

POINT CLOUD IN A WORLD OF DEEP LEARNING

Learn more about different neural network architectures for point clouds

Vladyslav Selotkin,
R&D Engineer

Volodymyr Karpiv,
R&D Lead

softserve

Three-dimensional (3D) data is readily available, and its usage is growing exponentially in different areas—from security and communication, to self-driving cars. There are multiple ways to represent 3D information: point cloud, voxel grid, mesh, and functions.

In this whitepaper, we'll focus on point cloud which is the base type to represent sensor data like LIDAR (light detection and ranging),

ToF (Time-of-Flight), structured light system, and many others.

Point clouds provide in-depth information without data simplification—unlike voxel grids, meshes, or a function representation that are hard to estimate for complex objects. The primary drawbacks of point cloud are sparsity of 3D space and unstructured data—which we will also address.



Point cloud

Voxel grid

Mesh

NEURAL NETWORKS FOR POINT CLOUD

Neural networks perform effectively in different domains. However, in the computer vision field, neural networks treat structured data like images. To apply neural networks to point clouds, other approaches and techniques should be developed. They differ from standard convolutional neural networks.

With some tweaks, neural networks can be used to solve the following tasks on point clouds, including but not limited to:

- Classification—detect a type of the point cloud
- Segmentation—split a point cloud to semantic parts
- Completion—fill missing values in a point cloud
- Reconstruction—generate a point cloud from a single or multiple 2D images
- Registration—stitch parts of the object into a single one



POINTNET/POINTNET++

Working with point clouds requires the use of symmetric functions (e.g. invariant to order of the arguments). Simple operations on scalars—like addition and multiplication—are a simple example of such functions. Fully connected networks are based on these simple operations.

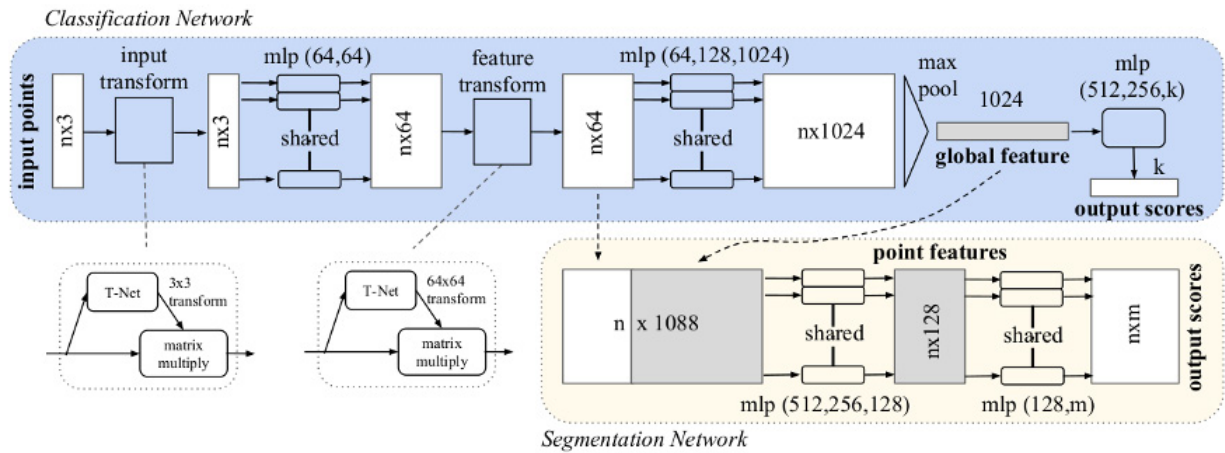
The first approach utilizing the idea of symmetric functions was presented by Charles R. Qi et al. in [PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation](#). Their idea was to use a pair of multilayer perceptrons—combined with the input transformation and feature transformation block—to extract a set of features from point clouds.

Input and feature transformation blocks are required to preserve transformation invariance of the neural network. These small subnetworks learn transformation matrices which can be used to transform point clouds to the canonical form in 3D and feature spaces accordingly.

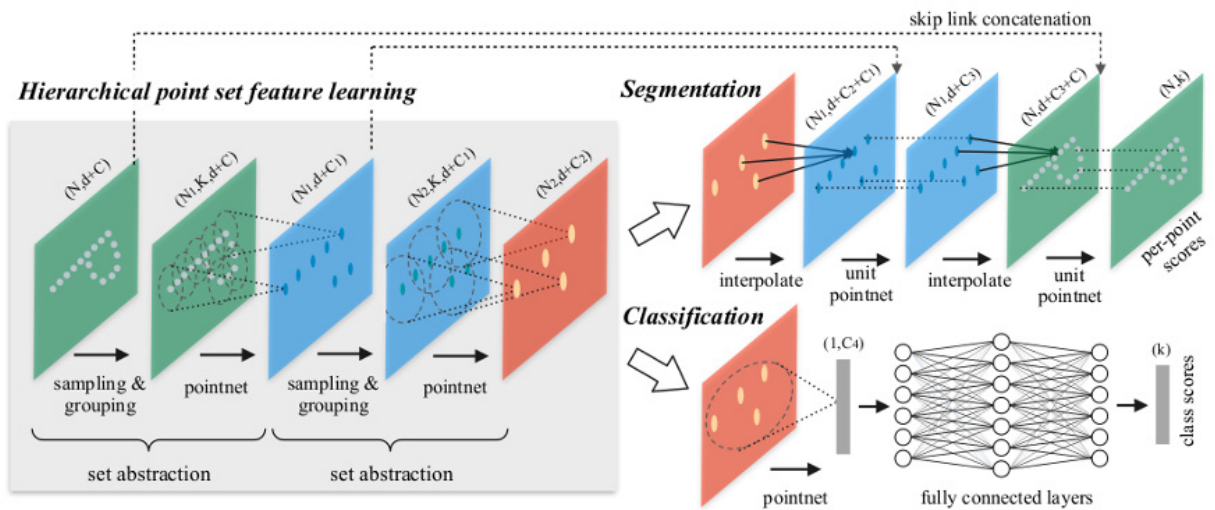
This network can be adopted effectively in any type of task.

PointNet++ is the extension of the classical PointNet and is based on Charles. R. Qi et al. added sampling and grouping layers to the neural network. It allowed them to learn local features of the point cloud by applying PointNet to the groups of point clouds. This helps managing point clouds with variable densities.





PointNet architecture from <http://stanford.edu/~rqi/pointnet/>



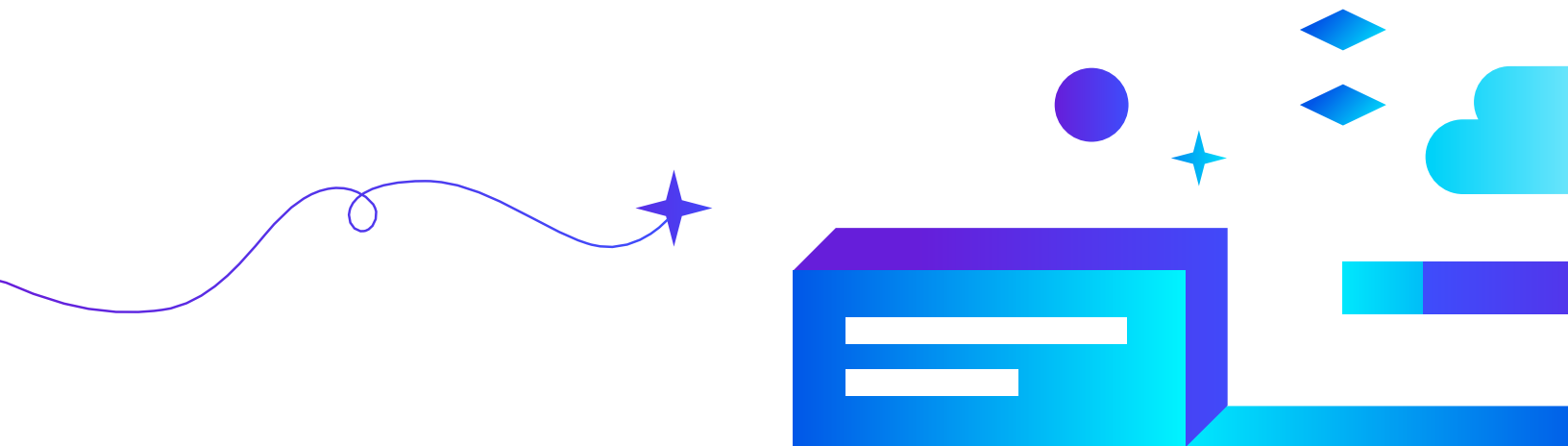
PointNet++ architecture from <http://stanford.edu/~rqi/pointnet2/>

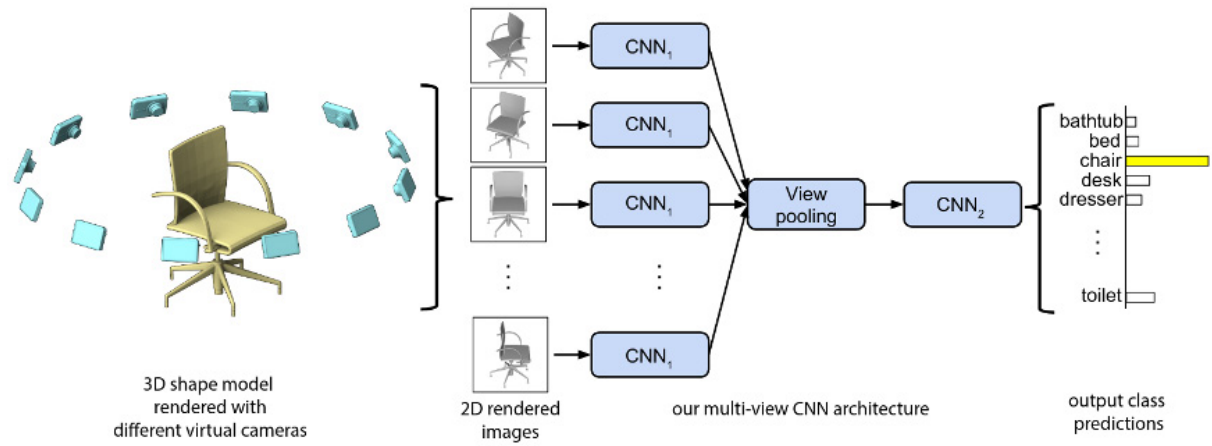
2D AND 3D CONVOLUTIONS

Convolutional neural networks show incredible performance in various image tasks. It should be noted, however, that point clouds can't be directly processed by classic CNN.

There are two convolution approaches for managing point clouds.

1. Generate a set of 2D images obtained from different views. 2D images are rendered from all sides. This patch of images then proceeds together through a regular CNN to generate a feature set which can then be used to predict the type of the point cloud. For example, take a look at the [Multi-view Convolutional Neural Networks for 3D Shape Recognition](#) paper.
2. Use 3D convolution operation or their analog. Various approaches were made to extend convolution to point clouds. To learn more, read the following publications [ELF-Nets: Deep Learning on Point Clouds Using Extended Laplacian Filter](#) and [PointConv: Deep Convolutional Networks on 3D Point Clouds](#). In these works, the authors suggested new types of layers based on convolution:
 - a. The first one is based on the 2D discrete Laplace operator. ELF employs two matrices: one for a center point and another for its neighboring points. The second matrix is scaled using a weighting function which is determined by the local configuration of the neighborhood.
 - b. The second algorithm is based on PointConv, which is an extension to the Monte Carlo approximation of the 3D continuous convolution operator. For each convolutional filter, it uses MLP to approximate a weight function, then applies a density scale to re-weight the learned weight functions.





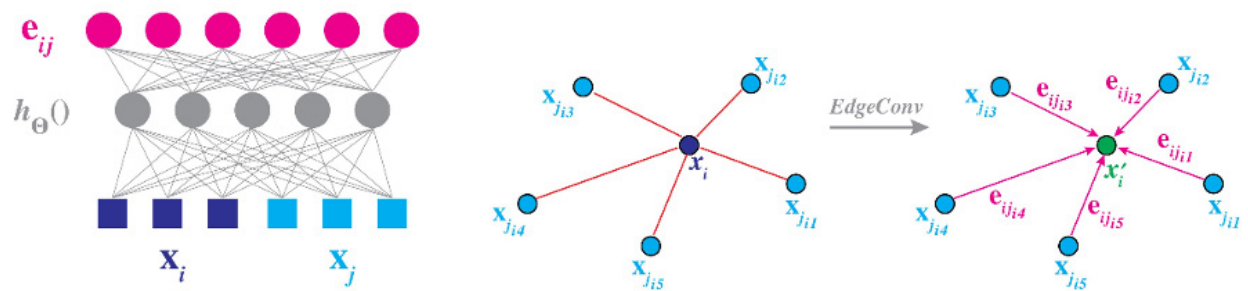
Multi-view CNN architecture from https://www.cv-foundation.org/openaccess/content_iccv_2015/papers/Su_Multi-View_Convolutional_Neural_ICCV_2015_paper.pdf

GRAPH CNN

A point cloud can be presented in multiple ways with the help of graphs. For instance, the whole point cloud can be presented as one unstructured graph. Then, different types of graph NN are applied to the point cloud. For example, Y. Zhang and M. Rabbat, authors of [Dynamic Graph CNN for](#)

[Learning on Point Clouds](#), used a Graph CNN for point cloud classification. This approach ensures robustness to rotation invariance.

It is also possible to represent point clouds as a set of local neighborhood graphs. Then, operations similar to convolution on several local graphs are applied to extract features.



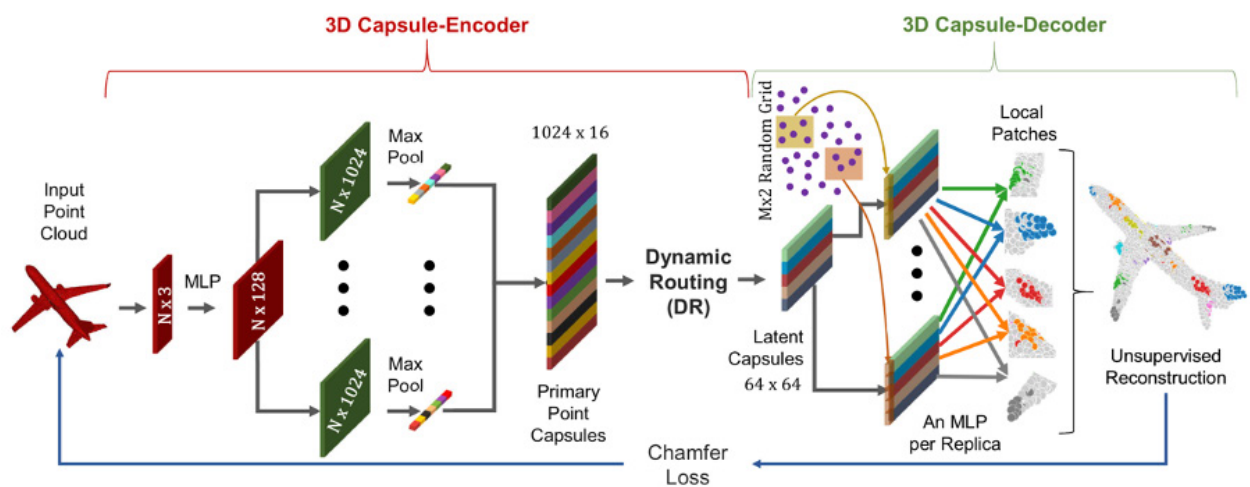
Example of Graph convolution operation from <https://arxiv.org/pdf/1801.07829.pdf>

3D POINT CAPSULE NETWORKS

Convolutional neural networks have revolutionized many fields of computer vision by reaching superhuman levels of performance. While exploiting the core benefits of CNNs, one should note existing drawbacks to this approach.

As argued by Geoffrey Hinton for many years, convolutional neural networks do not consider the important spatial relations of simple and complex objects. This is a direct consequence of the translational and rotational invariances of CNNs.

Capsule network, proposed by J. Hinton, is a deep learning building block that helps capture spatial hierarchies of the environment. Extracting important information from the data enables the sample efficiency of capsule networks with respect to their CNN counterparts.



3D Point Capsule Networks from <https://arxiv.org/pdf/1812.10775.pdf>

Y. Zhao et al. propose 3D capsule networks to enable richer representation than a one-dimensional latent embedding of common point cloud autoencoders. The embedded

latent capsules trained with dynamic routing achieve top results in feature extraction, segmentation, 3D reconstruction, interpolation, and replacement.

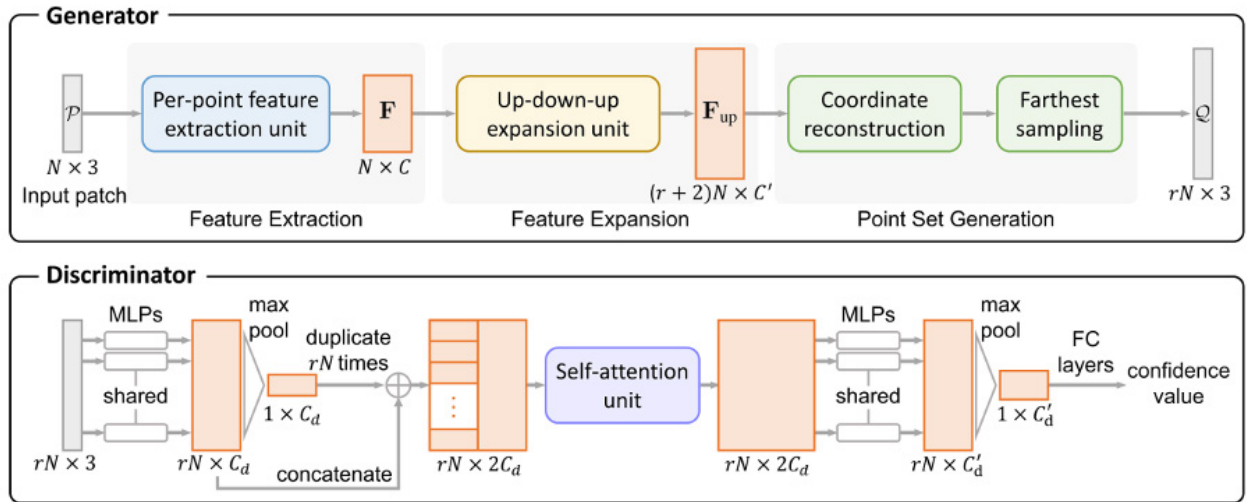
GANS AND AUTOENCODERS

Generative models could be used as a base approach to tackle not only the 3D reconstruction task, but also the derivative ones, such as classification, for example. For the moment, both methods are implemented in combination with the PointNet architecture and the other tricks to get max performance.

In their work [Learning Representations and Generative Models for 3D Point Clouds](#), P. Achlioptas et al. showed that GAN can learn a compact representation of the raw point cloud. This representation then can be used for classification, completion, generation, and other tasks.

Another way of applying GANs is to use them for point cloud upsampling. In the paper [PU-GAN: a Point Cloud Upsampling Adversarial Network](#) by R. Li et al., a new architecture was presented. This network fills tiny holes in point cloud data and even fixes the position of noisy points.





Generator and discriminator of the PU-GAN architecture from <https://arxiv.org/pdf/1907.10844.pdf>

IN CONCLUSION

We have reviewed different neural network architectures for point clouds. So far, there is no common vision on point cloud processing, as various approaches with both pros and cons arise. Our future work will be focused on reviewing algorithms for point clouds registration and benchmarking. Stay tuned for more!



ABOUT



VOLODYMYR KARPIV

Volodymyr Karpiv, Research Lead at SoftServe, is responsible for separating hype from really working stuff in the fields of artificial intelligence, computer vision, and signal processing. He is passionate about cross-disciplinary research and its further applications in the industry. Volodymyr's work interests are related to technologies that will have the highest impact in the future.



VLAD SELOTKIN

Vlad Selotkin is an Intermediate Research Engineer with more than two years' of experience in software development and research in the field of computer vision, deep learning, and signal and image processing. He has worked on many challenging projects including digital holography, compressive sensing, and neural networks for 3D reconstruction. Vlad has published and presented several papers, including works on creating an end-to-end solution for 3D recognition using structured-light technology.

ABOUT US

SoftServe is a digital authority that advises and provides at the cutting-edge of technology. We reveal, transform, accelerate, and optimize the way enterprises and software companies do business. With expertise across healthcare, retail, energy, financial services, and more, we implement end-to-end solutions to deliver the innovation, quality, and speed that our clients' users expect.

SoftServe delivers open innovation, from generating compelling new ideas, to developing and implementing transformational products and services.

Our work and client experience is built on a foundation of empathetic, human-focused experience design that ensures continuity from concept to release.

We empower enterprises and software companies to (re)identify differentiation, accelerate solution development, and vigorously compete in today's digital economy-no matter where you are in your journey.

Visit our [website](#), [blog](#), [LinkedIn](#), [Facebook](#), and [Twitter](#) pages.

NORTH AMERICAN HQ

201 W 5th Street, Suite 1550
Austin, TX 75703 USA
+1 866 687 3588 (USA)
+1 647 948 7638 (Canada)

EUROPEAN HQ

14 New Street
London EC2M 4HE
United Kingdom

info@softserveinc.com
www.softserveinc.com

softserve