

# ▣ Semantična segmentacija aerolaserskih oblakov točk in centriranje višin globalnih soseščin

Jernej Nejc Dougan<sup>1,3</sup>, Krištof Oštir<sup>2</sup>, Matej Kristan<sup>1</sup>

<sup>1</sup>Univerza v Ljubljani, Fakulteta za računalništvo in informatiko, Večna pot 113, Ljubljana,

<sup>2</sup>Univerza v Ljubljani, Fakulteta za gradbeništvo in geodezijo, Jamova cesta 2, Ljubljana,

<sup>3</sup>Flycom Technologies d.o.o., Ljubljanska cesta 24A, Kranj

nejc.dougan@flycom.si, kristof.ostir@fgg.uni-lj.si, matej.kristan@fri.uni-lj.si

## Izvelek

Aerolaserski oblaki točk so pomemben vir informacij v številnih prostorskih aplikacijah, kot na primer pri izdelavi digitalnih modelov terena ali kartiranju in popisu sredstev kritične infrastrukture. Semantična segmentacija se lahko uporablja v večini procesnih tokov obdelave aerolaserskih oblakov točk. V zadnjih letih najboljše rezultate za semantično segmentacijo in klasifikacijo dosegajo metode globokega učenja. Na kakovost segmentacije med drugim vpliva izbor soseščine točk in centriranje višine. V članku predstavimo in evalviramo različne metode za centriranje višin. Preizkuse smo izvedli na podatkovni zbirki ISPRS 3D Semantic Labelling, kjer smo s preprosto metodo centriranja najmanjše višine izboljšali rezultat za skoraj dva procenta.

**Ključne besede:** aerolasersko snemanje, globoko učenje, oblaki točk, semantična segmentacija

## Abstract

Aerial laser scanning point clouds are an important data source in many geospatial applications such as digital terrain model generation or asset mapping of critical infrastructure. Semantic segmentation can be used in the majority of point cloud processing pipelines. Current state-of-the-art methods for semantic segmentation and classification are based on deep learning. The quality of semantic segmentation depends also on the neighbourhood selection and elevation centering. In this paper, we propose and evaluate different methods for elevation centering. Experiments on ISPRS 3D Semantic Labelling show that the use of minimal elevation centering increases results by nearly two percent.

**Keywords:** Aerial laser scanning, deep learning, point clouds, semantic segmentation.

## 1 UVOD

Klasifikacija in semantična segmentacija oblakov točk aerolaserskega snemanja (ALS) sta pomembna problema, ki zahtevata znanje daljinskega zaznavanja, fotogrametrije in računalniškega vida. Številne prostorske aplikacije, na primer izdelava digitalnih modelov reliefa, zaznavanje stavb, rekonstrukcija stavb, kartiranje in popis sredstev kritične infrastrukture, temeljijo na obdelanih oblakih točk. Obdelava zajema razdelitev točk v različne razrede, na primer za izdelavo digitalnega modela terena je potrebno točke razdeliti v točke terena in ostale. Velika večina trenutno obstoječih postopkov temelji na me-

todah, ki ne temeljijo na stojnem učenju, na primer matematična morfologija [Mongus et al., 2014]. Velik uspeh metod strojnega učenja in predvsem globokega učenja v slikovni domeni [Krizhevsky et al., 2012] je spodbudil raziskave na področju uporabe globokega učenja za oblake točk. Tradicionalni postopki, temelječ na strojnem učenju, za semantično segmentacijo oblakov točk ALS temeljijo na ročno ustvarjenih značilnicah in klasifikatorjih. Pred kratkim so se za semantično segmentacijo in klasifikacijo oblakov točk začele uporabljati globoke nevronske mreže [Qi et al., 2017a, Qi et al., 2017b, Thomas et al., 2019] in trenutno dosegajo tudi najboljše rezultate. Vendar trenutno najbolj-

še metode ne naslavljajo vseh karakteristik oblakov točk ALS.

Oblaki točk ALS so obsežni, lahko obsegajo celotne države, za obdelavo jih je potrebno razdeliti v manjše soseščine. Izboru soseščine moramo posvetiti posebno pozornost, saj se velikosti objektov lahko razlikujejo za celotne velikostne razrede, višina nad terenom pa predstavlja eno izmed pomembnejših značilnic za uspešno segmentacijo [Niemeyer et al., 2014]. Določanje višine terena zahteva predhodno določitev točk terena. Posledično bi točke terena morali obravnavati ločeno. Z uporabo centriranja višin globalnih soseščin lahko dobimo dober približek višine nad terenom in se izognemo kompleksni arhitekturi za ločeno obravnavanje točk terena.

V tem članku predlagamo tri preproste metode centriranja višine globalnih soseščin in uporabo centriranih višin kot vhodnih značilnic mrežo.

## 2 PREGLED METOD GLOBOKEGA UČENJE ZA OBLAKE TOČK

Klasične metode nadzorovanega strojnega učenja za oblake točk ALS izkoriščajo ročno oblikovane značilnice. Pogosto uporabljene značilnice temeljijo na lastnih vrednostih in so linearnost, planarnost, razpršenost, omnivarianca, anizotropija, vsota lastnih vrednosti, sprememba ukrivljenosti [Weinmann et al., 2015], le-te opisujejo, kako se točke porazdeljujejo v okolici točke ocenjevanja. Zaradi svoje sposobnosti vključevanja kontekstualne informacije je eden izmed bolj uporabljenih pristopov pristop pogojno slučajnih polj (angl. conditional random fields, CRF) [Weinmann et al., 2015, Vosselman et al., 2017, Niemeyer et al., 2014].

Uspeh metod globokega učenja [LeCun et al., 2015] v preteklih letih je navdihnil nove raziskave klasifikacije in semantične segmentacije 3D oblakov točk. Ena izmed glavnih prednosti metod globokega učenja je njihova zmožnost učenja značilnic in posledično odpravljena potreba po ročnem oblikovanju letih. Metode globokega učenja za oblake točk lahko razdelimo v dve kategoriji: (i) projekcijske in (ii) direktne metode. Projekcijske metode projicirajo točke v regularne 2D ali 3D mreže. Ob postopku projekcije se del informacije izgubi, pojavljajo se neželeni artefakti diskretizacije prostora. Direktne metode delujejo neposredno na oblakih točk in tako niso izpostavljene omenjenim pomanjkljivostim projekcijskih metod. Ključni izzivi direktnih metod so odkrivanje

značilnic, definicija konvolucije in izbira soseščine. Trenutno najboljše so direktne metode, ki jih razdelimo v metode temelječe na: (i) točkovnih več-nivojski-perceptronih [Qi et al., 2017a, Qi et al., 2017b, Zhang et al., 2019], (ii) grafih [Wang et al., 2019, Landrieu and Simonovsky, 2018, Liu et al., 2019] in (iii) točkovnih konvolucijah [Thomas et al., 2019, Wang et al., 2018, Li et al., 2018, Wu et al., 2019].

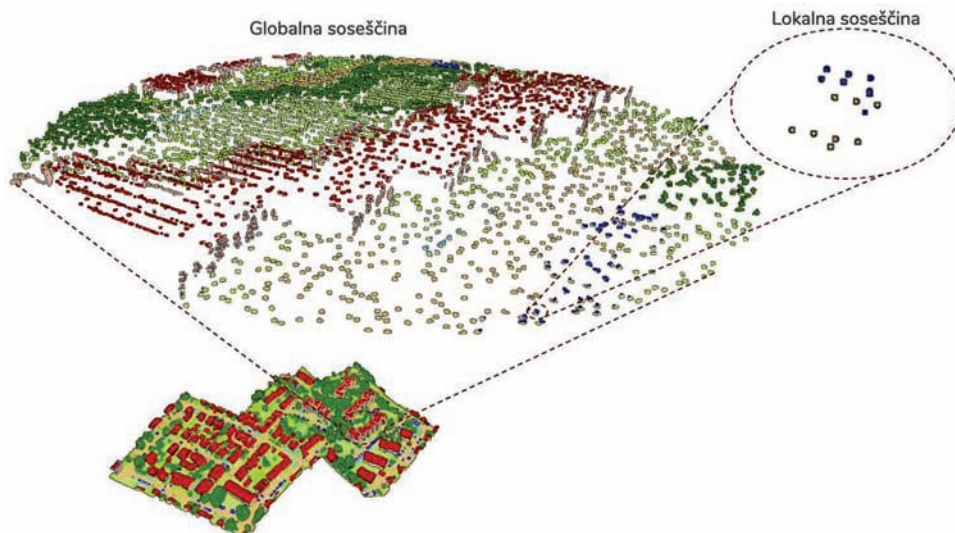
PointNet [Qi et al., 2017a] je bila prva globoka nevronska mreža za oblake točk, delujoča neposredno na točkah, ki je za odkrivanje značilnic uporabljala več točkovnih več-nivojskih-perceptronov (VNP). Trenutno najboljše rezultate dosegajo metode, ki temeljijo na točkovnih konvolucijah. Wang et al. [Wang et al., 2018] so na primer predlagali parametrično zvezno konvolucijo. Zvezna konvolucija za točke je definirana kot Monte-Carlo integracija parametrične funkcije, ki jo aproksimira VNP. Thomas et al. [Thomas et al., 2019] so predlagali novo konvolucijsko jedro definirano z jedrnimi točkami - Kernel Point Convolution (KPConv) in trenutno dosega najboljše rezultate na standardnih testih za semantično segmentacijo oblakov točk.

## 3 VIŠINA IN SOSEŠČINA

Oblaki točk ALS so praviloma preveliki, da bi jih lahko obdelovali naenkrat. Treba jih je razdeliti na manjša območja - soseščine. Izbor globalne in lokalne soseščine igra pomembno vlogo pri uspešnosti mreže, saj soseščina definira območni kontekst. Globalna soseščina je podmnožica točk iz celotnega oblaka točk, lokalna soseščina pa podmnožica točk za izračun enega koraka konvolucije (Slika 1). Izbira velikosti, centriranje višine in metode vzorčenja posamezne soseščine vplivajo na kakovost semantične segmentacije. Trenutno najboljša metoda KPConv [Thomas et al., 2019] uporablja fiksno sferično poizvedbo za globalno in lokalno soseščino. Višinsko so točke centrirane okoli točke poizvedbe. Velikost, oblika in vzorčenje soseščine so fiksni.

V izogib prevelike kompleksnosti mreže za določitev višine nad terenom predlagamo metode centriranja višin v globalnih soseščinah. Centrirane višine aproksimirajo višine nad terenom.

Globalna soseščina je podmnožica točk znotraj sfere s polmerom  $r$  in središčem v *središčni točki*. *Središčno točko* naključno izberemo iz množice vseh točk. Vse točke globalne soseščine centriramo po vseh treh prostorskih dimenzijah okoli *središčne toč-*



Slika 1: Izbor globalne in lokalne soseščina vpliva na kakovost semantične segmentacije.

ke. Višina je tako odvisna od izbrane središčne točke in je neuporabna kot značilnica. Zato predlagamo tri alternativne določitve globalne soseščine z uporabo centriranja višin: (i) centriranje s srednjo vrednostjo višine globalne soseščine, (ii) centriranje z najmanjšo vrednostjo višine globalne soseščine in (iii) centriranje z  $n$ -tim percentilom višin globalne soseščine. V ekstremnem primeru, kjer je teren popolnoma raven in velja  $z = 0$  za vse točke, sta pri uporabi centriranja z najmanjšo vrednostjo centrirana višina in višina nad terenom enaki. Osnovno metodo smo dodatno razširili tako, da kot vhodno značilnico sprejme tudi centrirano višino. Lokalno soseščino smo definirali kot  $k$ -najbližjih sosedov.

## 4 EVALVACIJA

### 4.1 Eksperimenti

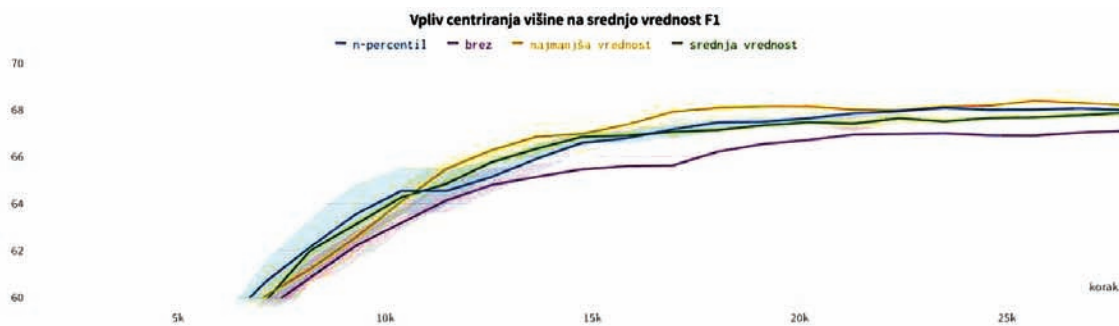
Za osnovno mrežo smo uporabili 5-nivojsko KP-Conv mrežo [Thomas et al., 2019] z deformabilnimi 15-točkovnimi jedri na zadnjih treh nivojih. Polmer globalne soseščine je 24 metrov, lokalno soseščino predstavlja 20 najbližjih sosedov. Parametre smo določili na podlagi preliminarne eksperimentov. Osnovni mreži smo modificirali načine centriranja višine globalnih soseščin. Celoten algoritem je razdeljen na dva dela: (i) branje in priprava podatkov ter (ii) učenje. Branje in pripravo podatkov izvaja centralna procesna enota (CPU), učenje poteka na grafično procesni enoti (GPU). Zaradi velike neuravnoteženosti zastopanosti razredov smo za kriterijsko

funkcijo uporabili uteženo križno entropijo (angl. weighted crossentropy) [Qi et al., 2017b]. Mrežo smo učili 500 epoh, kjer eno epoho sestavlja 50 korakov. Učenje smo izvedli na računalniku s procesorjem Intel Core i5-8400 in grafični kartici nVidia GTX 1080 Ti 11GB. Algoritem je implementiran v Python-u 3.6 z uporabo knjižnice Tensorflow 1.15.0. Povprečni čas učenja mreže in validacije je 3 ure, kjer en korak traja povprečno 350 milisekund. Znotraj enega koraka se obdela približno 72.200 točk.

Mrežo smo učili in testirali na podatkovni zbirki ISPRS 3D Semantic Labelling [Niemeyer et al., 2014]. Podatki so bili zajeti z instrumentom Leica ALS50, z višine 500 metrov nad terenom in vidnim poljem  $45^\circ$  [Cramer, 2010]. Podatki so označeni v devet semantičnih kategorij in razdeljeni v učno množico s 753.876 točkami in testno množico s 411.722 točkami. Pri izvedbi eksperimentov smo ohranili obstoječo razdelitev v učno in testno množico.

### 4.2 Rezultati

Za oceno rezultatov smo uporabili standardno proceduro na podatkih ISPRS 3D Semantic Labelling. Za vsako kategorijo posebej določimo oceno  $F1$  (Enaša 1), kjer  $TP$  predstavlja pravilno pozitivne,  $FP$  nepravilno pozitivne in  $FN$  nepravilno negativne segmentirane točke. Skupna ocena je srednja vrednost ocen  $F1$  vseh kategorij.



Slika 2: Linije predstavljajo glajeno srednjo vrednost metrike mF1. Območje v ozadju predstavlja razpon med najmanjšo in največjo vrednostjo posamezne skupine.

$$F1 = 2 * \frac{\frac{TP}{TP + FP} * \frac{TP}{TP + FP}}{\frac{TP}{TP + FP} + \frac{TP}{TP + FN}} \quad (1)$$

Vse metode za centriranje višine globalne soseščine izboljšajo rezultate semantične segmentacije oblakov točk (Slika 2). Najboljše rezultate smo dosegli z metodo centriranja z najmanjšo vrednostjo višine in uporabo višine kot značilnice. Podrobni rezultati so prikazani v Tabeli 1. ISPRS oblak točk je pretežno ravninski, posledično je centriranje z minimalno vrednostjo zelo dober približek dejanske višine nad terenom.

V članku smo obravnavali problem izbora globalnih soseščin in centriranja višine pri semantični segmentaciji ALS oblakov točk. Zanimala sta nas vpliv višinske informacije na kakovost semantične segmentacije in zasnova učinkovite in preproste metode za centriranje višin globalne soseščine brez uporabe višine nad terenom. Ugotovili smo, da je najučinkovitejše centriranje z upoštevanjem najmanjše višine in vključitev višinske informacije kot značil-

nice v mrežo. Centriranje z upoštevanjem najmanjše višine v dani podatkovni zbirki najverjetneje tudi najboljše aproksimira dejansko višino točk nad terenom. S predlagano metodo smo dosegli oceno 69,19 srednje vrednosti F1. Preprosta sprememba je izboljšala rezultat v primerjavi z metodo brez centriranja za skoraj 2 odstotka srednje vrednosti F1.

V prihodnjih raziskavah bomo obravnavali problem izbora soseščin celostno, kjer bomo raziskali dodatne faktorje kot so oblika, velikost in vzorčenje. Dodatno bi bilo smiselno preveriti vpliv centriranja višine na višinsko bolj razgibani podatkovni zbirki. Prav tako bi bilo smotrno nasloviti problem nizkih osamelcev. Preproste metode centriranja bo verjetno treba nadomestiti z naprednejšo metodo, kot na primer z uporabo morfoloških profilov za aproksimacijo višine terena.

Semantična segmentacija oblakov točk je kompleksen problem, ki ga v industriji še vedno rešujejo pretežno ročno oziroma polsamodejno, kar je zamudno in neučinkovito. Vsakršne izboljšave metod imajo tako takojšen vpliv tako na znanost kot tudi industrijo.

Tabela 1: Srednje vrednosti ocene F1 z uporabo različnih metod centriranja in brez. Uporaba višine kot značilnice je brezpredmetna, če višina ni centrirana.

Centriranje višine	uporaba višine kot značilnice	srednje vrednosti F1
brez	ne	67,38
najmanjša vrednost	ne	67,74
srednja vrednost	ne	68,02
percentil	ne	68,32
brez	da	67,35
srednja vrednost	da	68,03
percentil	da	68,12
najmanjša vrednost	da	69,19

## LITERATURA

- [1] [Cramer, 2010] Cramer, M. (2010). The DGPF-test on digital airborne camera evaluation - Overview and test design. *Photogrammetrie, Fernerkundung, Geoinformation*, 2010(2):73–82.
- [2] [Krizhevsky et al., 2012] Krizhevsky, A., Sutskever, I., and E. Hinton, G. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Neural Information Processing Systems*, 25.
- [3] [Landrieu and Simonovsky, 2018] Landrieu, L. and Simonovsky, M. (2018). Large-Scale Point Cloud Semantic Segmentation with Superpoint Graphs. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 4558–4567.
- [4] [LeCun et al., 2015] LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep Learning. *Nature*, 521:436–444. [Li et al., 2018] Li, Y., Bu, R., Sun, M., Wu, W., Di, X., and Chen, B. (2018). PointCNN: Convolution on X-transformed points. In Bengio, S., Wallach, H., Larochelle, H., Grauman, K., Cesa-Bianchi, N., and Garnett, R., editors, *Advances in Neural Information Processing Systems*, number NeurIPS, pages 820–830. Curran Associates, Inc.
- [5] [Liu et al., 2019] Liu, J., Ni, B., Li, C., Yang, J., and Tian, Q. (2019). Dynamic Points Agglomeration for Hierarchical Point Sets Learning. *IEEE International Conference on Computer Vision (ICCV)*, pages 7546–7555.
- [6] [Mongus et al., 2014] Mongus, D., Lukač N., and Žalik, B. (2014). Ground and building extraction from LiDAR data based on differential morphological profiles and locally fitted surfaces. *ISPRS Journal of Photogrammetry and Remote Sensing*, 93:145–156.
- [7] [Niemeyer et al., 2014] Niemeyer, J., Rottensteiner, F., and Soergel, U. (2014). Contextual classification of lidar data and building object detection in urban areas. *ISPRS Journal of Photogrammetry and Remote Sensing*, 87:152–165.
- [8] [Qi et al., 2017a] Qi, C. R., Su, H., Mo, K., and Guibas, L. J. (2017a). PointNet: Deep Learning on Point Sets for 3D Classification and Segmentation. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.
- [9] [Qi et al., 2017b] Qi, C. R., Yi, L., Su, H., and Guibas, L. J. (2017b). PointNet++: Deep Hierarchical Feature Learning on Point Sets in a Metric Space. *CoRR*, abs/1706.0.
- [10] [Thomas et al., 2019] Thomas, H., Qi, C. R., Deschaud, J.-E., Marcotegui, B., Goulette, F., and Guibas, L. J. (2019). KP-Conv: Flexible and Deformable Convolution for Point Clouds. *IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 6410–6419.
- [11] [Vosselman et al., 2017] Vosselman, G., Coenen, M., and Rottensteiner, F. (2017). Contextual segment- based classification of airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 128:354–371.
- [12] [Wang et al., 2018] Wang, S., Suo, S., Pokrovsky, W.-C. M. A., and Urtasun, R. (2018). Deep parametric continuous convolutional neural networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2589–2597.
- [13] [Wang et al., 2019] Wang, Y., Sun, Y., Liu, Z., Sarma, S. E., Bronstein, M. M., and Solomon, J. M. (2019). Dynamic graph Cnn for learning on point clouds. *ACM Transactions on Graphics*, 38(5).
- [14] [Weinmann et al., 2015] Weinmann, M., Schmidt, A., Mallet, C., Hinz, S., Rottensteiner, F., and Jutzi, B. (2015). Contextual classification of point cloud data by exploiting individual 3D neighbourhoods. *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2(3W4):271–278.
- [15] [Wu et al., 2019] Wu, W., Qi, Z., and Fuxin, L. (2019). PointConv: Deep convolutional networks on 3D point clouds. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019-June:9613–9622.
- [16] [Zhang et al., 2019] Zhang, Z., Hua, B.-S., and Yeung, S.-K. (2019). ShellNet: Efficient Point Cloud Convolutional Neural Networks using Concentric Shells Statistics. *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*.

■

**Jernej Nejc Dougan** je magistriral leta 2015 na Fakulteti za gradbeništvo in geodezijo Univerze v Ljubljani. Trenutno obiskuje doktorski študij na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Zaposlen je v podjetju Flycom Technologies d.o.o., kjer se raziskovalno ukvarja z metodami globokega učenja za obdelavo oblakov točk in drugimi analizami in obdelavami prostorskih podatkov.

■

**Krištof Oštir, prof. dr.**, je doktoriral leta 2000 na Fakulteti za gradbeništvo in geodezijo Univerze v Ljubljani. Kot predavatelj je zaposlen na Fakulteti za gradbeništvo in geodezijo Univerze v Ljubljani, kjer predava več do- in podiplomskih predmetov s področja geoinformatike, opazovanja Zemlje in obdelave podatkov. Glavno področje njegovega dela je optično in radarsko daljinsko zaznavanje. Opravljal je študije površja z radarsko interferometrijo, se ukvarjal z izdelavo digitalnih modelov višin, rabo in pokrovnostjo tal, po-obdelavo in mehko klasifikacijo. Ukvarja se z razvojem tehnologije malih satelitov za opazovanje Zemlje.

■

**Matej Kristan, izr. prof. dr.**, je doktoriral leta 2008 na Fakulteti za elektrotehniko Univerze v Ljubljani. Trenutno je član Laboratorija za umetne vizualne spoznavne sisteme (LUVSS) ter izredni profesor na Fakulteti za računalništvo in informatiko Univerze v Ljubljani. Njegovo raziskovalno področje obsega računalniški vid s poudarkom na vizualnem sledenju in semantični segmentaciji ter računalniškem vidu za avtonomne mobilne robote.