

IZDELAVA VISOKOLOČLJIVEGA SINTETIČNEGA ORTOFOTA NA OSNOVI POSNETKOV SENTINEL-2 Z UPORABO SUPER-LOČLJIVOSTNIH GENERATIVNIH NASPROTNIŠKIH MREŽ

Alen Mangafić, dr. Tomaž Žagar

Geodetski inštitut Slovenije

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

tomaz.zagar@gis.si

DOI: https://doi.org/10.3986/9789610504696_07

UDK: 528.7(497.41)

IZVLEČEK

Izdelava visokoločljivega sintetičnega ortofota na osnovi posnetkov Sentinel-2 z uporabo super-ločljivostnih generativnih nasprotniških mrež

V tem prispevku predstavljamo pristop globokega učenja za izdelavo visokoločljivega 2,5-metrskega sintetičnega ortofota na osnovi posnetkov Sentinel-2. Metoda izboljšave sloni na generativnih nasprotniških mrežah – zvrsti globokega učenja, ki sloni na dveh nevronskih mrežah, generatorju in diskriminatorju. Za povečanje ločljivosti izbranih posnetkov z izvajanjem podpikselske napovedi smo uporabili super-ločljivostne generativne nasprotniške mreže (SR-GAN). Generator, ki smo ga poimenovali SRĐAN, je primeren za izboljšavo vseh posnetkov Sentinel-2 severovzhodne Slovenije brez oblakov in snežne odeje.

KLJUČNE BESEDE

Sentinel-2, ortofoto, globoko učenje, nevronske mreže, generativne nasprotniške mreže

ABSTRACT

Creation of high-resolution synthetic orthophoto based on Sentinel-2 imagery using super-resolution generative adversary networks

In this paper we present a deep learning approach to produce high resolution, 2.5 meter synthetic orthophoto from Sentinel-2 imagery. The enhancement method is based on generative adversarial networks, which are a class of deep learning systems based on two neural networks, a generator and a discriminator. We used super-resolution generative adversary networks (SR-GAN) to increase the resolution of selected imagery by executing a sub-pixel prediction. The generator, which we named SRĐAN, is suitable for enhancing all the cloud-free and snow-free Sentinel-2 images that cover north-east Slovenia.

KEY WORDS

Sentinel-2, orthophoto, deep learning, neural networks, generative adversarial networks

1 Uvod

Ciklično aerofotografiranje Slovenije (v nadaljevanju CAS) zajema približno tretjino Slovenije vsako leto. Posledično je njegova časovna ločljivost 3 leta. Najpomembnejši izdelek CAS je ortofoto, ki ima prostorsko ločljivost do 0,25 m. Veliko projektov, ki temeljijo na zajemu prostorskih podatkov, je odvisnih od ortofota CAS. Stroga merila natančnosti zajema zahtevajo projekti, ki se ukvarjajo z vzdrževanjem dejanske rabe kmetijskih in gozdnih zemljišč (Interpretacijski ključ 2013) ali s posodabljanjem Državnega topografskega modela (Pravilnik o ... 2014). Ortofoto CAS sicer omogoča izpolnjevanje teh pogojev, ampak je časovna ločljivost oziroma semantična točnost in vsebinska popolnost zagotovljena le za triletno obdobje. Obstaja veliko primernih alternativ ortofotu CAS, ki ponujajo tudi enodnevne časovne ločljivosti. Satelitski posnetki WorldView-2 in 3 ponujajo pankromatske posnetke z ločljivostjo med 0,31 in 0,43 m v nadirju, konstelacija Pléiades pa prav tako izjemnih 0,5 m. Največja omejitev pogošte uporabe teh posnetkov je njihova cena. Drugi vir posnetkov z visoko časovno ločljivostjo so posnetki konstelacij Sentinel-2, Terra in Landsat-7 in 8. Ti posnetki so odprti podatki (brezplačni in prosto dostopni brez omejitev avtorskih pravic). Od naštetih je Sentinel-2 tista konstelacija, ki ima najboljšo prostorsko ločljivost in na območju Slovenije ponuja izdelke vsakih 2–5 dni (ESA 2020). Sentinel-2 zajame območje osrednje Slovenije enkrat na 5 dni, v zahodnem in vzhodnem delu Slovenije pa se orbite prekrivajo, tako da v 5 dneh ti območji zajame dvakrat (Pehani 2020).

V okviru projekta Lidarsko podprte prostovoljske geografske informacije za ugotavljanje topografskih sprememb, ki ga financira ARRS (L2-1826), raziskujemo tudi uporabnost podatkov Sentinel-1 in Sentinel-2 za namen posodabljanja zbirke podatkov Državnega topografskega modela (DTM). Izvirni podatki posnetkov Sentinel-2 imajo prostorsko ločljivost do 10 metrov. Njihova fotointerpretacija in samodejna klasifikacija sta zaradi slabše absolutne položajne natančnosti in prostorske ločljivosti lahko slabši oziroma neprimerni kot dopolnilni vir za nekatere topografske objektne tipe, kot so na primer infrastrukturni objekti (MPC Team 2020). Obstoječa metodologija zajema topografskih podatkov za posodabljanje zbirke podatkov DTM sloni na uporabi kombinacije izdelkov CAS in Laserskega skeniranja Slovenije (LSS). Največji potencial podatkov Sentinel-2 v tem kontekstu je ta, da so lahko vir za nekatere kontrole in vsebinske popravke za približno dve tretjini Slovenije, ki ni posneta vsako leto, ter za preostalo tretjino območja, ki je posneta v drugem časovnem obdobju. To pomeni, da lahko z uporabo posnetkov Sentinel-2 hitro identificiramo najnovije spremembe ter s tem prehitimo ortofoto CAS z obdobjem snemanja treh let.

Raziskavo smo razdelili na dve nalogi: prvo, ki raziskuje možnosti izboljšave ločljivosti podatkov Sentinel-2, in drugo, ki ocenjuje prostorsko in semantično natančnost le-teh. V tem prispevku predstavljamo algoritem kognitivne izboljšave posnetkov Sentinel-2, ki ne sloni na optičnem ostrenju, ampak na podpikselski napovedi z uporabo modela globokega učenja – področja strojnega učenja. Rezultat izboljšave je sintetični ortofoto 2,5-metrске ločljivosti generiran iz posnetkov Sentinel-2.

2 Metoda izdelave modela

2.1 Generativne nasprotniške mreže

Podpikselske analize kot je na primer spektralno ločevanje MESMA (Dennison in Roberts 2003) zahtevajo poznavanje končnih pikslov (angleško *endmember*). Končne piksele v terminologiji strojnega učenja lahko predstavljajo označbe mešanih vrednosti znotraj piksla, koncept izvajanja spektralnega ločevanja pa je lahko analogen nadzorovanemu strojnemu učenju. V nasprotju s spektralnim ločevanjem ima predlagani model nenadzorovano osnovo učenja (in nadzorovano posodabljanje uteži nevronske mreže) le pri svoji izgradnji, potem pa ga lahko uporabljamo hitro in brez spreminjanja na ostalih podatkih Sentinel-2. Metoda izboljšave sloni na generativnih nasprotniških mrežah (GAN), ki

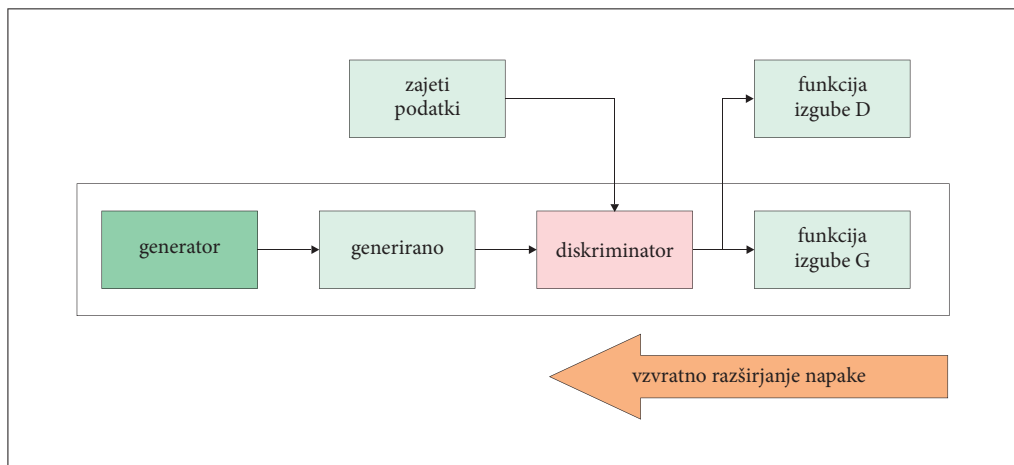
so veja globokega učenja sloneča na dveh nevronskih mrežah: generatorju in diskriminatorju. Mreži iterativno tekmujeta med seboj z namenom generacije novih podatkov in diskriminacije oziroma zavračanja ali sprejemanja le-teh.

Pri učenju GAN so vhodni podatki slike, ki jih želimo generirati. Rezultat učenega klasifikatorja pa so umetno zgrajene slike, ki so čim bolj podobne vhodnim podatkom. V fazi učenja generator izdeluje slike, ki jih diskriminator oceni s primerjanjem z vhodnimi podatki. Rezultata iteracije učenja sta funkciji izgube generatorja in diskriminatorja, ki se vzvratno razširjata ter posodabljata naslednjo iteracijo učenja (slika 1). Izgradnja optimalnega modela se konča, ko generator pride do take stopnje, da generira slike, ki jih diskriminator oceni kot resnične (Goodfellow s sodelavci 2014). Poznamo različne kategorije GAN, ki se razlikujejo glede na namen. Za namene kognitivne izboljšave ločljivosti slik (iz manjše v večjo) je Ledig s sodelavci (2016) predlagal arhitekturo superločljivostnih generativnih nasprotniških mrež (SR-GAN), ki so zelo učinkovite v kombinaciji s tehnikami prenesenega učenja nevronskih mrež (Ledig s sodelavci 2016; Salimans s sodelavci 2016; Nagano in Yohei 2018). Te imajo arhitekturo, ki vsebuje več milijonov naučenih konvolucijskih nevronskih mrež, medsebojno povezanih z milijardami povezav kot so VGGNet, Xception, ResNet in Inception.

Model SR-GAN, ki je namenjen kognitivni izboljšavi posnetkov Sentinel-2, so izdelali na francoskem Nacionalnem inštitutu za prostorske raziskave in tehnologije okolja in kmetijstva (Cresson 2019). Referenčni podatki za izboljšavo so bili pankromatsko ostreni posnetki Spot 6 in 7, ki imajo ločljivost 1,5 m. Rezultate so sicer objavili v obliki spletnega servisa, ampak jih niso ovrednotili ali objavili izvorne kode oziroma opisali arhitekture. V nadaljevanju bomo predstavili ozadje našega generatorja za izboljšavo ortofota, ki smo ga poimenovali SRĐAN.

2.2 SRĐAN: podatki, arhitektura in izgradnja generatorja

Naše študijsko območje obsega severovzhodno Slovenijo. To območje smo izbrali, ker je bilo na njem v letu 2016 izvajano CAS in ker se na tem območju križata dva različna pasova Sentinel-2, kar pomeni večjo gostoto snemanja. Za to območje se je CAS ponovno izvajal leta 2019. Tako smo lahko preverili uporabnost modela tudi za podatke po snemanju leta 2016. Raznolikost pasov Sentinel-2 je pomembna, ker smo želeli izdelati čim bolj univerzalen generator, ki ni tako občutljiv na radiometrične in fenološke razlike. SRĐAN je generator za izboljšavo podatkov Sentinel-2, ki je zgrajen na osnovi

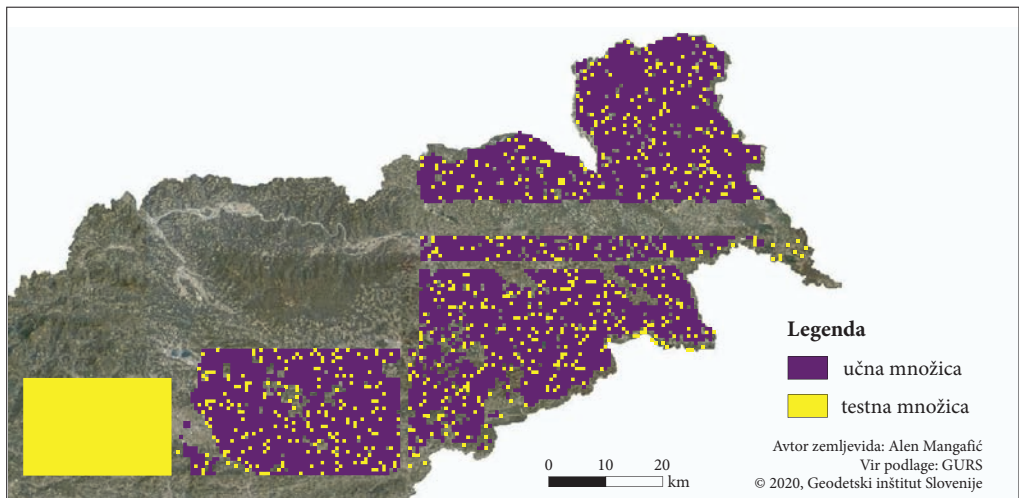


Slika 1: Dinamika učenja generativnih nasprotniških mrež.

6073 parov ortofoto – Sentinel-2 (slika 2). Predobdelava vhodnih podatkov obsega pridobivanje podatkov Sentinel-2, ki so časovno najbližji datumu snemanja CAS. Datumi snemanja izbranega ortofoto vzorca so bili 16. 4., 21. 4. in 22. 5. 2016. Izbrani podatki Sentinel-2 pa so zajeti 13. 4., 26. 4. in 23. 5. 2016.

Sestavili smo pare skladnih podatkov, tako da smo izločili vse podatke Sentinel-2, ki imajo oblake, sence oblakov, radiometrične napake ali pa zaradi časovne razlike nastanka posnetka niso vsebinsko skladni (na primer vegetacija na njivah na posnetku Sentinel-2, ki je ni na ortofotu ali obratno). Vsebinsko skladnost smo preverili s fotointerpretacijo. Pred tem smo podatke Sentinel-2 radiometrično obdelali ter jih dodatno georeferenciali. Georeferenciranje je pomembno za generativne nasprotniške mreže, saj so sestavljene iz blokov konvolucijskih nevronske mreže, ki so občutljive na rotacije, premike in razlike merila (Goodfellow, Bengio in Courville 2016). Arhitekturo modela (slika 3) smo prevzeli iz SR-GAN arhitekture, ki jo predlagajo Ledig s sodelavci (2016) in Ahirwar (2019). Uporabljena arhitektura definira funkcijo izgube kot evklidsko razdaljo med vhodnimi podobami in podobami pridobljenimi iz konvolucijskih mreže, ki so izvedene s prenosom znanja iz mreže VGG19 (Simonyan in Zisserman 2015). Izračun uteži modela se začne tako, da generator izdelava raster z naključnimi vrednostmi, ki ga diskriminator oceni s funkcijo izgube nato pa se uteži iterativno izboljšujejo.

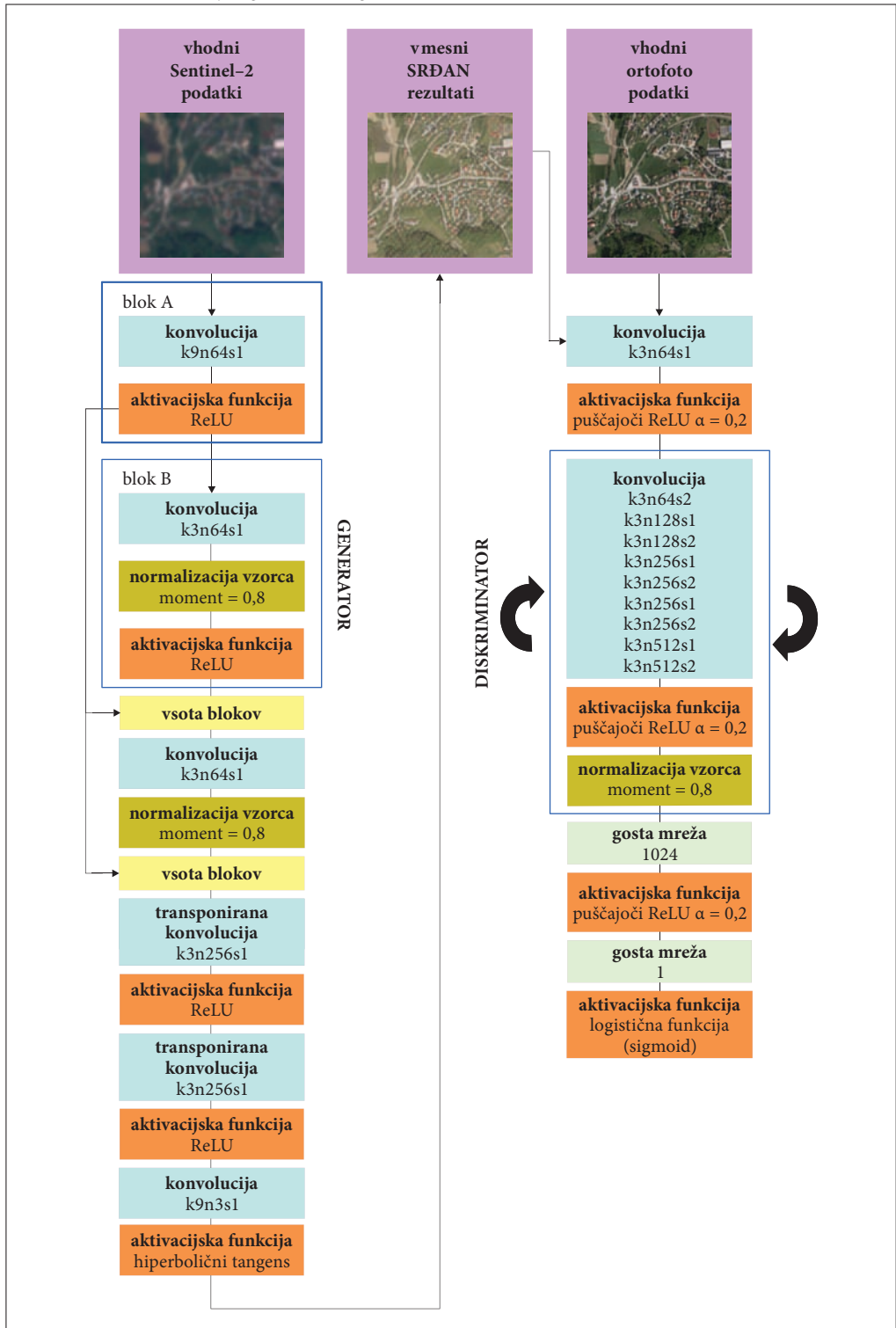
Ko je model zgrajen, lahko izdelamo sintetično ortofoto sliko samo s pomočjo izračunanih uteži modela in posnetka Sentinel-2. Slika 3 prikazuje (hiper)parametre konvolucij, aktivacijskih funkcij (ReLU, puščajoči ReLU, hiperbolični tangens in sigmoid), normalizacije in goščanja. Parameter konvolucije k določa velikost konvolucijskega filtra, ki se izvaja na sliki velikosti $n \times n$, medtem ko s določa korak premikanja konvolucijskega filtra. Uporaba aktivacijskih funkcij zmanjšuje možnost dogodka »izgina ali eksplozije« gradienta (angleško *vanishing, exploding gradient*), vendar pa ne zagotavlja, da se znotraj nevronske mreže v fazi učenja ne pojavijo uteži z zasičenimi vrednostmi (0 ali 1). Za rešitev tega problema lahko uporabimo tehniko paketne normalizacije pred ali po aktivaciji posameznega skritega sloja (Géron 2019). Moment je hiperparameter, ki se uporablja za eksponentna premikajoča povprečja. Dodelili smo mu vrednost 0,8, ker je ta v literaturi empirično določena kot boljša za generacijo novih slik z GAN (Purkait 2019). Uporaba aktivacijske funkcije puščajočega ReLU pripomore tudi k temu, da nevroni ne obtičijo v iteracijah (tako imenovani mrtvi nevron). Hiperparameter α puščajočega ReLU določa naklon



Slika 2: Študijsko območje.

Slika 3: Arhitektura generatorja in diskriminatorja. ►

Izdelava visokoločljivega sintetičnega ortofota na osnovi posnetkov Sentinel-2 z uporabo ...



(gradient) »puščanja« za vse uteži manjše od 0 (Géron 2019). Preizkusili smo pet različic modela, katerih parametri so prikazani v preglednici 1.

Preglednica 1: Preizkušeni modeli.

parametri	SRĐAN 1	SRĐAN 2	SRĐAN 3	SRĐAN 4	SRĐAN 5
velikost podvzorca na iteracijo	1	8	10	14	3
število iteracij	30.000	50.000	50.000	100.000	10.000
naključna izbira podvzorca	DA	DA	DA	DA	DA
zrcaljenje 50 % naključnih slik	NE	NE	NE	DA	DA

3 Rezultati

Skladnost rezultatov z izvornimi posnetki Sentinel-2 smo ocenili z indeksom strukturne skladnosti (SSIM; Wang s sodelavci 2004) na testni množici 6073 rastrov (preglednica 2).

Od vseh preizkušenih generatorjev se je najbolj izkazal model SRĐAN 5. Glede na to, da generator preslika en piksel posnetka Sentinel-2 v 16 pikselov sintetičnega ortofota in s tem za štirikrat izboljša prostorsko ločljivost, je srednja vrednost (aritmetična sredina) SSIM 0,62 odlična. SSIM ortofota in izvorne posnetka Sentinel-2 je slabši od vrednosti modela SRĐAN 5, kar je pričakovano, saj rezultati modela SRĐAN 5 izhajajo iz posnetka Sentinel-2 in ne iz ortofota. Cilj spreminjanja parametrov modela je bil dobiti SSIM vrednosti podobne ali višje od SSIM vrednosti ortofota. Pri modelu SRĐAN 5 sta najmanjša tudi standardni odklon in koeficient variacije, kar pomeni, da so vrednosti rezultatov modela SRĐAN 5 najbližje srednji vrednosti. Najvišja minimalna vrednost SSIM kaže, da je njegov najslabši rezultat boljši od ostalih najslabših. Na sliki 4 so prikazani vhodni podatki, rezultat (sintetični ortofoto) in ortofoto, ki ima funkcijo semantične kontrole rezultatov.

Zaradi procesa dekonvolucije (transponirane konvolucije) lahko na rezultatih opazimo močno izraženo teksturo v obliki šahovnice, ki so jo opisali Odena, Dumoulin in Olah (2016).

4 Sklep

SRĐAN je primeren za izboljšavo vseh posnetkov Sentinel-2 severovzhodne Slovenije brez oblakov in snežne odeje posnetih v obdobju snemanja CAS. Izboljšana ni le prostorska ločljivost, temveč tudi barve in tekstura. Rezultati so se izkazali kot izjemno primerni za fotointerpretacijo nekaterih topografskih

Preglednica 2: Statistika indeksov strukturne skladnosti

SSIM	SRĐAN 1	SRĐAN 2	SRĐAN 3	SRĐAN 4	SRĐAN 5	ortofoto
srednja vrednost (aritmetična sredina)	0,59	0,42	0,49	0,28	0,62	0,51
standardni odklon	0,09	0,11	0,10	0,10	0,07	0,08
koeficient variacije	0,15	0,26	0,19	0,37	0,12	0,16
maksimalna vrednost	0,90	0,87	0,90	0,74	0,84	0,96
minimalna vrednost	0,32	0,19	0,26	0,11	0,33	0,13



Slika 4: Primeri vhodnega podatka, rezultata in kontrolnega ortofota.

objektnih tipov, kot so ceste in pozidani objekti. Teksturna izboljšava odlično diskriminira drevesa, grmičevja in trave, kar je osnova za fotointerpretacijo tako gozdov kot ožjih mejic. V nadaljevanju bomo nadgradili SRĐAN model z dodatno generativno nasprotniško mrežo kot na primer CycleGAN (Zhu s sodelavci 2017) ter tako poskusili izločiti efekt šahovnice. Novi SRĐAN bo zgrajen na podlagi ortofoto posnetkov od junija 2016 naprej. Generalirali bomo sintetične ortofote velikosti DTK5 sekcij ter zanje izvedli kontrolo kakovosti v skladu s standardi za preverjanje kakovosti izdelkov CAS (Oven, Tršan in Bric 2019).

Glede na spodbudne rezultate predstavljenega modela, smo mnenja, da je ob ustreznem upoštevanju sprememb v okolju aplikativni potencial sintetičnih ortofotov velik in lahko podkrepi tako posodabljanje različnih prostorskih evidenc v skoraj realnem času (za posnetke Sentinel-2 na 2 do 5 dni) kot tudi odločitvene procese terenskih služb, ki se ukvarjajo s kontrolo okolja.

ZAHVALA: Raziskava je bila opravljena v okviru aplikativnega raziskovalnega projekta L2-1826, ki ga sofinancirajo Javna agencija za raziskovalno dejavnost Republike Slovenije, Geodetska uprava Republike Slovenija in Ministrstvo za obrambo.

5 Viri in literatura

- Ahirwar, K. 2019: Using SRGANs to generate photo-realistic images. Medmrežje: <https://github.com/PacktPublishing/Generative-Adversarial-Networks-Projects/tree/master/Chapter05> (4. 2. 2020).
- Cresson, R. 2019: Enhancement of Sentinel-2 images at 1.5m. IRSTEA. Medmrežje: <https://mdl4eo.irstea.fr/2019/03/29/enhancement-of-sentinel-2-images-at-1-5m/> (4. 2. 2020).
- Dennison, P. E., Roberts, D. A. 2003: Endmember selection for multiple endmember spectral mixture analysis using endmember average RSME. *Remote Sensing of Environment* 87, 2-3. DOI: [https://doi.org/10.1016/S0034-4257\(03\)00135-4](https://doi.org/10.1016/S0034-4257(03)00135-4)
- ESA, 2020. User guides, Sentinel-2 MSI, Revisit and coverage, Sentinel online. Medmrežje: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/revisit-coverage> (20. 4. 2020).
- Géron, A. 2019: *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow*. Sebastopol.
- Goodfellow, I., Bengio, Y., Courville, A. 2016: *Deep learning*. Cambridge, Massachusetts, London.
- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y. 2014: Generative adversarial networks. arXiv:1406.2661 [cs.LG].
- Interpretacijski ključ, 2013. Podroben opis metodologije zajema dejanske rabe kmetijskih in gozdnih zemljišč. Direktorat za kmetijstvo, Služba za register kmetijskih gospodarstev, Ministrstvo za kmetijstvo, gozdarstvo in prehrano. Ljubljana.
- Ledig, C., Lucas, T., Ferenc, H., Jose, C., Andrew, C., Alejandro, A., Andrew, A., Alykhan, T., Johannes, T., Zehan, W., Wenzhe, S. 2016: Photo-realistic single image super-resolution using a generative adversarial network. arXiv 1609.04802 [Cs, Stat].
- MPC Team, 2020. Sentinel-2 L1C data quality report issue 47. ESA. Medmrežje: https://sentinel.esa.int/documents/247904/685211/Sentinel-2_L1C_Data_Quality_Report (3. 2. 2020).
- Nagano, Y., Yohei, K. 2018: SRGAN for super-resolving low-resolution food images. *Proceedings of the Joint Workshop on Multimedia for Cooking and Eating Activities and Multimedia Assisted Dietary Management, CEA/MADiMa '18, Stockholm*. DOI: <https://doi.org/10.1145/3230519.3230587>
- Odena, A., Dumoulin, V., Olah, C. 2016: Deconvolution and checkerboard artifacts. *Distill*. DOI: <https://doi.org/10.23915/distill.00003>
- Oven, K., Tršan, S., Bric, V. 2019: Kontrola kakovosti izvedbe projekta »Ciklično aerofotografiranje Slovenije 2017-19«. Zbirno poročilo celotnega cikla in dopolnitev tehnične dokumentacije za nov razpis CAS 2020-22, Geodetski inštitut Slovenije. Ljubljana.
- Pehani, P. 2020. Osebni vir, 5. 3. 2020.

- Pravilnik o katalogu topografskih podatkov in topografskem ključu. Uradni list Republike Slovenije 54/2014. Ljubljana.
- Purkait, N. 2019: Hands-on Neural Networks with Keras: Design and Create Neural Networks Using Deep Learning and Artificial Intelligence Principles. Birmingham.
- Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., Chen, X. 2016: Improved Techniques for Training GANs. arXiv 1606.03498 [cs.LG].
- Simonyan, K., Zisserman, A. 2015: Very deep convolutional networks for large-scale image recognition. arXiv 1409.1556 [cs.CV].
- Wang, Z., Bovik, A. C., Sheikh, H. R., Simoncelli, E. P. 2004: Image quality assessment: from error visibility to structural similarity. IEEE Transactions on Image Processing 13-4. DOI: <https://doi.org/10.1109/TIP.2003.819861>
- Zhu, J.-Y., Park, T., Isola, P., Efros, A. A. 2018: Unpaired image-to-image translation using cycle-consistent adversarial networks. arXiv:1703.10593 [cs]