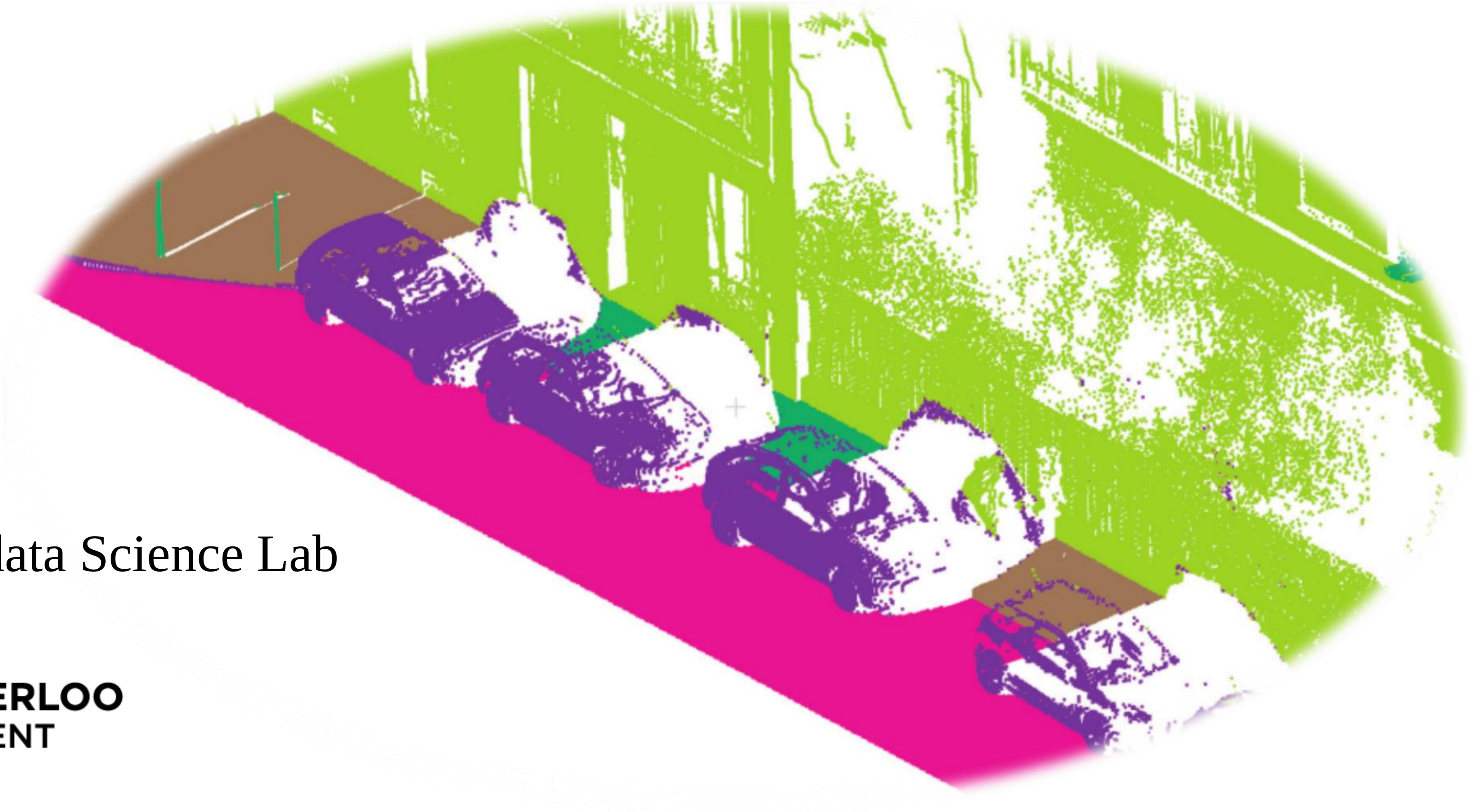
Deep Learning on Large-scale Point Clouds for 3D Segmentation



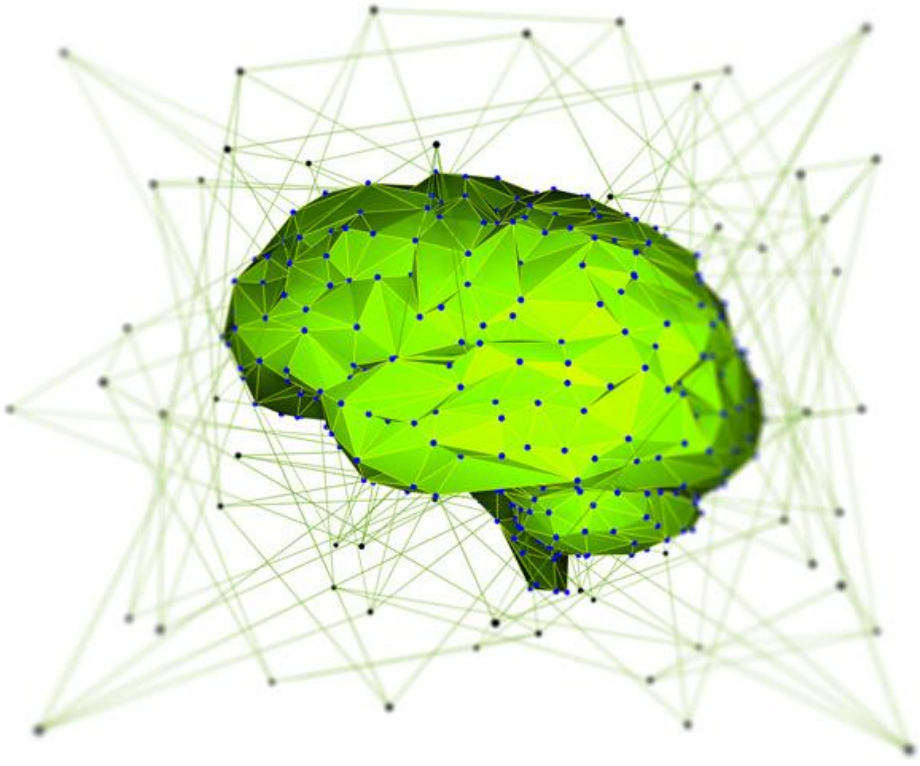
27/03/2019

Presented by: Ying Li

Mobile Sensing And Geodata Science Lab

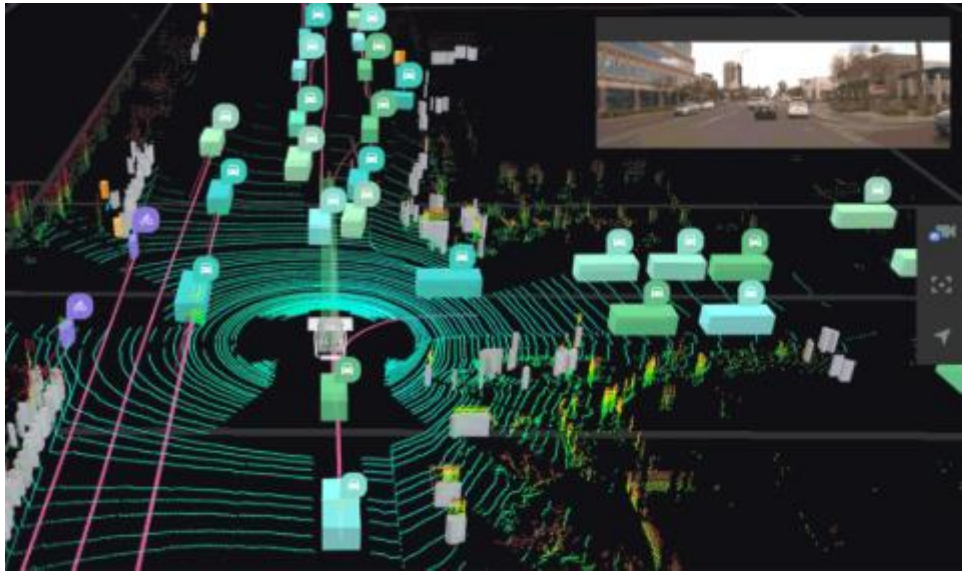
UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

CONTENT

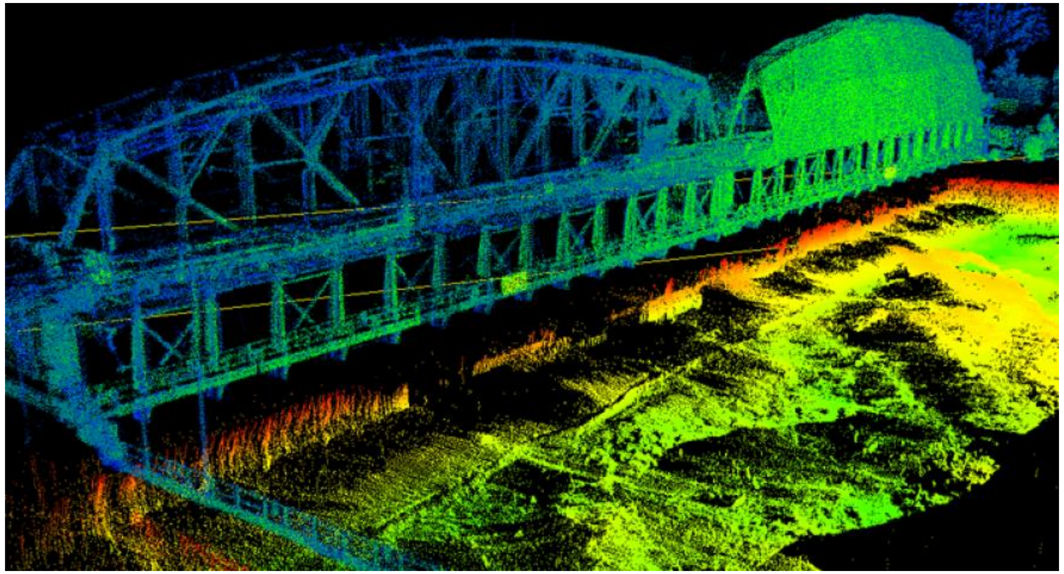


- Introduction
- Challenges
- Model requirements
- Deep learning networks
- Conclusion

Introduction



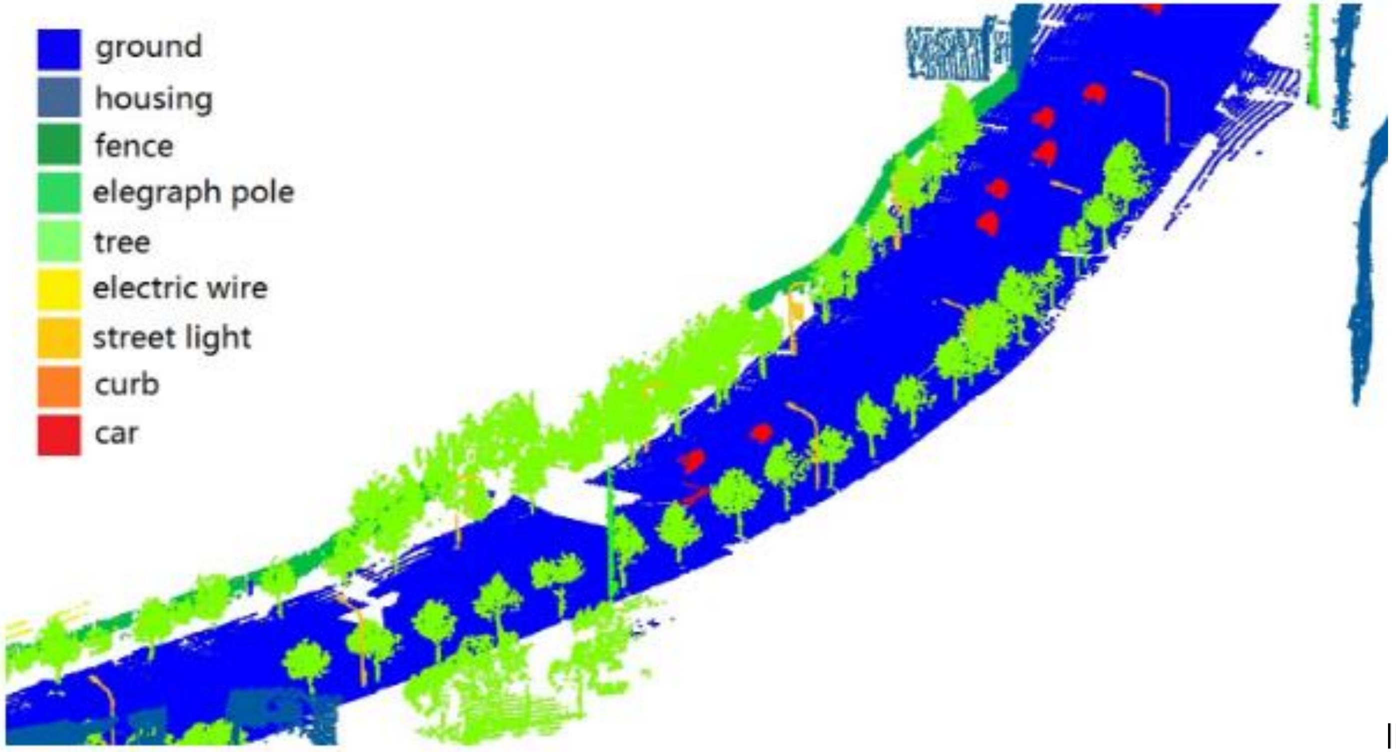
Autonomous Vehicle



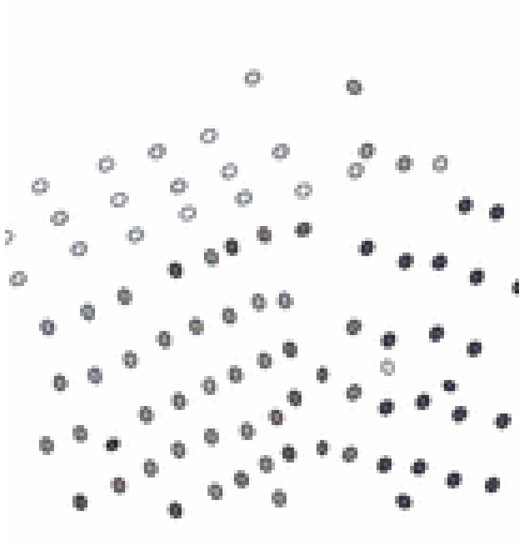
3D Reconstruction

Problem Definition

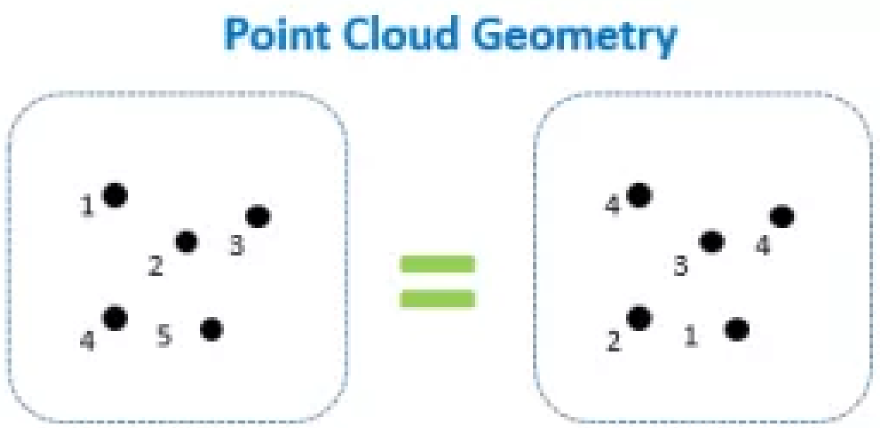
3D Segmentation: Given an arbitrary point cloud data, the goal of segmentation task is to group points with similar geometric attributes into homogeneous region. These regions are labeled with a semantic class, such as ground, tree, building.



Challenges on deep learning models



Unstructured data

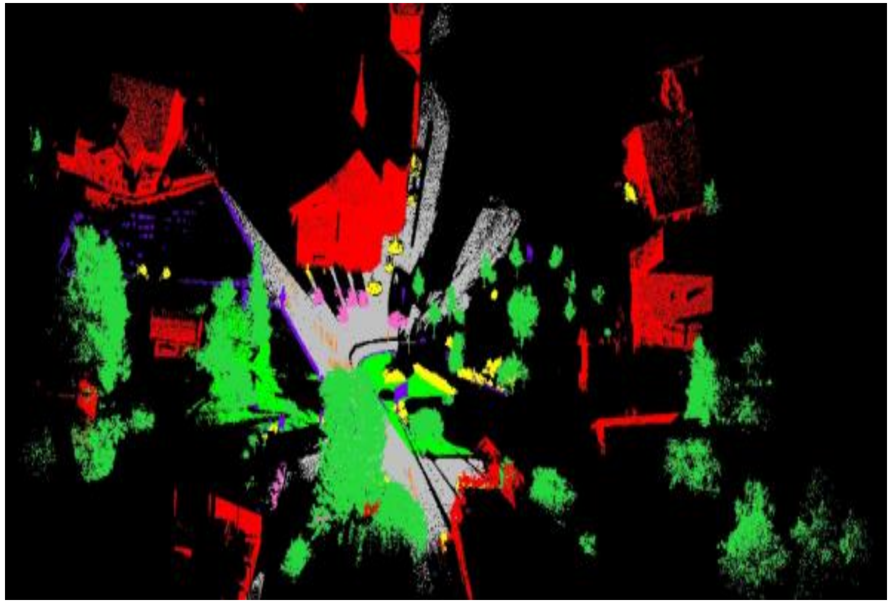


Unordered data

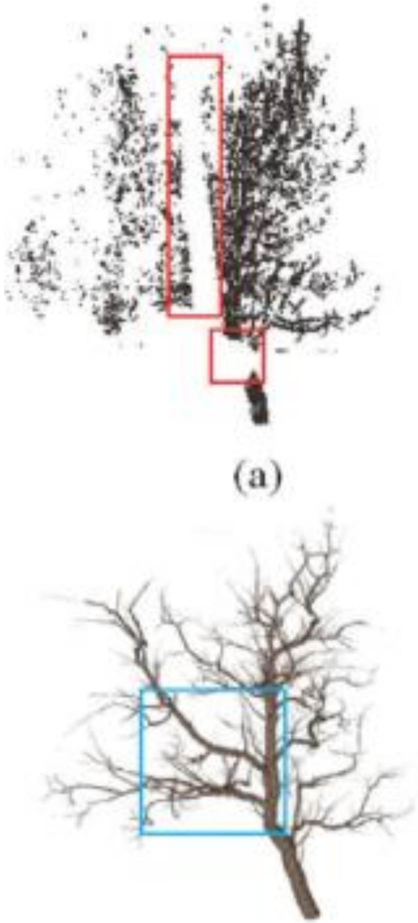
Name	Number of points
Oakland	1.61M
Semantic3D	1660M
Paris-rue-Madame	20M
IQmulus	12M
Paris-Lille-3D	143.1M

Changed number of points

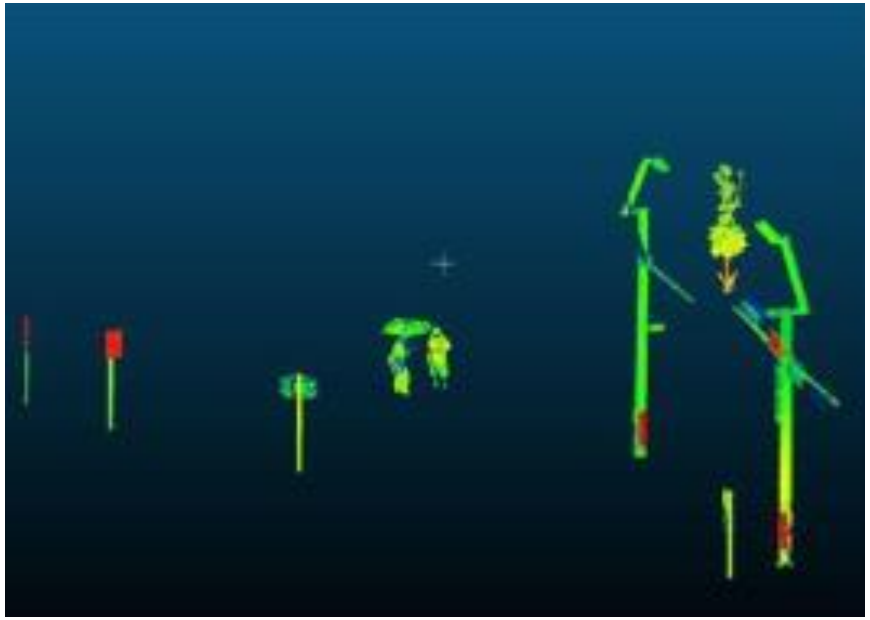
Challenges on point cloud data



Noisy and unevenly distributed data

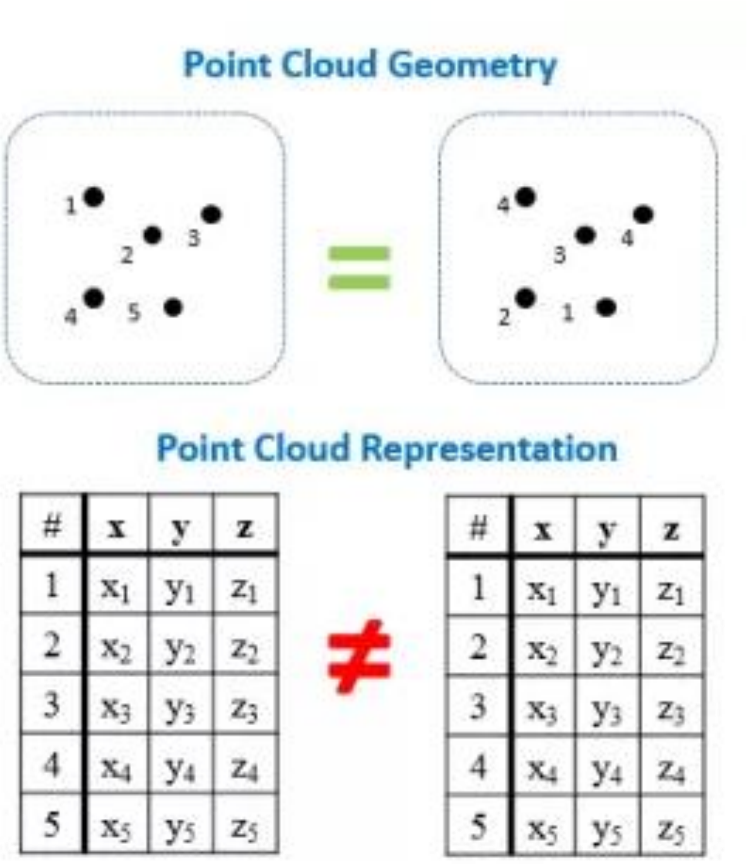


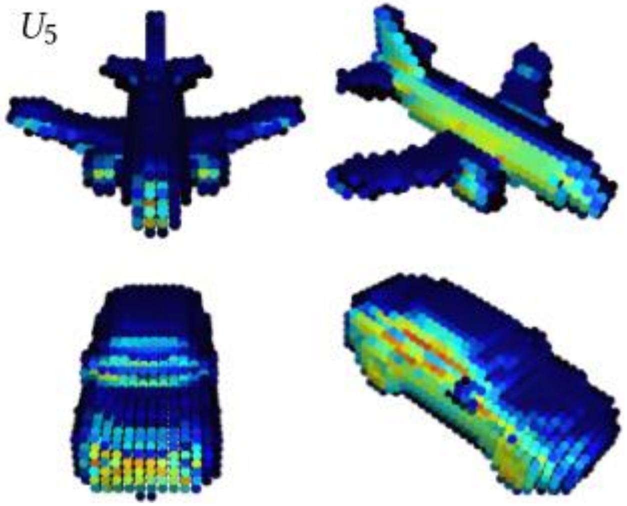
Missing data



Similar objects interference

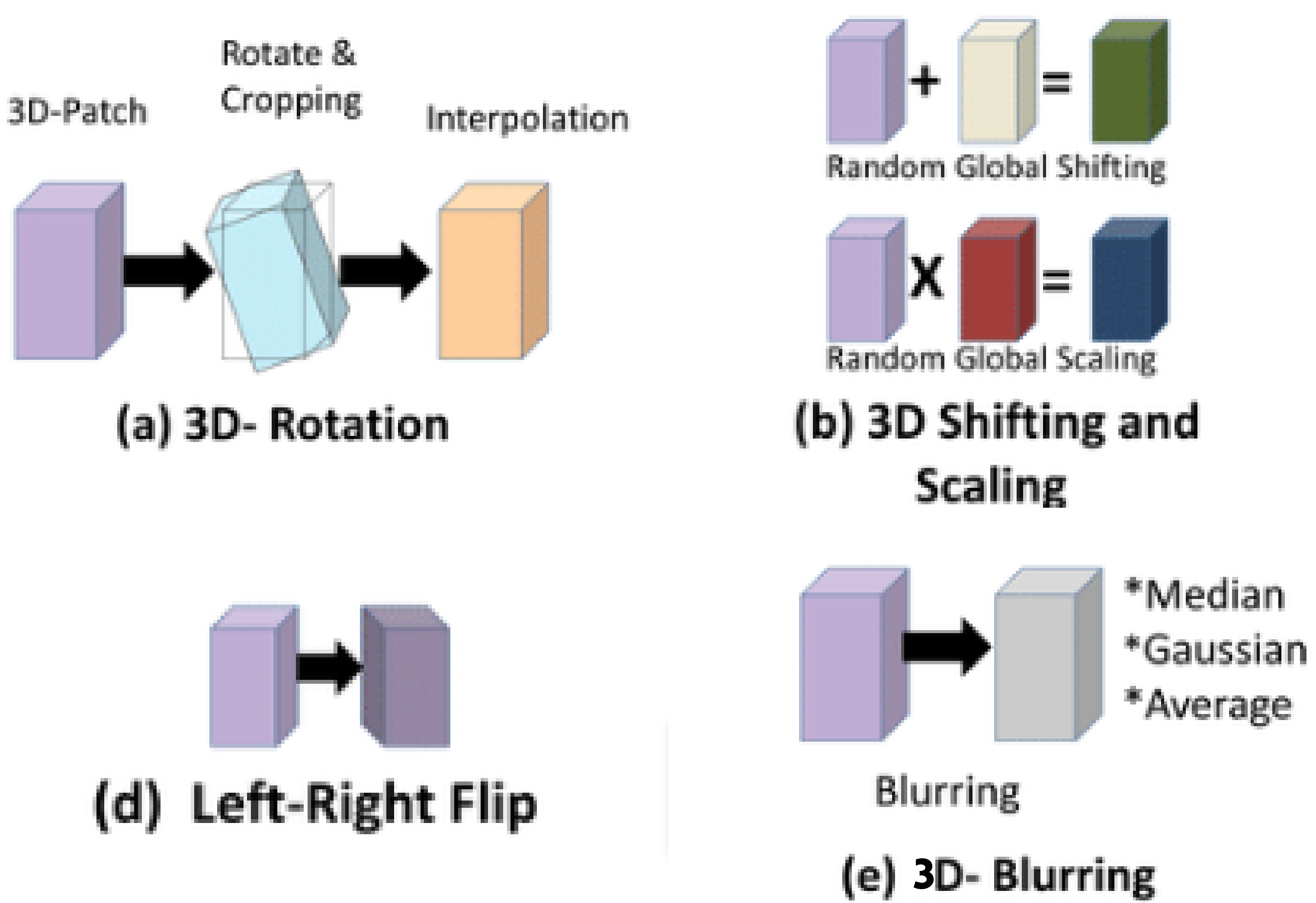
Models requirements

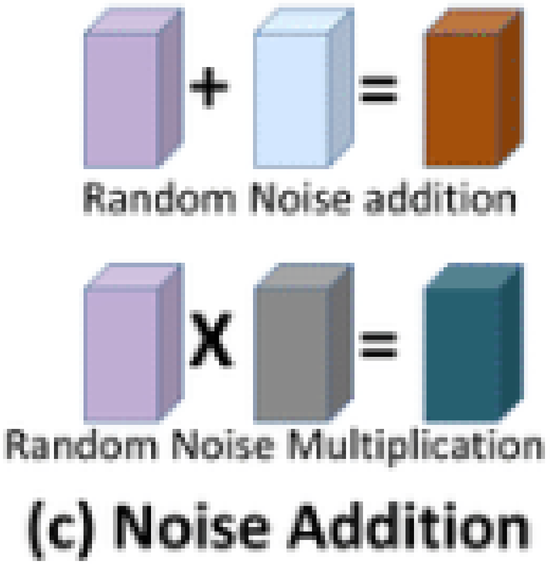




- Rotation invariance
- Translation invariance
- Permutation invariance
- Scale invariance
- Density invariance
- Local and global feature learning

Data augmentation

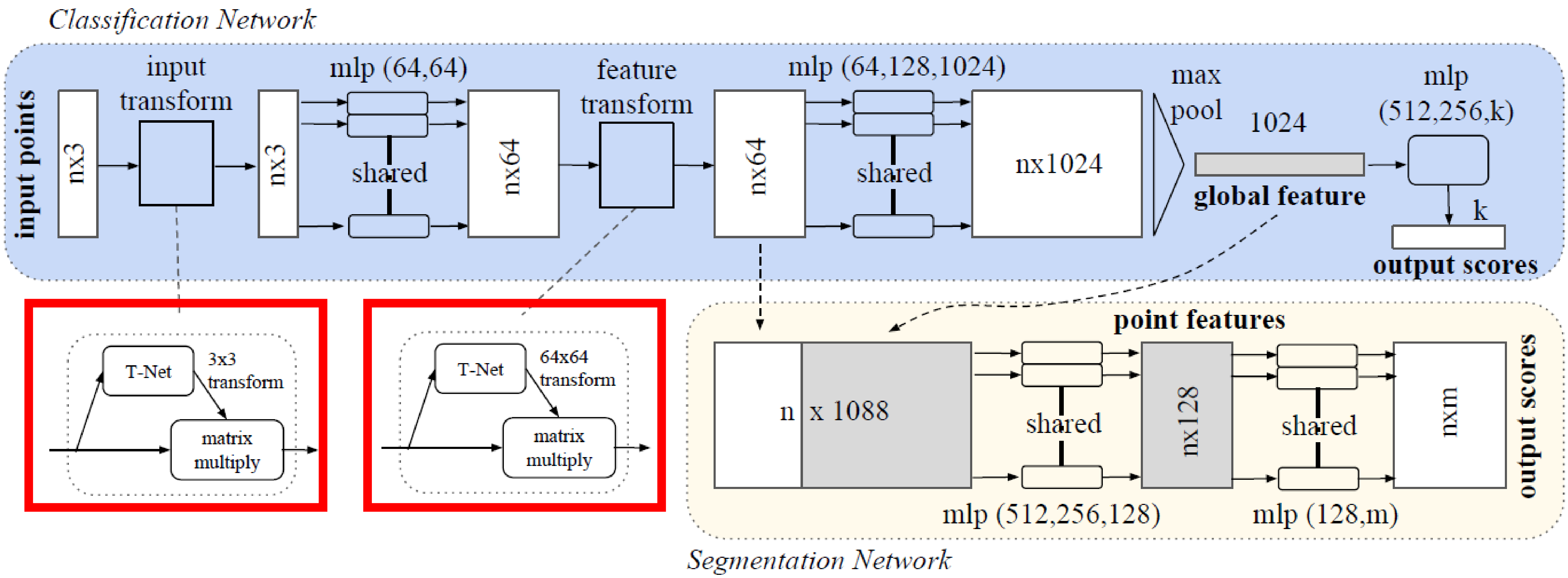




- **Rotation invariance**
- **Translation invariance**

Lung Cancer Tumor Region Segmentation Using Recurrent 3D-DenseUNet

PointNet Architecture



UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PointNet: Deep learning on point sets for 3d classification and segmentation.

Joint alignment network

Objection

- Permutation invariance => symmetric function
- arbitrarily approximate any continuous set function => universal approximators

$$f(x_1, x_2, \ldots, x_n) \equiv f(x_{\pi_1}, x_{\pi_2}, \ldots, x_{\pi_n}), \quad x_i \in \mathbb{R}^D$$

Symmetric function example:

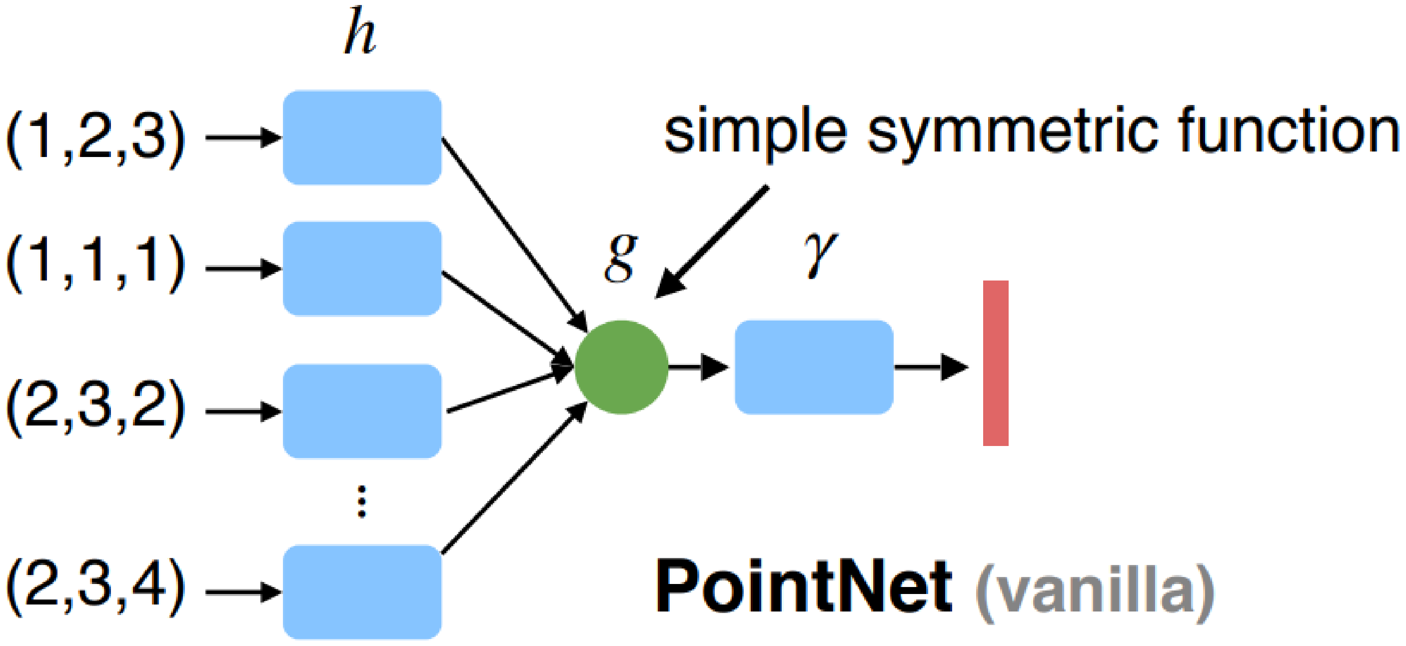
$$f(x_1, x_2, \ldots, x_n) = \max\{x_1, x_2, \ldots, x_n\}$$
$$f(x_1, x_2, \ldots, x_n) = x_1 + x_2 + \ldots + x_n$$

...

Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*.

Joint alignment network

$f(x_1, x_2, \ldots, x_n) = \gamma \circ g(h(x_1), \ldots, h(x_n))$ is symmetric if g is symmetric

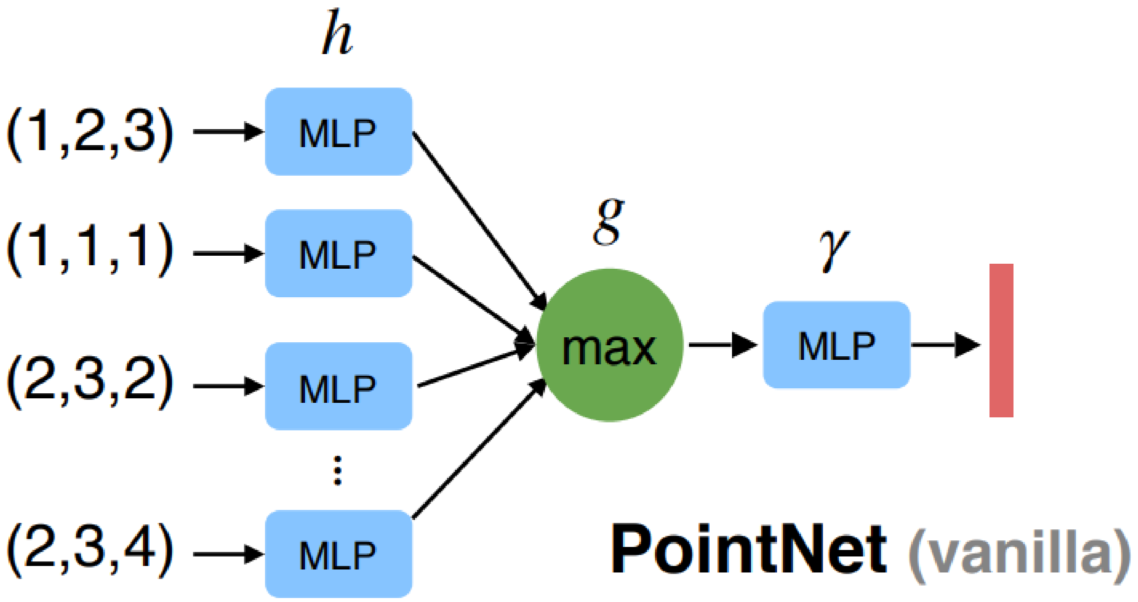


UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*.

Joint alignment network

- Permutation invariance => max-pooling
- arbitrarily approximate any continuous set function =>MLP

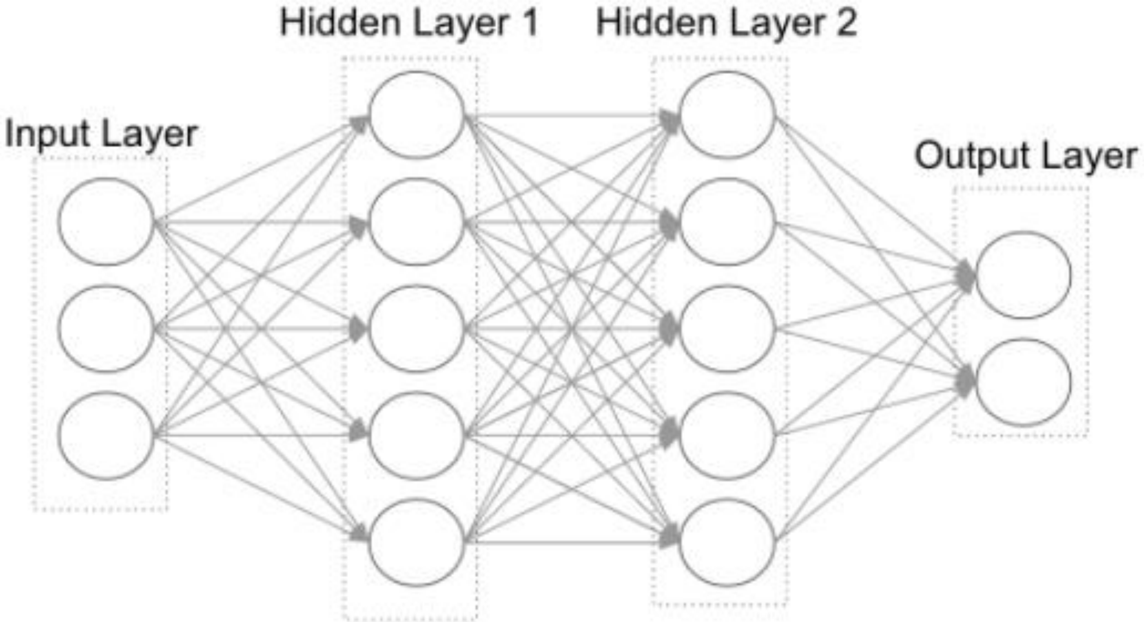


UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

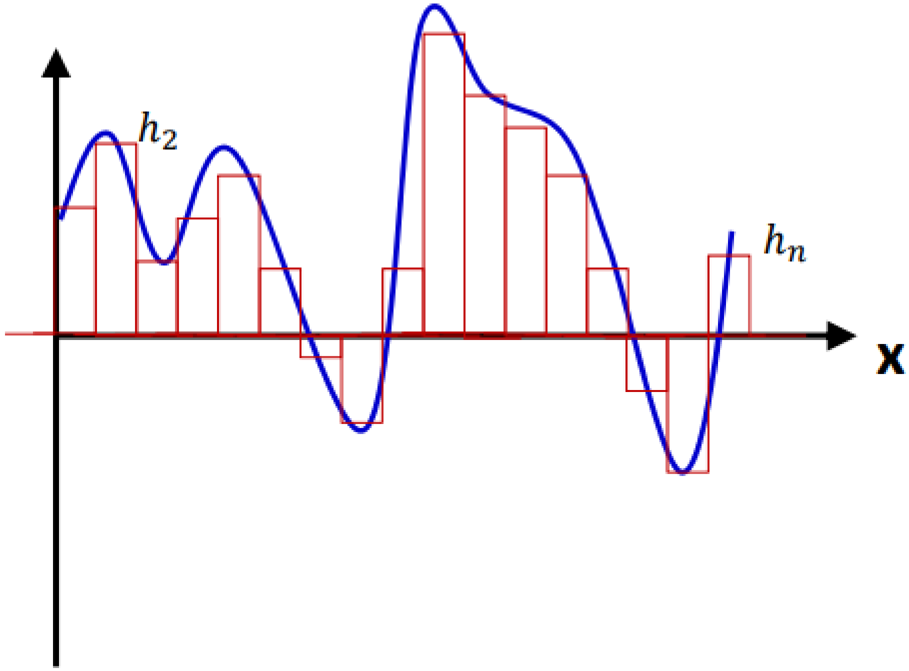
Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*.

Multi-layer perceptron (MLP)

The **universal approximation theorem** states that a feed-forward network with a single hidden layer containing a finite number of neurons can approximate continuous functions on compact subsets of $\mathbf{R}^n$, under mild assumptions on the activation function



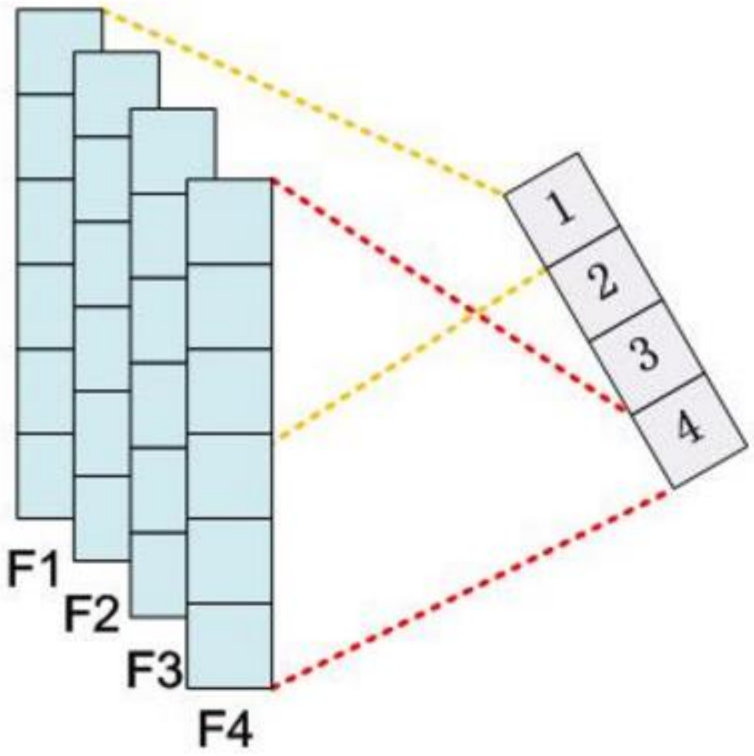
MLP



Balázs Csanád Csáji (2001) Approximation with Artificial Neural Networks; Faculty of Sciences; Eötvös Loránd University, Hungary

Methods: Max-pooling

Max-pooling: uses the maximum value from each of a cluster of neurons at the prior layer

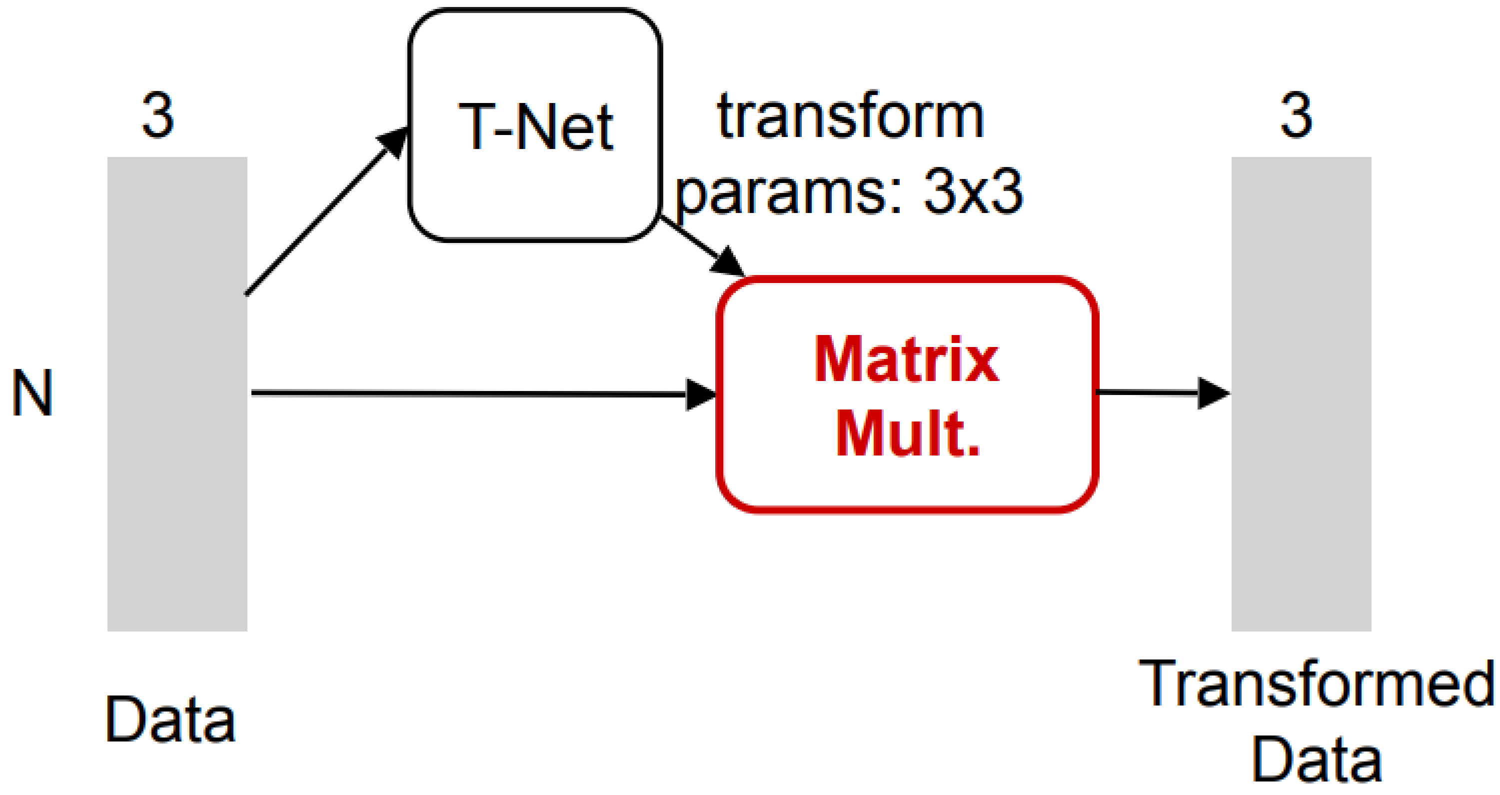


Pros:

- Translation invariance: learned features can be successfully extracted regardless of the locations of these features
- Lower in dimensions, and improve the results (avoid over-fitting)

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436-444.

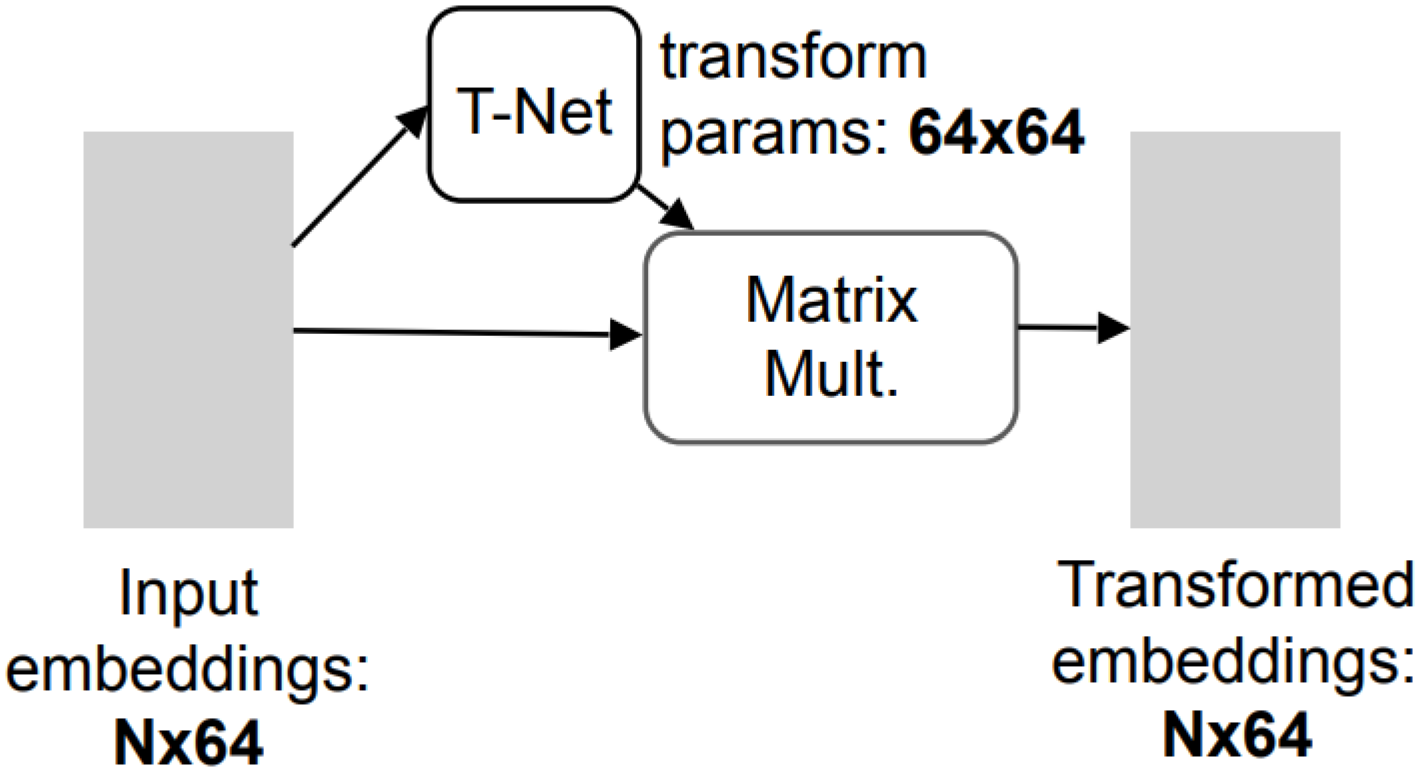
Joint alignment network



UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*.

Embedding Space Alignment



Regularization:

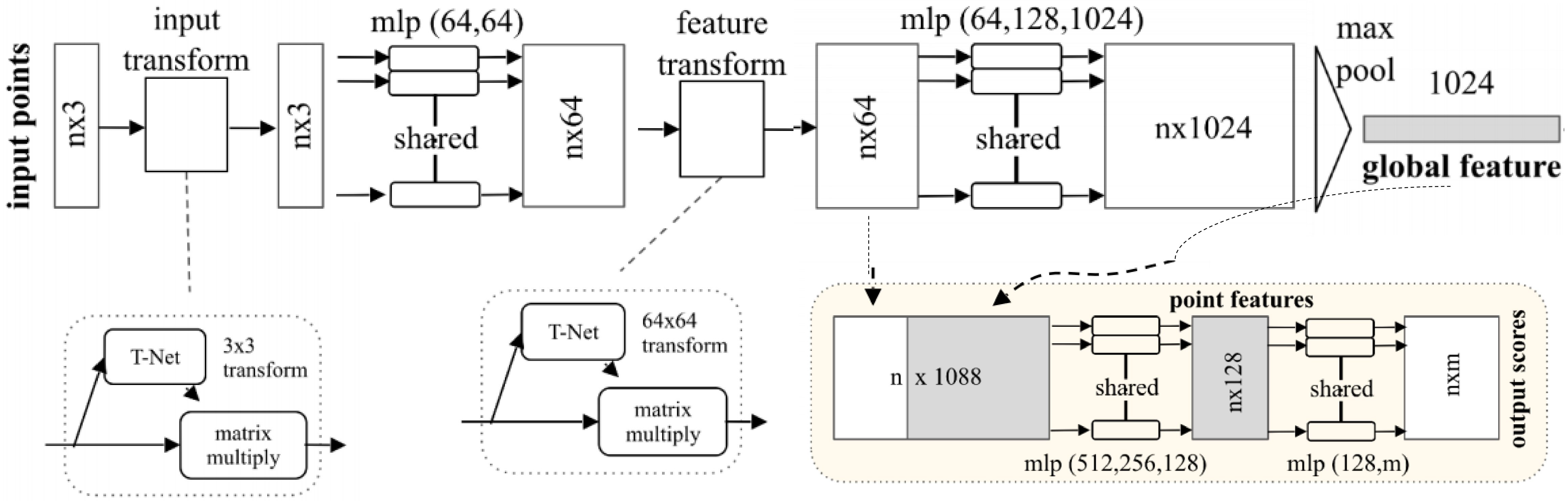
Transform matrix A 64x64 close to orthogonal:

$$L_{reg} = \|I - AA^T\|_F^2$$

UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*.

PointNet



Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*.

PointNet

Pros:

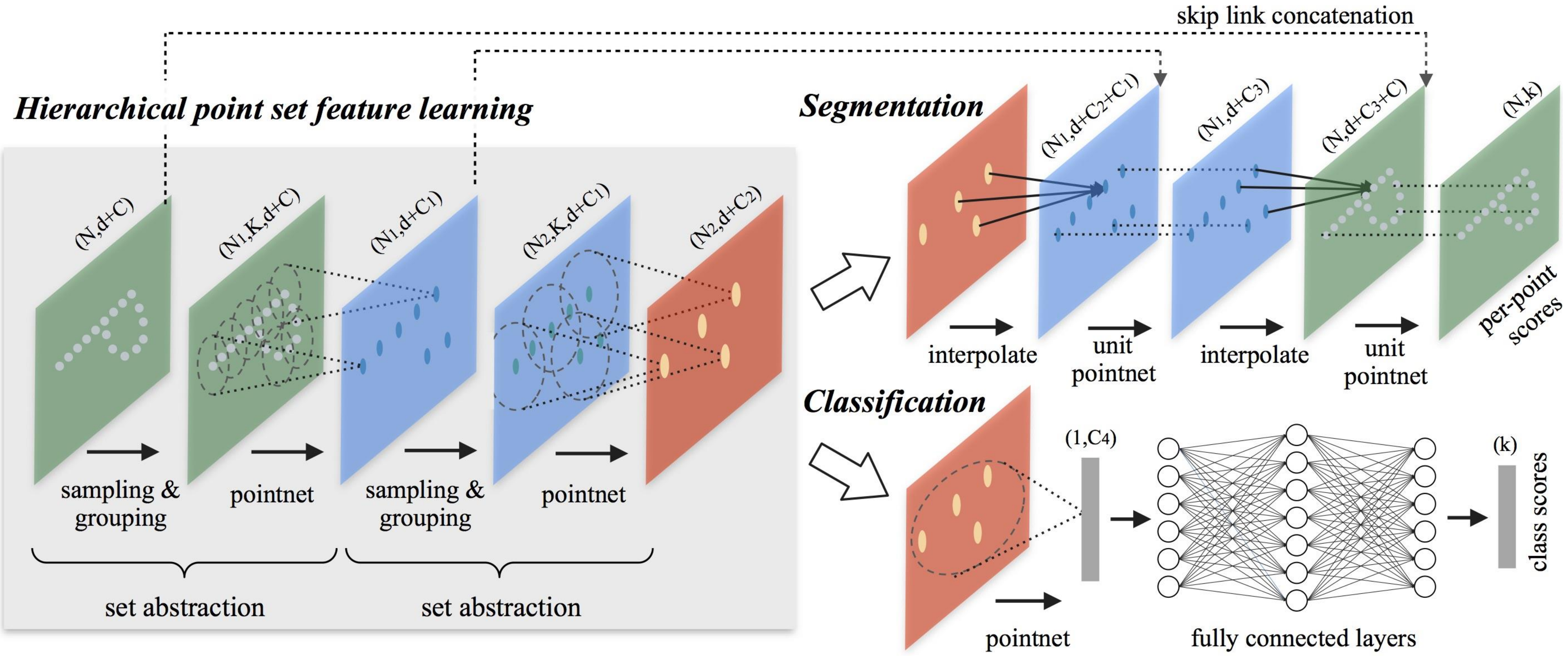
- solve the point cloud permutation problem
- Extract global feature

Cons:

- Not capture local structure
- Not scale invariance
- Not density invariance
- Not suitable for large-scale point cloud data

Pointnet: Deep learning on point sets for 3d classification and segmentation. *arXiv preprint arXiv:1612.00593*.

PointNet++ Architecture



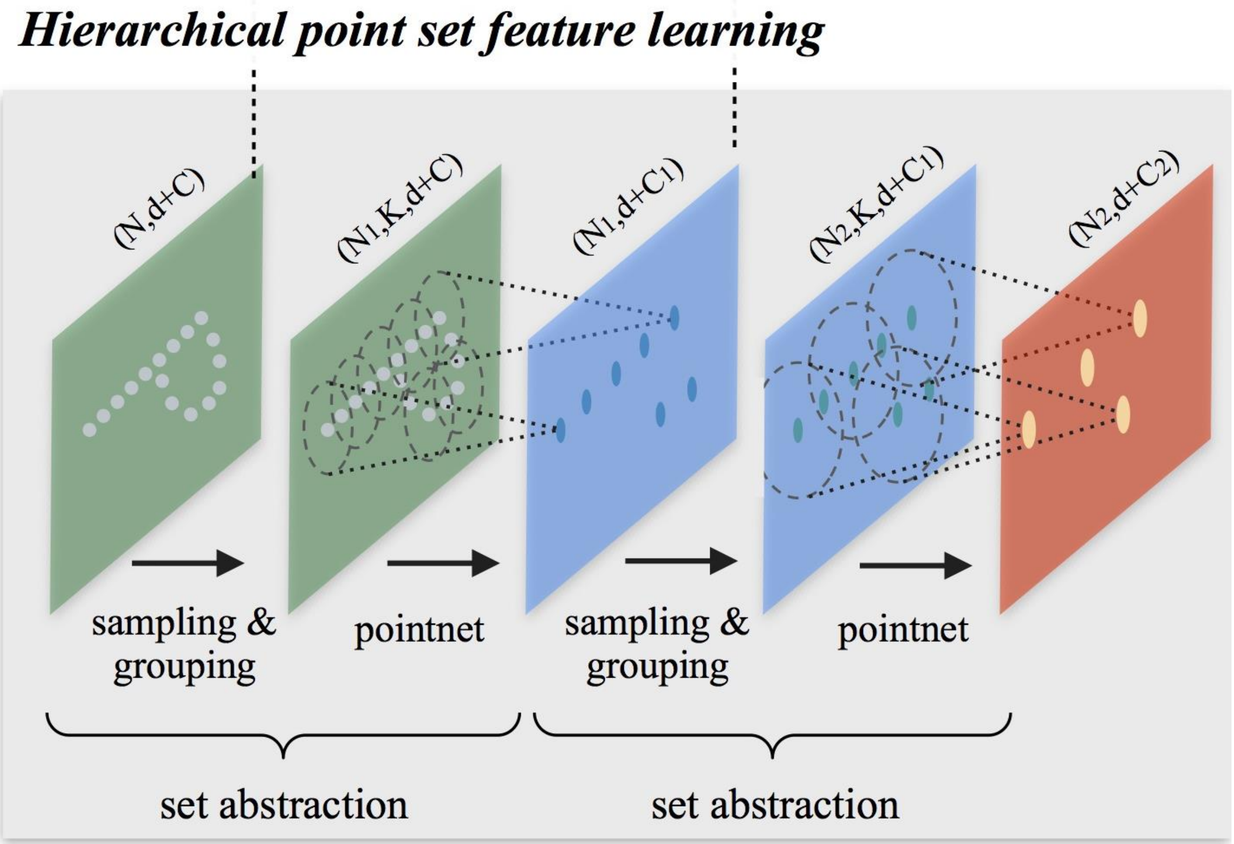
UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Hierarchical Point Set Feature Learning (set abstraction level)

Objection

- Scale invariance => sampling
- Density invariance => grouping
- Permutation invariance => PointNet



UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

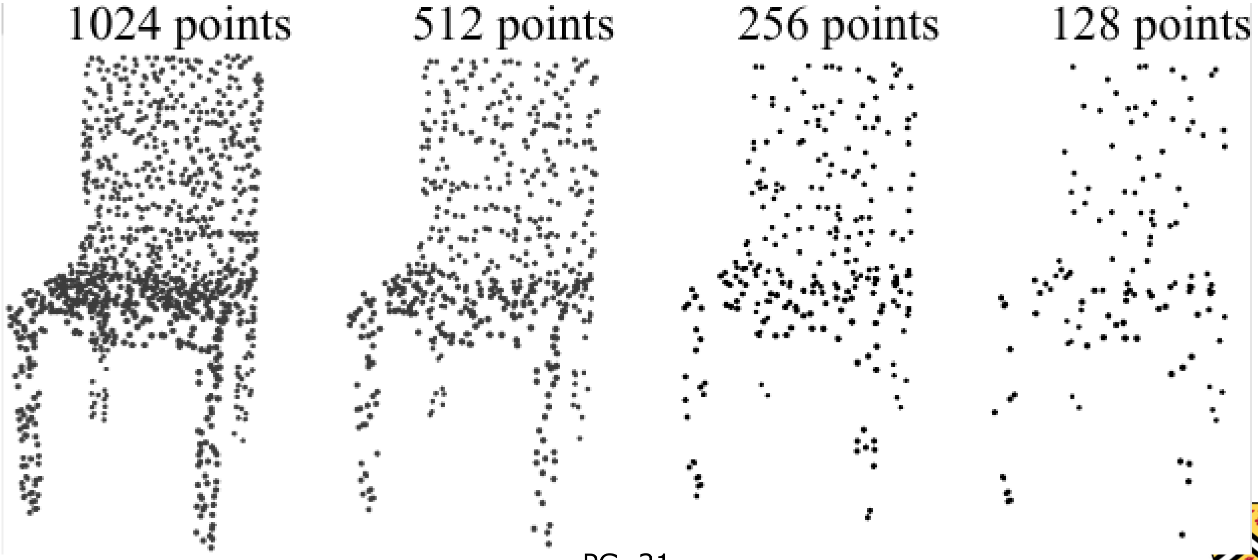
PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Point clouds sampling –scale invariance

Farthest point sampling (FPS): Given input points {x1, x2, ..., xn}, FPS is iteratively used to choose a subset of points {xi1 distance) from the set {xi1, xi2 , ..., xim}, such that xij is the most distant point (in metric , xi2 , ..., xij−1} with regard to the rest points.

Pros:

- Evenly cover the entire point set given the same number of centroids.
- Sampling strategy generates receptive fields in a data dependent manner.

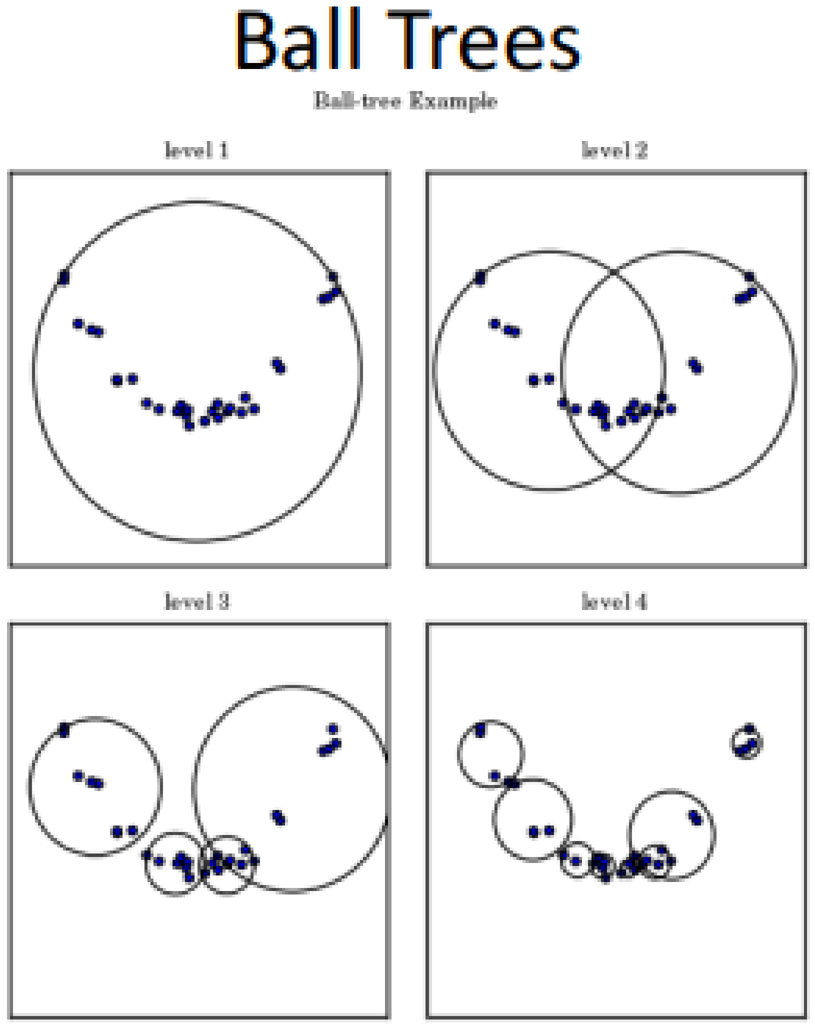


UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Point clouds grouping

Ball query: finds all points that are within a radius to the query point (an upper limit of K is set in implementation).
Pros: guarantees a fixed region scale thus making local region feature more generalizable across space.
Cons: not robust to noisy points



UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

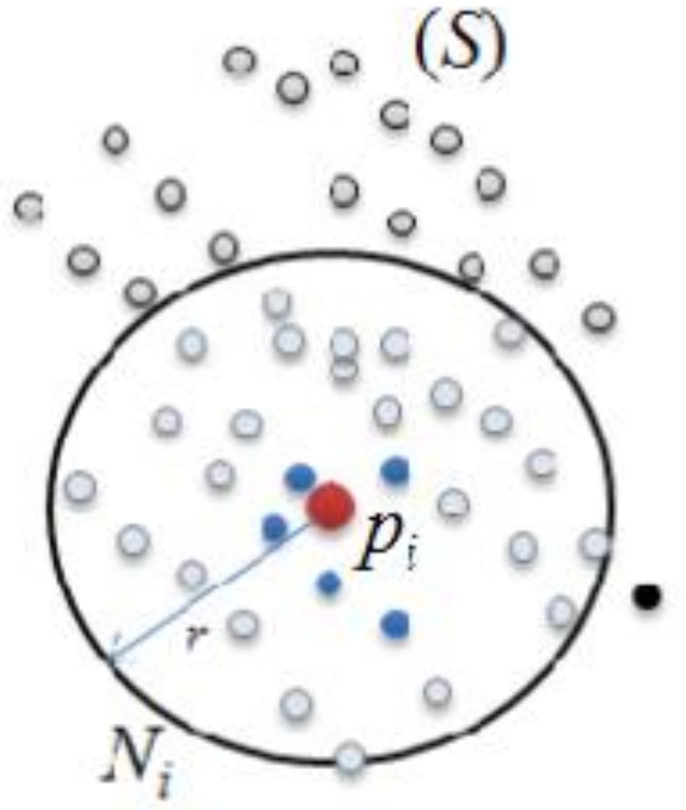
Point clouds grouping

K nearest neighbor (KNN): finds a fixed number of neighboring

Pros: robust to noisy points

Cons: not scale-invariance

high computation cost



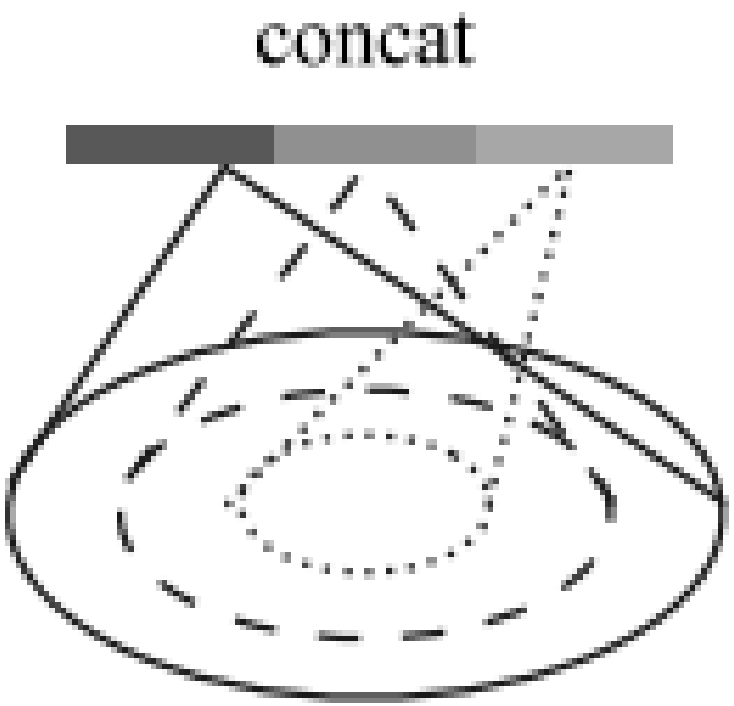
UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Multi-scale grouping (MSG)-density invariance

MSG: grouping layers with different scales followed by according PointNets to extract features of each scale. Features at different scales are concatenated to form a multi-scale feature. => random input dropout

Pros: simple
Cons: not density-invariance
high computation cost



UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

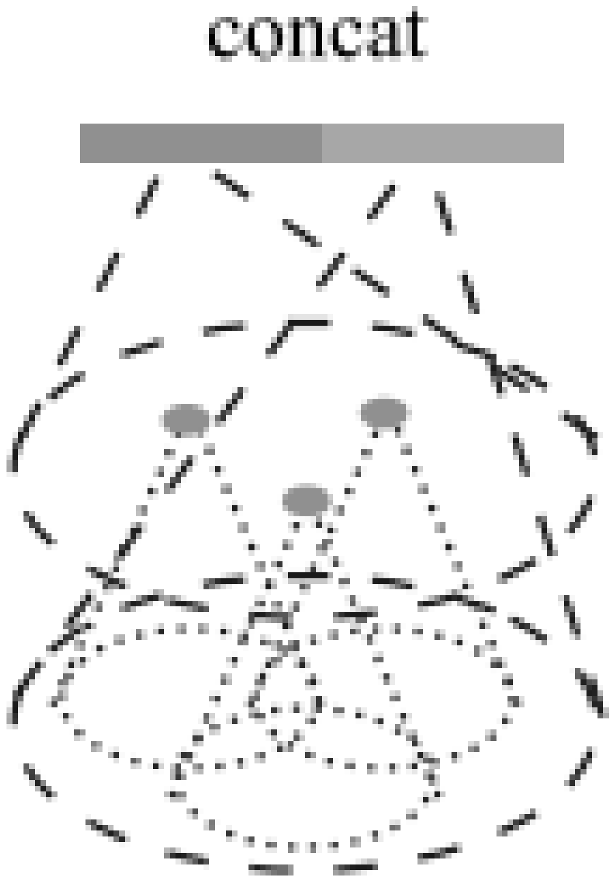
PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Multi-resolution grouping (MRG)- density invariance

MSG: One feature is obtained by summarizing the features at each subregion from the lower level. The another feature is obtained by directly processing all raw points in the local region using a single PointNet.

Pros: density-invariance
computation efficient

Cons: features are constrained by the number of neighbors

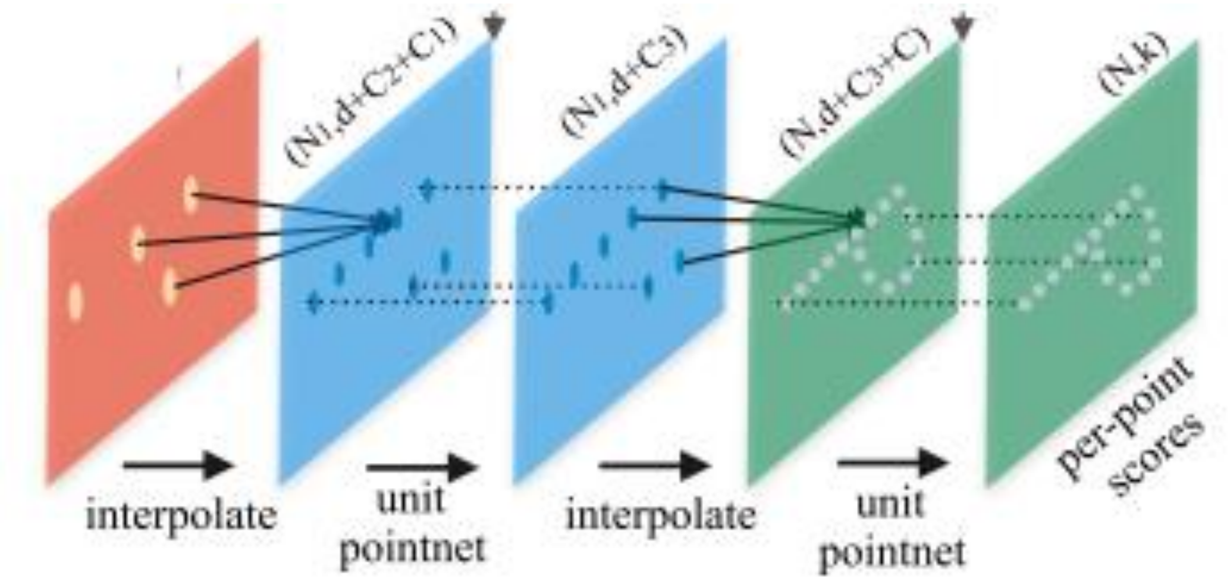


UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Point Feature Propagation

Objection: obtain point features for all the original points from subsampled point set

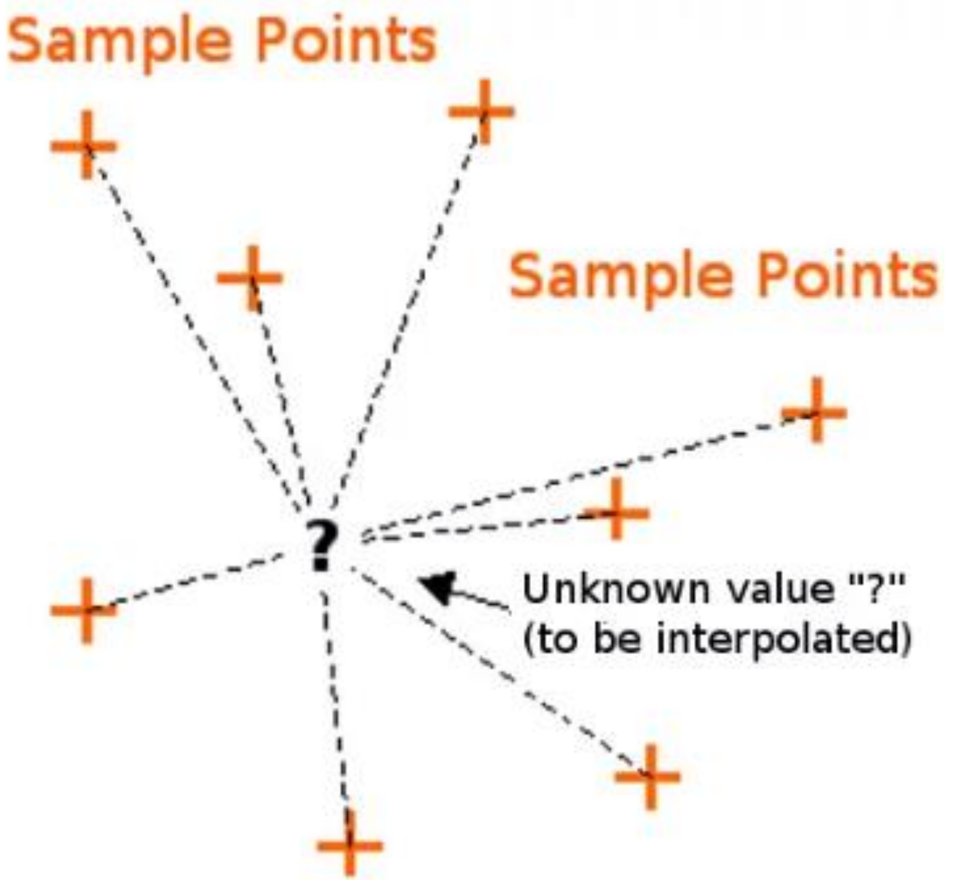


PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Point clouds interpolation

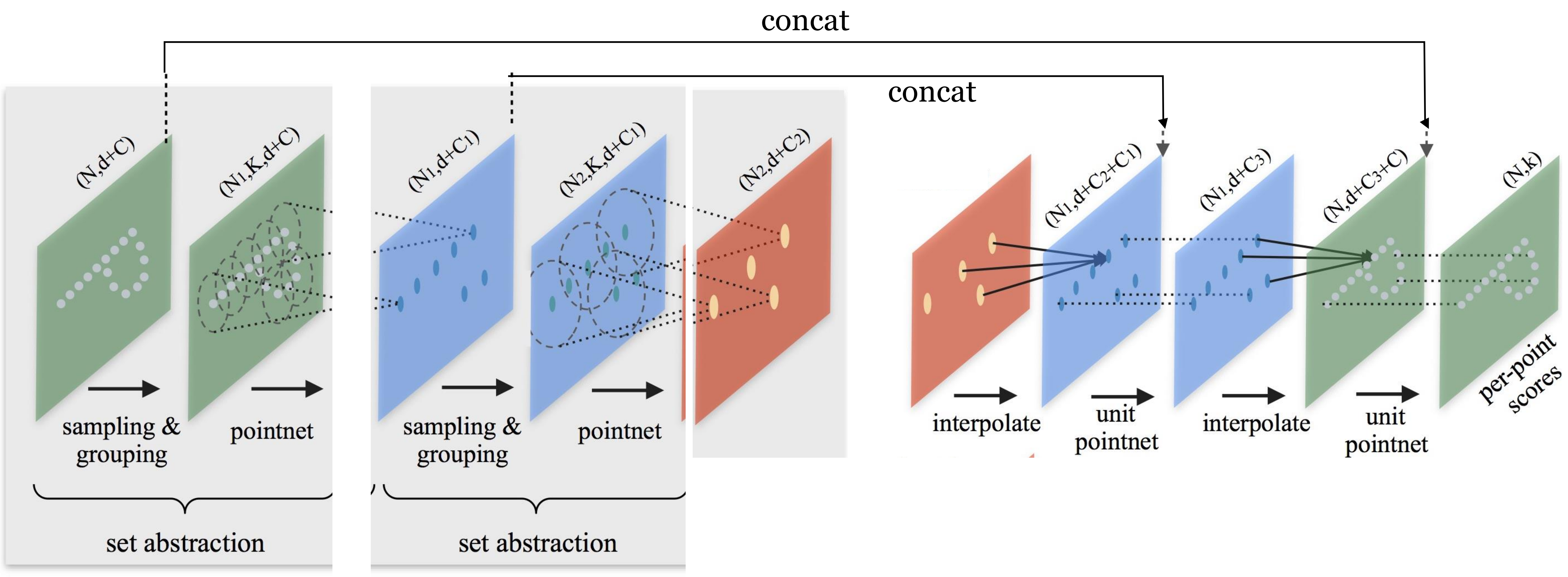
Inverse distance weighted average: multivariate interpolation with a known scattered set of points. it resorts to the inverse of the distance to each known point based on KNN when assigning weights

Pros: will not produce estimated values outside neighborhoods
Cons: not robust to noise



https://en.wikipedia.org/wiki/Inverse_distance_weighting

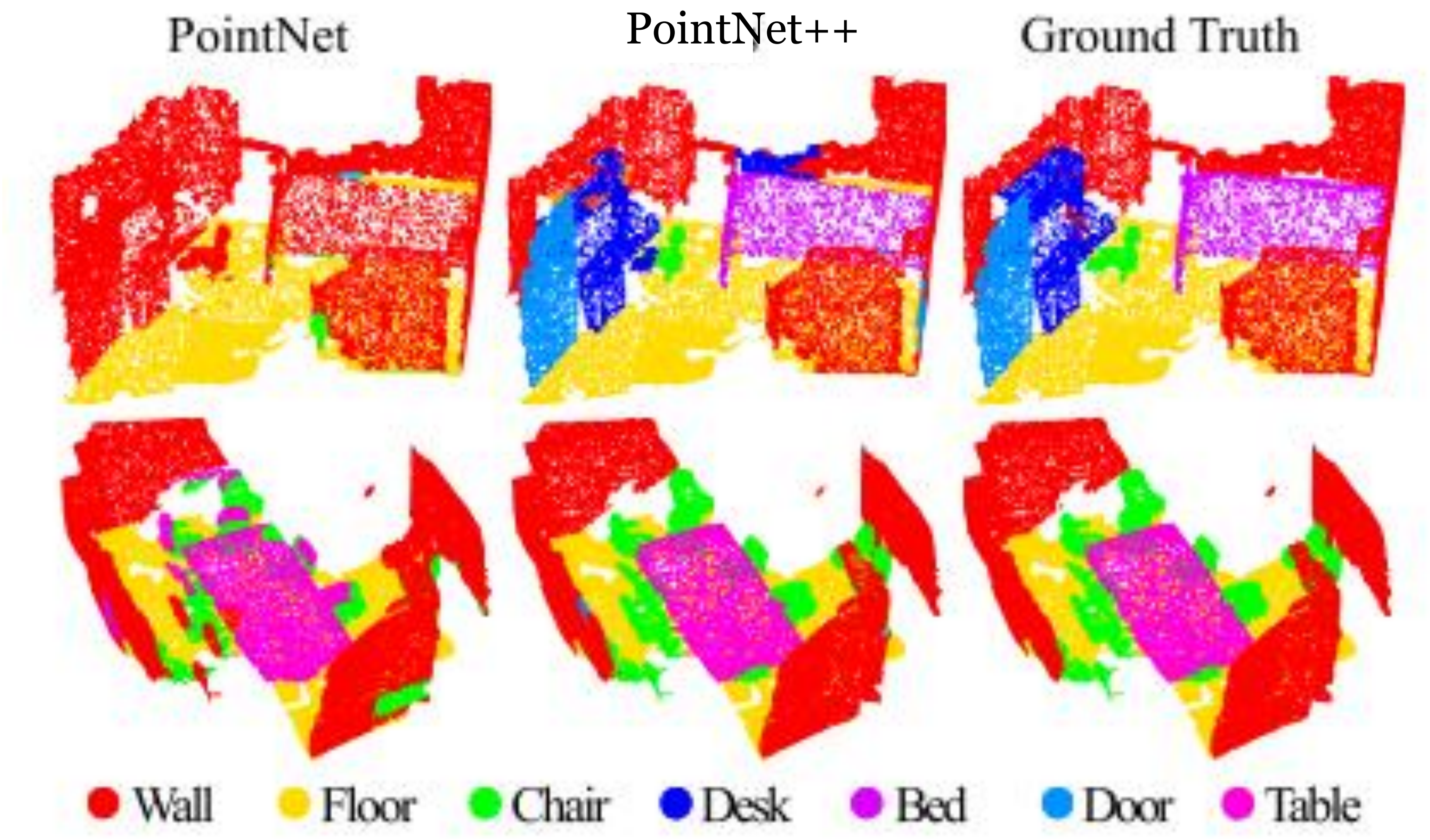
PointNet++ Architecture



UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Segmentation results



UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

PointNet++

Pros:

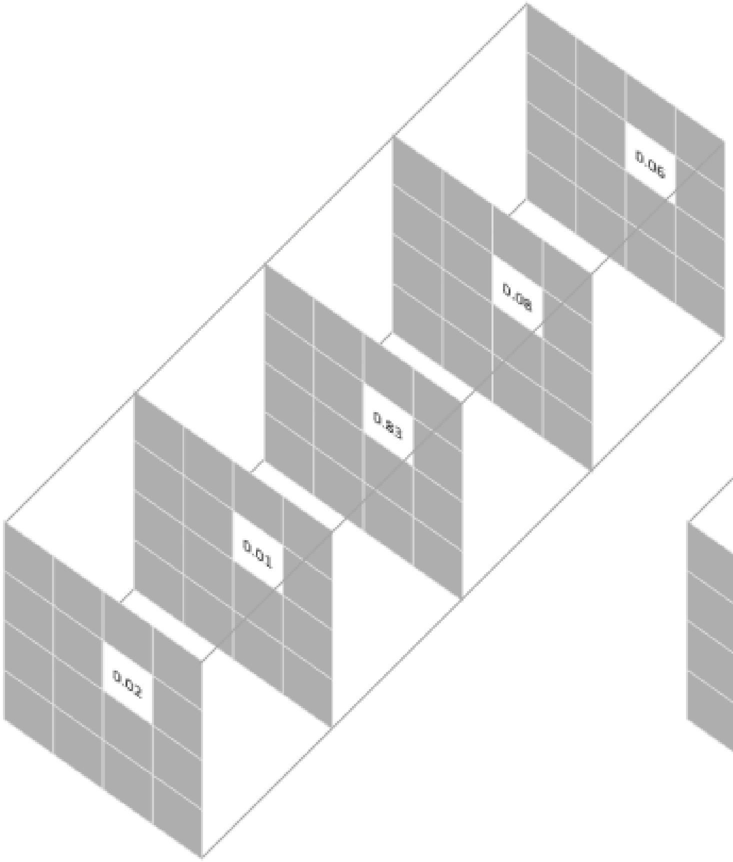
Orientation invariance
Rotation invariance
Translation invariance
Permutation invariance
Scale invariance
Density invariance
Local and global feature learning

Cons:

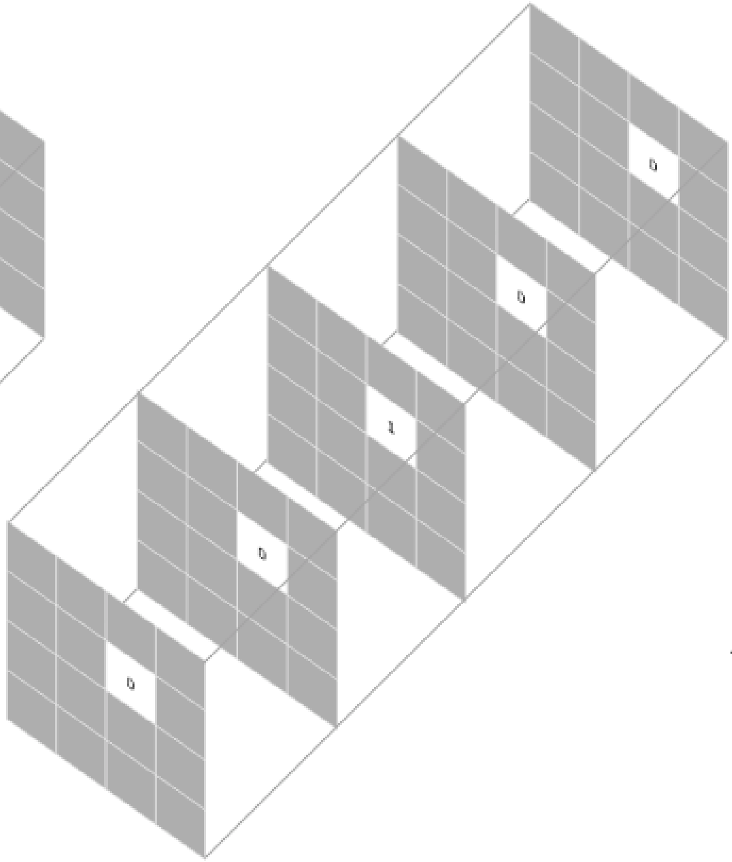
- Points neighbor spatial relationships are not exploited

PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space

Loss function for 3D segmentation



Prediction for a selected pixel



Target for the corresponding pixel

Pixel-wise loss is calculated as the log loss, summed over all possible classes

$$-\sum_{classes} y_{true} \log\left(y_{pred}\right)$$

This scoring is repeated over all **pixels** and averaged

https://www.jeremyjordan.me/semantic-segmentation/#loss

PUBLIC MLS DATASETS

Dataset	Laser scanner	Camera	Year	Data size	Path	Creator
KITTI Vision Benchmark	Velodyne HDL-64E laser scanner	PointGray Flea2 grayscale and color cameras	2011	180G	~ 50 km	KIT & University of Toronto
New College Dataset	LMS 291-S14 lasers scanning	Point Grey BumbleBee 20 Hz grayscale	2009	30GB	2.2km	Oxford
Malaga Dataset 2013	SICK scanners	Point Grey Research's Bumblebee 2 stereo camera	2013	90.22 G	36.8 km	University of Málaga
Ford Campus Vision and Lidar Dataset	Velodyne HDL-64E lidar & Riegl LMS-Q120 lidar	Point Grey Ladybug3 omnidirectional camera	2009	~100 GB	-	Ford & University of Michigan

UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

PUBLIC MLS DATASETS

Dataset	Laser scanner	Camera	Year	Data size	Path	Creator
NCLT Dataset	Velodyne HDL-32E 3D lidar & Hokuyo planarb lidars	Ladybug3 omnidirectional camera	2012	~1000 GB	147.4 km	Ford & The University of Michigan North Campus
Paris-Lille-3D	Velodyne HDL-32E LiDAR	-	2017	8.8G	1.94km	PSL Research University
Apollo scape	VUX-1HA laser scanners	VMX-CS6 camera system	2018	~10G	-	Baidu
OXFORD ROBOTCAR DATASET	LMS-151 2D LIDAR & SICK LD-MRS 3D LIDAR	Point Grey Bumblebee XB3 trinocular stereo camera, Point Grey Grasshopper2 monocular camera	2017	23.15TB	1010.46km	University of Oxford

UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

EVALUATION METRICS

$$IoU_i = \frac{c_{ii}}{c_{ii} + \sum_{j \neq i} c_{ij} + \sum_{k \neq i} c_{ki}}$$

$$\overline{IoU} = \frac{\sum_{i=1}^{N} IoU_i}{N}$$

$$\mathrm{OA} = \frac{\sum_{i=1}^{N} c_{ii}}{\sum_{j=1}^{N} \sum_{k=1}^{N} c_{jk}}$$

N is number of object classes, c_{ij} is the number of points from ground-truth class i predicted as class j

Hackel et al. arXiv preprint arXiv:1704.03847 (2017).

Conclusion

- PointNet proposes a Joint alignment network solving input point cloud permutation problem
- PointNet++ extract the local and global feature of point clouds

Improvements:

- High level neighbor relationship feature extraction
- more effective sampling methods

UNIVERSITY OF WATERLOO
FACULTY OF ENVIRONMENT

THANK YOU
Q&A