

SAMODEJNA ZAZNAVA PROMETNIC Z GLOBOKIM UČENJEM NA PODATKIH CIKLIČNEGA LASERSKEGA SKENIRANJA SLOVENIJE

Alen Mangafić, Primož Kete

Geodetski inštitut Slovenije

alen.mangafic@gis.si, ORCID: <https://orcid.org/0000-0002-7718-5969>

primoz.kete@gis.si

DOI: https://doi.org/10.3986/9789610508885_14

UDK: 528.8.044.6:625.7(497.4)

IZVLEČEK

Samodejna zaznava prometnic z globokim učenjem na podatkih Cikličnega laserskega skeniranja Slovenije

Razvili smo model za samodejno zaznavanje prometnic na podlagi laserskih oblakov točk. Naš cilj je bil razviti model globokega učenja, ki bi bil sposoben zanesljivo razpoznati linijske objekte na rastrskih podlagah izpeljanih iz laserskih oblakov točk. Za preizkus smo uporabili podatke Cikličnega laserskega skeniranja Slovenije (CLSS) iz leta 2023. Za ta namen smo uporabili mini-U-Net, poenostavljeno različico U-Net arhitekture, ki omogoča hitrejše učenje in inferenco ter zmanjšanje računskih zahtev. Model temelji na konvolucijskih nevronske mrežah, ki so zmožne učinkovito prepoznavati geometrijo in značilnosti objektov, ter jih razvrščati v določene razrede. Razvili smo tudi metodo za samodejno izdelavo učne množice, ki avtomatizira pridobivanje oznak, kar pospeši proces in zmanjša produkcijske stroške. Poleg tehničnih vidikov smo preučili tudi praktično uporabnost modela v realnih scenarijih, vključno s hitrostjo delovanja, enostavnostjo implementacije in morebitnimi omejitvami pri uporabi. Naš cilj je bil razviti zmogljiv model, ki bi bil praktično uporaben v proizvodnih primerih na državni ravni.

KLJUČNE BESEDE

CLSS, lidar, oblak točk, strojno učenje, semantična segmentacija, detekcija

ABSTRACT

Automatic road detection with deep learning on Cyclic Laser Scanning of Slovenia data

We developed a model for automatic detection of road networks based on laser point clouds. Our goal was to develop a deep learning model, capable of reliably recognizing linear objects on raster layers derived from laser point clouds. For this test we used data from the Cyclic laser scanning of Slovenia from 2023. For this purpose, we used mini-U-Net, a simplified version of the U-Net architecture, which enables faster learning and inference and reduces computational requirements. The model is based on convolutional neural networks, capable of effectively recognizing the geometry and features of objects and classifying them into specific classes. We also developed a method for automatic generation of training datasets, automating the acquisition of labels, speeding up the process and reducing production costs. In addition to technical aspects, we also examined the practical usability of the model in real-world scenarios, including speed and ease of operation and potential limitations in use. Our aim was to develop a powerful model, that would be practically applicable in production cases at the national level.

KEY WORDS

CLSS, lidar, point cloud, machine learning, semantic segmentation, detection

1 Uvod

Izdelali smo model za samodejno zaznavanje prometnic na podlagi laserskih oblakov točk. Naš cilj je bil razviti model globokega učenja, ki bi bil sposoben zanesljivo razpoznati prometnice na rastrskih podlagah izpeljanih iz laserskih oblakov točk. Za namene naše naloge smo izbrali mini-U-Net, poenostavljeno različico U-Net arhitekture (Ronneberger, Fischer in Brox 2015). Ta vsebuje manj plasti in manj parametrov od izvorne različice U-Net in se pogosto uporablja v primerih, kjer je pomembna hitrost, imamo omejene računske zmogljivosti ali imamo na voljo manjše zbirke podatkov (Khanal in Estrada 2020). S poenostavljanjem U-Net v mini-U-Net zmanjšujemo določene zmogljivosti in kompleksnosti mreže, vendar še vedno ohranjamo ključne lastnosti pri nalogah segmentacije, ki slonijo na zmogljivosti kodirnika in dekodirnika in njihovimi povezavami. Mini-U-Net temelji na konvolucijskih nevronskih mrežah. Konvolucijske nevronske mreže so zmožne učinkovito prepoznati geometrijo in značilnosti objektov na slikah ter jih razvrščati v določene razrede. Pri tem izvajajo postopek semantične segmentacije oziroma vsakemu slikovnemu elementu dodelijo pripadajoči razred in omogočijo natančno identifikacijo objektov na sliki (Long, Shelhamer in Darrell 2015; Khanal in Estrada 2020).

Kot velja za konvolucijske mreže tipa FCN (angleško *Fully Convolutional Network*), se lahko U-Net in mini-U-Net učita na vhodnih podatkih različnih ločljivosti (Cresson 2020), to je iz različnih virov oziroma časov zajema. To omogoča uporabo podatkovnih množic, ki imajo trajnostno zasnovo, oziroma lahko obstoječe podatkovne množice nadgrajujemo z novimi podatki za globoko učenje. Prav tako mini-U-net omogoča semantično segmentacijo slik različnih velikosti. Učeni model lahko uporabimo na rastrih originalne velikosti, denimo ortofoto velikosti lista državne topografske karte v merilu 1 : 5000, ne da bi ga delili na podmnožice manjših rastrov. Izbrali smo pristop nadzorovanega učenja, kar pomeni, da smo model učili na podatkih, ki so vsebovali pare slik (rastrov) in vektorskega sloja prometnic. Vsaka slika je bila podprta z referenčno sliko, ki je vsebovala natančne oznake linijskih objektov. Tako je model lahko učinkovito prilagajal svoje parametre, da bi čim bolj ustrezal pričakovanim rezultatom.

Dodatno smo razvili metodo za samodejno izdelavo podatkovne množice za globoko učenje. Ocenili smo model s vidika stabilnosti, uspešnosti in kakovosti rezultatov. Poleg tehničnih vidikov, smo preučili tudi praktično uporabnost modela v realnih scenarijih, vključno s hitrostjo delovanja, enostavnostjo implementacije in morebitnimi omejitvami pri uporabi. Naš cilj je bil razviti model, ki bi bil natančen, kot sta na primer kompleksnejša modela ResNet (He s sodelavci 2015) in U-Net, ter tudi praktično uporaben v produkcijskih primerih na državnem nivoju.

2 Metode dela in podatki

U-Net, iz katerega je izpeljan mini-U-Net, se je prvotno pojavil kot arhitektura za segmentacijo slik, zlasti v biomedicinskih aplikacijah, vendar se lahko uporablja tudi za druge naloge segmentacije (Ronneberger, Fischer in Brox 2015). Mini-U-Net sestavljajo tri komponente: kodirnik, odgovoren za izvleček značilnk preko konvolucijskih plasti in vzorčenja; most, ki ohranja prostorske informacije med kodirnikom in dekodirnikom; dekodirnik, ki zvišuje značilke nazaj na izvorno ločljivost slike. Preskakovalne povezave olajšujejo pretok informacij med ustrežajočimi plastmi kodirnika in dekodirnika, kar izboljšuje natančnost segmentacije. Končna plast dekodirnika uporablja *softmax* aktivacijo (Goodfellow, Bengio in Courville 2016) za ustvarjanje maske segmentacije.

Glavna razlika med modeli zasnovanimi na arhitekturi FCN (kot je mini-U-Net) in tradicionalnimi konvolucijskimi mrežami s polno-povezanimi nivoji je ta, da vmesni rezultat FCN ni le klasifikacija vhodne slike, temveč maska značilnk. To je pomembno za proces semantične segmentacije, saj značilke izhodne maske FCN ohranjajo lokacijske informacije značilnk predhodnih, povezanih nivojev (Smole 2019).

Mini-U-Net izvaja semantično segmentacijo slik sestavljenih iz treh dimenzij, kot so npr. slike sestavljene iz rdečega, zelenega in modrega kanala (RGB). Prednosti tega, da se lahko mini-U-Net uči na učnih množicah sestavljenih iz rastrov različnih ločljivosti, so:

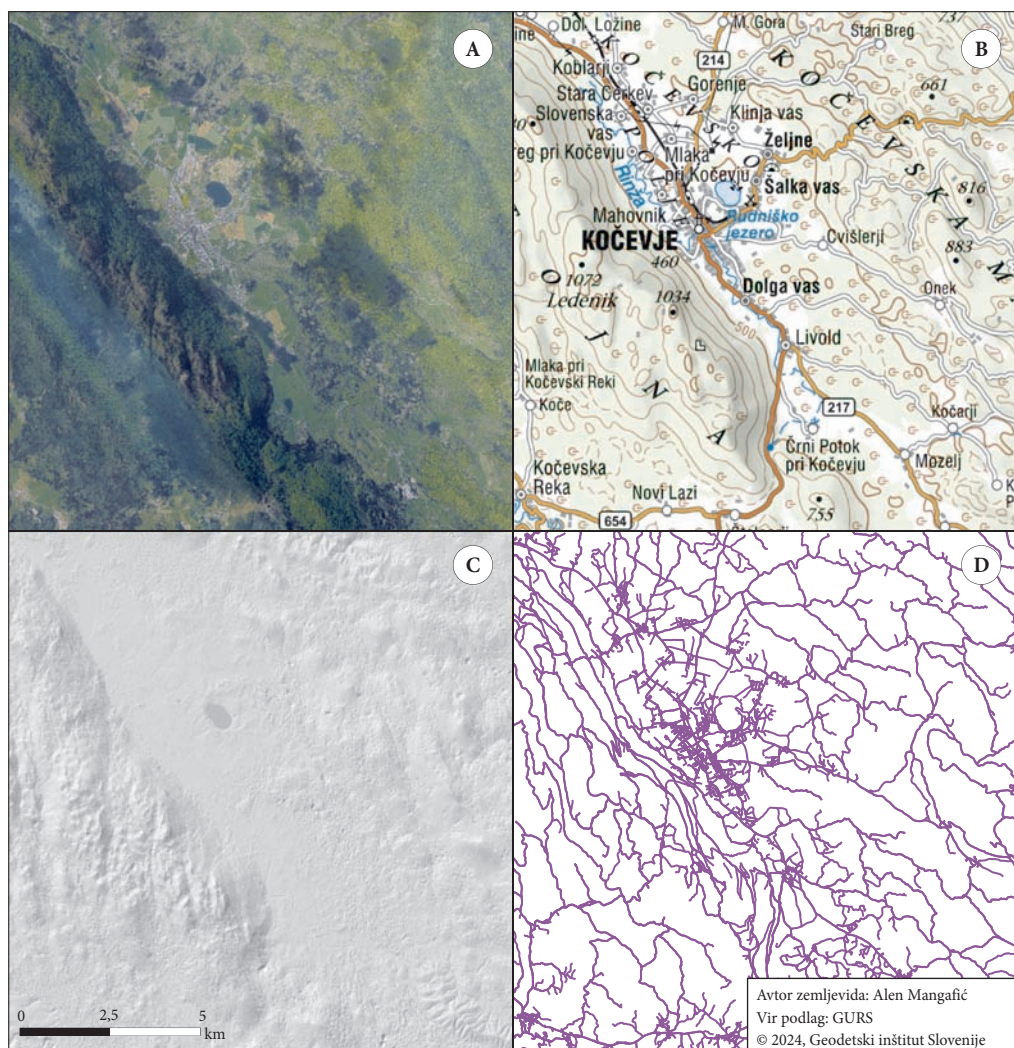
- Model lahko uporabi celotno sliko kot vhod za učenje in napovedovanje. To pomeni, da lahko obravnava slike različnih velikosti, ne da bi bile te potrebne predhodne obdelave ali izrezovanja. To omogoča boljšo uporabo raznolike učne množice, kar lahko poveča splošnost in prilagodljivost modela na različne okoliščine in vrste podatkov.
- Učno množico sestavljeno iz verificiranih parov (vektorski sloj prometnic se prilega slikam) lahko posodabljammo s podatki različnih ločljivosti, brez izločanja starih parov. To omogoča zasnovo dolgoročne, »trajnostne« učne množice.
- Model omogoča kakovostno zaznavo linijskih objektov in zmožnost učenja. Določeni linijski objekti so zelo široki in slike z manjšo ločljivostjo omogočajo integracijo širših objektov v učno množico; ločljivost vhodnih slik je navzgor omejena s strojnimi zmožnostmi računalnika (npr. zmožnost pomnilnika grafične kartice).

Test zaznavanja prometnic smo izvedli na območju v okolici Kočevja (slika 1) na podlagi lidar-skih podatkov Cikličnega laserskega skeniranja Slovenije (CLSS) iz 2023. Uporabili smo georeferencirane oblake točk (GKOT). Na testnem območju se pojavljajo različni tipi reliefa in pokritosti tal. Pri pripravi metodologije smo želeli zagotoviti primerljivo kakovost zaznave med različnimi tipi reliefa in pokritosti tal. Območje smo izbrali zaradi gostega vegetacijskega pokrova, saj so prometnice pod krošnjami največji izziv za samodejno zaznavanje. Povprečna gostota GKOT na študijskem območju znaša 22 točk/m² (med 363 in 1095 m n. v.). Zajem aerolaserskih podatkov je potekal v maju in juniju 2023, ko je bilo drevje že olistano. Študijsko območje je sestavljeno iz 210 listov hierarhične mreže SIHM1000; vsak list pokriva površino 1 km². Učni vektorski sloj prometnic z oznakami razredov smo sestavili s pomočjo geometrij obstoječih slojev prometnic. Uporabili smo prometnice iz Državnega topografskega modela (DTM) in Zbirnega katastra gospodarske javne infrastrukture (ZKGJI).

Informacije, ki jih model uporablja za prepoznavanje lastnosti objektov (v nadaljevanju značilke) smo izbrali na osnovi študije zaznavanja gozdnih cest iz lidarskih oblakov točk na osnovi metode naključnih gozdov, ki jo je izvedla Buján s sodelavci (2021). V svoji raziskavi so razdelili prometnice na več kategorij ter za vsako od njih ocenili pomembnost. V našem primeru smo izbrali najbolj vplivne značilke ter iz njih sestavili trirazsežne matrike za učenje, validacijo in testiranje modela. Dodali smo še optični model in dva reliefna modela po lastni presoji. Naš končni model je sestavljen iz sedmih modelov. V nadaljevanju so podani modeli in uporabljane značilke:

- Optični model; GKOT CLSS vsebuje barvne vrednosti popolnega ortofota (POF) narejenega s slikovnim ujemanjem aerofotografij, ki so zajete sočasno z aerolaserskim skeniranjem:
 - rdeč,
 - zelen in
 - moder kanal;
- reliefni model 1:
 - digitalni model reliefa;
 - povprečna intenziteta odbojev na tleh;
 - število odbojev.
- reliefni model 2:
 - digitalni model reliefa;
 - naklon terena;
 - usmerjenost terena.
- gola tla:
 - normirani digitalni model površja prvih odbojev ($nDMP_{PO}$);
 - povprečna intenziteta prvih odbojev znotraj enega standardnega odklona (I_{PO});

- homogenost $nDMP_{PO}$ (obratni diferencialni moment sivinskih nivojev; Haralick, Shanmugam in Dinstein 1973).
- nizka vegetacija:
 - $nDMP_{PO}$;
 - I_{PO} ;
 - varianca sivinskih nivojev I_{PO} (Haralick, Shanmugam in Dinstein 1973).
- asfaltirane ceste:
 - povprečna intenziteta prvih odbojev nižjih od 4 m znotraj enega standardnega odklona (I_{MV});
 - I_{PO} ;
 - povprečna intenziteta zadnjih odbojev znotraj enega standardnega odklona (I_{ZO}).



Slika 1: Študijsko območje: ortofoto (A), Državna pregledna karta 1: 250.000 (B), analitično senčen GKOT CLSS (C), vektorski sloj prometnic (D).

- makadamske ceste:
 - standardni odklon nadmorske višine;
 - I_{ZO} ;
 - normirani digitalni model površja najvišjih vrednosti odbojev na nizki vegetaciji nižji od 4 m.

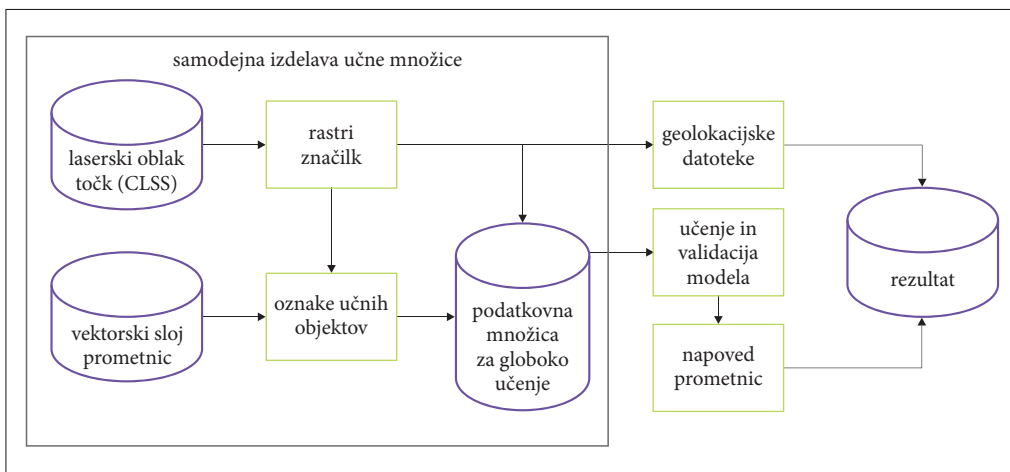
Avtomatizirali smo izdelavo podatkovne množice za globoko učenje, saj je ročni zajem oznak pravi-
loma dolgotrajen in finančno zahteven proces. Pri ročnem zajemu je največja težava označevanje točk
aerolaserskih podatkov, ki se nahajajo pod krošnjami. To zahteva podrobno fotointerpretacijo različnih sen-
čenih višinskih modelov, kot tudi terenske ogleda. Za pripravo podatkovne množice za globoko učenje smo
linijsko geometrijo kombiniranih prometnic razširili za 2 m (slika 1, D). GKOT smo reklasificirali v dva razreda:
prometnice pod vegetacijo (ASPRS razredi 3, 4 in 5) in prometnice na utrjenem površju (ASPRS razredi
2 in 6) ter naredili podatkovno množico oznak iz slikovnih elementov, ki se nahajajo na geometriji raz-
širjenih prometnic. Poleg razreda tal (APSRs razred 2), smo vključili tudi razred stavb (ASPRS razred 6),
da so se modeli lažje naučili ločevati prometnice okoli stavb zaradi nenadne topografske spremembe.

Učenju modela smo namenili 170 (81 %), validaciji 20 (9,5 %) in testiranju 20 listov (9,5 %) mreže
SIHM1000. Stabilnost in uspešnost posameznih modelov smo ocenili na osnovi učnih in validacijskih
funkcij izgub. Za funkcije izgube smo izbrali kategorično navzkrižno entropijo (angleško *categorical
cross-entropy*), pogosto uporabljeno mero za izračun uspešnosti večrazredne semantične segmentaci-
je nevronske mreže, ki so aktivirane s funkcijo *softmax*. Vse modele smo učili 30 epoch, kar pomeni,
da so modeli 30 krat šli čez celotno učno in validacijsko množico.

Končni rezultat je unija rezultatov vseh modelov. Za razliko od vmesnih rezultatov modelov je kon-
čni rezultat binarna maska zaznanih prometnic. Rezultati globokega učenja so slike brez geolokacijskih
metapodatkov. Za prostorsko umeščanje rezultatov, smo uporabili geolokacijske datoteke izpeljane iz vhod-
nih (geolociranih) rastrov značilk. Poleg tehničnih vidikov, smo preučili tudi praktično uporabnost modela
v realnih scenarijih, vključno s hitrostjo delovanja, enostavnostjo implementacije in morebitnimi ome-
jitvami pri uporabi.

Končni rezultat smo primerjali s testnimi podatki ter ocenili skladnost z uporabo naslednjih metrik
(Haghighi 2018):

- absolutna točnost: delež pravilno pozitivnih napovedanih slikovnih elementov glede na oznake iz
vektorskega sloja prometnic;
- priklic: razmerje pravilno pozitivnih napovedi in vsote pravilno pozitivnih in napačno negativnih napovedi;
- delež napačno pozitivnih napovedi.



Slika 2: Zaporedje samodejnih procesov.

Diagram samodejnega procesa lahko vidimo na sliki 2.

Modeliranje smo izvajali z visokonivojsko knjižnico Keras. Procese smo paralelizirali na grafični procesni enoti NVIDIA GeForce RTX 2080 SUPER (8 GB) in računalniku s procesorjem Intel® Core™ i9-990X (3,50 GHz x 20) in pomnilnikom od 64 GB RAM na OS z Linux jedrom 6.5.0-17.generic.

3 Rezultati in njihovo ovrednotenje

Kot glavni rezultat lahko izpostavimo, da se je učenje tako zastavljenih modelov izkazalo kot stabilno. Na sliki 3 lahko vidimo spremembe točnosti in validacijske točnosti nekaterih modelov skozi epohe, prav tako tudi vrednosti izgube in validacijske izgube.

Na vseh primerih lahko vidimo, da sta točnost in validacijska točnost nad 97 %. Podobnost točnosti in validacijske točnosti kaže na to, da so modeli (zelo) uspešni tudi na podatkih, ki jih med učenjem niso upoštevali. Večina modelov ima podobno razliko med vrednostmi izgube in validacijske izgube. Model sestavljen iz matrik DMR – intenziteta – število odbojev (reliefni model 1) ima večjo razliko. Velika razlika med izgubo in validacijsko izgubo lahko kaže na prenaučenos modela (angleško *overfitting*). Tovrstni model se dobro prilagaja učnim podatkom, ampak naučeno slabše posplošuje na novih podatkih. Končne vrednosti izgube so zelo majhne in zadovoljive v večini primerov. Na osnovi teh metrik lahko trdimo, da je reliefni model 1 prenaučen. Pri večini modelov se razlike stopnjujejo po deseti epohi. To pomeni, da bi bila morebitna optimizacija modela učenje samo do deset epoh. Druga optimizacijska rešitev, bi lahko bila implementacija regularizacijskih tehnik, kot so L1, L2 ali dodajanje plasti izpada (angleško *dropout*; Goodfellow, Bengio in Courville 2016).

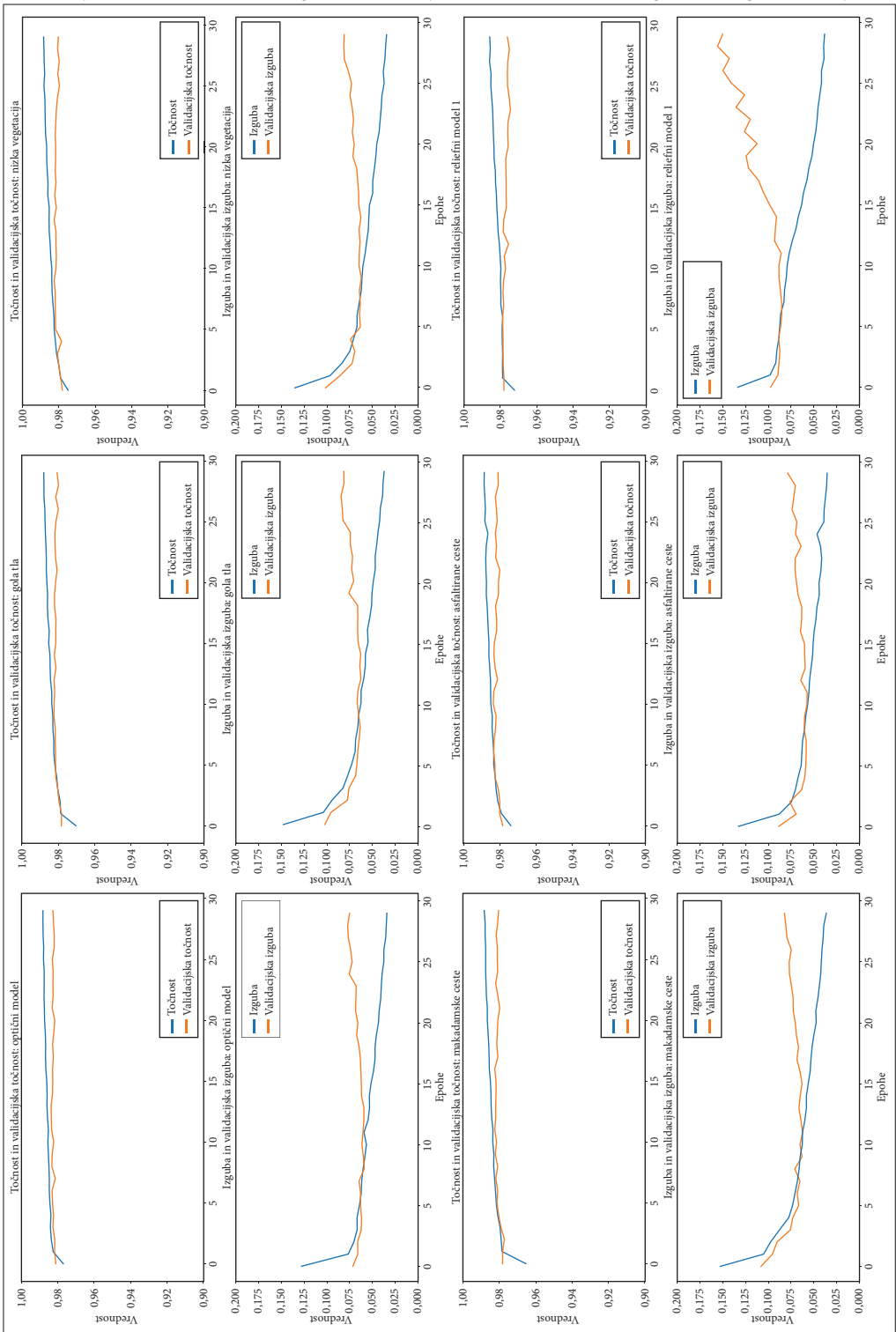
Trajanje učenja ene epohe je bilo med 40 in 58 s za vse modele; le prve epohe so trajale med 127 in 138 s. To pomeni, da smo za učenje vseh modelov potrebovali približno 16 ur. Trajanje napovedi je pri vseh modelih trajalo približno enako: posamezen model je potreboval 225 s za napoved 20 testnih listov (20 km²). To pomeni, da bi za zaznavo prometnic na celotnem območju Slovenije v enakem okolju potrebovali približno 18,5 dni. Priprava značilk traja med 3 in 10 s na list (osnovne značilke so hitrejše od kompleksnejših značilk, kot sta homogenost in varianca sivinskih nivojev).

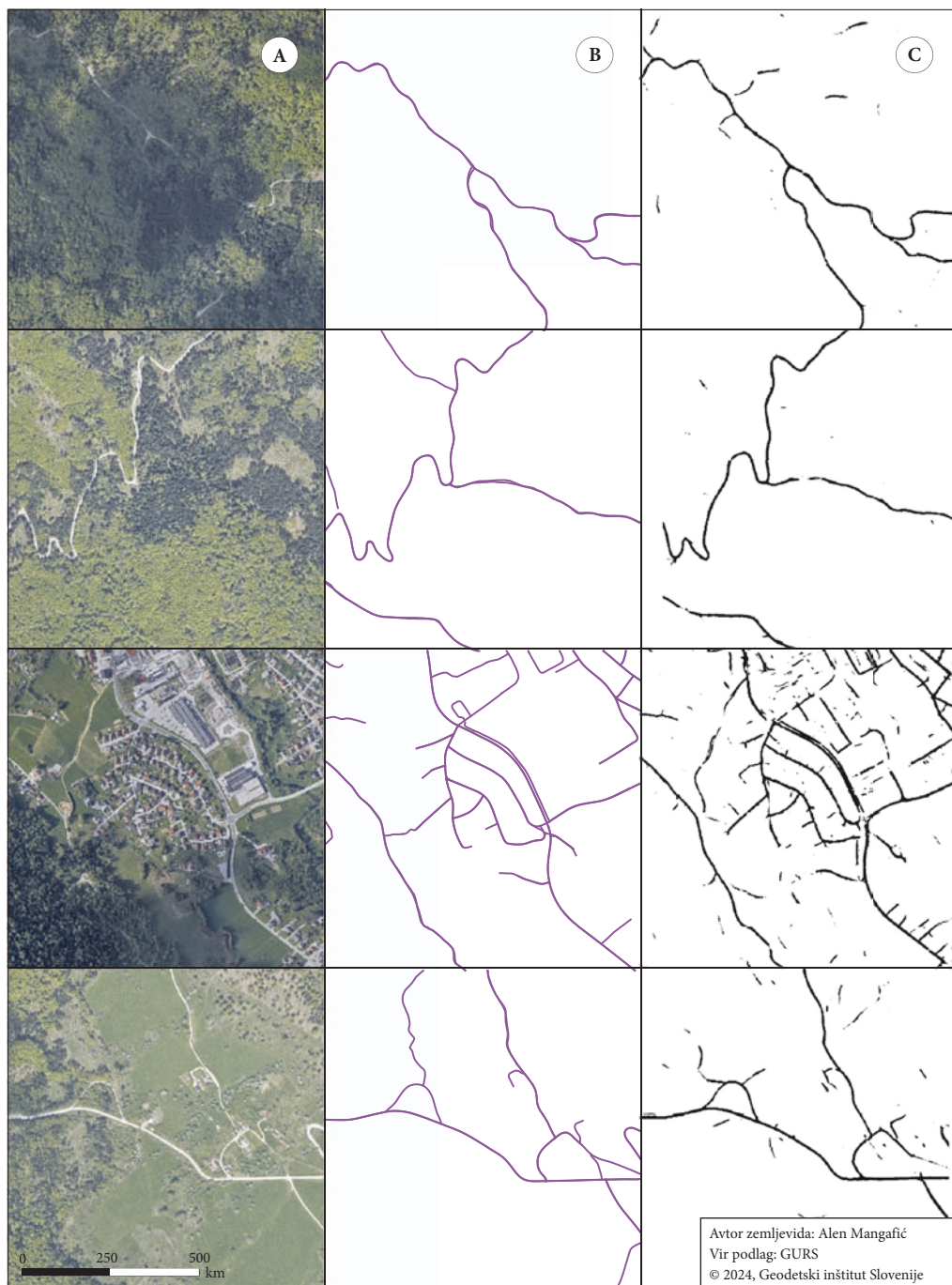
Končni rezultat (slika 4) je dosegel skupno točnost nad 99,8 % in priklic v vrednosti od 60 %. Skupna točnost in priklic v tem primeru nista zanesljiva, saj je večina upoštevanih vrednosti iz razreda ozadja (razred, kjer je odsotnost prometnic), ki je v veliki večini. Dejstvo je, da smo testno množico primerjali z vektorskim slojem oznak, ki nima evidentiranih vseh prometnic ter da je vsak slikovni element, napovedan kot prometnica brez oznake (torej vsaka nova prometnica), označen kot napačno negativna napoved, ki v tem primeru znaša 38 %. Stopnja napačno pozitivnih napovedi znaša 0,001 %, kar pomeni, da je končni rezultat zelo uspešen pri razlikovanju prometnic od ostalega. Primere rezultatov modeliranja lahko vidimo na sliki 4.

Opazimo lahko, da je preizkušeni model uspešno zaznal prometnice pod krošnjami, ter da se te razlikujejo od evidentiranih. Zato menimo, da gre za uspešen model. Opazne so prekinitev v zaznanih prometnicah, ki so najverjetneje posledica odstopanj v vhodnih podatkih. Sklepamo, da določene prometnice niso zaznane zaradi njihove širine: proces samodejne izdelave učne množice upošteva širino prometnic 2 m povsod. V nadaljevanju bi lahko spremenili proces izdelave učne množice z upoštevanjem različnih širin odmikov od osi (na primer z upoštevanjem atributa širine osi iz različni podatkovnih baz). Na urbanih območjih lahko vidimo, da so odseki cest nepovezani. Temu bi se lahko izognili z izboljšavo metodologije izdelave učnih podatkov: modeli so učeni le na linijskih objektih in zaradi tega določena križišča, parkirišča in podobne površine, ki so široke in imajo ploskovno geometrijo, niso klasificirana kot prometnica.

Slika 3: Točnosti in izgube skozi epohe. ► (str. 175)

Samodejna zaznava prometnic z globokim učenjem na podatkih Cikličnega laserskega skeniranja ...





Slika 4: popolni ortofoto (A), vektorski sloj prometnic (B) in rezultati modeliranja (C).

4 Sklep

Kvaliteta aerolaserskih podatkov ključno vpliva na kvaliteto rezultatov zaznave po tukaj preizkušeni metodi. Podatki, ki so bili uporabljeni za izvedbo testa, so bili podatki novega cikličnega aerolaserskega skeniranja v Sloveniji (CLSS), ki se od prejšnjih loči po višji gostoti točk od predhodnega (LSS) in vsebnosti RGB komponent pripisanih točkam v GKOT. Razvili smo proces avtomatiziranega pridobivanja oznak, kar pospešuje celoten proces in hkrati zmanjšuje produkcijske stroške. Predlagana metoda omogoča zaznavo osnovne geometrije prometnic, za semantično ločitev v posamezne vrste prometnic (npr. pot, boljši in slabši kolovoz, cesta, gozdna vlaka) pa je še vedno potrebna nadaljnja ročna obdelava ali interpretacija. Rezultati so primerni za osnovni opozorilni sloj, na osnovi katerega se lahko preverja stanje evidentiranih objektov (sprememba poteka prometnice), identificira nove ali porušene prometnice ter pomaga ročnemu zajemu pri interpretaciji in identifikaciji objektov ob novem zajemu na območjih pod vegetacijo.

5 Viri in literatura

- Buján, S., Guerra-Hernández, J., González-Ferreiro E., Miranda, D. 2021: Forest road detection using LiDAR data and hybrid classification. *Remote Sensing* 13. DOI: <https://doi.org/10.3390/rs13030393>
- Goodfellow, I., Bengio Y., Courville, A. 2016. *Deep Learning*. MIT Press.
- Haghighi, S., Jasemi, M., Hessabi S., Zolanvari, A. 2018: PyCM: Multiclass confusion matrix library in Python. *Journal of Open Source Software* 3. DOI: <https://doi.org/10.21105/joss.00729>
- He, K., Zhang X., Ren S., Sun J. 2015: Deep residual learning for image recognition. *arXiv*. arXiv:1512.03385. DOI: <https://doi.org/10.48550/arXiv.1512.03385>
- Haralick, R.M., Shanmugam, K., Dinstein, I. 1973: Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3. DOI: <https://doi.org/10.1109/TSMC.1973.4309314>
- Khanal, A., Estrada, R. 2020: Dynamic deep networks for retinal vessel segmentation. *Frontiers in Computer Science* 2. DOI: <https://doi.org/10.3389/fcomp.2020.00035>
- Long, J., Shelhamer, E., Darrell, T. 2015: Fully convolutional networks for semantic segmentation. *arXiv*:1411.4038. DOI: <https://doi.org/10.48550/arXiv.1411.4038>
- Ronneberger, O., Fischer P., Brox, T. 2015: U-Net: Convolutional networks for biomedical image segmentation. *arXiv*:1505.04597. DOI: <https://doi.org/10.48550/arXiv.1505.04597>
- Smole, T. 2019: Globoko učenje za segmentacijo in klasifikacijo cestišča. Magistrsko delo: Univerza v Ljubljani, Fakulteta računalništvo in informatiko. Ljubljana.