**Проблемы классификация видов деревьев по данным LiDAR с помощью модели глубокого обучения**

**Issues of tree species classification from LiDAR data using deep learning model**

**Sakharova K. Elizaveta, Nurlyeva D. Dana, Fedorova A. Antonina, Kanev I. Anton, Valery I. Terekhov**

Сахарова Елизавета Константиновна, Нурлыева Дана Джалилевна, Федорова Антонина Алексеевна, Канев Антон Игоревич, Терехов Валерий Игоревич

Bauman Moscow State Technical University, Moscow, Russia

**Аннотация**

Учет насаждений деревьев является важной задачей для экологии и обустройства городских парков. Это влияет и на расчет поглощения углекислого газа. При этом разные виды деревьев обладают своими особенностями, и актуальной является задача их классификации. В настоящее время используются различные датчики для решения этой задачи. Наиболее часто исследователи применяют данные LiDAR в своих работах. При этом используются различные методы классификации, в том числе глубокое обучение. Одной из современных моделей глубокого обучения на данных облака точек является PointNet. Поэтому авторы данной работы применили ее для классификации видов деревьев. Россия обладает своим определенным набором видов деревьев, поэтому важно проводить исследование видах, произрастающих на данной местности. Авторы собрали набор данных, содержащий виды деревьев, характерные для европейской части России, и провели его разметку. Полученные результаты выявили проблему недостаточного количества данных для высокой точности обучения. Также были выявлены проблемы с ручной разметкой экземпляров по данным облака точек.

Accounting for tree plantations is an important task for forest inventory, the ecology and arrangement of city parks. This also affects the calculation of carbon dioxide absorption. At the same time, different types of trees have their own characteristics, and the problem of their classification is urgent. Various sensors are currently used to meet this challenge. Researchers most often use LiDAR data in their work. At the same time, various classification methods are used, including deep learning. One of the modern models of deep learning on point cloud data is PointNet. Therefore, the authors of this work applied it to classify tree species. Russia has its own specific set of tree species, so it is important to conduct research on the species growing in the area. The authors collected a dataset containing tree species typical for the European part of Russia and carried out its labeling. The results obtained revealed the problem of insufficient data for high training accuracy. Also, problems were identified with manual marking of instances based on point cloud data.

**Ключевые слова**

LiDAR, tree species classification, deep learning

LiDAR, классификация видов деревьев, глубокое обучение

**Introduction**

At the moment all over the world one of the most discussed and urgent problems is the problem of ecology. As it is known, there are a huge number of different factors that affect the state of nature around us. One such factor is the carbon footprint that anyone leaves behind. The carbon footprint is the aggregate of all greenhouse gas emissions that directly or indirectly accompany any activity of a person, organization, or produced by a product, event, city, in terms of carbon dioxide. A carbon footprint is the total amount of greenhouse gases (including carbon dioxide and methane) that are generated by our actions.

The average carbon footprint for a person in the United States is 16 tons, one of the highest rates in the world. Globally, the average is closer to 4 tons. To have the best chance of avoiding a 2℃ rise in global temperatures, the average global carbon footprint per year needs to drop under 2 tons by 2050 [1].

In order to achieve this goal, the Paris Agreement was adopted. The Paris Agreement is a legally binding international treaty on climate change. It was adopted by 196 Parties at COP 21 in Paris, on 12 December 2015 and entered into force on 4 November 2016. Its goal is to limit global warming to well below 2, preferably to 1.5 degrees Celsius, compared to pre-industrial levels. To achieve this long-term temperature goal, countries aim to reach global peaking of greenhouse gas emissions as soon as possible to achieve a climate neutral world by mid-century [2].

Under this agreement, countries are encouraged to take action to implement and support activities related to forest conservation, sustainable forest management and increasing forest carbon stocks in developing countries. [8] That is why it becomes necessary to monitor the situation with forest plantations and how much carbon will be absorbed by this or that part of the forest plantation. Accurate characterization of forest species and their spatial distribution is critical for sustainable forest management and for ecological and environmental protection. The estimation of biomass, carbon content, species diversity, and the condition of a forest community requires precise individual-tree information by species [4]. In the case of urban areas, tree species classification is gaining increasing attention for safety studies, noise modeling, and environmental and ecological analysis because trees play a critical role in urban ecosystems for the maintenance of environmental quality, aesthetic beauty of urban landscape, and social service for inhabitants [6].

The number of studies focusing on tree species classification has constantly increased over the last 35 years which is well-supported by the general trend of increased publication activity [7]. An almost exponential increase can be seen between the periods 2005–2010 and 2010–2015, driven by the increased availability of hyperspectral and airborne LiDAR data. Both data sources have been frequently applied in a forest inventory context with tree species being one of the most popular target variables besides total growing stock volume and biomass [7].

Existing methods for classifying forest species from remote sensing data are mostly based on the spectral information from forest canopies. Despite vegetation cover classifi- cation successes at the stand- and landscape-level, the accuracy of individual-tree classification remains low. This is attributed mainly to: the effects of several spectral and spatial factors on sur- face reflectance values of forest canopies and limitation in the spectral and spatial configurations of image sensors [4].

Various types of lidars are already being used to classify tree species. For the past decade, researchers have examined the potential to use data from airborne LiDAR to classify forest stand types or individual species[4]. Later mobile LiDAR has attracted much attention for urban vegetation detection and model- ling because it acquires data at a much higher point density and more complete data coverage than an airborne LiDAR system and at a higher efficiency than a terrestrial LiDAR system [6].

Once the data has been received, there are many ways to process and analyze it. For example, the Finnish company Arbonaut Oy Ltd uses LiDAR systems for forest inventory with the method based on sparse Bayesian regression for modeling forest characteristics. Keep in mind that this system is used in Finland, where there is a smaller variety of trees that are planted in straight lines. Such plantings are easier to analyze, unlike forest plantations in the Russian Federation, where the order and structure of forests is more chaotic, and the diversity of species is much greater [9].

There are many different algorithms for classifying trees. At first, the most widely used classification techniques included supervised maximum likelihood classifiers and unsupervised clustering (K-means, ISODATA). These methods were easily applicable. Later, non-parametric decision tree based classifiers and neural networks emerged as an alternative to the other classifiers. These classifiers do not require the input data to be normally distributed. Some recent studies using mixed sets of input variables have preferred the use of non-parametric machine learning methods like RF or SVM[7]. Deep learning and neural networks are best suited for such tasks. In our case, we have sufficient computing power. Therefore, this study will consider such an algorithm.

There are more and more different techniques that implement LiDAR technology. LiDAR is a remote sensing technology. LiDAR technology uses the pulse from a laser to collect measurements. These are used to create 3D models and maps of objects and environments. With the help of this technology, it becomes possible to automatically shoot forest tracts and receive point clouds at the output, combined with real images obtained from an additional camera. Further, based on the obtained point clouds, it is possible to automatically analyze the types of trees in the forest.

LiDAR (Light Detection and Ranging) is a remote sensing technology that uses light in the form of a pulsed laser to measure variable distances to an object. By emitting several thousand infrared laser pulses per second, as well as recording the location and orientation of the device in space, a dense three-dimensional point cloud is formed that reflects the surrounding space. The growing availability of LiDAR has generated great interest among natural resource managers, as LiDAR can be used to measure tree characteristics [14].

There are various models for recognizing LiDAR point clouds. But the most effective are deep learning models. The point cloud model architecture can be divided into two broad groups: direct and indirect. Direct methods use the point cloud itself as input data for the neural network. To use indirect methods, it is necessary to first translate the point cloud into another form of data representation (3D voxel grids or image collections) [21]. Image collections mean a sequence of 2D images captured from different angles. Next, convolutional neural networks (CNN) are applied to them, after which the result of the markup is projected back into three-dimensional space. Voxel-based methods transform irregular and sparse point clouds into regular 3D grids [15]. Direct approaches are more efficient, because when generating voxel grids and images, an error appears, which affects the quality of the analysis of point clouds. Let's look at several different models with a direct and indirect approach: SnapNet, SEGCloud, PointNet, and PointNet ++.

The SnapNet input includes a large number of RGB point cloud imagery and a depth map. SnapNet generates a series of RGB-D image pairs that are projections from different angles for the original point cloud and performs pixel marking. Next, a reverse projection of the received marks in 3D takes place [22]. SEGCloud combines the methods of classical machine learning and deep architectures, that is, this model uses voxel grids and raw 3D points obtained using trilinear interpolation [23]. PointNet works directly with the point cloud [3]. This network has shown impressive results in indoor 3D object recognition and semantic segmentation. [16] PointNet ++ uses PointNet recursively to explore local features. PointNet ++ has been successfully applied for tree identification on point clouds [17]. Therefore, it was decided to use PointNet to classify tree species by point cloud.

**PointNet classification of LiDAR data**

A point cloud is the simplest way to represent an object as a set of single points in a X, Y and Z geometric coordinate system, but when it includes colour information the coordinate system becomes 4D. Points clouds are created by scanning objects or their structure with 3D sensors such as a LiDAR sensor and a RGBD Camera. There are usually two problems to be solved in deep learning for 3D point cloud: classification and segmentation.

The first approach that processes unstructured point cloud data using convolutional neural networks is called PointNet [3]. The architecture of the PointNet is shown in Fig 1. The classification network consists of two transformation networks, each of which has a shared multi-layer perceptron (MLP) with layer output sizes (64, 64) and (64, 128, 1024) and a max pooling.

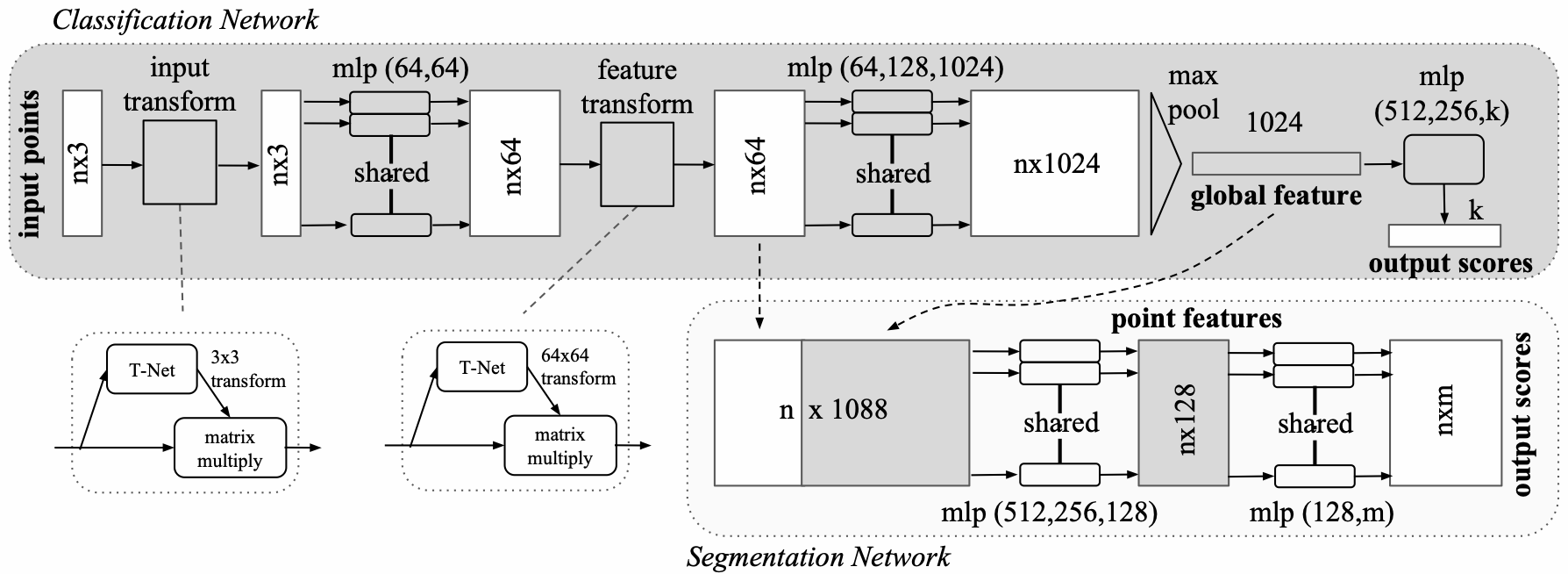
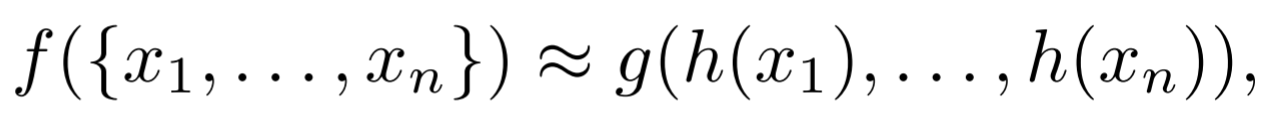


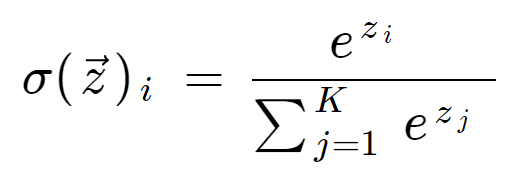
Figure 1. PointNet architecture

This method takes scattered and unordered point cloud data consisting of N points and processes it separately by N! permutations provided that the model is invariant to these changes. A single point is not considered in isolation. On the contrary, the point interacts with its nearest neighbours and they often carry information useful for classification. Therefore PointNet uses two symmetric functions (Eq. 1) that make the model robust to transformations, the output of these functions still the same as an input:

(1)

where g(x) is a max pooling layer and h(x) is a multilayer perceptron.

The MLPs are feature transformations that map independently each of the N points from one dimension to another. The last one transformation network also has two fully connected layers with output sizes (512, 256, k), where *k* is a number of classes. To predict the class of a given point cloud the model in its final layer has a softmax activation function, which is often used for multi-class classification problems. The softmax function, shown in equation 2, normalizes outputs and the sum of obtained probabilities equals 1.

(2)

It is important to minimize the model’s error rate, so an Adam optimizer is used, which combines the advantages of RMSProp (Root Mean Square Propagation) and AdaGrad (Adaptive Gradient Algorithm).

**Dataset collection issues**

To identify trees, a survey of the area in the forest was carried out using the LiDar GeoSLAM ZEB-HORIZON apparatus. With a measuring range of up to 100 m, the ZEB-HORIZON is perfect for surveying open spaces. It scans 300,000 points per second with an accuracy of 1-3 cm. For preprocessing the LiDAR data, the GeoSLAM Hub + Draw software was used, which converts the original 3D scanning data into the pcd format for further processing [12].

The shooting was carried out in the central region of European Russia. This survey resulted in point cloud images of trees. Individual trees were manually selected from the full terrain images and placed in separate files using 3D Forest software. 3D forest is an open source software for lidar data segmentation, visualization, measurement and export of various tree parameters [5]. After all the transformations, 261 tree images were obtained.

After collecting the files, it was necessary to mark up the data, to determine the type of each tree. To do this, the resulting dataset was manually viewed and each tree was assigned its own view. In total, 10 species have been identified: aspen, willow, chestnut, poplar, spruce, birch, linden, maple, pine, oak. For all species, the corresponding species names in Latin were selected: Populus tremula, Salix, Castánea, Populus, Picea, Betula, Tilia, Acer, Pinus, Quercus. During the marking process, for each type of tree its unique criteria were identified. Each species was determined based on the external characteristics of trees, such as: the volume of the trunk, the crown of the tree, the direction of growth of branches, etc.

The data were collected in the winter season, which made it possible to shoot trees without leaves and make the crown more informative, since the leaves reduce the accuracy [4]. However, marking is a rather time-consuming process, since there was often not enough expert knowledge to unambiguously determine the type of wood [4]. In the form of point clouds, it is rather difficult to determine which species a particular tree belongs to.

For a more complete picture, it is necessary to have an image of the captured tree itself. Therefore, it was decided to carry out the marking several times separately from each other and then form the final sample from those trees where the species is uniquely identified. Since it was not possible to accurately classify manually all the images of trees, only 253 trees were included in the final sample. Figure 2 shows examples of trees from left to right: birch, spruce, poplar.



Figure 2. Examples of trees

In order to display each tree for further markup, the graph\_objects, mplot3d, open3d libraries were used. The result is a point cloud that can be viewed from any side, which has improved and simplified the process of manually classifying trees.

**Experiment and results**

Two classes (Castánea and Salix) were excluded from the dataset, which contained an insufficient number of copies for training. The number of instances for the rest of the classes is presented in Table 1. The original data was converted to mesh format [11], which allows displaying a point cloud as a polygonal model. Further, the data was converted to 2048 point format for transfer to the model. The dataset was divided into training and test samples in a ratio of 80% to 20%. To improve the quality of training, examples were shuffled and then jittered by normal distribution.

Table 1. Number of instances in the dataset of each class

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Tree | Aspen | Poplar | Spruce | Birch | Linden | Maple | Pine | Oak |
| Species | Populus tremula | Populus  alba | Picea  abies | Betula  pendula | Tilia  cordata | Acer  platanoides | Pinus  sylvestris | Quercus  robur |
| Quantity | 17 | 33 | 48 | 35 | 8 | 26 | 66 | 20 |

The implementation from the keras site [13] was taken as a basis, which showed an accuracy of 73% on tbe ModelNet10 dataset. Training done in Google Colab using GPU. The optimizer has been changed to SDG with a learning rate of 0.001 instead of Adam. During the training process, 50 epochs were performed and the classification accuracy was achieved on the training data set of 35% and on the test data set - 32% after the 5th epoch (Figure 3). For 50 epochs, it was not possible to obtain a higher accuracy for the test. The accuracy on the test and training samples has similar values, which means that the model is not overfitted.

|  |  |
| --- | --- |
|  |  |
| Figure 3. Confusion matrix and model accuracy by epochs | |

The accuracy turned out to be well below 77.5% [4] and 86% [6]. This is due to the fact that the number of copies in the dataset is less than in these works. In work [4] there were also 10 classes, but the number of copies in each was 5000, that is, two orders of magnitude more than in our study. In a study [6], the number of examples in each class ranged from 300 to 400. This is one order of magnitude more, but the authors studied only 4 classes. The problem of a small number of copies was considered in [7], the quality of the dataset markup also affects classification. The segmentation quality of the 3D Forest software also affects the classification results.

In the resulting error matrix (Figure 3), it can be seen that the resulting model best recognizes the Betula class (birch) from the test dataset. The confusion matrix shows that most of the errors are evenly distributed in the matrix between the classes. But there is a strong correlation between the classes Betula and Picea - the model erroneously recognizes spruce as birch. However, it is difficult to draw a final conclusion due to the imbalance of classes: the Pinus (pine) class has the maximum amount of data is 66, while Tilia (linden) has only 6 examples.

**Conclusion**

The results obtained during investigation showed good capabilities of the PointNet model for classifying tree species even on a small amount of data. But at the same time, problems were identified with manual labeling of the dataset with point clouds of trees. The amount of data collected for training and testing was also insufficient. In future studies, it is planned to significantly increase the number of examples for each class. To improve the quality of the markup, it is planned to use data from several sources, namely cameras and LiDAR.

**Acknowledgement**

The authors thank Bauman Moscow State Technical University (BMSTU) for the LiDAR provided and Sergey Ustinov, Evgeny, and Sergey Chumachenko from BMSTU for help in data collection.

**References**

1. https://www.nature.org/en-us/get-involved/how-to-help/carbon-footprint-calculator/
2. https://unfccc.int/process-and-meetings/the-paris-agreement/the-paris-agreement
3. https://arxiv.org/pdf/1612.00593.pdf
4. Classification of tree species based on structural features derived from high density LiDAR data / Jili Li, Baoxin Hu, Thomas L. Noland / Agricultural and Forest Meteorology 171–172 (2013) 104–114
5. 3D Forest: An application for descriptions of three-dimensional forest structures using terrestrial LiDAR Jan Trochta1\*, Martin Krůček1,2, Toma ́ sˇ Vrsˇ ka1, Kamil Kra ́ l [https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0176871]
6. Deep learning-based tree classification using mobile LiDAR data / Haiyan Guan, Zheng Ji, Yongtao Yu, Jonathan Li, / Remote Sensing Letters · September 2015 / 864-873
7. Fabian Ewald Fassnacht Review of studies on tree species classification from remotely sensed data / Fabian Ewald Fassnacht, Hooman Latifi, Krzysztof Stereńczak, Aneta Modzelewska, Michael Lefsky, Lars T. Waser, Christoph Straub, Aniruddha Ghosh // Remote Sensing of Environment. – [United States](https://www.scimagojr.com/journalrank.php?country=US) , 2016. – С. 65-87.
8. Парижское соглашения от 2015 г. //Организация Объединённых наций. – 2015.
9. Eroshenkova D. A. et al. Automated determination of forest-vegetation characteristics with the use of a neural network of deep learning //International Conference on Neuroinformatics. – Springer, Cham, 2019. – С. 295-302.
10. Chumachenko S. I. et al. Simulation modelling of long-term stand dynamics at different scenarios of forest management for coniferous–broad-leaved forests //Ecological Modelling. – 2003. – Т. 170. – №. 2-3. – С. 345-361.
11. открытое ПО https://www.meshlab.net/
12. характеристики лидара https://geoslam.com/solutions/zeb-horizon/
13. https://keras.io/examples/vision/pointnet/
14. Hudak, A.T.; Evans, J.S.; Smith, A.M.S. LiDAR utility for natural resource managers. Remote Sensing 2009, 1, 934–951.
15. Multi Voxel-Point Neurons Convolution (MVPConv) for Fast and Accurate 3D Deep Learning Authors: [Wei Zhou](https://arxiv.org/search/?searchtype=author&query=Zhou%2C+W), [Xin Cao](https://arxiv.org/search/?searchtype=author&query=Cao%2C+X), [Xiaodan Zhang](https://arxiv.org/search/?searchtype=author&query=Zhang%2C+X), [Xingxing Hao](https://arxiv.org/search/?searchtype=author&query=Hao%2C+X), [Dekui Wang](https://arxiv.org/search/?searchtype=author&query=Wang%2C+D), [Ying He](https://arxiv.org/search/?searchtype=author&query=He%2C+Y) https://arxiv.org/pdf/2104.14834.pdf
16. https://arxiv.org/pdf/2105.04169.pdf / PillarSegNet: Pillar-based Semantic Grid Map Estimation using Sparse LiDAR Data / Juncong Fei1, Kunyu Peng, Philipp Heidenreich1, Frank Bieder, and Christoph Stiller
17. Прикольная статья, но не знаю куда впихнуть https://arxiv.org/pdf/2006.05560.pdf / Tree Annotations in LiDAR Data Using Point Densities and Convolutional Neural Networks / Ananya Gupta, Jonathan Byrne, David Moloney, Simon Watson, Hujun Yin
18. PointNet++: https://arxiv.org/abs/1706.02413v1 / PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space / [Charles R. Qi](https://arxiv.org/search/cs?searchtype=author&query=Qi%2C+C+R), [Li Yi](https://arxiv.org/search/cs?searchtype=author&query=Yi%2C+L), [Hao Su](https://arxiv.org/search/cs?searchtype=author&query=Su%2C+H), [Leonidas J. Guibas](https://arxiv.org/search/cs?searchtype=author&query=Guibas%2C+L+J)
19. Relation-Shape Convolutional Neural Network for Point Cloud Analysis: https://arxiv.org/abs/1904.07601v3 / Yongcheng Liu, Bin Fan, Shiming Xiang, Chunhong Pan
20. Point Transformer: https://arxiv.org/abs/2012.09164v1 / Hengshuang Zhao, Li Jiang, Jiaya Jia, Philip Torr, Vladlen Koltun
21. Zhang J. et al. A Review of Deep Learning-based Semantic Segmentation for Point Cloud (November 2019) //IEEE Access. — 2019.

# SnapNet: 3D point cloud semantic labeling with 2D deep segmentation networks [AlexandreBoulchaJorisGuerryaBertrandLe SauxaNicolasAudeberta](https://www.sciencedirect.com/science/article/abs/pii/S0097849317301942#!) https://www.sciencedirect.com/science/article/abs/pii/S0097849317301942

1. SEGCloud: Semantic Segmentation of 3D Point Clouds[Lyne P. Tchapmi](https://arxiv.org/search/cs?searchtype=author&query=Tchapmi%2C+L+P), [Christopher B. Choy](https://arxiv.org/search/cs?searchtype=author&query=Choy%2C+C+B), [Iro Armeni](https://arxiv.org/search/cs?searchtype=author&query=Armeni%2C+I), [JunYoung Gwak](https://arxiv.org/search/cs?searchtype=author&query=Gwak%2C+J), [Silvio Savarese](https://arxiv.org/search/cs?searchtype=author&query=Savarese%2C+S) https://arxiv.org/abs/1710.07563