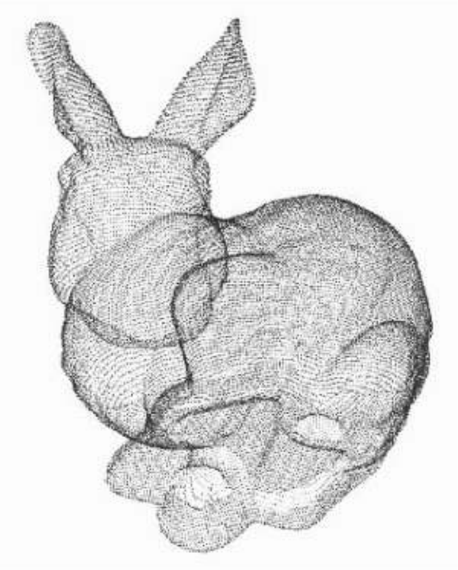
Deep Learning for Augmented Reality

Vincent Lepetit

École des Ponts
ParisTech

Processing Point Clouds

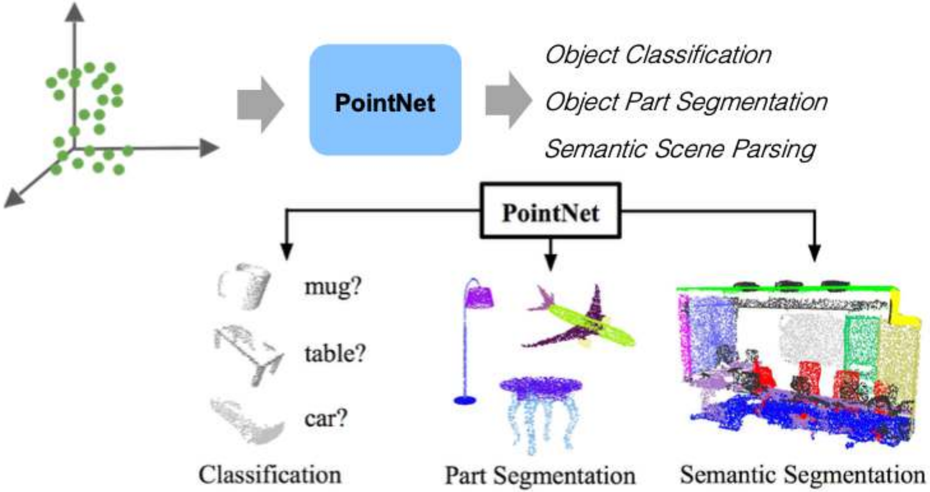


The problem with point clouds

PointNet

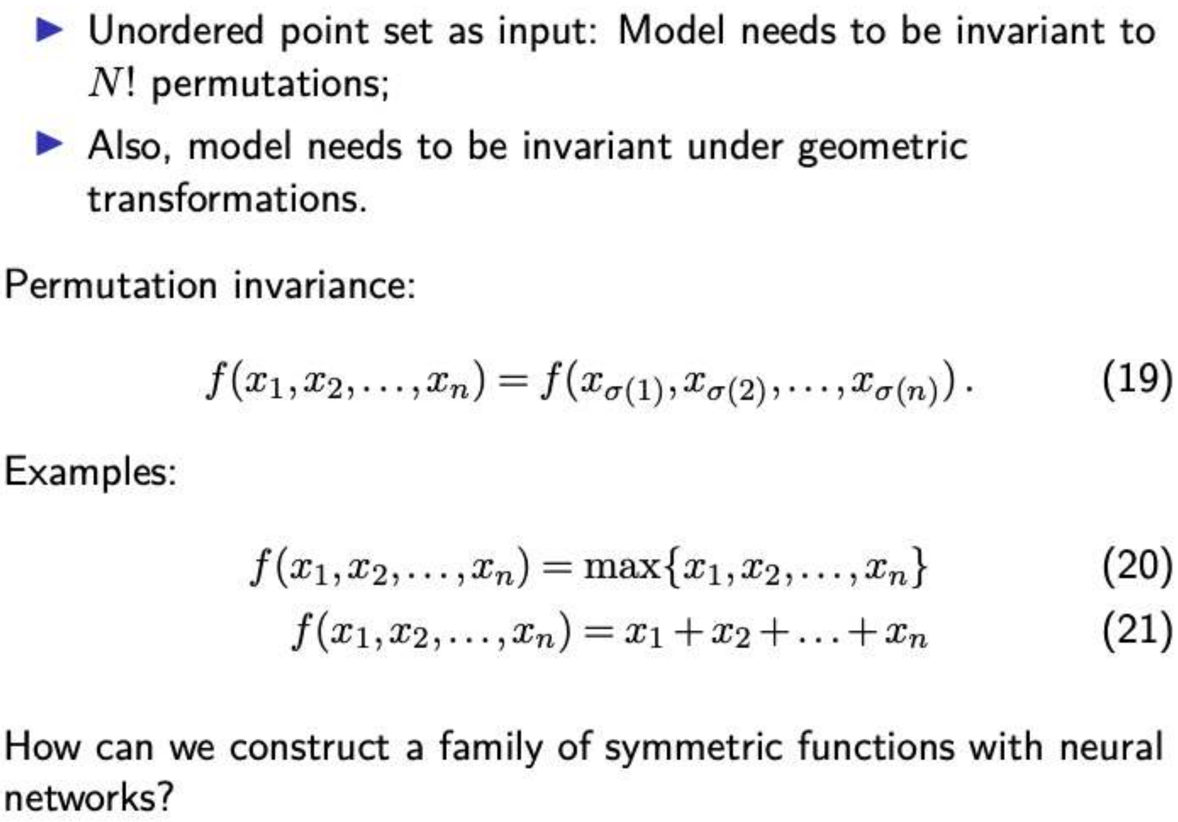
End-to-end learning for **scattered, unordered** point data

Unified framework for various tasks



C.R. Qi et al. “PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation”. In: *Conference on Computer Vision and Pattern Recognition*. 2017.

PointNet



- Unordered point set as input: Model needs to be invariant to $N!$ permutations;
- Also, model needs to be invariant under geometric transformations.

Permutation invariance:

$$f(x_1, x_2, \ldots, x_n) = f(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)}) \,. \tag{19}$$

Examples:

$$f(x_1, x_2, \ldots, x_n) = \max\{x_1, x_2, \ldots, x_n\} \tag{20}$$

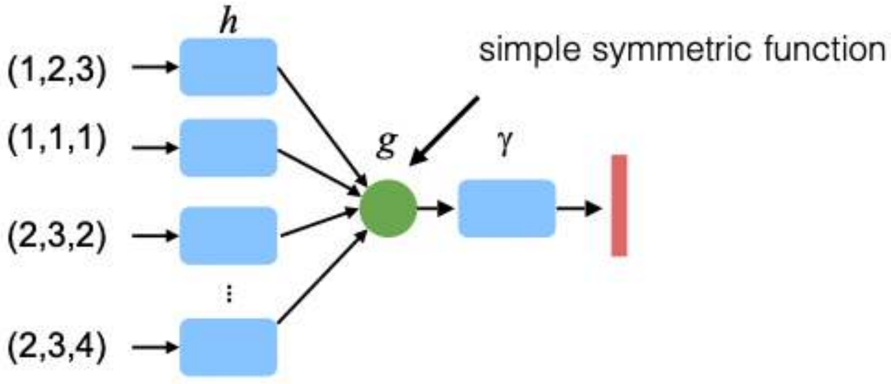
$$f(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \ldots + x_n \tag{21}$$

How can we construct a family of symmetric functions with neural networks?

PointNet: Invariance to ordering

$$f(x_1, x_2, \ldots, x_n) = \gamma(g(h(x_1), h(x_2), \ldots, h(x_n)))$$

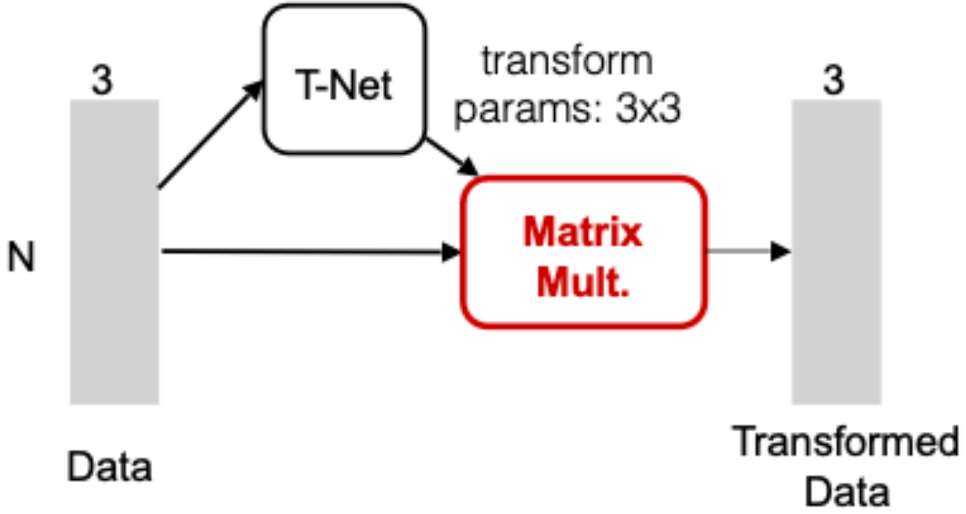
is symmetric if g is symmetric.



Use neural networks for h, γ, and max-pooling for g.

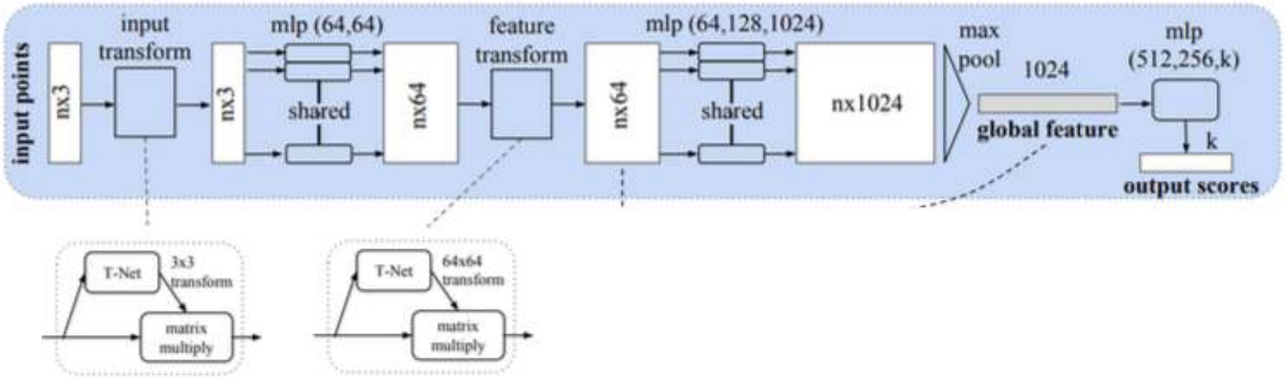
PointNet: Intuition

PointNet: Invariance to geometric transformation



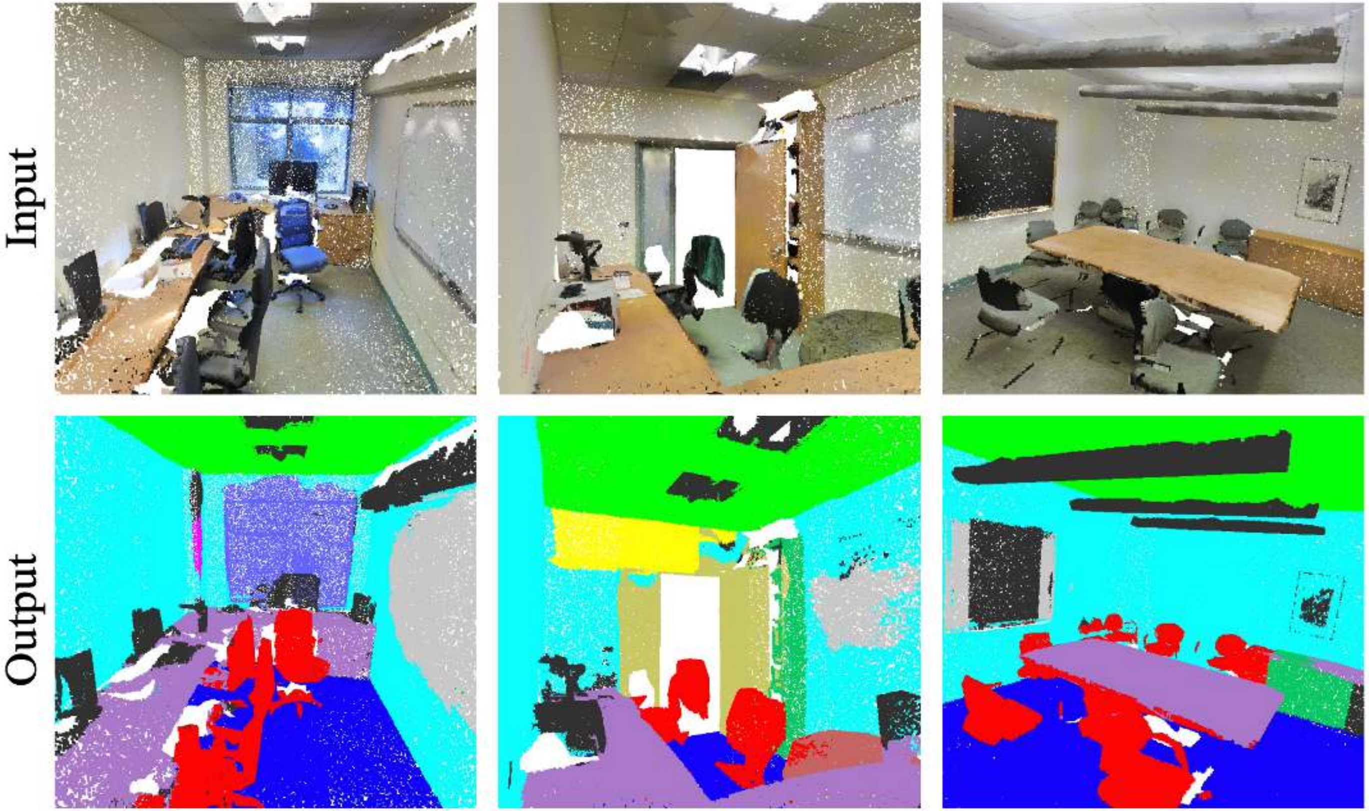
PointNet: Full network

Classification Network

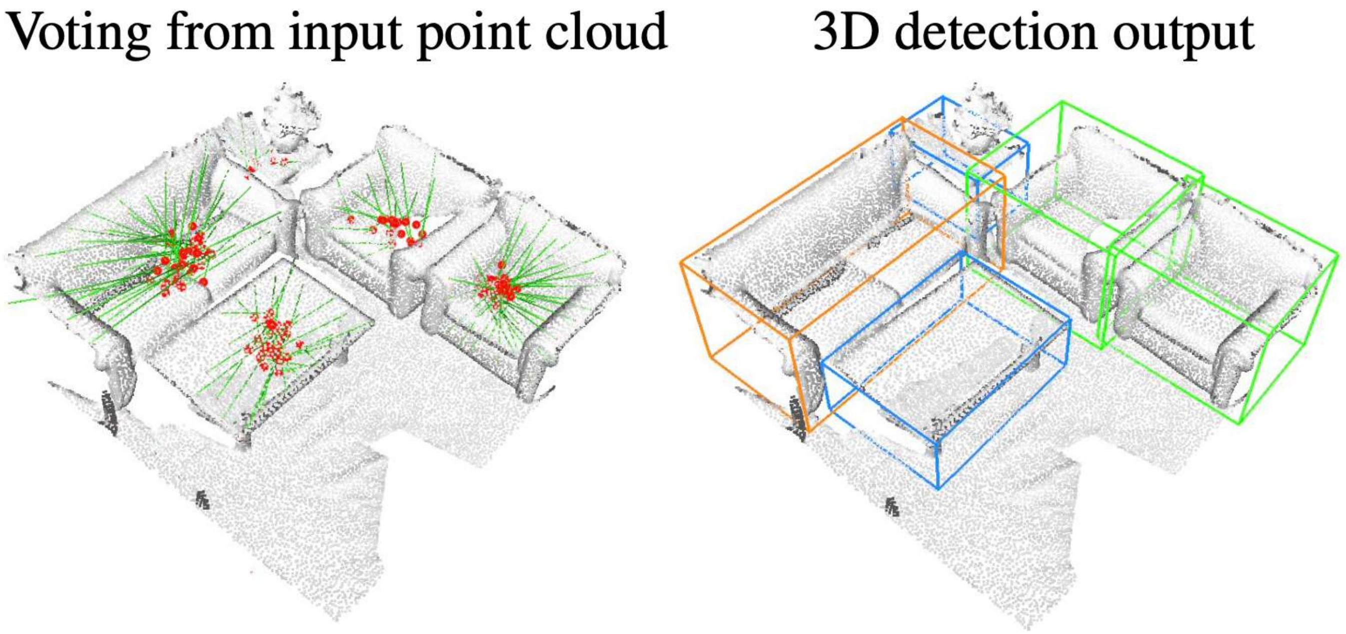


n: number of points;
k: number of possible objects;
m: number of possible segments.

PointNet: some results



VoteNet

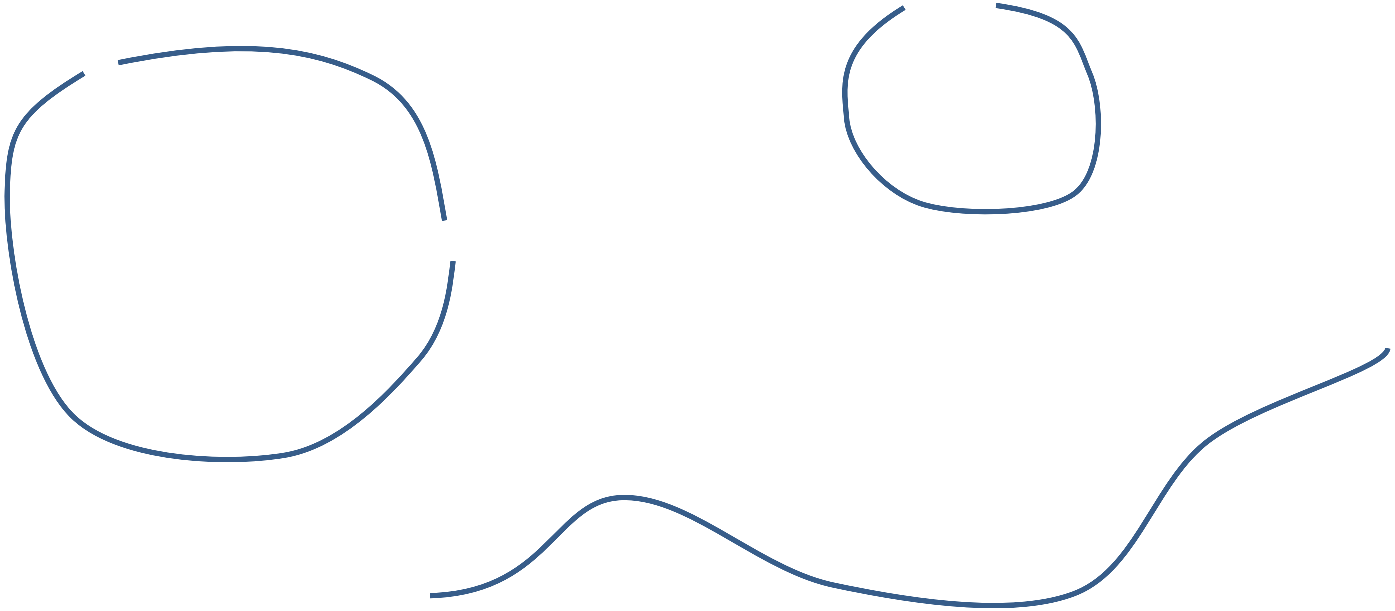


Voting from input point cloud

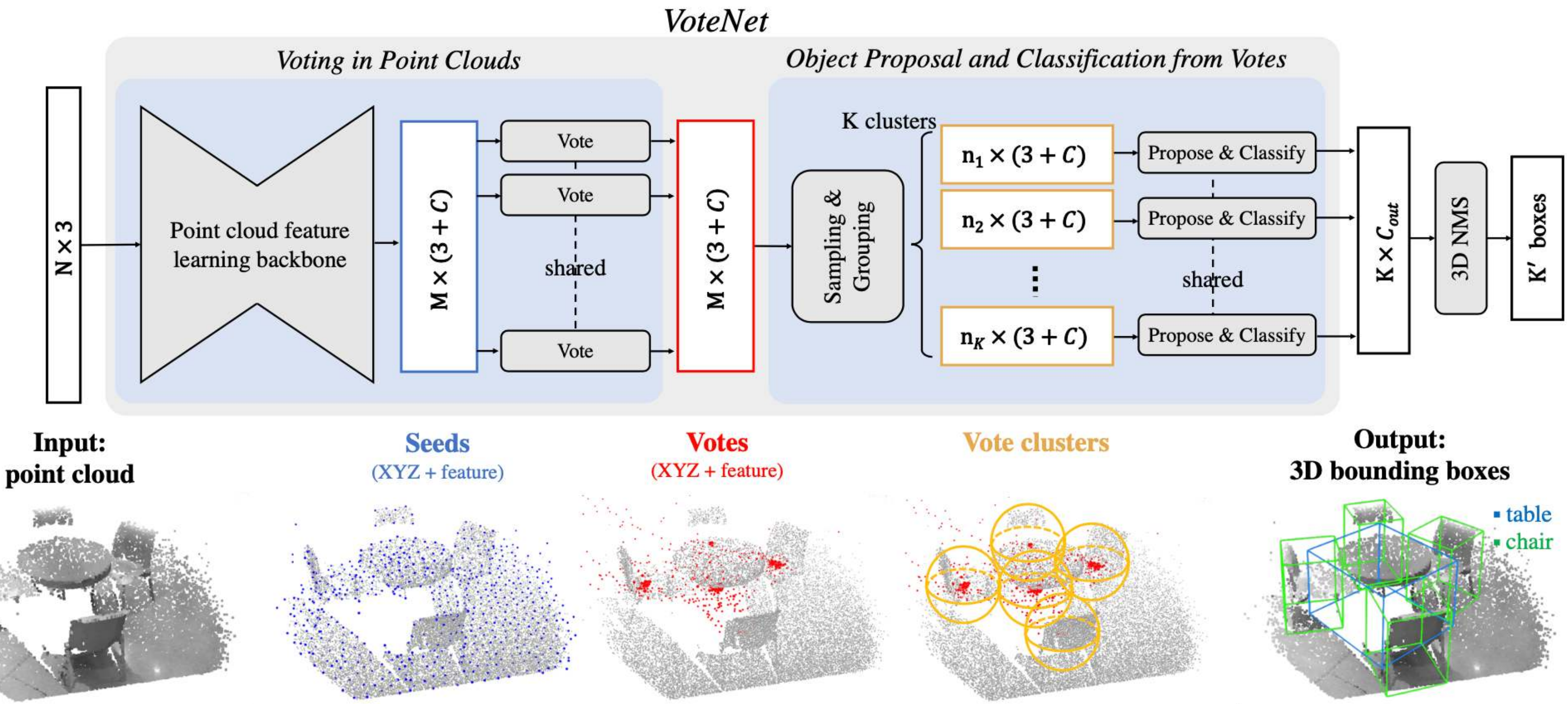
3D detection output

Qi, Charles R and Litany, Or and He, Kaiming and Guibas, Leonidas J}, title = {Deep Hough Voting for 3D Object Detection in Point Clouds}, booktitle = {Proceedings of the IEEE International Conference on Computer Vision}, year = {2019}

Hough Transform (not Deep Learning)

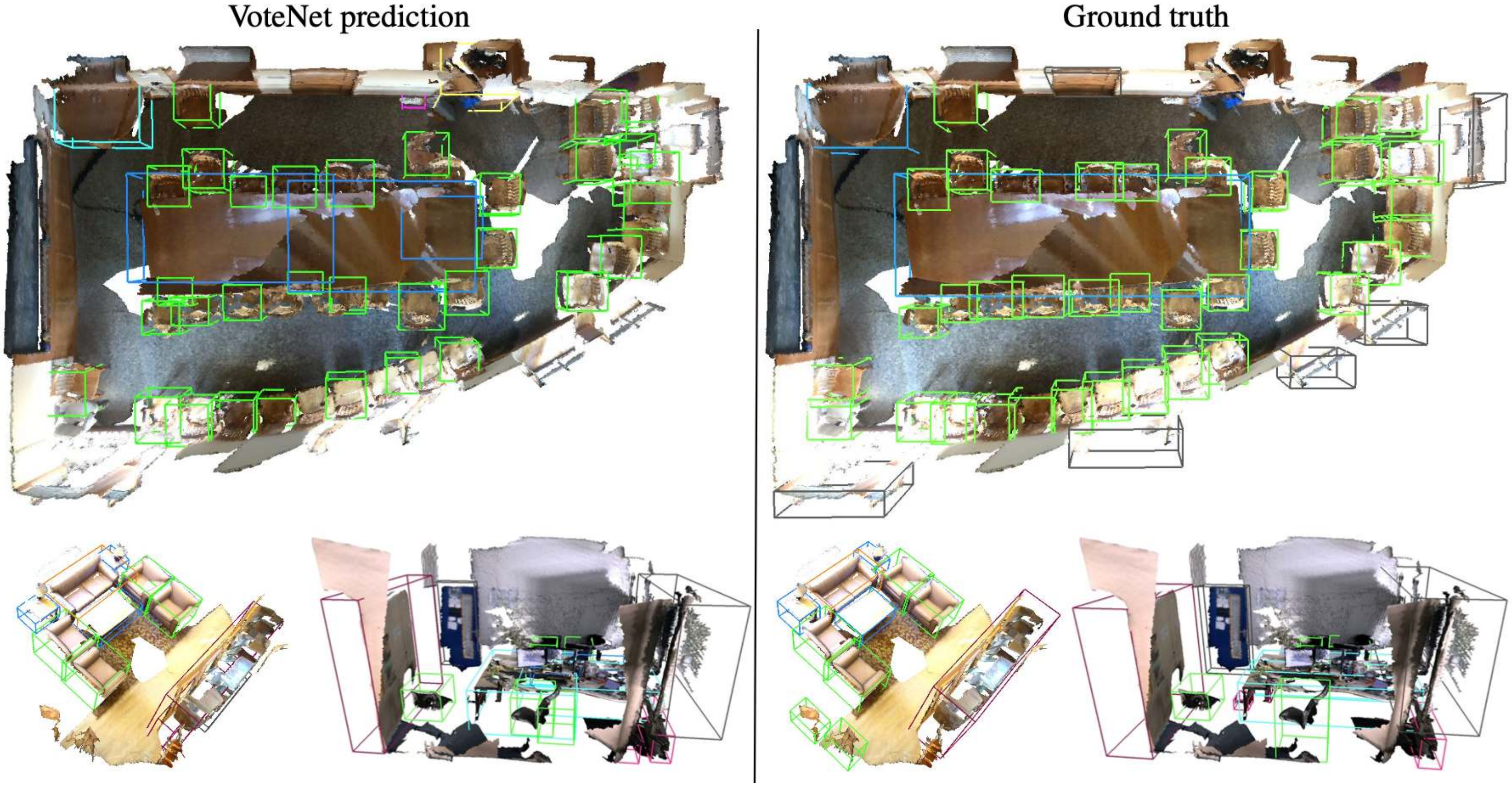


VoteNet



École des Ponts ParisTech

VoteNet: Some results



MinkowskiNet

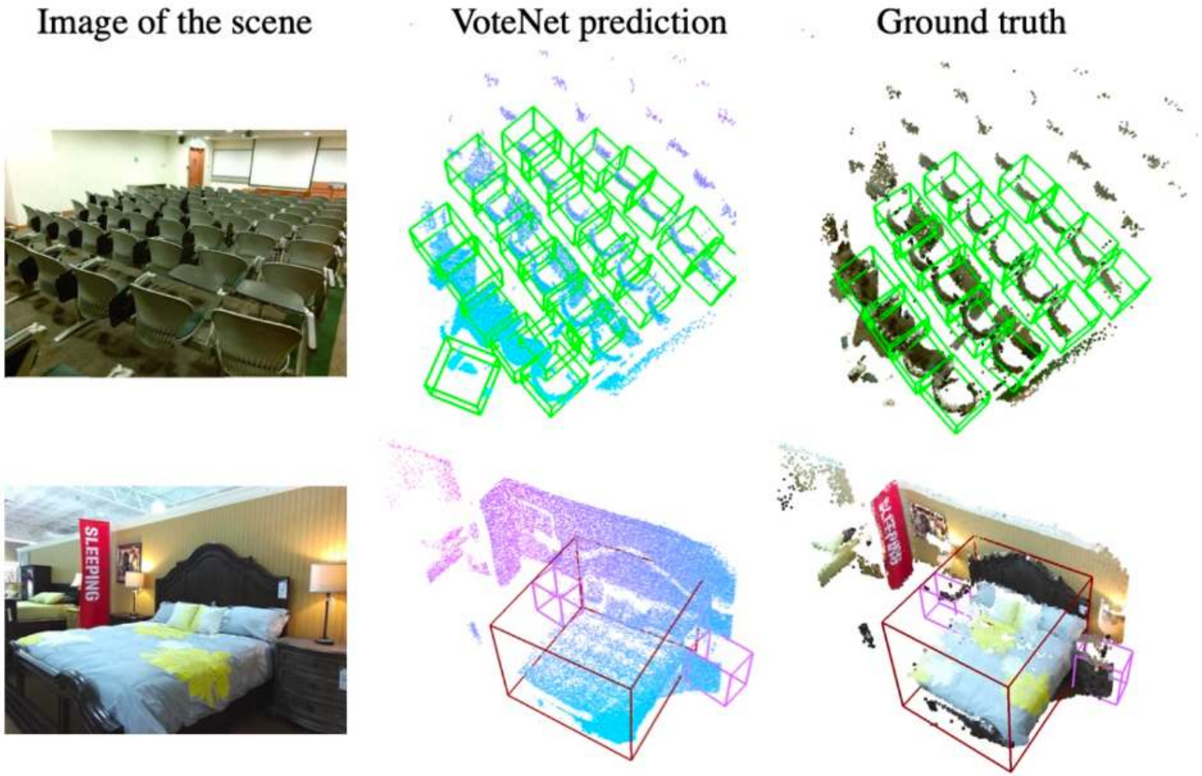


Image of the scene

VoteNet prediction

Ground truth