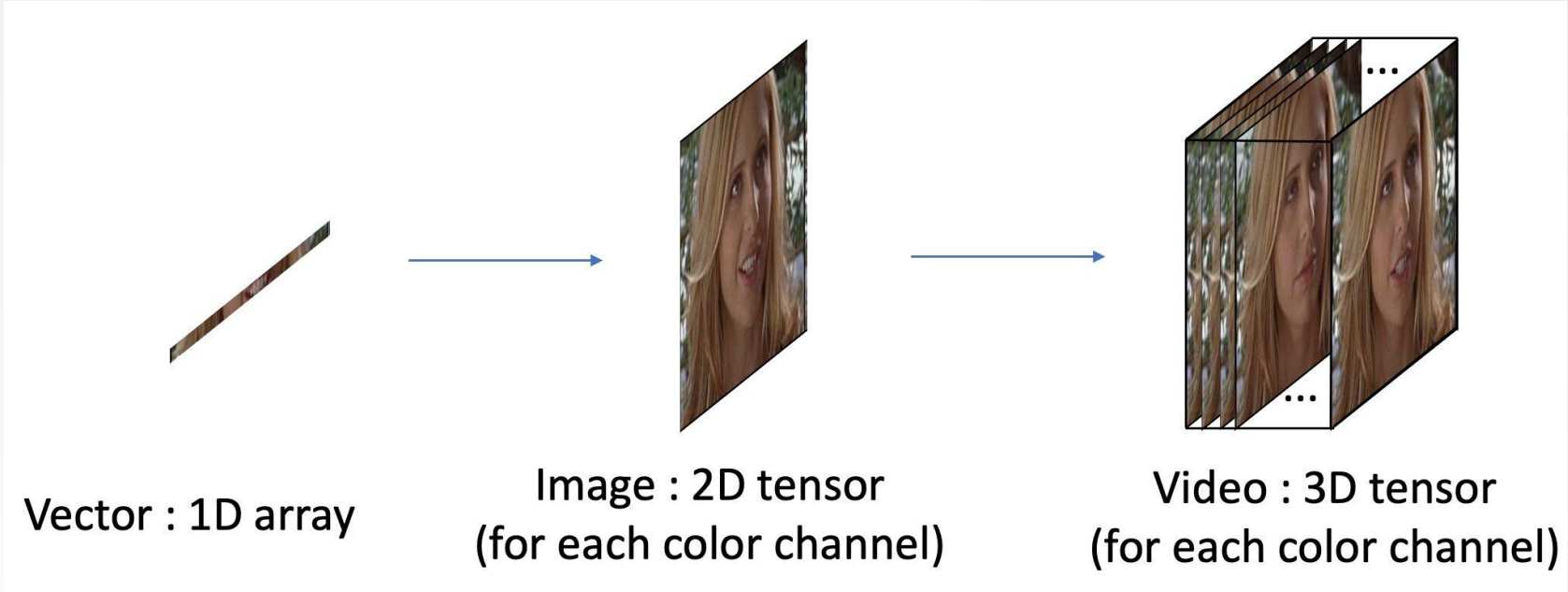
30.04.2024 Computer Vision Seminar 23/24

Agenda

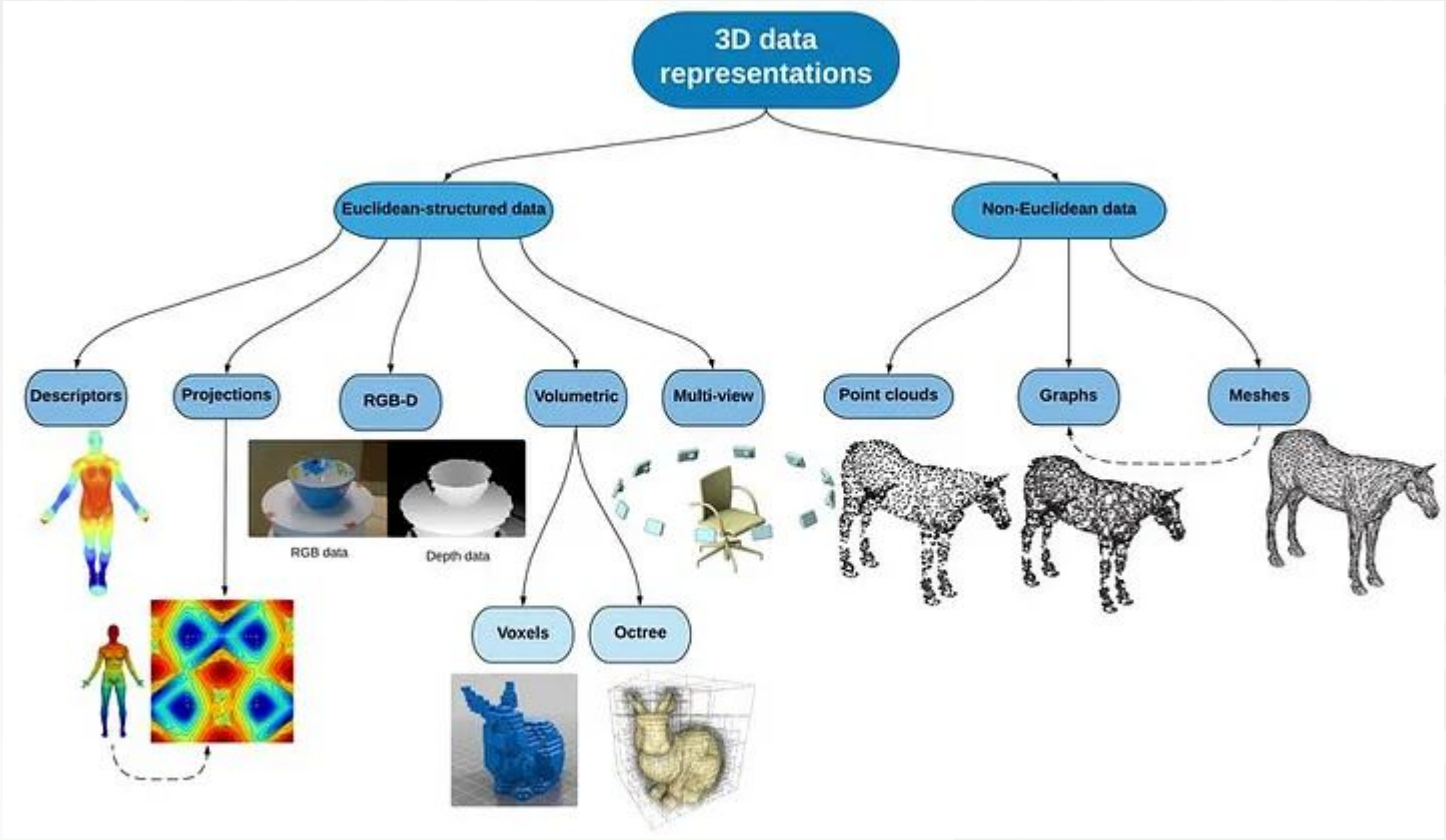
1. Wizja komputerowa w 2D a 3D
2. Reprezentacje danych 3D
3. Typy segmentacji
4. Segmentacja z użyciem DL

[3D Machine Learning] - 3D data representations — Antoine Toisoul

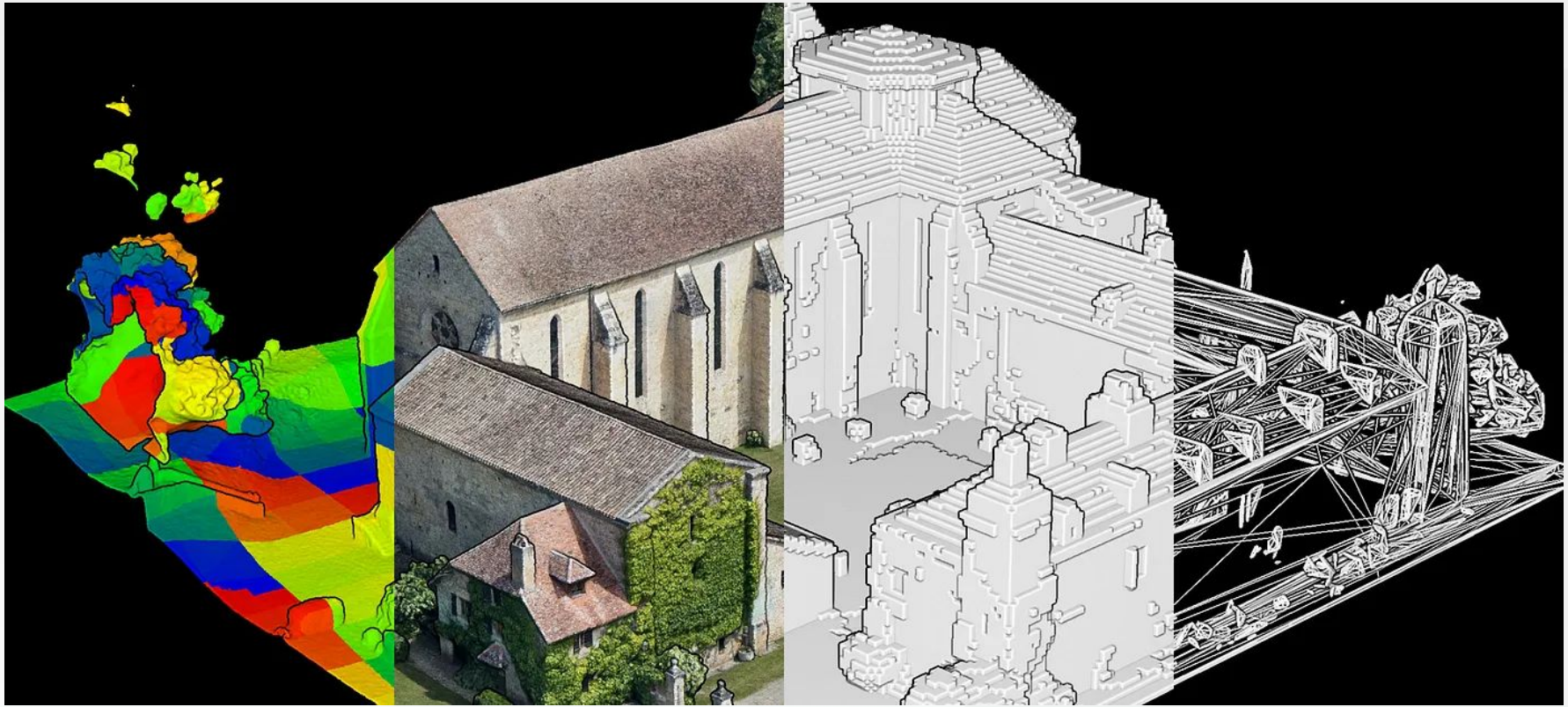
Tensory



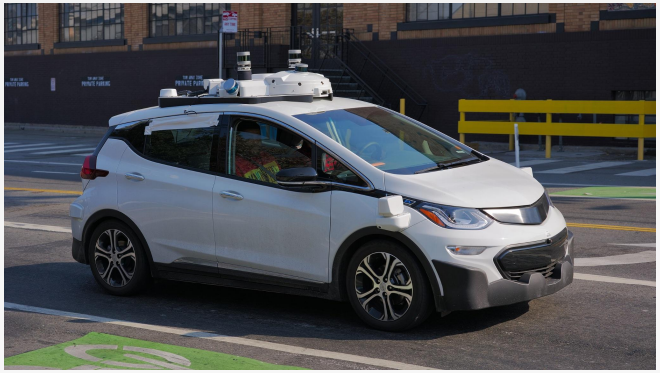
From Pixels to 3D Shapes: An Overview of 3D Data Representations | by Dedeepya Lekkala | Medium

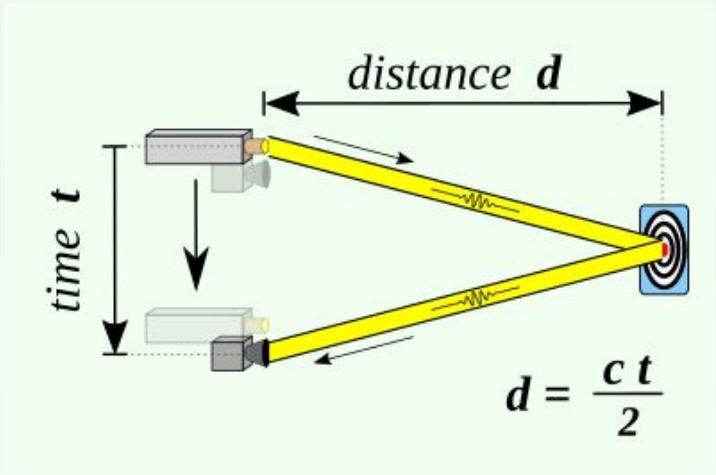


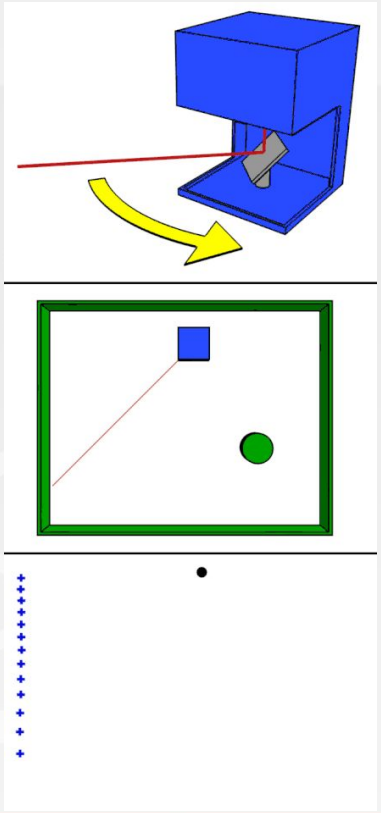
How to represent 3D Data



LiDAR – light detection and ranging



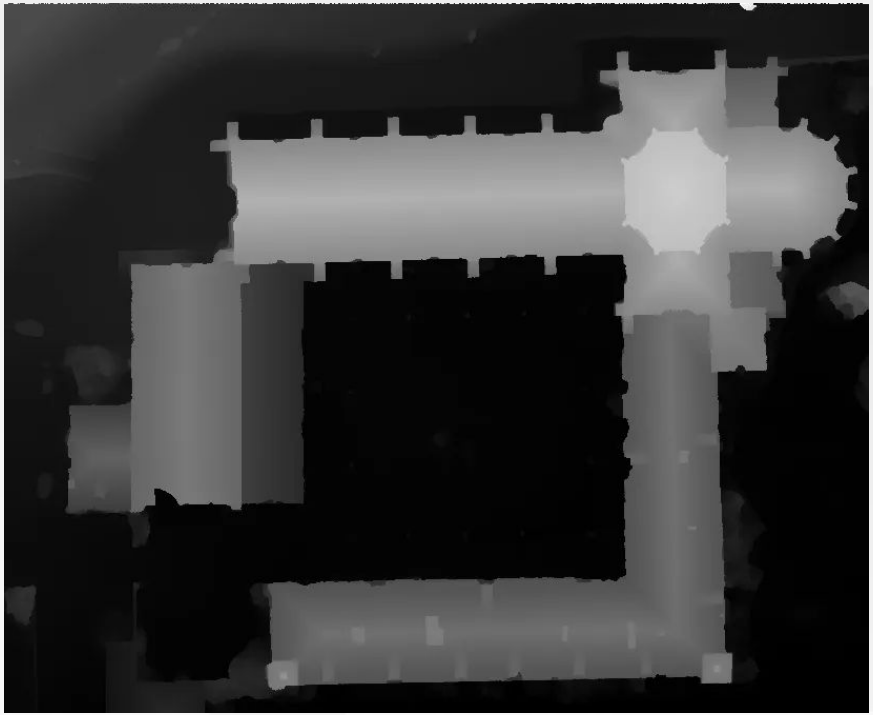




Lidar - Wikipedia

Depth map

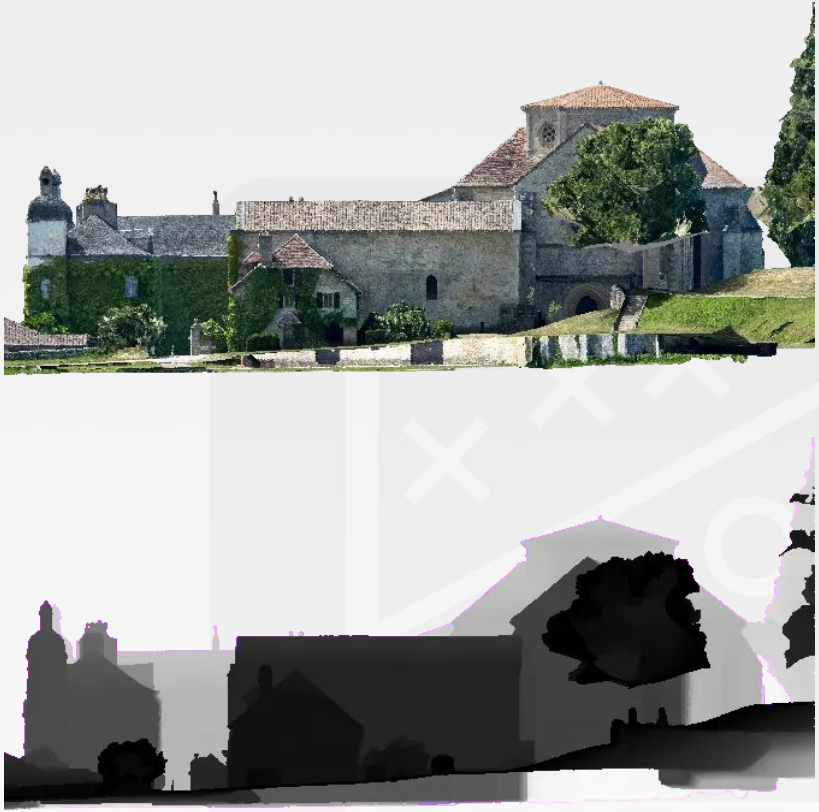
- Zdjęcie/kanał
- Jasność pikseli oznacza głębię (im jaśniej, tym bliżej znajduje się obiekt)



How to represent 3D Data

RGB-D

- Połączenie informacji o kolorze (RGB) z głębią (D)
- Informacja 2,5D
- Prosta, popularna reprezentacja



How to represent 3D Data

Projekcje

- Istnieje bardzo dużo projekcji tego samego obiektu
- Sferyczne/birds-eye
- Strata informacji w wyniku projekcji

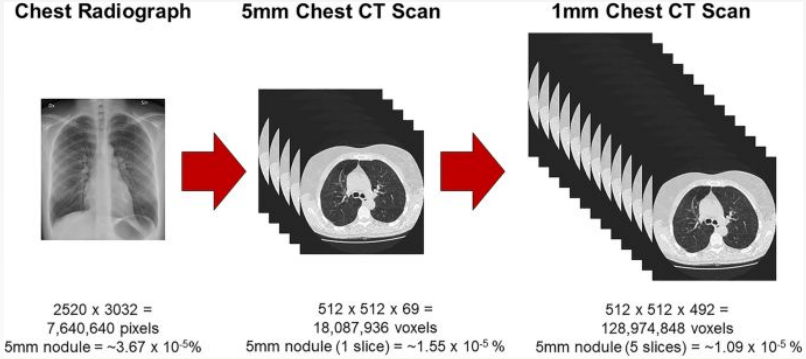
Multi-view

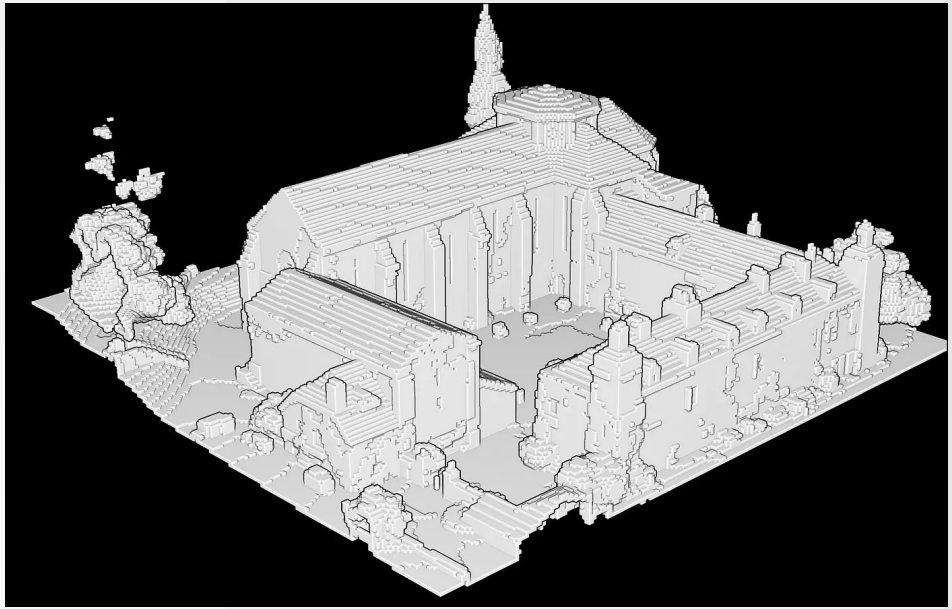
- Sposób na szum, niepełną informację, okluzję (occlusion)
- Trudne pytanie: ile potrzeba takich ujeć?

How to represent 3D Data

Woksele (voxels, volumetric picture element)

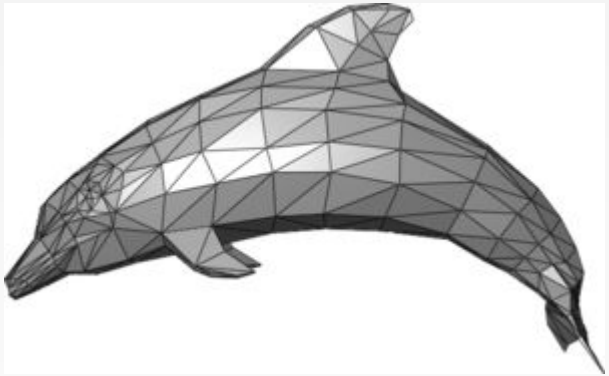
- Odpowiednik piksela w 3D
- Przedstawienie sceny jako trójwymiarowej tablicy
- Przechowują informację o tym, co jest w środku bryły (solid modelling)





Siatka (mesh)

- Wykorzystuje wielokąty połączone ze sobą krawędziami
- Zwykle proste wielokąty, np. trójkąty



Chmury punktów

- Zbiór punktów w układzie trójwymiarowym
- Współrzędne X, Y, Z
- Dane zebrane np. przez LiDAR
- Szybkie renderowanie
- Brak struktury (trudne do przetworzenia przez CNN)

Equivalence of Representations

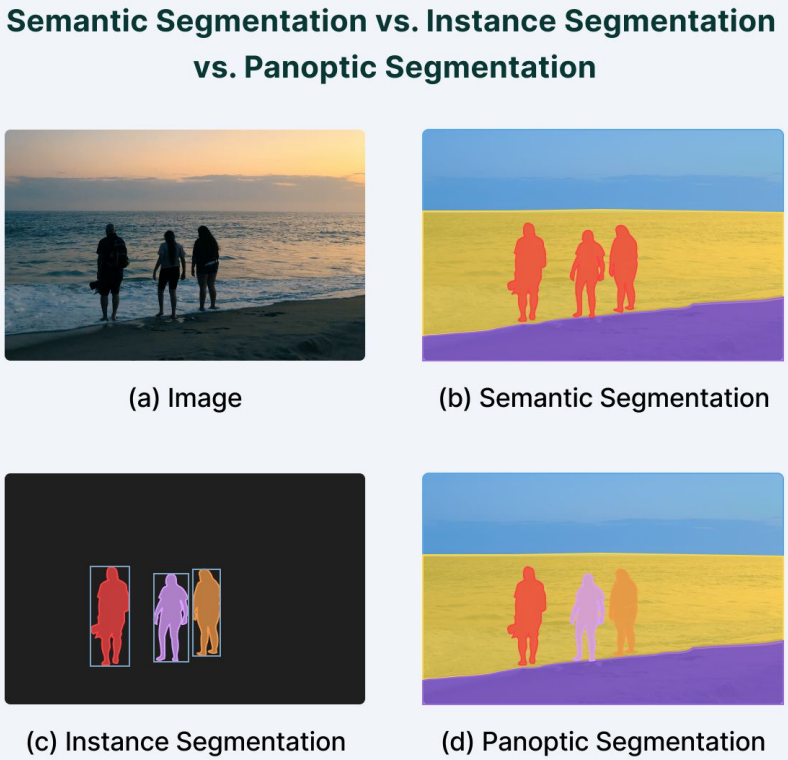
- Thesis:
  - Each fundamental representation has enough expressive power to model the shape of any geometric object
  - It is possible to perform all geometric operations with any fundamental representation!
- Analogous to Turing-Equivalence:
  - All computers today are turing-equivalent, but we still have many different processors

Computational Differences

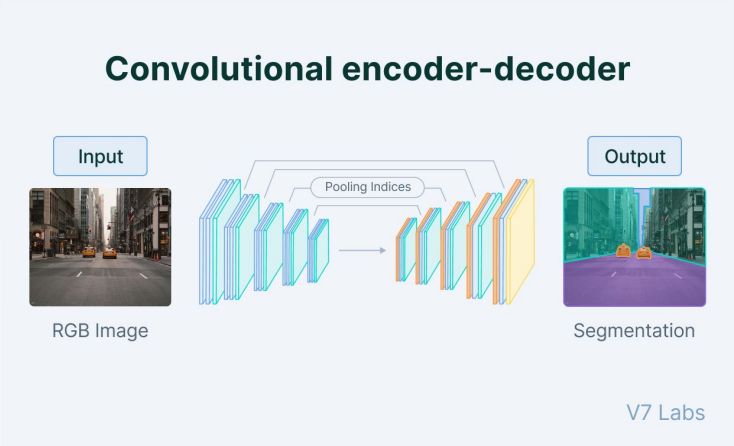
- Efficiency
  - Combinatorial complexity (e.g. O(n log n))
  - Space/time trade-offs (e.g. z-buffer)
  - Numerical accuracy/stability (degree of polynomial)
- Simplicity
  - Ease of acquisition
  - Hardware acceleration
  - Software creation and maintenance
- Usability
  - Designer interface vs. computational engine

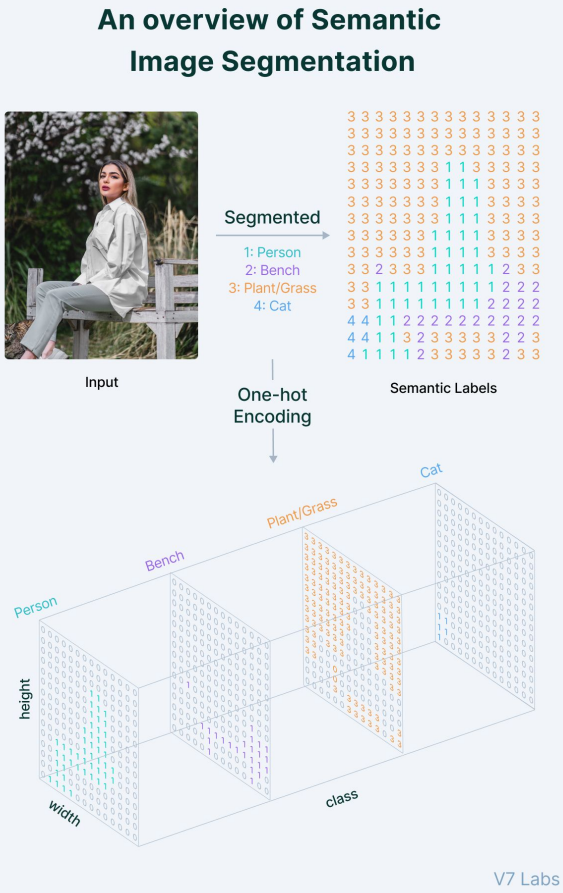
Typy segmentacji

[Image Segmentation: Deep Learning vs Traditional [Guide]]



Segmentacja z użyciem głębokich sieci neuronowych

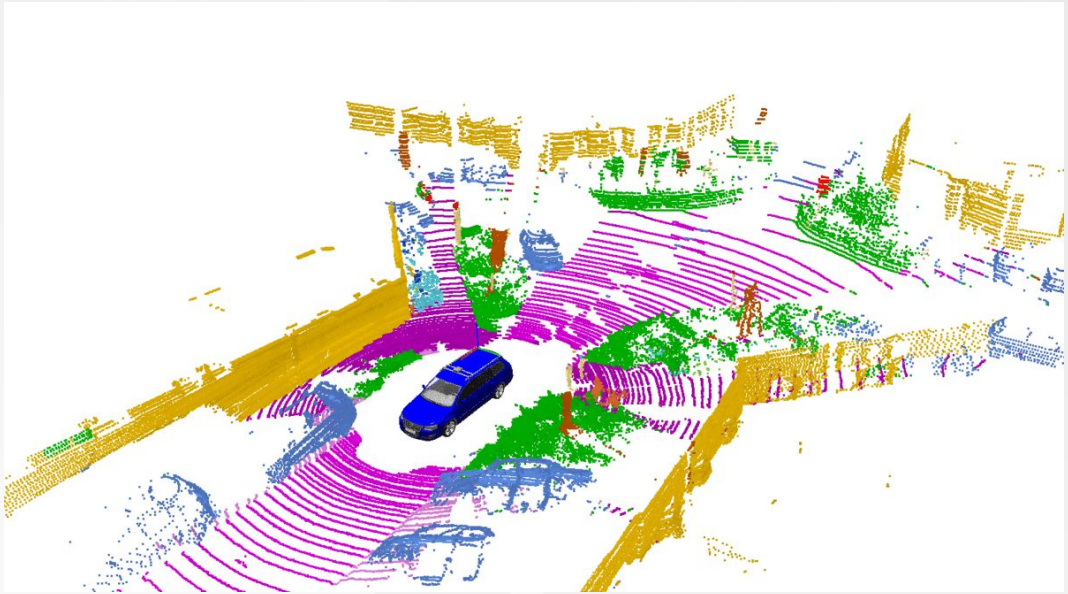




[Image Segmentation: Deep Learning vs Traditional [Guide]]

https://github.com/PRBonn/lidar-bonnetal?tab=readme-ov-file

SemanticKITTI dataset



RangeNet++: Fast and Accurate LiDAR Semantic Segmentation

RangeNet++

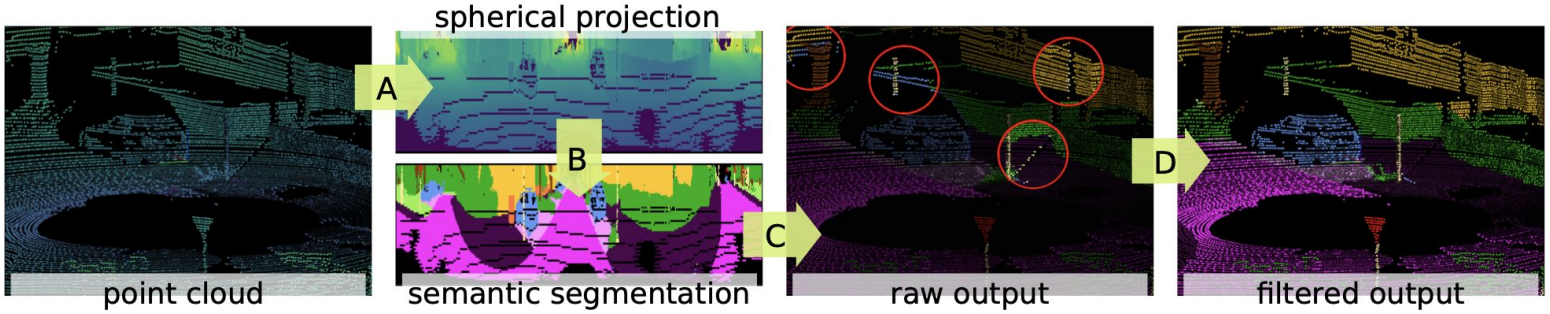
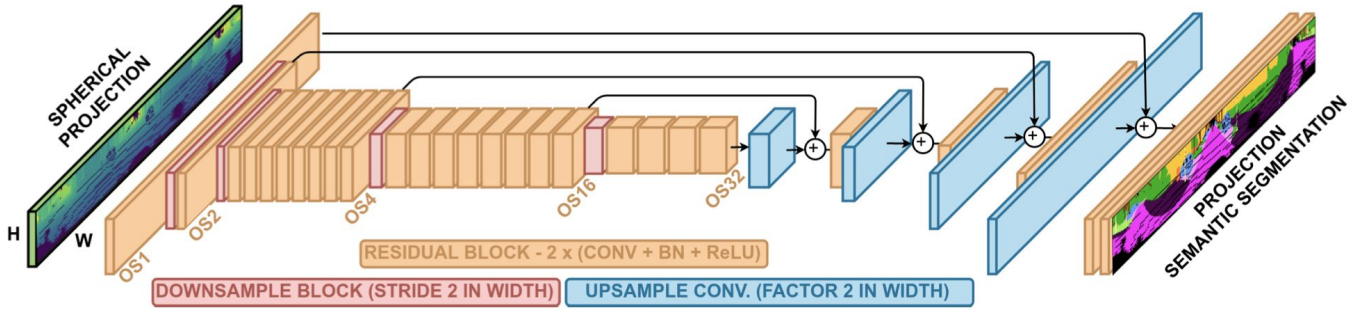
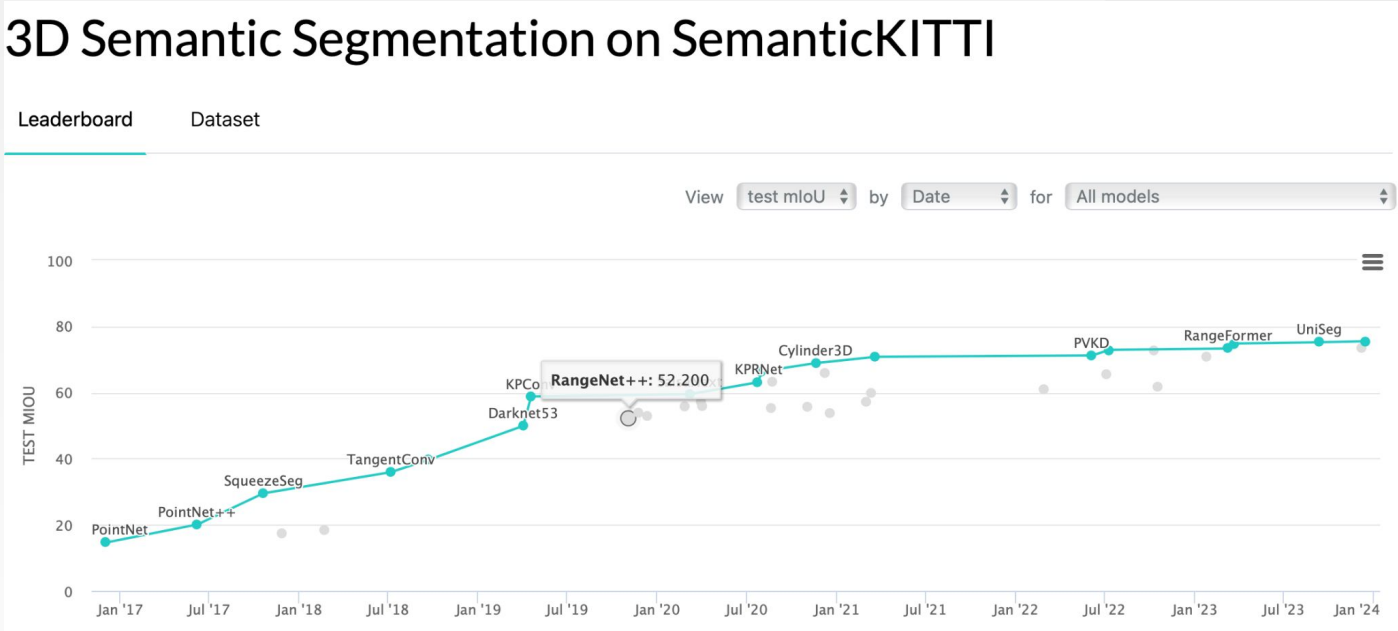


Fig. 2: Block diagram of the approach. Each of the arrows corresponds to one of our modules.



https://arxiv.org/pdf/2312.10035

SOTA semantic segmentation



Materiały

- Overview of 3D Object Representations
- From Pixels to 3D Shapes: An Overview of 3D Data Representations | by Dedeepya Lekkala | Medium
- How to represent 3D Data
- Image Segmentation: Deep Learning vs Traditional [Guide]
- An Intuitive Introduction to Point Net