



Investigation of methods for semantic segmentation on 3D images

Presenter: M. Tamjidi

Supervisor: Dr. A. Mousavinia

Adviser: Dr. B. Nasihatkon

March 2024

outline

- Introduction
- 2D Segmentation
- 3D Semantic Segmentation
- Future Trends
- Conclusion

Why segmentation?

- Scene understanding
- Medical scans
- Remote sensing

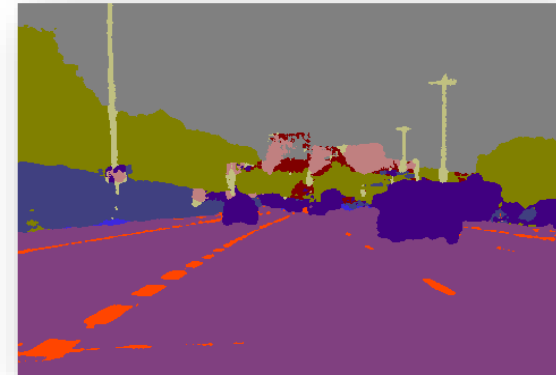


Which one is semantic?

- Semantic VS Others
 - Processing uncountable stuff
 - Annotating All pixels
 - Analyzing upon perception



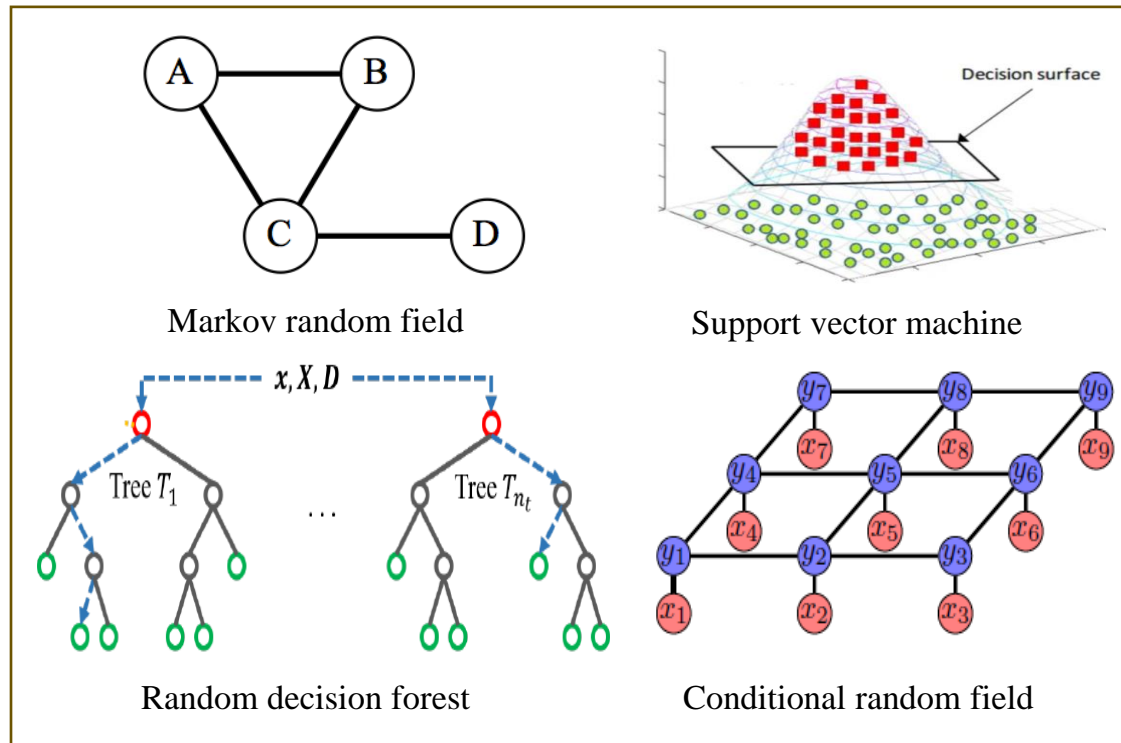
↓
Segmentor



Outline

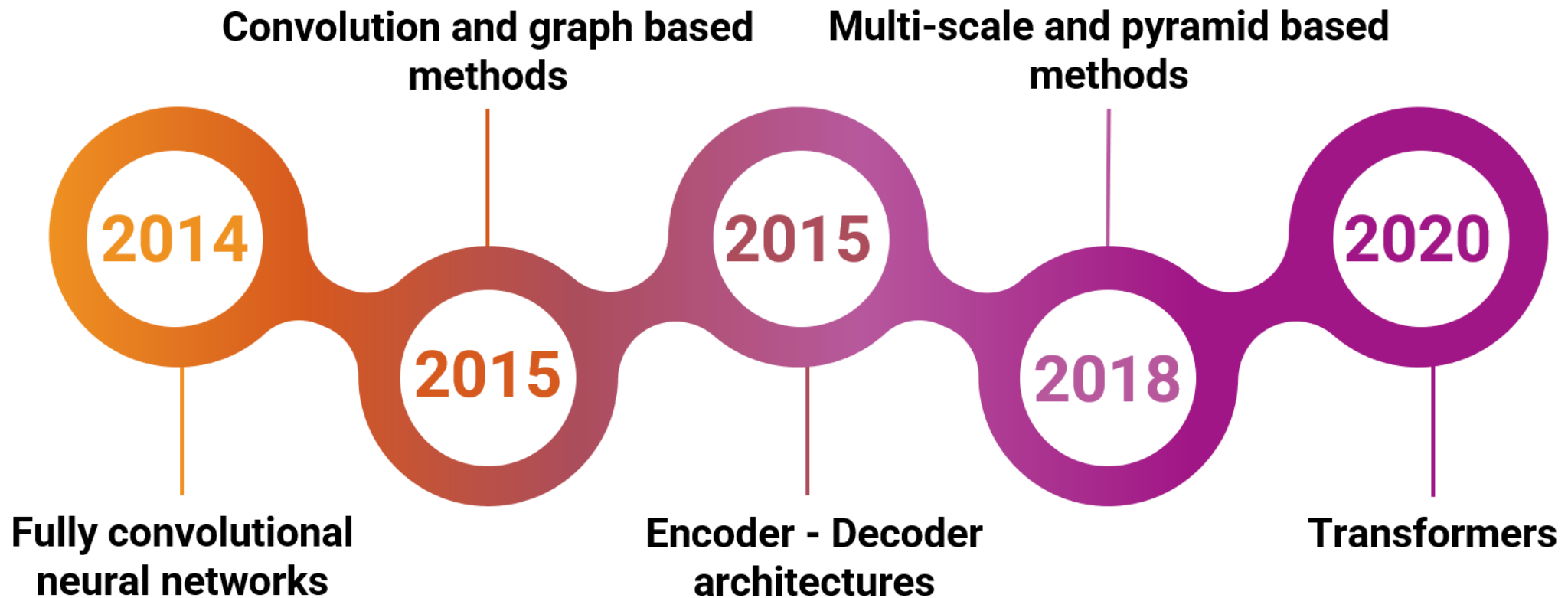
- Introduction
- 2D Segmentation
 - Traditional Approaches
 - Deep Learning based methods
- 3D Semantic Segmentation
- Future Trends
- Conclusion

Traditional Approaches



- Probabilistic graphical models
 - Unable to extract local feature
 - Limited Modeling Capability
 - Used in conjunction with deep learning

Deep learning based methods

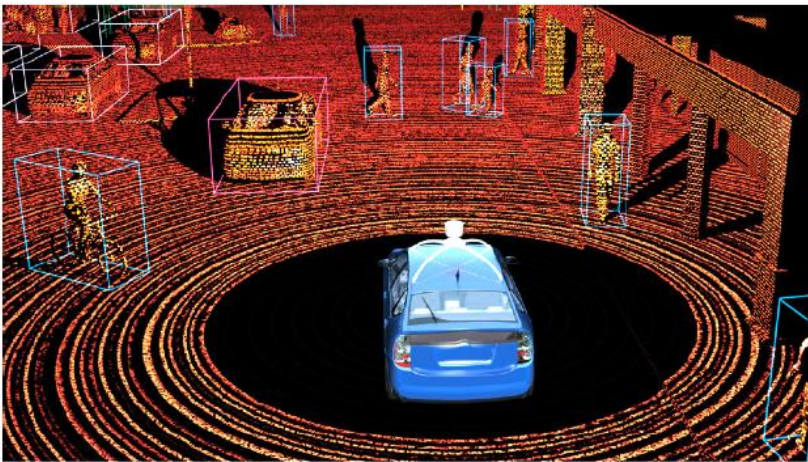


outline

- Introduction
- 2D Segmentation
- Semantic segmentation on 3D images
 - Emerging 3D Applications
 - 3D Data & Point Cloud
 - Point cloud semantic segmentation
- Future Trends

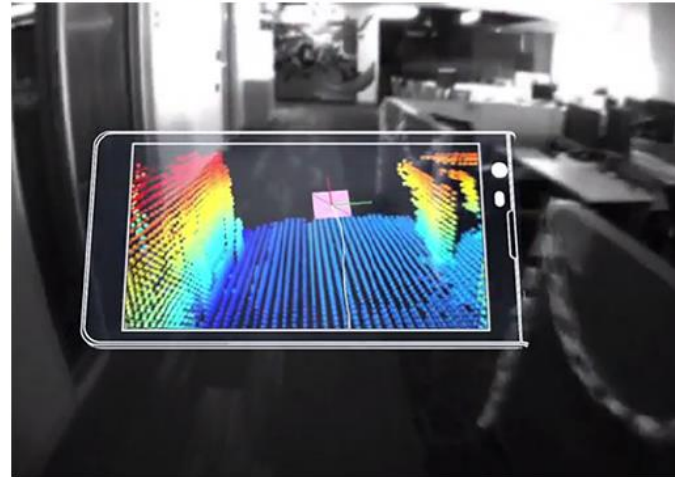
Emerging 3D Applications

Robot Perception



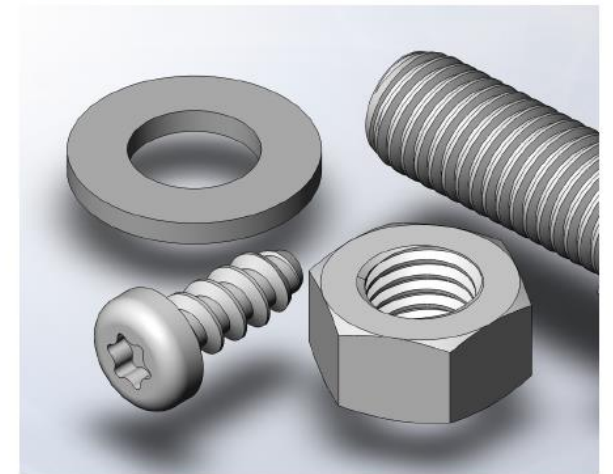
source: Scott J Grunewald

Augmented Reality



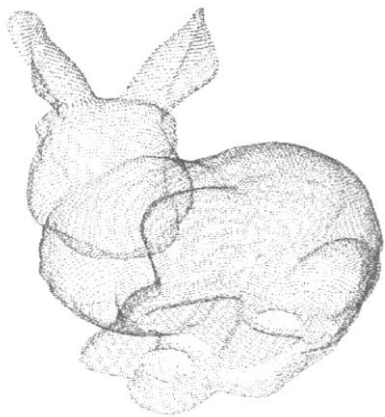
source: Google Tango

Shape Design

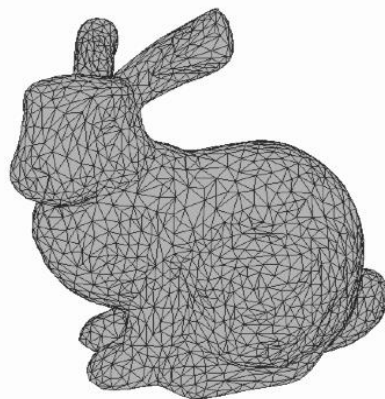


source: solidsolutions

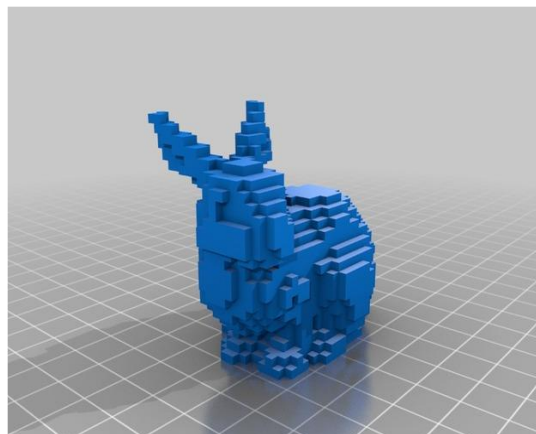
3D Representation



Point cloud



Mesh



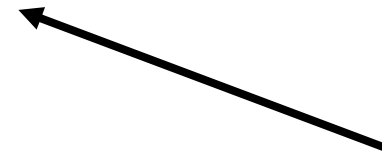
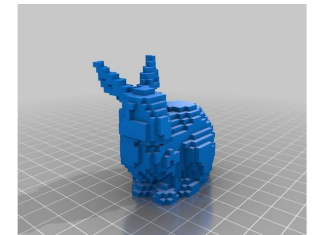
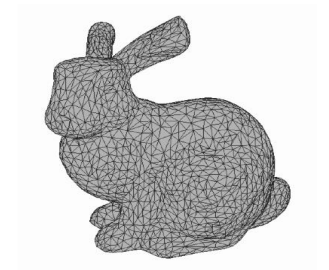
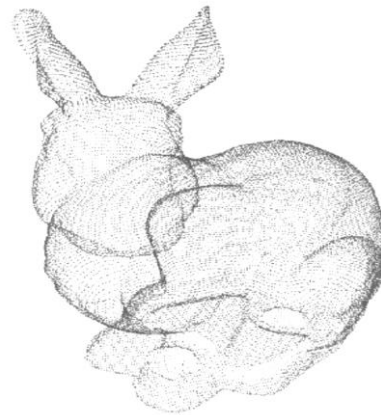
3D voxel grids



RGBD image

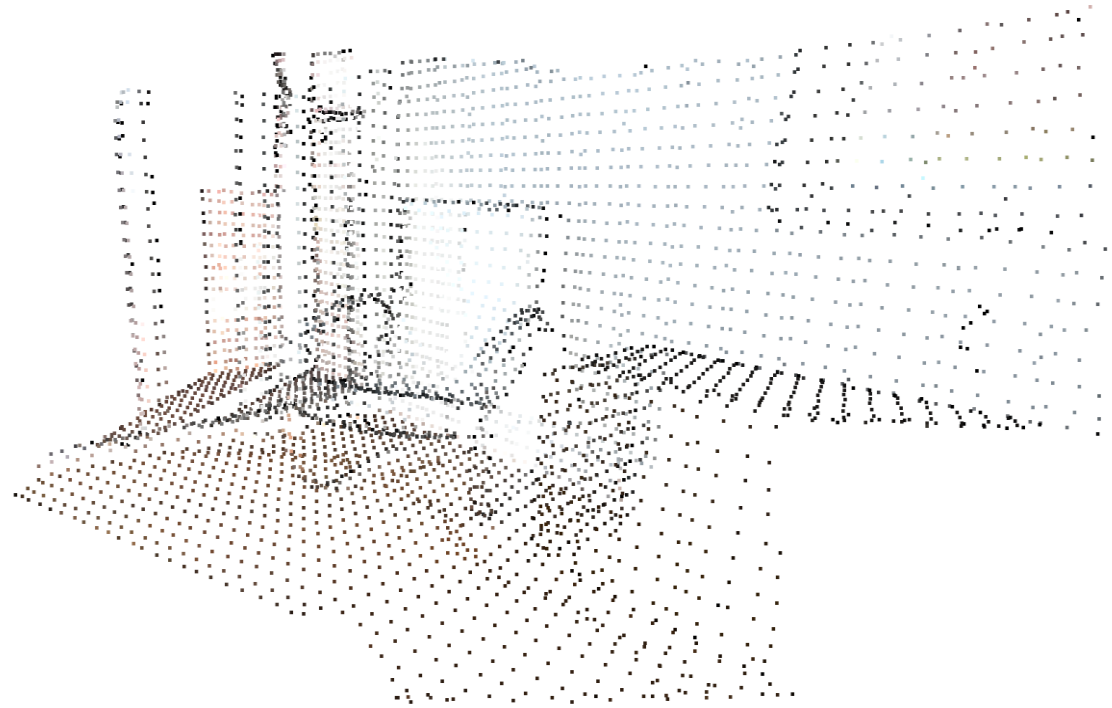
3D Representation

- ✓ Point cloud is close to raw sensor data
- ✓ Usually preferred over other representation



Point Cloud

- Non-grid structure data
- Unordered
- In continuous space
- Permutation invariant

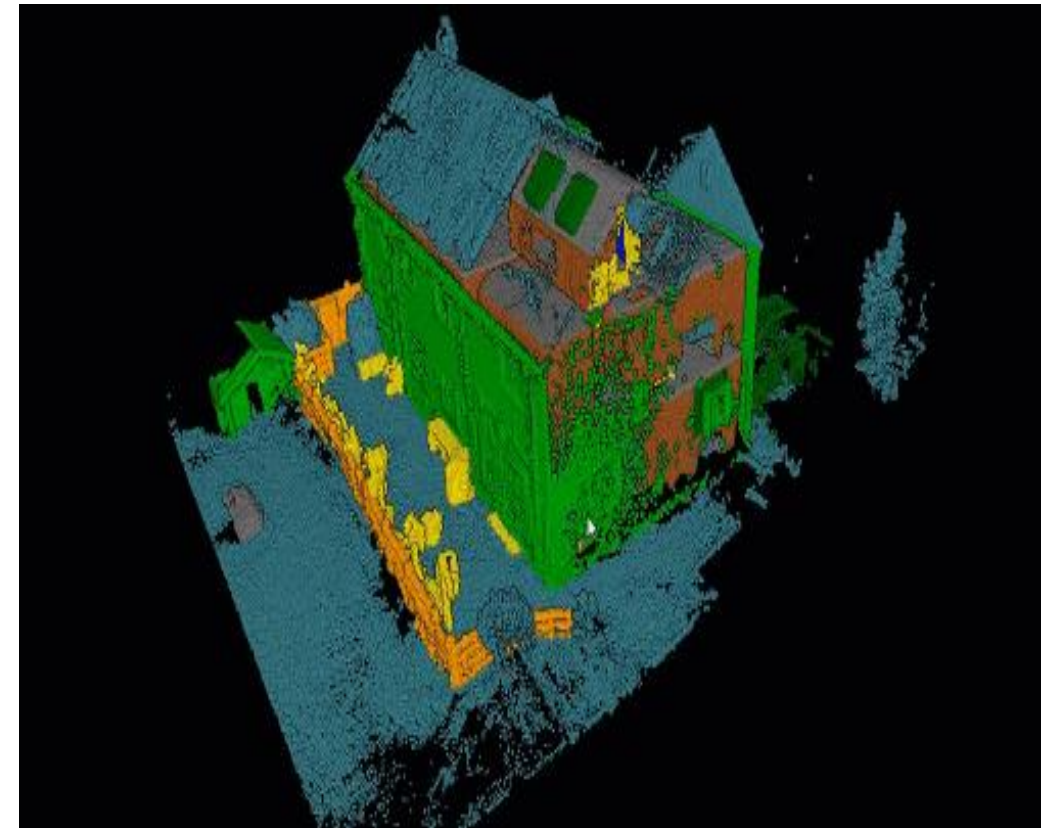
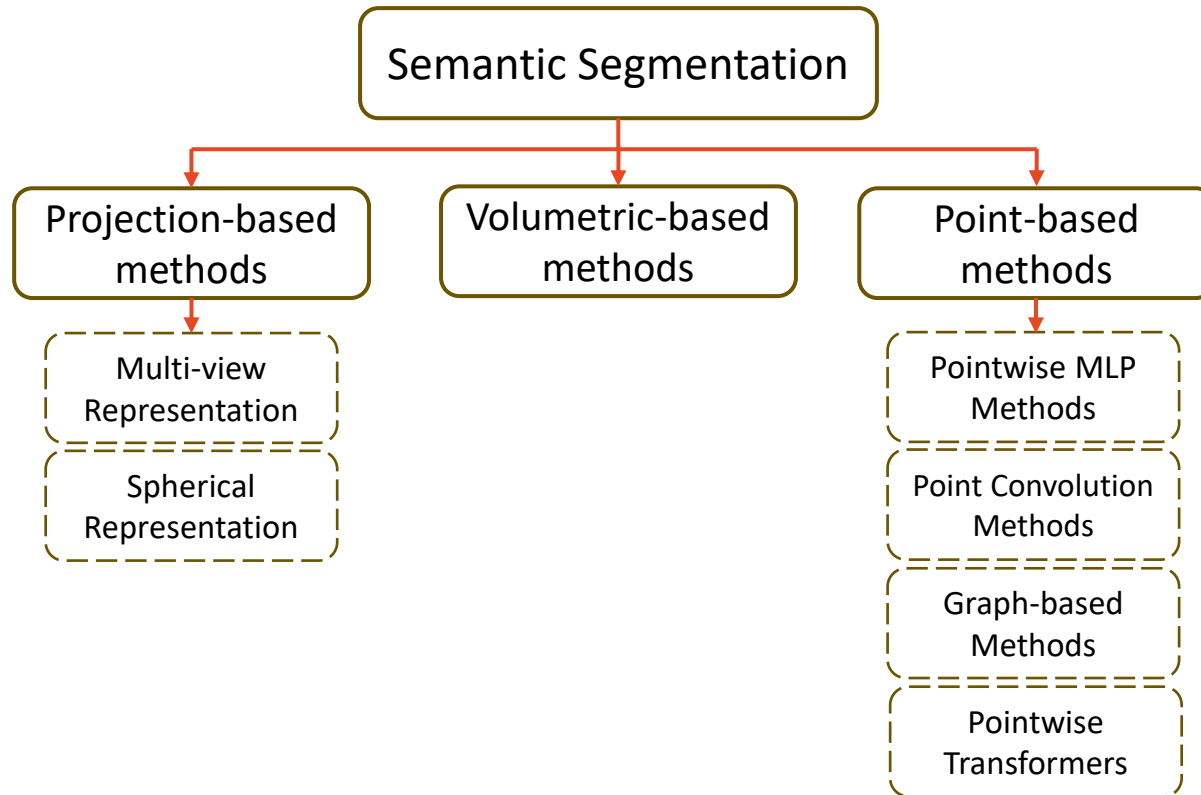


Point cloud visualization by OPEN3D library

How to apply convolution on point cloud?

Point cloud semantic segmentation

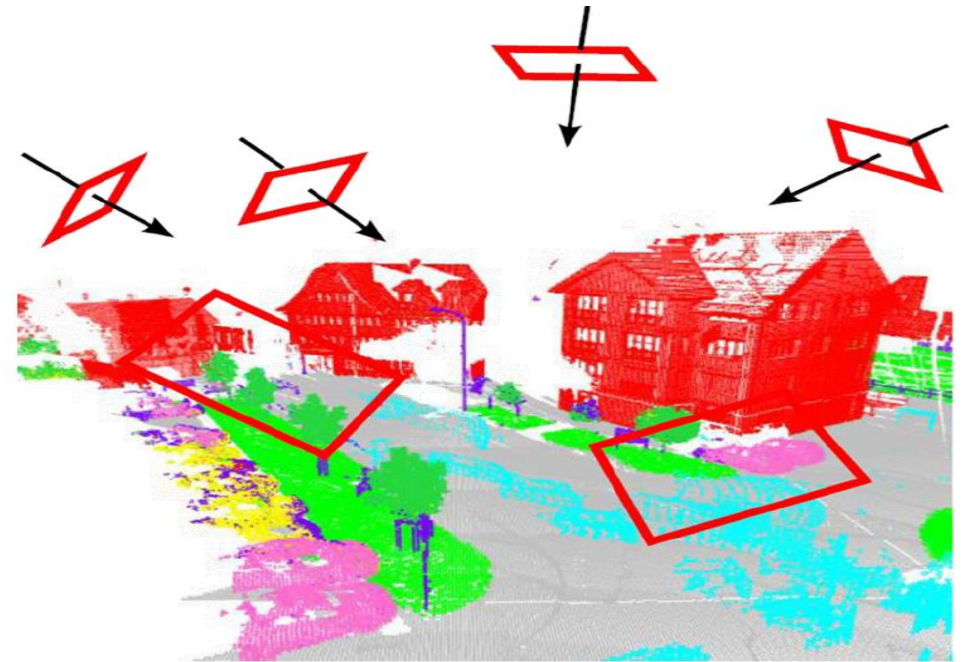
Recent Deep learning based methods



<https://pointcloudproject.com/the-future-of-3d-point-clouds-a-new-perspective>

Projection-based methods

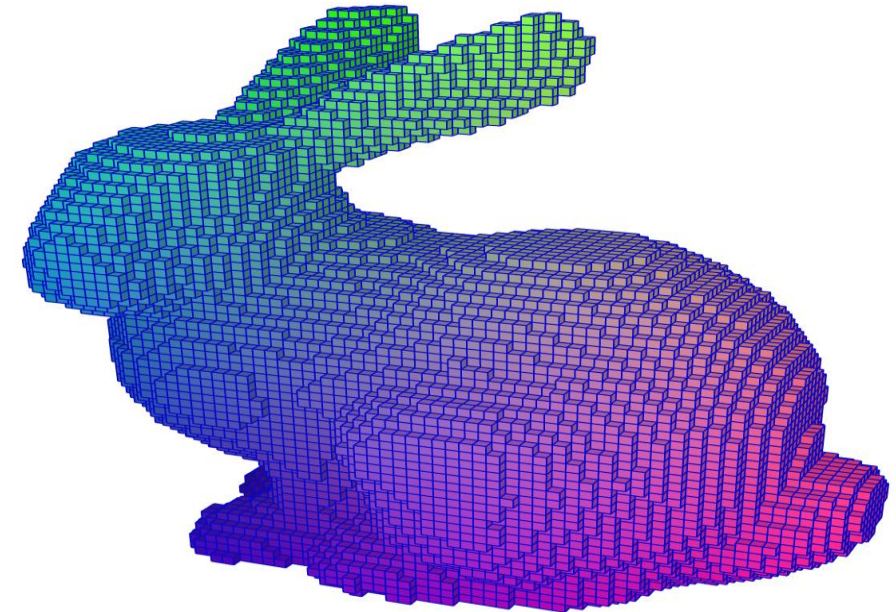
- Multi-view and Spherical Representation
 - Deep projective 3D semantic segmentation (Lawin et al.)
 - Tangent Convolution
 - Squeezeseg



What is the attribute of a specific view?

Volumetric-based methods

- Mapping points to voxel grids
 - Huang et al.
 - Graham et al.
- Challenges
 - Quantization error
 - Heavy computation of 3D convolution

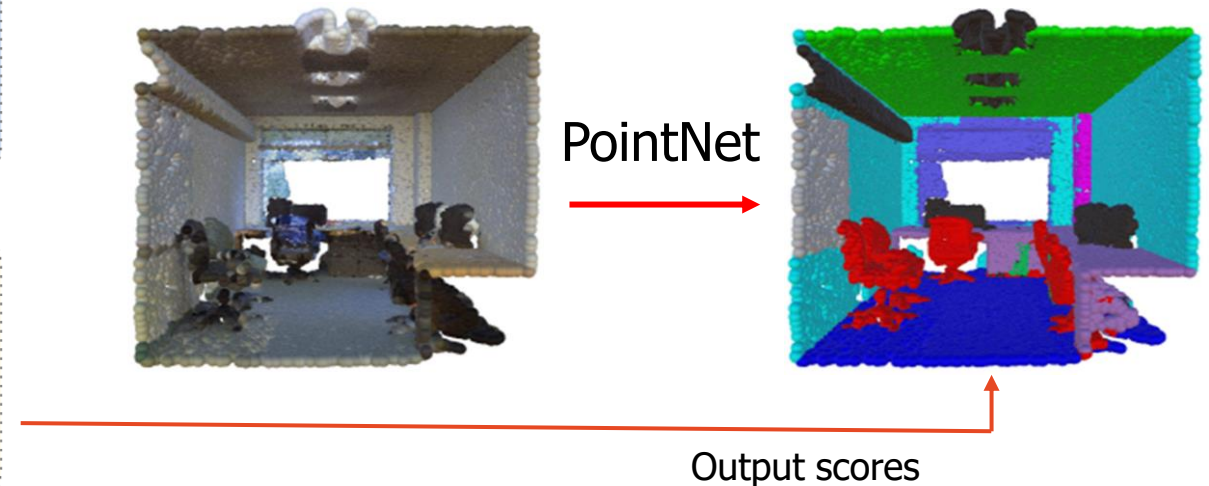
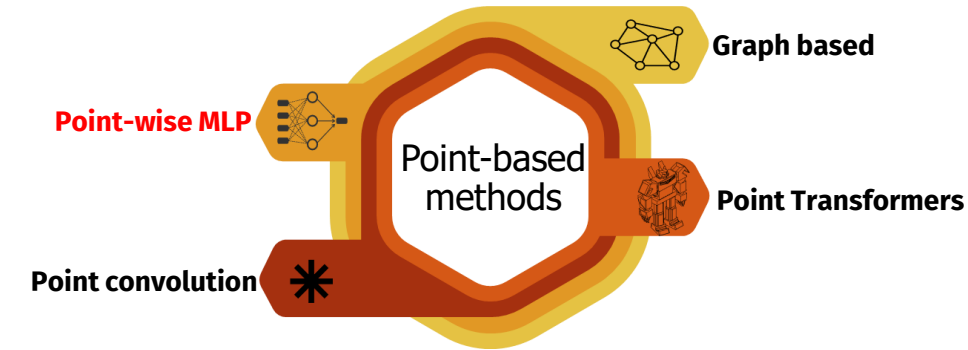
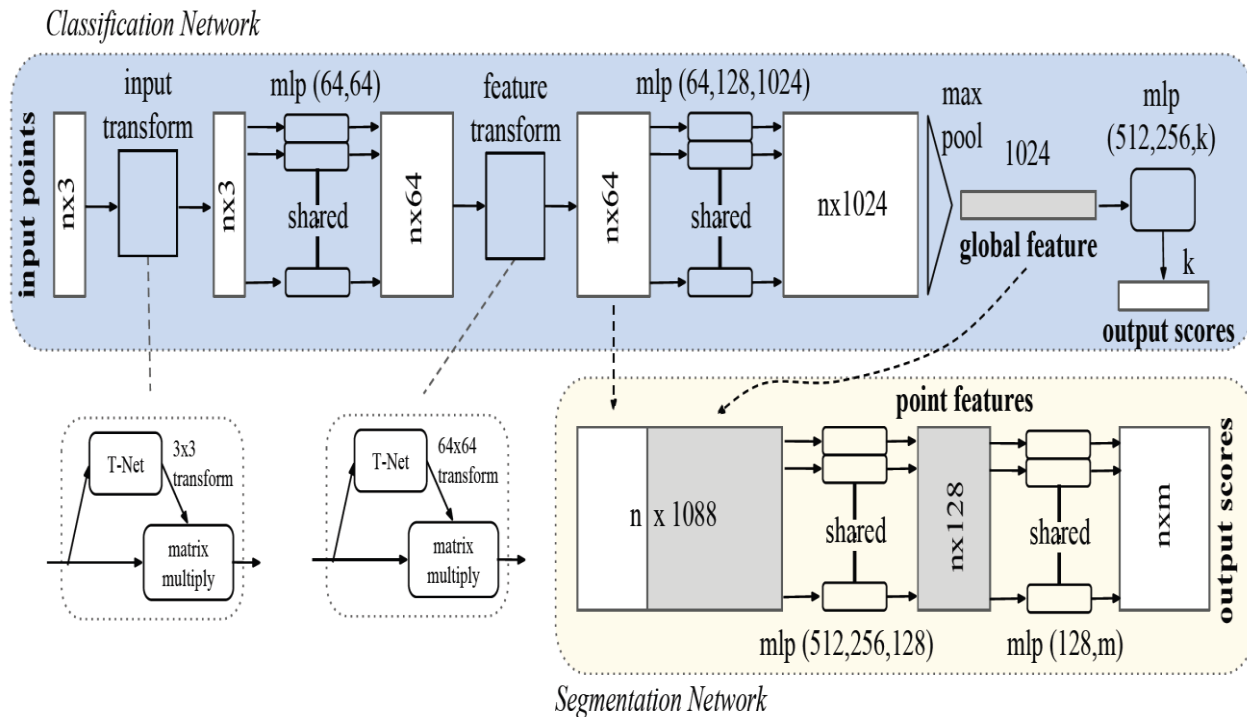


1- Huang, J., & You, S.(2017) Point cloud labeling using 3d convolutional neural network. In 2016 23rd International Conference on Pattern Recognition (ICPR) (pp. 2670-2675). IEEE.

2-Graham, B., Engelcke, M., & Van Der Maaten, L. (2026). 3d semantic segmentation with submanifold sparse convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 9224-9232).

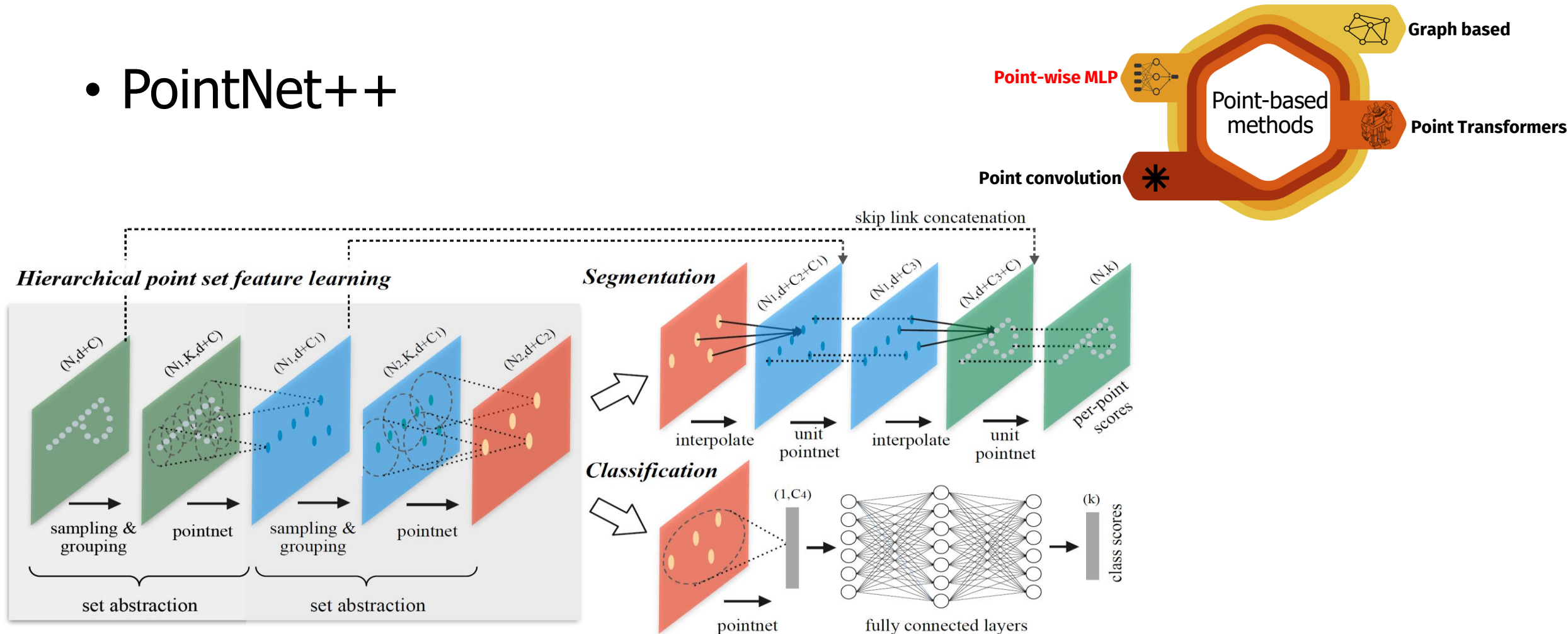
Point-based methods

- PointNet



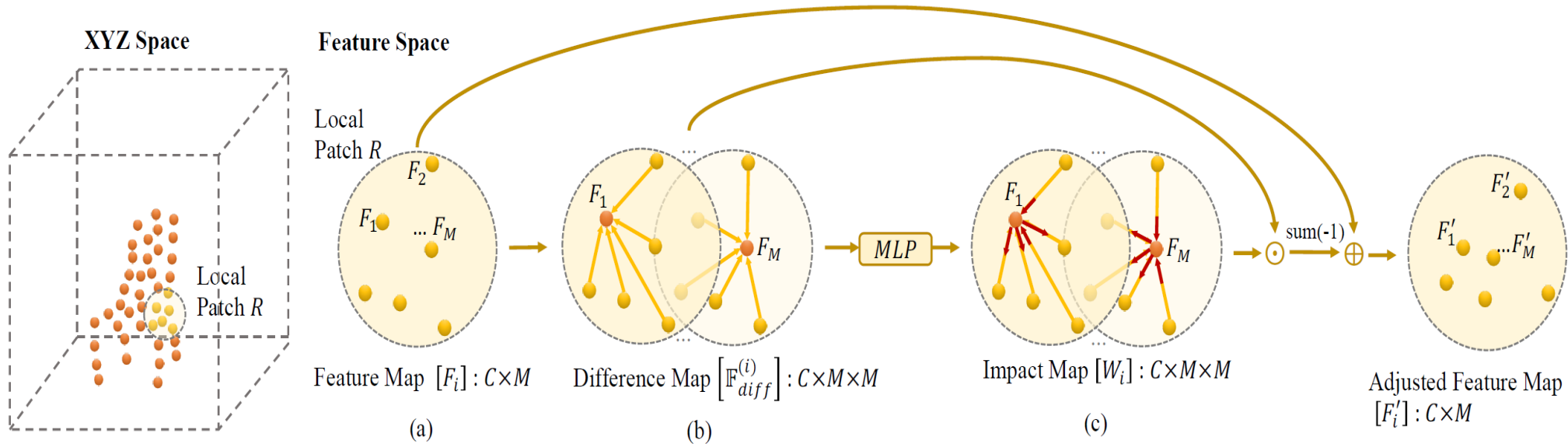
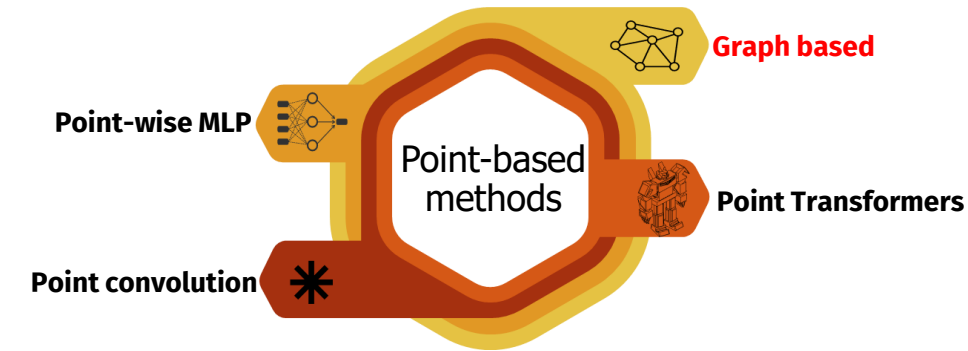
Point-based methods

- PointNet++



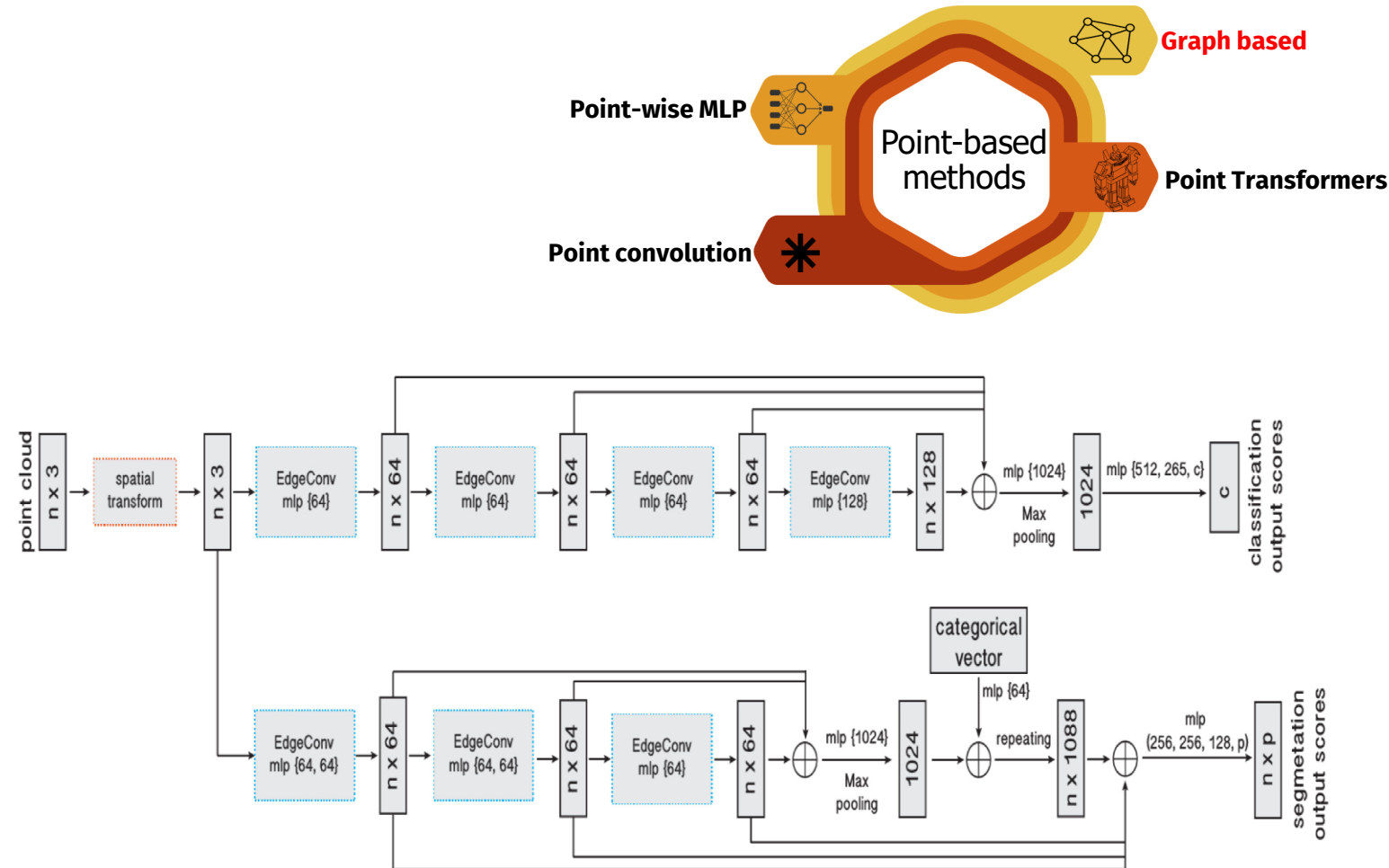
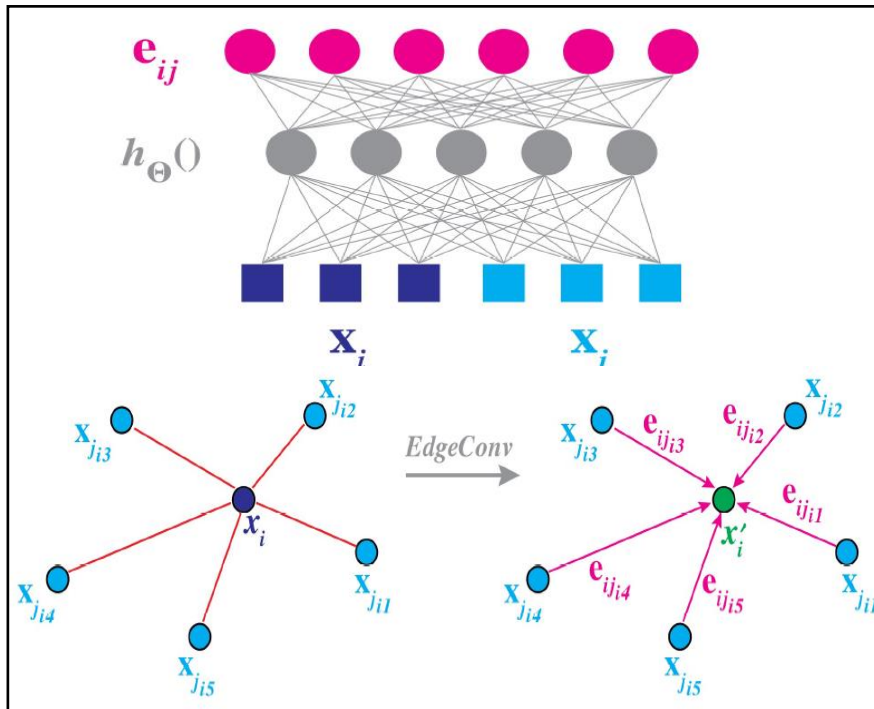
Point-based methods

- PointWeb



Point-based methods

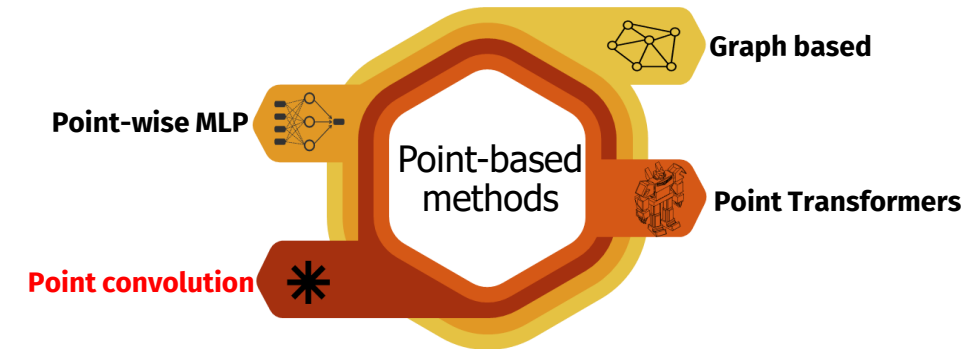
- DGCNN



Point-based methods

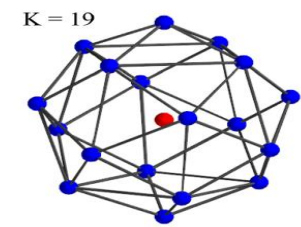
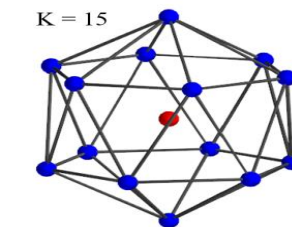
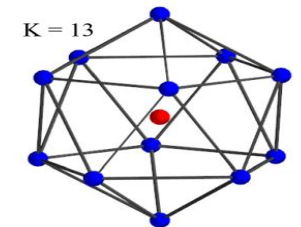
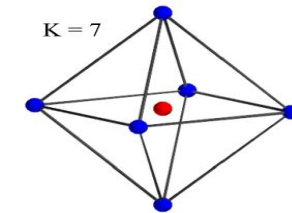
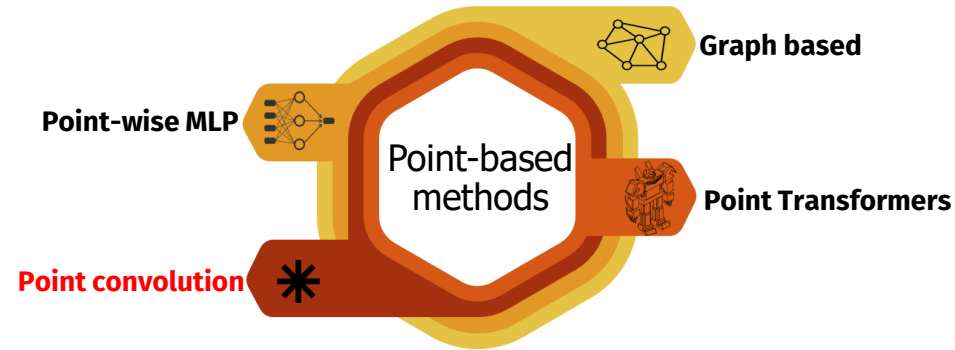
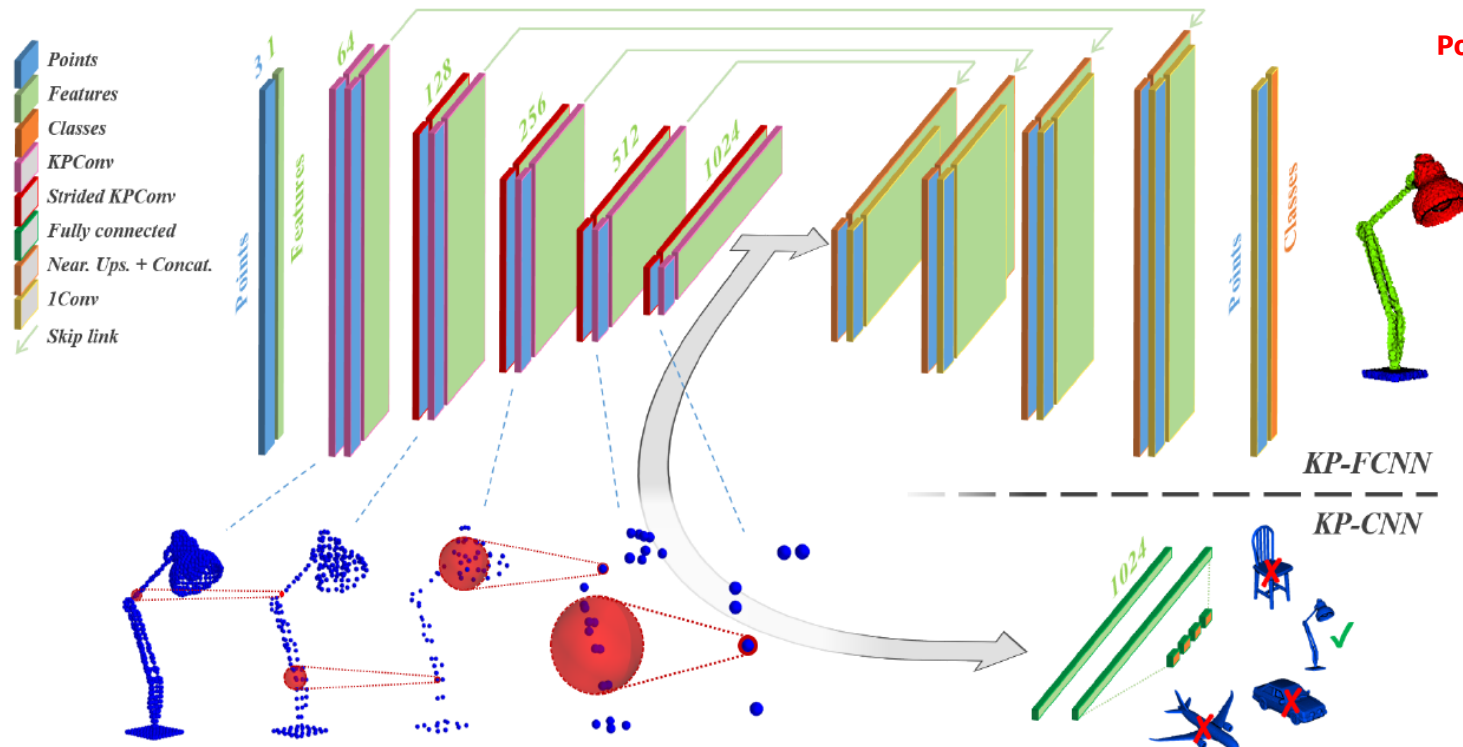
Adaptation of convolution on 3D Point Cloud

- PCNN
- PointCNN
- PCCN
- PointConv
- KPConv



Point-based methods

- KPConv



Point-based methods

- Point Cloud Transformer (PCT)
- Point Transformer

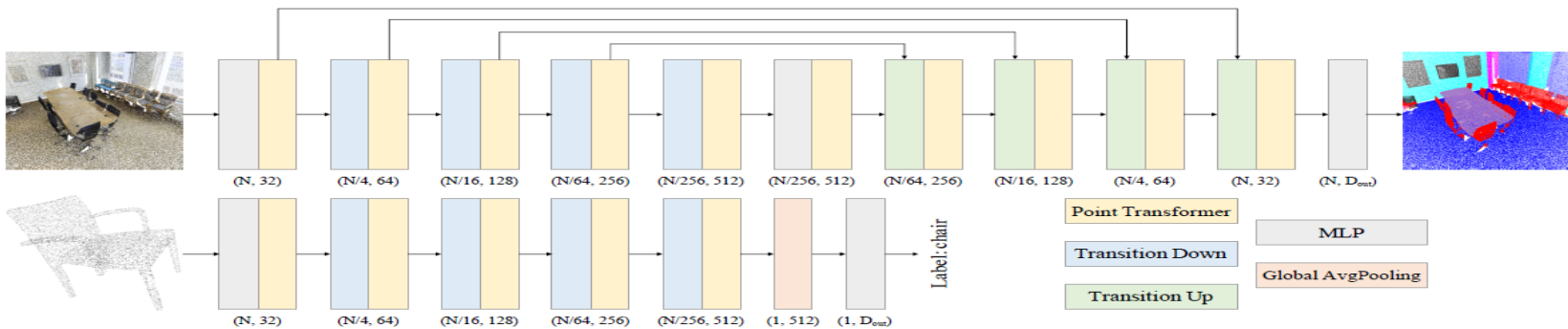
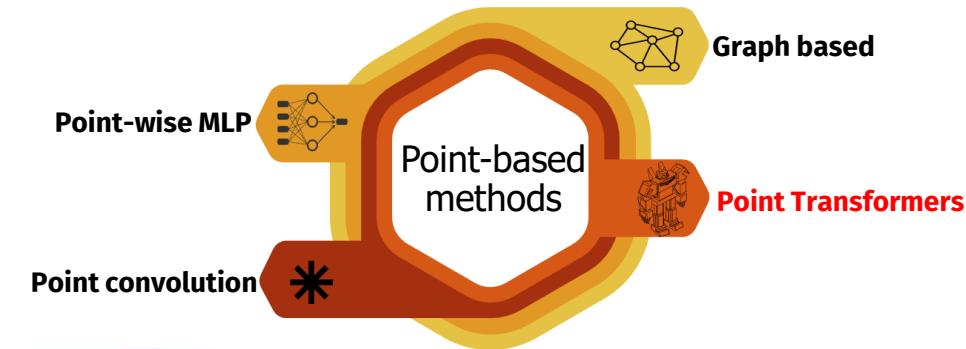
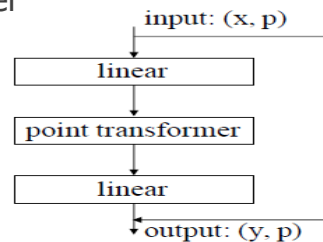
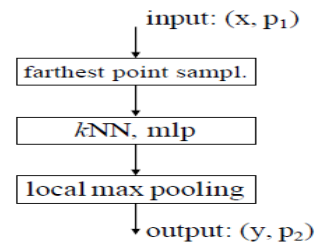


Figure 3. Point transformer networks for semantic segmentation (top) and classification (bottom).

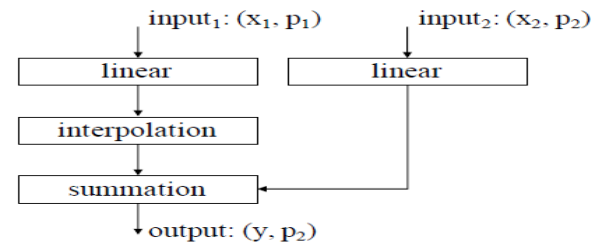
Point transformer



(a) point transformer block



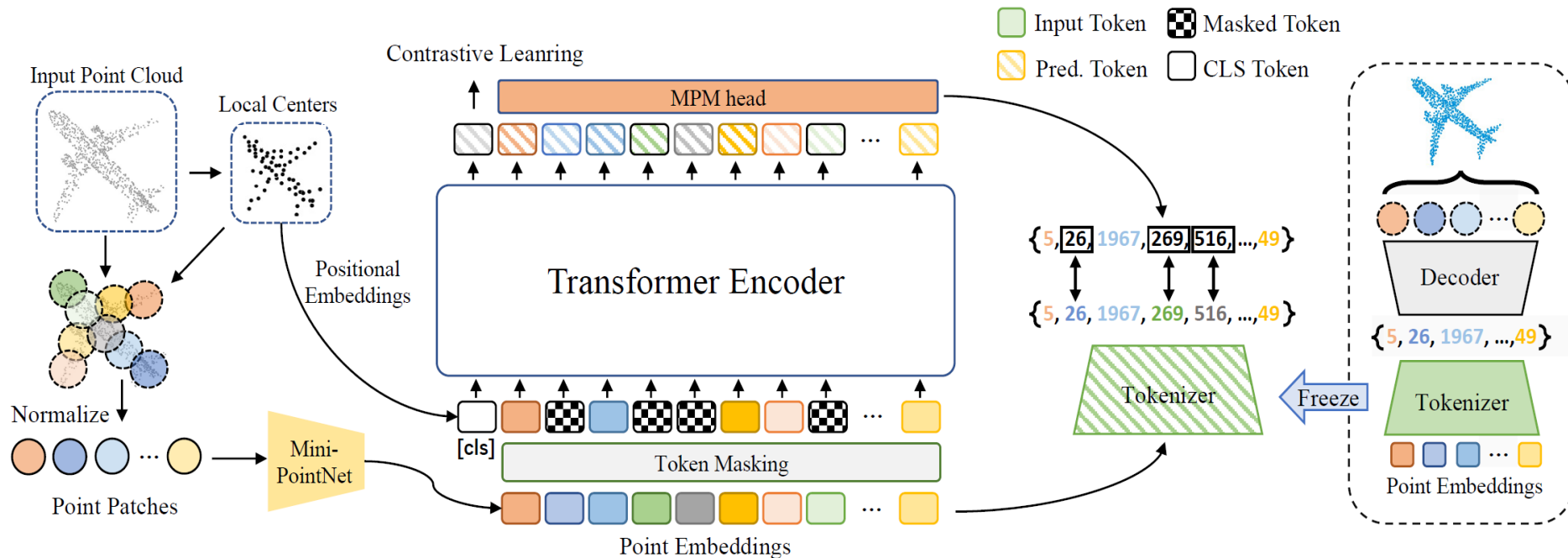
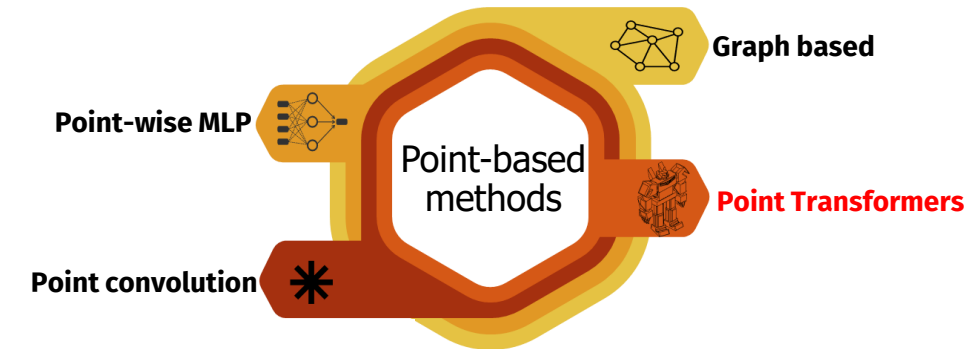
(b) transition down



(c) transition up

Point-based methods

- Point Transformer V2
- Point-BERT

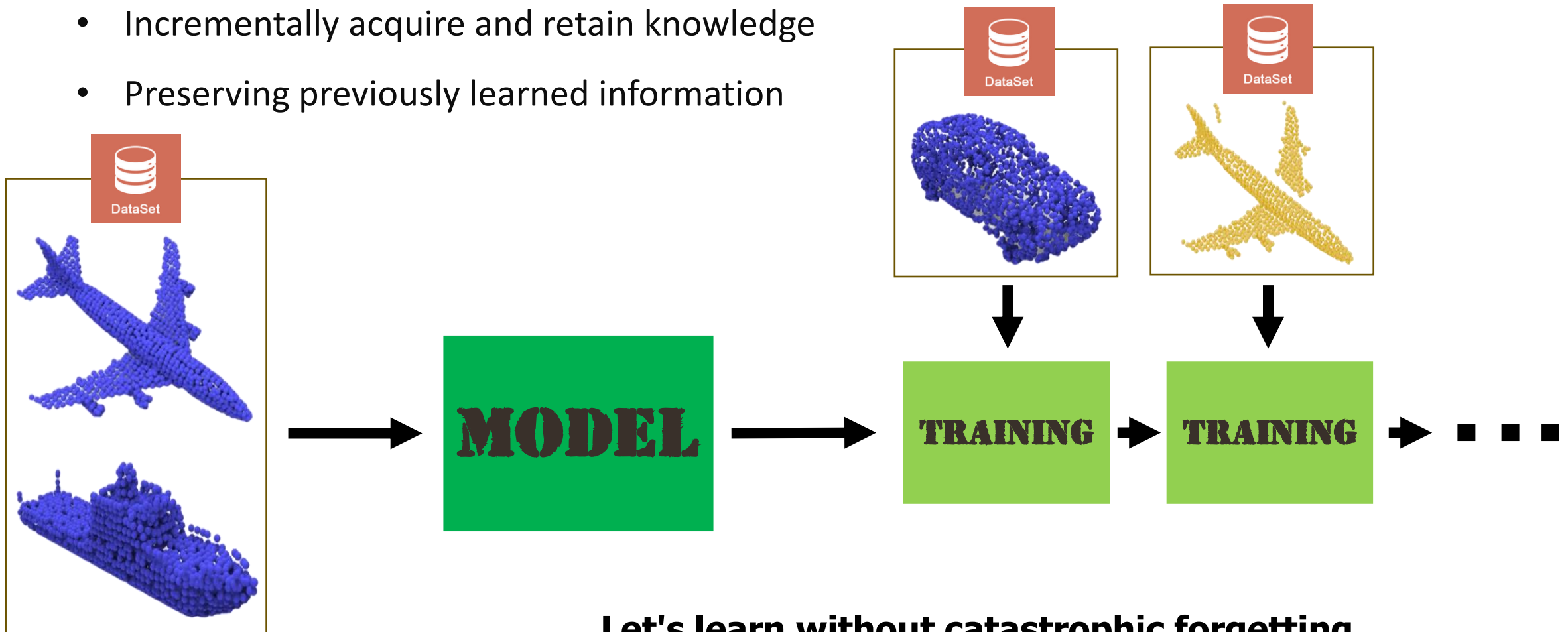


outline

- Introduction
- 2D Segmentation
- 3D Semantic Segmentation
- Future Trends
 - Incremental learning
 - Test-time Training on 3D point cloud
 - Test-time domain adaptation
 - Zero-shot and few-shot segmentation
- Conclusion

Incremental learning

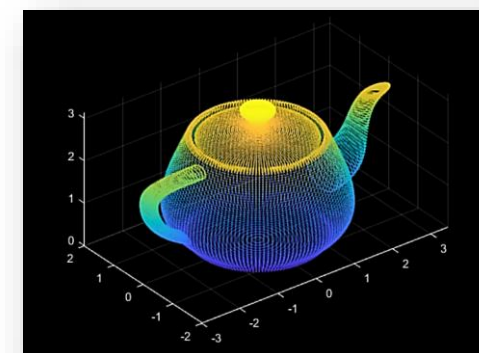
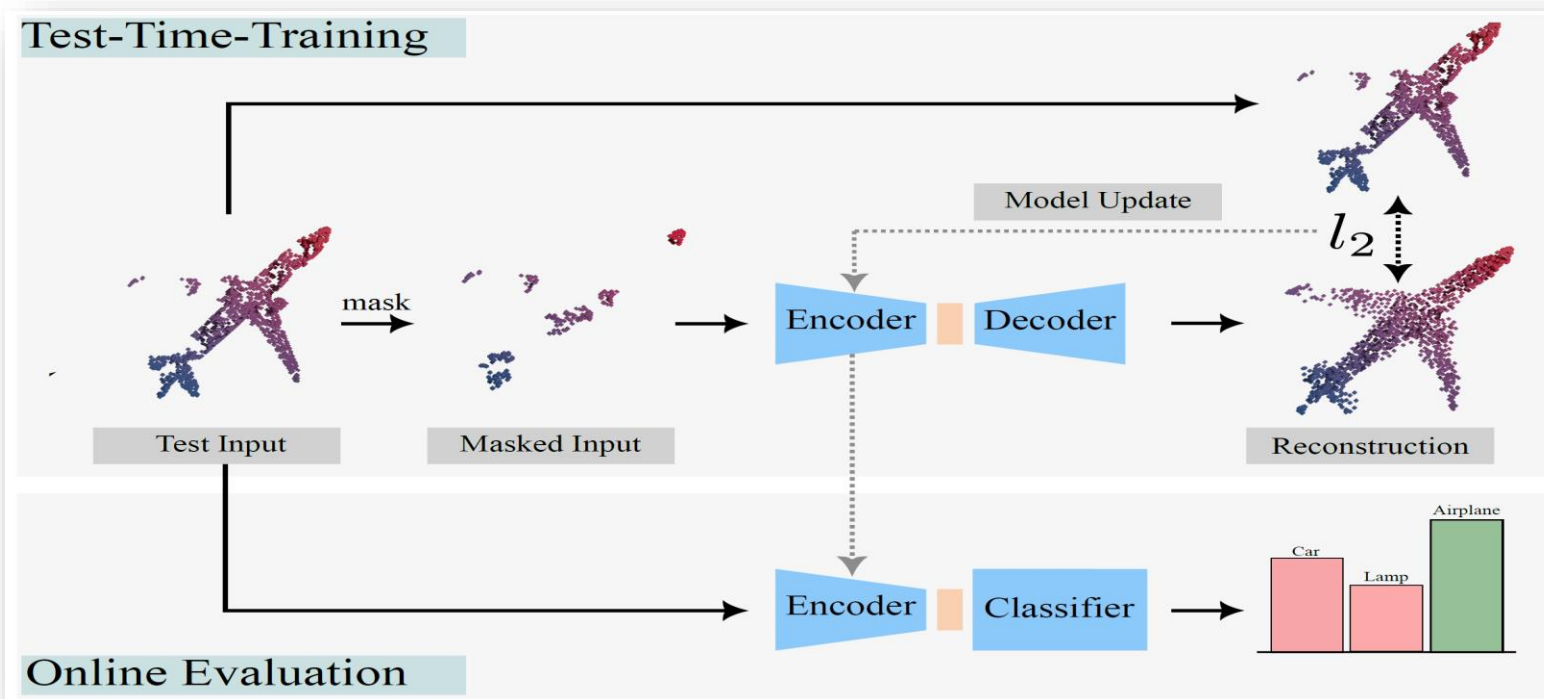
- Incrementally acquire and retain knowledge
- Preserving previously learned information



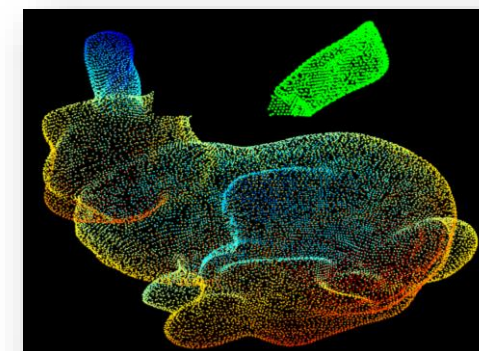
Let's learn without catastrophic forgetting

Test-Time Training on 3D point cloud

- Masked Auto Encoders are online 3D learners

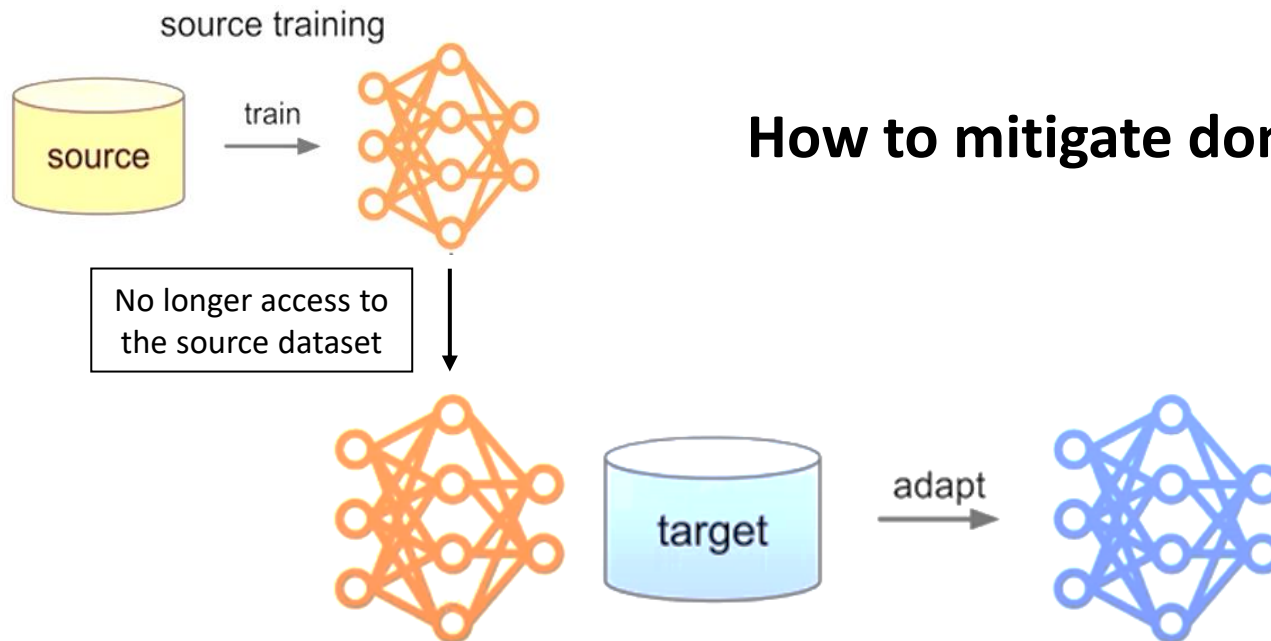


Synthetic to real



Test-Time Domain Adaptation

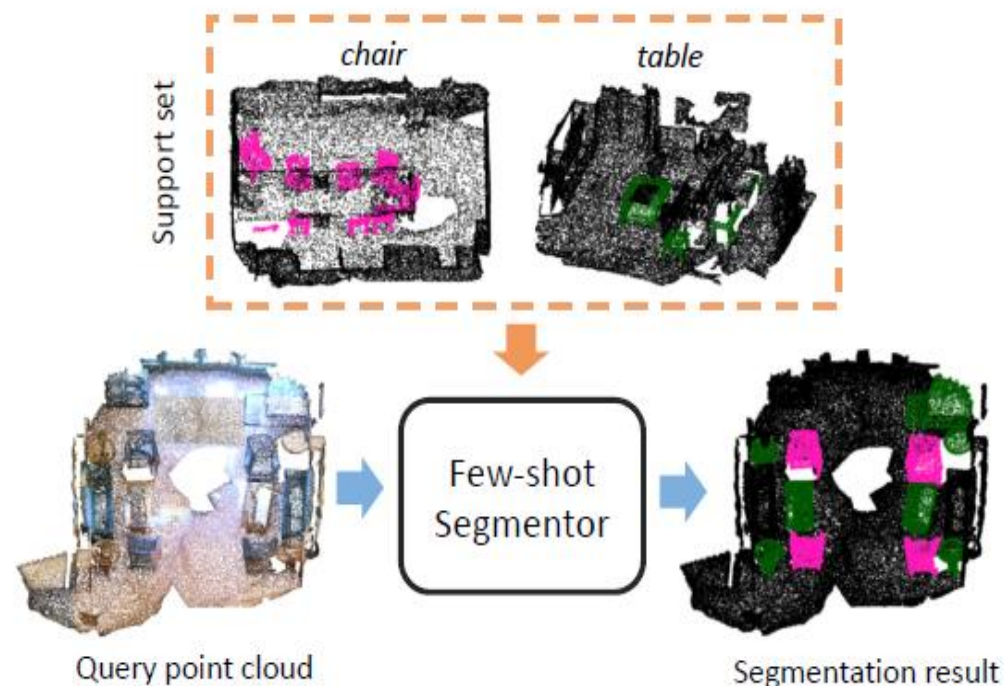
- The need for adopting model to the target dataset
- Source dataset is no longer available



How to mitigate domain shift?

Zero-shot and Few-shot segmentation

What happens if confronting new classes?



1- Cheraghian, Ali, et al. "Zero-shot learning on 3d point cloud objects and beyond." International Journal of Computer Vision 130.10 (2022):

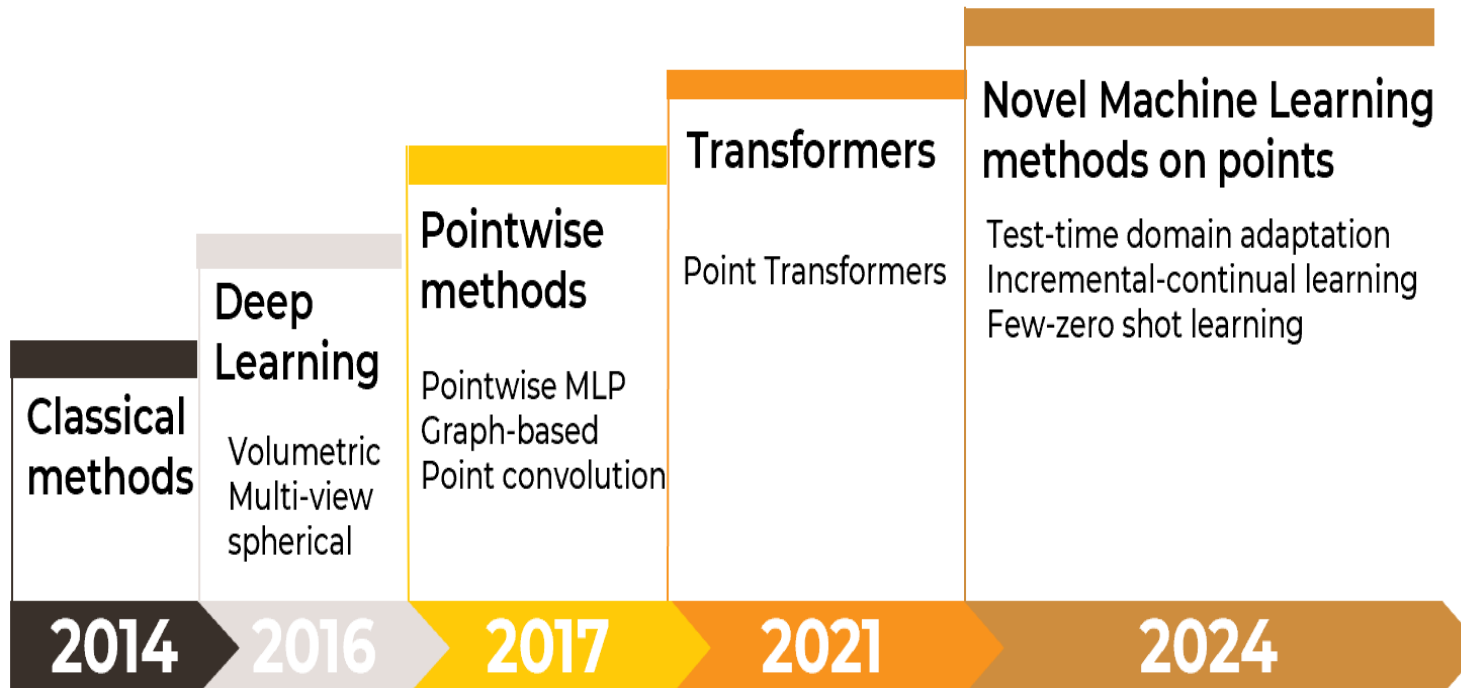
2- Zhao, Na, Tat-Seng Chua, and Gim Hee Lee. "Few-shot 3d point cloud semantic segmentation." Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.

outline

- Introduction
- 2D Segmentation
- 3D Semantic Segmentation
- Future Trends
- Conclusion

Conclusion

Semantic segmentation on point cloud is challenging!



Method	mAcc	mIOU
PointNet	48.98	41.09
PointNet++	59.8	-
DGCNN	84.1	56.1
PointCNN	63.86	57.27
PCCN	67.0	58.0
PointWeb	66.64	60.28
PCT	67.01	61.33
KPConv	72.8	67.1
Point Transformer	76.5	70.4
Point Transformer V2	78.0	71.6

Evaluation metrics of methods on S3DIS

Thanks for your attention



Scan this QR code to access
the presentation materials

Feel free to contact me

Mehrant.0611@gmail.com
Telegram: @Mttnt

