DOPAD CHYBOVOSTI TRÉNOVACÍCH DAT NA VÝSTUP STROJOVÉHO UČENÍ PRO KLASIFIKACI MRAČEN BODŮ ELEKTRICKÉHO VEDENÍ

Daniel BERAN[1], Nan LI [2], Norbert PFEIFER [2]

[1] Západočeská univerzita, Fakulta aplikovaných věd, Katedra geomatiky, Univerzitní 22, Plzeň, Česká republika
dberan[at]kgm.zcu.cz

[2] TU Wien, Department of Geodesy and Geoinformation, Research Division Photogrammetry, Wiedner Hauptstraße 8, Vienna, Austria
photo-sek[at]geo.tuwien.ac.at

doi: https://doi.org/10.31490/9788024843988-21

Abstrakt

Jednou z mnoha aplikací ULS (UAV-borne laser scanning) je inspekce elektrického vedení. Nicméně s výstupem LiDAR sběru dat (mračen bodů) přichází i potřeba data automaticky klasifikovat, neboli sémanticky segmentovat, za účelem navazující analýzy. Metod pro automatickou klasifikaci mračen bodů bylo představeno nemalé množství, mnoho z nich s využitím strojového učení. Motivací tohoto výzkumu je nutná podmínka strojového učení v podobě referenčních (trénovačích) dat pro učení modelu - konkrétně dopad chybovosti v klasifikaci referenčních dat na přesnost modelované klasifikace výstupů modelu. K zjištění dopadu chybovosti trénovacích dat na výstup strojového učení pro klasifikaci mračen bodů elektrického vedení jsme použili metodu klasifikačních a regresních stromů (CART) implementovanou v programu Opals. V rámci výzkumu byly testovány datové sady s různou mírou a různým typem chybovosti referenčních dat a jejich vliv na výslednou přesnost byl porovnán s daty, které nebyly použity pro samotné učení modelu.

Abstract

One of the applications of ULS (UAV-borne laser scanning) lies in power line inspection. However, with LiDAR data (i.e. point clouds) comes the need for reliable automatic classification, also called semantic segmentation, of data which allows further analysis of gathered data. Vast number of possible methods for automatic classification of point clouds have been proposed and implemented, many of which depend on machine learning. Motivation for this research is the need for pre-classified data for training of machine learning models, specifically the impact of label accuracy/error in the pre-classified data used for machine learning classification. To find out what is the impact of error levels of labels on machine learning classification of power line point clouds we have used the method of Classification and regression trees (CART) using Opals software. During this research several tests were conducted with various levels and types of error in class labelling of training data and the results were compared with correctly labelled data to calculate confusion matrices and thus evaluate the impact of different error levels.

Klíčová slova: mračno bodů; LiDAR; ULS; strojové učení; elektrické vedení

Keywords: point cloud; LiDAR; ULS; machine learning; power line

EXTENDED ABSTRACT

For the past decades the utilization of LiDAR technology in various disciplines increased. More recently the decreasing cost of hardware allowed LiDAR technology, previously exclusive for airplane surveying, to expand into the realm of UAVs and thus creating the ULS (UAV-borne laser scanning). One of the possible applications of ULS lies in power line inspection. However, with LiDAR data (i.e. point clouds) comes the need for reliable automatic classification, also called semantic segmentation, of data which allows further visual and mathematical analysis of gathered data.

Vast number of possible methods for automatic classification of point clouds have been proposed and implemented, many of which depend on machine learning [1] [2]. The advantage of machine learning is clearly the ability to find correlations between values of point features (e.g. orientation of normal vector, planarity, omnivarence etc.) and the classes required for the output, without the need for a user to define these correlations manually. One of the disadvantages is the need for precise training data, i.e. manually labelled points. All machine learning workflows are therefore dependent on the quality of manual classification of data samples used for model training. A problem of this approach lies in the fact that error free manually labelled data are hard to acquire and usually require multiple overviews since it is easy for a human eye to miss mislabelled parts of the point cloud. This struggle is common for all machine learning applications even those with more developed research history such as image classification [3] [4]. The focus of this work is to determine the relation between level of error in training data and the quality of the classification.

Motivation for this research is the aforementioned need for pre-classified data for training of machine learning models, specifically the impact of label accuracy/error in the pre-classified data used for machine learning classification. Data used in this research consisted of ULS point clouds of a power line network. Classes defined in this dataset are: environment, conductor, pylon, and insulator. These classes are present in every tile of the dataset, however with vastly heterogeneous distribution: environment (40%), conductor (40%), pylon (18%), insulator (2%).

Testing dataset composed of 0.5 million points was trained and evaluated using Opals software [5]. It utilises CART (Classification and regression trees) [6] method on the basis of the software package R. First a test with correctly labelled data was conducted. Randomly selected 50% of the points from the training dataset were used for training the model. The model was evaluated with the other half of points in the training dataset using a normalized confusion matrix (see Table 1).

ref\estim.	1	14	15	16	Sum_ref	EoO	Completeness
1	39.6	0.0	0.1	0.0	39.7	0.3	99.7
14	0.0	40.5	0.0	0.0	40.5	0.1	99.9
15	0.2	0.0	18.2	0.1	18.5	1.7	98.3
16	0.0	0.0	0.1	1.2	1.2	7.2	92.8
Sum_estim	39.8	40.5	18.4	1.2	100.0	---	---
EoC	0.6	0.1	1.2	6.5	---	---	---
Correctness	99.4	99.9	98.8	93.5	---	---	---

Table. 1. Self-evaluation of model from correctly labelled data

This model was further evaluated against another correctly labelled point cloud that was not used in the training of the model, results of this evaluation can be seen in Table 2 and Figure 1. The overall accuracy of the estimated classification was 97.4% with completeness for each class as follows: environment (99.75%), conductor (99.36%), pylon (91.88%), insulator (80.14%). This error distribution among classes corresponds with pylon and insulator being the less represented classes in the dataset.

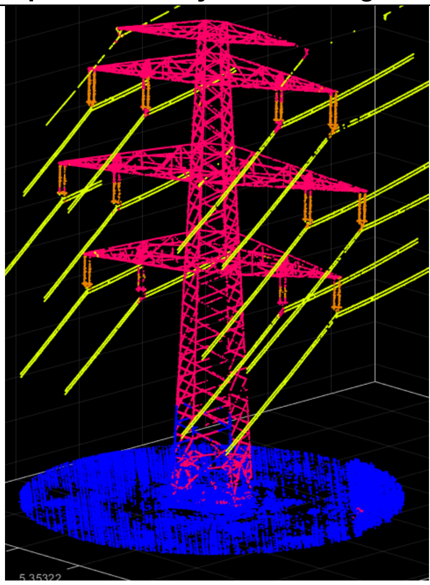

Figure. 1. Estimated classification using model from correctly labelled data

	ref\estim	environment	conductor	pylon	insulator
100% correct label	environment	**99,75%**	0,00%	0,25%	0,00%
	conductor	0,10%	**99,36%**	0,34%	0,20%
	pylon	4,81%	0,19%	**91,88%**	3,11%
	insulator	2,01%	1,49%	16,35%	**80,14%**

Table. 2. Evaluation of model from correctly labelled data

Further three more models were trained using the same method with 5%, 10% and 20% of random label error in the same training dataset. These three models were evaluated against the same dataset as in the previous case and the results can be seen in Tables 4,5 and 6.

	ref\estim	environment	conductor	pylon	insulator
5% random error	environment	**99,39%**	0,00%	0,61%	0,00%
	conductor	0,07%	**99,45%**	0,30%	0,18%
	pylon	5,01%	0,25%	**94,35%**	0,39%
	insulator	1,82%	2,12%	17,33%	**78,73%**

Table. 3. Evaluation of model from data with 95% label accuracy

	ref\estim	environment	conductor	pylon	insulator
10% random error	environment	**87,92%**	0,00%	12,08%	0,00%
	conductor	0,10%	**99,48%**	0,30%	0,12%
	pylon	4,96%	0,32%	**94,36%**	0,36%
	insulator	2,08%	2,51%	16,89%	**78,51%**

Table. 4. Evaluation of model from data with 90% label accuracy

	ref\estim	environment	conductor	pylon	insulator
20% random error	environment	**99,17%**	0,00%	0,83%	0,00%
	conductor	0,11%	**99,40%**	0,36%	0,13%
	pylon	4,83%	0,24%	**93,54%**	1,39%
	insulator	3,08%	1,99%	31,36%	**63,57%**

Table. 5. Evaluation of model from data with 80% label accuracy

The results indicate that up to 20% of random errors in labelling are not crucial for estimating the classification tree model as well as the overall accuracy. However, it is obvious that the confusion increases as can be seen in the off diagonal elements. As it was to be expected, the least represented class (insulators) suffers the most from this random error with the completeness dropping to 64%. It shall also be noted that the label errors are random distributed spatially and class-wise. Randomly distributed error is the reason for such small decrease in overall accuracy, because the average split value of used feature in the tree stays more or less the same. This model gives an indication but is not necessarily representative for label errors occurring due to manual labelling.

This type of human-like type errors was also tested. See Figure 2 and Table 6 for results. We can see that this type of error decreases the accuracy of represented class significantly (decrease by 10%). This is interesting when compared with result in Table 4. where also 10% of class insulators was wrongly classified (same as the rest of the dataset in Table 4) but accuracy of the class decreased only by 1.5%.

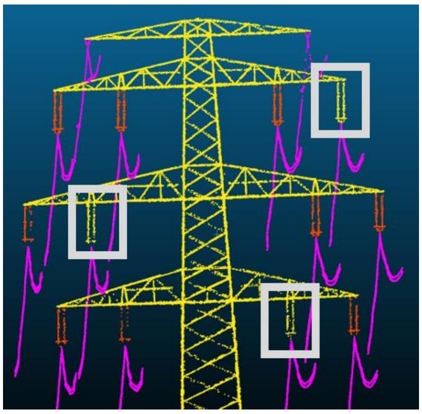

Figure. 2. Human-like type error (i.e. insulators labelled as pylon)

	ref\estim	environment	conductor	pylon	insulator
10% of insulators as pylon (0.2% of all points)	environment	**98,51%**	**0,01%**	**1,48%**	**0,00%**
	conductor	**0,10%**	**99,43%**	0,31%	0,16%
	pylon	**5,31%**	0,17%	**94,41%**	0,11%
	insulator	**1,99%**	2,51%	25,21%	**70,28%**

Table. 6. Evaluation of model from data with human-like type error (i.e. insulators labelled as pylon)

It is fair to state that these are preliminary results and further testing is planned. The aim of these tests is to help us design further more in depth testing. Firstly, a set of different tiles need to be tested as 0.5 million points is not a large testing datasets. We would also like to focus more on different types of human-like errors that occur during the process of manual labelling. At last we need to test more methods model training to test if generalization of our findings for different methods of machine learning is possible.

References:

[1] Yan, Wai Yeung, Ahmed Shaker, and Nagwa El-Ashmawy. "Urban land cover classification using airborne LiDAR data: A review." Remote Sensing of Environment 158 (2015): 295-310.

[2] Antonarakis, A. S., Keith S. Richards, and James Brasington. "Object-based land cover classification using airborne LiDAR." Remote Sensing of environment 112.6 (2008): 2988-2998.

[3] Li, Junnan, et al. "Learning to learn from noisy labeled data." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2019.

[4] Frénay, Benoît, and Michel Verleysen. "Classification in the presence of label noise: a survey." IEEE transactions on neural networks and learning systems 25.5 (2013): 845-869.

[5] Orientation and Processing of Airborne Laser Scanning, TU Wien, Department of Geodesy and Geoinformation, Research Division Photogrammetry and Remote Sensing, https://opals.geo.tuwien.ac.at/

[6] Breiman, L., Friedman, J., Stone, C. J., & Olshen, R. A. (1984). Classification and regression trees. CRC press.