

Detekcija vegetacijskih krajinskih elementov v podatkih LiDAR in ortofoto z nevronske mreže

Štefan Horvat¹, Štefan Kohek¹, Danijel Ivajnskič^{2,3}, Damjan Strnad¹

¹Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko

²Univerza v Mariboru, Fakulteta za naravoslovje in matematiko

³Univerza v Mariboru, Filozofska fakulteta

E-pošta: stefan.horvat@student.um.si

Detection of vegetation landscape elements in LiDAR and orthophoto data using a neural network

Abstract. *Vegetation landscape elements are small patches of trees, hedges, and other vegetation types, which importantly contribute to landscape and biotic diversity. They have many positive effects on wildlife and ecology in general but are threatened owing to the increasing human impact on the cultural environment. Indeed, their conservation status is of high importance. Thus, many efforts are orientated towards (semi)automated vegetation landscape elements monitoring. With the evolution of GIS, acquiring and processing remote data has never been easier. In this research, we use a binary segmentation neural network model for detection of vegetation landscape elements from LiDAR and orthophoto data, which could be the first step towards continuous monitoring automation. Using the data of a region in north-east Slovenia to train the model, we achieved 86 % precision and 92 % recall on a test set.*

1 Uvod

Vegetacijski krajinski elementi so območja vegetacije, med katere prištevamo mejice, obvodno drevnino, otoke grmovja in drevja ter samostoječe grme in drevesa. Predstavljajo zatočišče ali prostor za domovanje in iskanje hrane številnim živalskim vrstam, predvsem pticam in drugim koristnim zatiralcem škodljivcev na okoliških kmetijskih površinah. Preprečujejo ali omilijo lahko tudi erozijo in izsuševanje tal, ter imajo na splošno ugoden vpliv na ekološke in naravovarstvene kazalce na danem območju. Zaradi tega sta popis in spremljanje razvoja krajinskih elementov ena izmed ključnih ukrepov za ohranjanje biotske raznovrstnosti kulturne krajine, kar sodi med prioritete naloge prihodnje kmetijske politike v Sloveniji [1].

Za določitev področij vegetacijskih krajinskih elementov v podatkih oddaljenega zaznavanja, kot so oblaki točk LiDAR in digitalni ortofoto, so danes na voljo različna orodja geografskih informacijskih sistemov (GIS). Ta vključujejo potrebne algoritme za izpeljavo ustreznih območij vegetacije v obliki geo-referenciranih mnogokotnikov, vendar celoten postopek še vedno vključuje veliko ročnega dela in nadzora posameznih korakov [2]. Samodejna detekcija krajinskih elementov v

podatkih oddaljenega zaznavanja je prvi korak na poti k delni avtomatizaciji postopkov, ki bi olajšali tako spremljanje stanja kot sledenje spremembam.

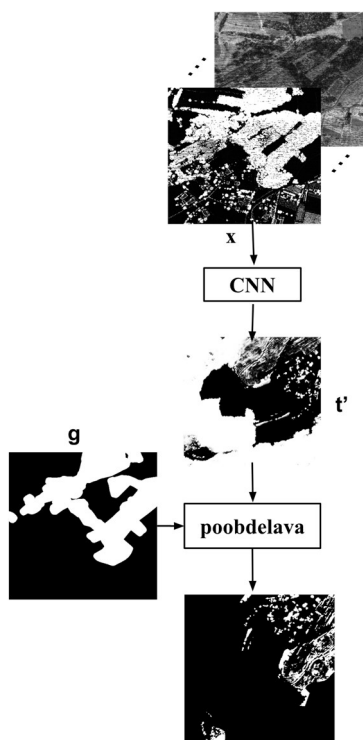
V pričujoči raziskavi smo detekcijo krajinskih elementov izvajali s pomočjo napovednega modela, katerega osnovni del je konvolucijska nevronska mreža (CNN). Ta kot vhod prejme digitalni ortofoto posnetek izbranega območja in pripadajoče sloje značilnic, izpeljanih iz posnetka LiDAR istega območja, oboje v prostorski ločljivosti 1 m. Na izhodu nevronska mreža tvori sliko enakih dimenzij kot na vhodu, iz katere se z binarizacijo in odstranitvijo gozda določijo območja krajinskih elementov, ki jih je v nadaljevanju mogoče pretvoriti v vektorsko obliko.

Sloj vegetacijskih krajinskih elementov za celotno Slovenijo je bil pred kratkim izdelan s strani ZRC SAZU in Gozdarskega inštituta Slovenije [2], zato smo enako metodologijo uporabili za validacijo našega napovednega modela. Podoben obstoječi produkt je vektorski sloj Small Woody Features (SWF), ki je brezplačno dostopen preko storitve Copernicus Land Monitoring [3], a vsebuje samo zaplate lesne vegetacije s površino nad 200 m² v prostorski ločljivosti 5 m. S tem so izločeni številni manjši vegetacijski krajinski elementi, z njimi pa tudi možnosti za podrobno ocenitev stanja ter spremljanje razvoja.

2 Metodologija in podatki

2.1 Opis postopka

Postopek določanja območij vegetacijskih krajinskih elementov v daljinsko zajetih podatkih s pomočjo nevronske mreže lahko povzamemo v treh korakih (slika 1). Prvi korak je pretvorba podatkov iz izvorne oblike v obliko, ki jo na vhodu sprejema nevronska mreža. Vhodni podatki predstavljajo opazovano področje kot skupek digitalnih ortofoto posnetkov in štirih rastrskih slojev, izpeljanih iz meritev LiDAR. V drugem koraku uporabimo prehodno naučen napovedni model za tvorbo izhodne sivinske slike, v kateri višje vrednosti predstavljajo večjo verjetnost obstoja krajinskega elementa na pripadajoči lokaciji. V zadnjem koraku izvedemo poobdelavo napovedi iz drugega koraka, s katero najprej sivinsko sliko pretvorimo v binarno obliko pri izbranem pragu, nato pa iz nje s pomočjo znane maske gozdnih površin izločimo gozdna območja. Rezultat poobdelave je binarna segmentirana slika, v kateri so z belo barvo označena območja krajinskih elementov.



Slika 1: Postopek detekcije vegetacijskih krajinskih elementov: rastrske sloje digitalnega ortofoto in značilnic LiDAR (x) pošljemo skozi CNN, ki tvori sivinsko sliko napovedi (t'). To v poobdelavi binariziramo in iz nje s pomočjo maske g izločimo znana območja gozda.

Za učenje CNN je potrebno pripraviti učno množico, ki je sestavljena iz ustrezno oblikovanih vhodnih podatkov in pripadajočih pričakovanih izhodov. Napredek napovednega modela med samim učenjem ocenjujemo s standardnimi metrikami, uspešnost naučenega modela pa merimo na neodvisni testni množici. Podrobnejši opis postopka je v podpoglavju 2.3.

2.2 Podatki

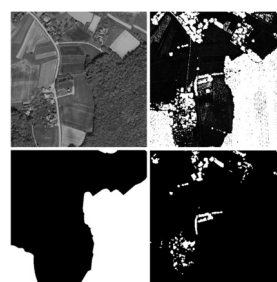
Za izvedbo eksperimenta smo uporabili digitalne ortofoto posnetke Slovenije [4] in oblake točk LiDAR [5], pri čemer smo zaradi časovne uskladitve vseh potrebnih slojev uporabili podatke iz leta 2014. Dodatno smo za določitev gozdne maske uporabili sloj rabe tal iz istega leta [6], za določitev ciljnih vrednosti pri učenju modela pa sloj vegetacijskih krajinskih elementov, ki je bil za izbrano območje izdelan z enako metodologijo in na podlagi istih podatkov LiDAR kot referenčni sloj ZRC SAZU [2], a brez poenostavitve geometrije.

Digitalne ortofoto posnetke smo prevzorčili na prostorsko ločljivost 1 m s kubično interpolacijo. Iz oblakov točk LiDAR smo s pomočjo programske knjižnice lidR [7] generirali rastrske sloje značilnic na metrski mreži. Izmed vseh značilnic, ki jih omenjena knjižnica podpira, smo pri učenju nevronske mreže uporabili štiri: maksimalno višino in njen standardni odklon ter maksimalno in povprečno intenziteto točk v posamezni celici. Binarno masko gozda g smo tvorili z rasterizacijo mnogokotnikov z oznako rabe 2000 iz sloja rabe tal, masko pravih vrednosti t pa z rasterizacijo sloja krajinskih elementov, za kar smo uporabili orodje

QGIS [8]. Za potrebe učenja CNN smo maske združili v eno polje y z naslednjo operacijo:

$$y = -g + (1 - g) \times t. \quad (1)$$

V eksperimentu smo se omejili na območje Goriškega, kjer smo izločili pet območij velikosti 4 km². Od teh smo tri uporabili za učenje modela, eno za sprotno validacijo in optimizacijo parametrov, eno pa za končno testiranje naučenega modela. Za tvorbo učnih primerov smo vsako od območij razdelili na manjša področja velikosti 0,25 km². Učno množico smo povečali na 192 primerov z rotacijami posameznih slik za 90, 180 in 270 stopinj. Vsak učni primer tako sestoji iz vhodnega tenzorja x dimenzije 500×500×7 in izhodnega tenzorja y dimenzije 500×500×1. Primer dveh vhodnih slojev, gozdne maske in maske ciljnih vrednosti za en učni vzorec prikazuje slika 2.



Slika 2: Primeri rastrskih slojev za isto geografsko območje, ki se uporabljajo pri učenju CNN – digitalni ortofoto (levo zgoraj), sloj maksimalnih višin (desno zgoraj), maska gozda (levo spodaj) in ciljne vrednosti (desno spodaj).

2.3 Nevronska mreža

Naloga nevronske mreže je v našem primeru binarna segmentacija vhodnih podatkov na območja vegetacijskih krajinskih elementov in ostalih elementov okolja. Ker je izhod modela slika enake širine in višine kot podatkovni sloji na vhodu, smo za segmentacijo uporabili arhitekturo CNN, ki je sestavljena iz samih konvolucijskih plasti. Preskušali smo arhitekture z različnim številom konvolucijskih plasti (med 3 in 6) in vrednostmi hiperparametrov. Kot končni model smo izbrali arhitekturo s štirimi plastmi:

- konvolucijska plast z 32 filtri velikosti 7×7,
- konvolucijska plast s 64 filtri velikosti 5×5,
- konvolucijska plast s 64 filtri velikosti 3×3,
- izhodna konvolucijska plast z enim filtrom velikosti 1×1.

Skrite plasti nevronske mreže uporabljajo aktivacijsko funkcijo ReLU, izhodna plast pa sigmoidno aktivacijsko funkcijo, ki za vsak piksel izhodne slike izračuna vrednost med 0 in 1. Vse plasti ohranjajo dimenzijo vhodnega tenzorja, tako da ga razširijo z ustreznim številom vrstic in stolpcev ničelnih vrednosti.

Učenje CNN poteka s prilagajanjem filtrov glede na izgubno funkcijo binarne prečne entropije. Ker pa se izgubna funkcija in metrike uspešnosti privzeto izračunavajo nad vsemi piksli izhodne slike, nas pa zanimajo samo piksli negozdskih površin, smo implementirali maskirane metrike in izgubno funkcijo. Te iz podanega tenzorja y najprej izluščijo masko gozda

g in ciljne vrednosti t , nato pa izračun izgube oz. metrike omejijo na interesne regije z množenjem z inverzom maske gozda. Ker vegetacijski krajinski elementi v splošnem predstavljajo manjši delež v maskah ciljnih vrednosti, je število pozitivnih in negativnih primerov neuravnoteženo. Posledica tega je lahko pristransko učenje, pri katerem napovedni model optimizira delovanje z osredotočanjem na eno vrsto primerov. Zaradi tega smo uporabili uteženo funkcijo izgube, pri kateri težo napak v piksljih, ki pripadajo krajinskim elementom, obtežimo z utežjo w . V eksperimentih smo uporabili empirično vrednost $w=0,7$. Maskirana funkcija izgube L se tako izračuna po enačbi:

$$L = \text{mean}((1 - w) \times (1 - t) \circ x + w \times t \circ x), \quad (2)$$

kjer \circ označuje množenje tenzorjev po elementih, mean izračuna srednjo vrednost tenzorja, x pa je maskirani tenzor napak prečne entropije med napovedmi t' in ciljnim vrednostmi t po posameznih piksljih:

$$x = (1 - g) \circ (t \circ \log t' + (1 - t) \circ \log(1 - t')). \quad (3)$$

3 Rezultati

3.1 Opis eksperimenta

Nevronsko mrežo smo implementirali s pomočjo ogrodja Tensorflow [9]. Učenje je potekalo na računalniku s procesorjem Ryzen 9, grafično kartico Nvidia GeForce RTX 3060 in 64 GB sistemskega pomnilnika. Učenje smo izvajali 500 epoh v paketih po šest učnih primerov, za optimizacijo pa smo uporabili algoritem adam s privzetimi vrednostmi hiperparametrov.

Optimizacijo arhitekture modela in parametrov učenja smo izvajali glede na rezultate na validacijski množici. Pri izbiri modela smo preizkušali več različic arhitekture, kjer smo spreminjali število konvolucijskih plasti, število in velikosti filtrov ter aktivacijske funkcije skritih plasti. Poskusili smo tudi z vključevanjem združevalnih (maxpool) plasti in razširjene konvolucije, kar pa ni privedlo do izboljšanja rezultatov. Rezultate zato poročamo za opisano arhitekturo s štirimi konvolucijskimi plastmi, ki smo jo uporabili za končno učenje in testiranje modela na neodvisni testni množici.

Ocenjevanje napredka pri učenju in sposobnost posploševanja mreže na validacijski oz. testni množici smo preverjali s pomočjo standardnih metrik natančnosti (precision, P) in priklica (recall, R) ter njunega harmoničnega povprečja F_1 . Pri izračunu metrik podobno kot pri izgubni funkciji maskiramo gozdne predele, saj nas zanima uspešnost modela samo na območjih, ki se v končni sliki ne izločijo. Sivinska slika na izhodu mreže se najprej pretvori v binarno glede na izbrani prag, ki je privzeto enak 0,5. S primerjavo dobljene binarne slike in ciljnih vrednosti se nato določijo:

- število pravilno napovedanih pikslov, ki pripadajo iskanim krajinskim elementom (true positive, TP),
- število napačno napovedanih pikslov, ki dejansko pripadajo iskanim krajinskim elementom (false negative, FN) in

- število napačno napovedanih pikslov, ki niso iskani krajinski elementi (false positive, FP).

Vrednosti metrik se nato izračunajo kot:

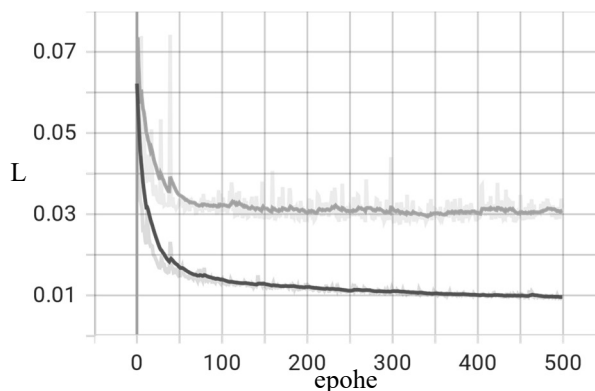
$$P = \frac{TP}{TP + FP}, \quad (4)$$

$$R = \frac{TP}{TP + FN}, \quad (5)$$

$$F1 = \frac{2 \cdot P \cdot R}{P + R}. \quad (6)$$

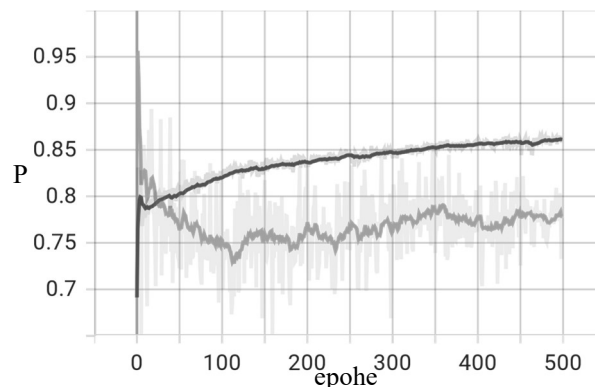
3.2 Rezultati in diskusija

Slika 4 prikazuje graf vrednosti funkcije izgube na učni in validacijski množici po epohah učenja. Opazimo lahko, da se izguba na učni množici zmanjšuje skozi celotno učenje, medtem ko se njen napredek na validacijski množici konča med epohama 200 in 300. Končna napaka je na validacijski množici pričakovano višja kot na učni množici, razlike med njima pa nismo uspeli zmanjšati z regularizacijo.

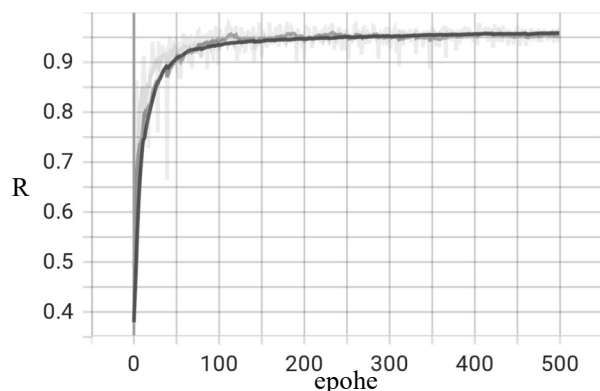


Slika 4: Vrednost izgube skozi epohe učenja na učni (temna krivulja) in validacijski (svetla krivulja) množici.

Na slikah 5 in 6 vidimo spreminjanje natančnosti in priklica na učni in validacijski množici med učenjem. Opazimo lahko, da se krivulji priklica skoraj prekrivata, medtem ko je pri natančnosti razlika bolj opazna. Največje težave smo imeli s tem, da je bil priklic običajno visok, natančnost pa nizka. Z optimizacijo smo našli zadovoljivo konfiguracijo, ki smo jo po zaključku učenja uporabili za napovedovanje na neodvisni testni množici, kjer je dosegla natančnost 86 %, priklic 92 % in vrednost F_1 89 %.

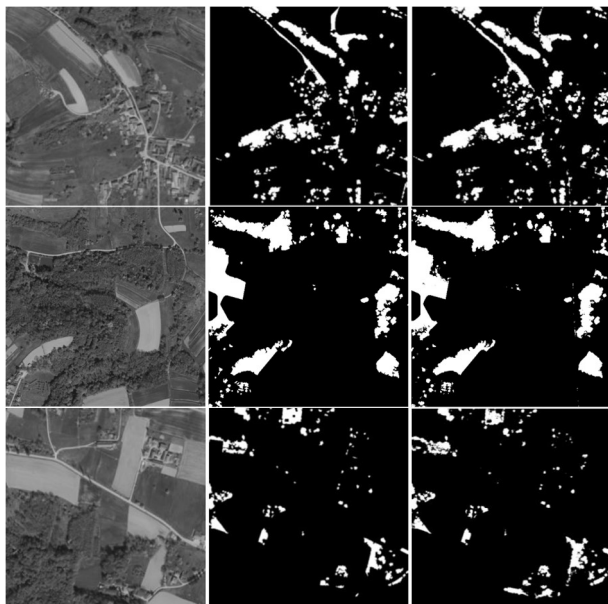


Slika 5: Natančnost na učni (temna krivulja) in validacijski (svetla krivulja) množici skozi epohe učenja.



Slika 6: Priklic na učni (temna krivulja) in validacijski (svetla krivulja) množici skozi epohe učenja.

Uspešnost modela smo ocenili tudi z vizualno primerjavo napovedi in pravih vrednosti, katerih primere prikazuje slika 7. Ugotovili smo, da ima model občasno težave z napačno označitvijo objektov, ki niso vegetacijski krajinski elementi (npr. vinogradi in sadovnjaki).



Slika 7: Primeri napovedi - ortofoto področja (levo), pravilne vrednosti (sredina) in napovedi (desno).

4 Sklep

V raziskavi smo uporabili lastno konvolucijsko nevronske mreže za izdelavo napovednega modela, ki tvori binarno masko vegetacijskih krajinskih elementov na podlagi posnetkov ortofoto in LiDAR. Izbrana arhitektura je za izbrano regijo dosegla natančnost nad 80 % in priklic nad 90 %. Kot dobro orodje se je izkazala tudi lastna implementacija metrik in izgubne funkcije, ki lahko pomembno prispeva k manj pristranskemu učenju in evalvaciji modela.

Ena glavnih pomanjkljivosti trenutnega pristopa je omejitev na izbrano regijo (v našem primeru Goričko). Za posplošitev modela na celotno Slovenijo bi bilo potrebno pripraviti precej bolj obsežno učno množico, ki

bi zajemala posnetke v različnih časovnih obdobjih in pri različnih vremenskih pogojih. V tem primeru je pomembno zagotoviti tudi čim boljše časovno poravnano podatk ortofoto in LiDAR, zato da razlike med njimi zaradi sprememb v naravi niso prevelike.

Obstoječ pristop je mogoče razširiti na več načinov. Ena od možnih izboljšav je uporaba dodatnih rastrskih slojev iz značilnic izpeljanih iz podatkov LiDAR. Rezultate bi lahko izboljšali tudi tako, da bi uporabili učenje s prenosom znanja, pri katerem na drugih podatkovnih zbirkah prednaučene modele segmentacije prilagodimo za detekcijo vegetacijskih krajinskih elementov. Vsekakor pa je avtomatizacija monitoringa vegetacijskih krajinskih elementov na prostorsko zelo heterogenih območjih, kakršna je Slovenija, ob upoštevanju časovne neskladnosti potrebnih vhodnih podatkov, zapletena naloga. Ne glede na to so predstavljeni rezultati korak bliže k rešitvi te problematike, ki bo v prihodnji perspektivi Skupne kmetijske politike EU vse bolj pridobivala na pomenu.

Zahvala: Raziskavo sta sofinancirala Javna agencija za raziskovalno dejavnost Republike Slovenije in Ministrstvo za kmetijstvo, gozdarstvo in prehrano v sklopu ciljnega raziskovalnega projekta V4-2018: Krajinske značilnosti in ukrepi bodoče kmetijske politike v Sloveniji.

Literatura

- [1] Skupna kmetijska politika 2023-2027, <https://www.gov.si/zbirke/projekti-in-programi/skupna-kmetijska-politika-po-letu-2020/>
- [2] Ž. Kokalj, L. Stančič, K. D. Noumonvi, A. Kobler, »Testiranje možnosti in izvedba kartiranja krajinskih struktur, pomembnih za biotsko raznovrstnost in blaženje podnebnih sprememb z daljinskim zaznavanjem : končno poročilo«, ZRS SAZU, 2020.
- [3] Small Woody Features, <https://land.copernicus.eu/pan-european/high-resolution-layers/small-woody-features>
- [4] Geodetska uprava Republike Slovenije, Digitalni ortofoto, 2014, <https://www.e-prostor.gov.si/zbirke-prostorskih-podatkov/topografski-in-kartografski-podatki/ortofoto/> [dostop 1. 7. 2022]
- [5] Ministrstvo za okolje in prostor RS, LiDAR, 2014, <http://www.evode.gov.si/index.php?id=69> [dostop 1. 7. 2022]
- [6] Ministrstvo za kmetijstvo, gozdarstvo in prehrano RS, Raba tal, 2014, <https://rkg.gov.si/vstop/> [dostop 1. 7. 2022]
- [7] J. Roussel, D. Auty, N. C. Coops, P. Tompalski, T. R. Goodbody, A. S. Meador, J. Bourdon, F. de Boissieu, A. Achim, »lidR: An R package for analysis of Airborne Laser Scanning (ALS) data«, *Remote Sensing of Environment*, 251, 112061, 2020.
- [8] QGIS.org, 2022. QGIS Geographic Information System. QGIS Association. <http://www.qgis.org> [dostop 1. 7. 2022]
- [9] M. Abadi et al., »TensorFlow: Large-scale machine learning on heterogeneous systems«, 2015, <https://www.tensorflow.org/> [dostop 1. 7. 2022]