

Napovedovanje senčenja v geoprostorskih podatkih z globokim modelom U-Net

Tjaša Repič, Domen Kavran, Domen Mongus, Niko Lukač

Univerza v Mariboru, Fakulteta za elektrotehniko, računalništvo in informatiko,
Koroška cesta 46, 2000, Maribor, Slovenia

E-pošta: { tjasa.repic1@um.si, domen.kavran1@um.si, domen.mongus@um.si, niko.lukac@um.si }

Shadow Prediction in Geospatial Data Using a Deep U-Net Model

The following work presents a deep learning approach for predicting shadow masks using terrain and position of the Sun. By employing a U-Net architecture with an auxiliary branch that encodes solar azimuth and elevation angles, the model can account for illumination geometry. The network is trained using a numerically stable binary cross-entropy loss function, improving convergence and prediction sharpness. Training is supported by a dataset comprised of approximately 200,000 terrain patches, derived from digital surface models (DSMs) and labeled with shadow masks. The dataset reflecting seasonal and diurnal variations in urban settings is used to evaluate the proposed method. Results indicate consistent performance across training and validation, showcasing strong predictive capabilities and stable learning dynamics. The model effectively captures shadow patterns shaped by complex spatial and lighting conditions.

1 Uvod

Zaznavanje senčenih območij je pomembna naloga na področju daljinskega zaznavanja, saj sence pogosto zakrijejo podrobnosti površja. To lahko vpliva na natančnost algoritmov na področjih, kot so 3D rekonstrukcije površja [1], prepoznave majhnih površinskih objektov [2], analize rabe površja [3], ocenjevanje sončnega obsevanja [4, 5, 6] in termalne obremenitve [7]. Oblika in razporeditev senc je odvisna od geometrijskih lastnosti površja in trenutne pozicije sonca na obzorju (višinski in azimutni kot) [8, 5]. Kompleksnejši relief povzroča bolj raznolike sence, medtem ko se s premikanjem sonca skozi dan (in letne čase) spreminja tudi njihova razporeditev na Zemljinem površju. Kombinacija digitalnega modela površja (angl. Digital Surface Model, DSM) in informacij o poziciji Sonca omogoča različnim algoritmom natančno določanje območij v senci [3].

Pri višjih ločljivostih DSM-ja imajo algoritmi simuliranja senc na podlagi metanja žarkov visoko računsko zahtevnost izvajanja, kljub implementacijam na grafičnih procesnih enotah (angl. Graphics Processing Units, GPUs) [5]. Rezultat je 2.5 dimenzionalna maska senčenja za podani DSM. Ena takšnih metod je bila

uporabljena v [6], kjer se je za potrebe pravilne ocenitve solarnega potenciala moralo upoštevati prisotnost senčenih območij. Te so prepoznali z uporabo algoritma za sledenje žarkom (angl. ray-casting). Predlagane so bile tudi različne metode, ki ne slonijo na uporabi DSM ali prostorskih informacij Sonca. Metoda, predlagana v [9], vhodni posnetek površja najprej pretvori iz RGB v HSV barvni prostor. Sledi izračun razmerja med H in V. Za tem sledi sprememba intenzitete izračunanih razmerij s potenčno funkcijo. V zadnjem koraku se s pragovno funkcijo razloči med senčenimi in nesenčenimi območji. Avtorji v [10] predlagajo metodo, ki večspektralni posnetek Zemeljskega površja najprej pretvori v RGB posnetek. Temu sledi izračun izboljšanega indeksa senc (angl. Improved Shadow Index, ISI), ki izhaja iz barvnega prostora YCbCr (pretvorjen iz posnetka RGB) in dodatnega bližnjega-infrardečega sloja (angl. Near-Infrared, NIR). Medtem je posnetek RGB segmentiran z Mean-Shift algoritmom gručenja. Pridobljena segmentacijska maska je nato uporabljena za regijsko rekonstrukcijo ISI. Sledi pragovna funkcija, ki rekonstruiran ISI pretvori v masko senčenih območij.

V zadnjih letih so se predlagale različne metode prepoznave senčenih območij z uporabo globokega učenja, ki skušajo biti hitrejše od simulacijskih metod, hkrati pa doseči visoko natančnost. V [11] so avtorji predlagali metodo, ki v prvem koraku uporablja gručenje nad značilkami pikslov. V drugem koraku se uporabi algoritem za združevanje značilk z upoštevanjem kontekstnih informacij. Te so namenjene za kasnejšo boljso usmeritev modela v prepoznavo senčenih območij ob uporabi pozornosti znotraj predlagane arhitekture globoke nevronske mreže. Nenadzorovano ogrođje za prepoznavo senčenih območij, predlagano v [12], uporablja modul za generiranje psevdo oznak (angl. pseudo label generation, IPG) prisotnosti senc, modul za postopno učenje (angl. curriculum learning, CL) in modul za samostojno učenje (angl. self-training, ST). Modula CL in ST sta naučena različnih oblik senc in skušata izboljšati kakovost generiranih psevdo oznak modula IPG. Zgolj ta je po končanem učenju uporabljen za prepoznavo senčenih območij nad novimi podatki.

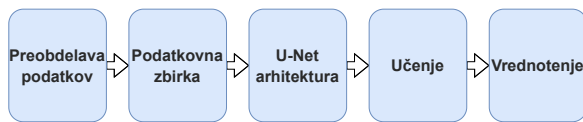
V članku predlagamo novo metodo za prepoznavo senčenih območij z uporabo globoke arhitekture U-Net [13], ki temelji na kombinaciji DSM, izračunanih

geometrijskih značilnostih površja ter pozicije sonca. Novost metode je v uporabi trokanalnega vhodnega tenzorja, kar omogoča bogatejšo reprezentacijo vhodnih podatkov in natančnejšo prostorsko lokalizacijo senc, saj model poleg višin uporablja tudi naklon in orientacijo, zato lažje sklepa o osvetlitvi glede na položaj Sonca in natančnejše napove sence. Razvit model zagotavlja hitro sklepanje in je tudi hitrejši od klasičnega simulacijskega pristopa na podlagi metanja žarkov, hkrati pa ohranja visoko natančnost.

V naslednjem poglavju je predstavljena predlagana metoda. V tretjem poglavju predstavimo rezultate eksperimentov in ključne ugotovitve, v četrtem pa zaključimo članek.

2 Metodologija

Na sliki 1 so prikazani vsi koraki predlaganega metodološkega poteka, ki zajemajo pripravo vhodnih podatkov, oblikovanje podatkovne zbirke in pripravo nalagalnika, implementacijo arhitekture U-Net, učenje modela ter končno vrednotenje rezultatov.



Slika 1: Delovni tok modela U-Net za napoved senčenja.

2.1 Predobdelava podatkov

Za učenje modela uporabimo večslojne podatke, ki kot celota omogočajo napovedovanje senčnih območij na površju. Osnovni sloj predstavlja DSM, definiran kot G , ki vsebuje podatke o nadmorski višini. G je predstavljen kot raster velikosti $G_w \times G_h$ (tj. širina in višina) enakoraznih celic določene prostorske ločljivosti G_{res} (npr. $G_{res} = 1$ pri 1 m širine celice). V fazi predobdelave ustvarimo dodatna sloja naklona β in orientacije γ površja za vsako celico v G , na podlagi izračuna gradienta:

$$\beta = \arctan \left(\sqrt{\left(\frac{\partial G}{\partial x} \right)^2 + \left(\frac{\partial G}{\partial y} \right)^2} \right), \quad (1)$$

$$\gamma = \arctan 2 \left(-\frac{\partial G}{\partial x}, \frac{\partial G}{\partial y} \right), \quad (2)$$

$$\frac{\partial G}{\partial x} \approx \frac{G(x+1, y) - G(x-1, y)}{2G_{res}}, \quad (3)$$

$$\frac{\partial G}{\partial y} \approx \frac{G(x, y+1) - G(x, y-1)}{2G_{res}}, \quad (4)$$

kjer sta x in y prostorski koordinati izbrane celice [4]. Vključeni so tudi podatki o legi Sonca za vsako uro v letu, podani v obliki sončnih kotov (višina in azimut), katere vrednosti so izračunane na podlagi algoritma SPA (angl. Solar Position Algorithm, SPA) [8]. Ti podatki so ključni za simulacijo osvetlitve površja v različnih

obdobjih leta. Kot ciljne vrednosti se uporabljajo binarne maske senčenja, ki so bile pridobljene na podlagi obstoječega dela avtorjev [5].

V fazi predobdelave so vhodni sloji združeni v enoten vhodni tenzor oblike [3, G_H , G_W], pri čemer vsak izmed treh kanalov predstavlja en sloj prostorskih podatkov. Vsi sloji so med seboj usklajeni glede na prostorsko poravnavo in dimenzije. Celoten vhodni prostor se nato razdeli na manjše odseke dimenzij 256×256 pikselov. Vsakemu odseku se glede na časovni žig pripne ustrezna senčna maska, ki prikazuje, kateri deli površja so v senci glede na trenutni položaj Sonca. Celoten nabor podatkov je razdeljen na tri podmnožice: učno, validacijsko in testno, in sicer v razmerju 60:20:20. Med vsako epoko učenja se podatki naključno premešajo z namenom zmanjšanja tveganja za prenaučnost.

2.2 Napovedni model globokega učenja U-Net

U-Net predstavlja eno izmed najbolj uveljavljenih arhitektur na področju semantične segmentacije slik, zlasti tam, kjer je potrebna visoka prostorska natančnost. Prvotno je bil zasnovan za segmentacijo biomedicinskih slik, [13] zaradi učinkovitosti pri sočasnem zajemanju globalnih vsebinskih vzorcev in lokalnih prostorskih značilnosti pa se je arhitektura U-Net uveljavila tudi na področjih, kot so daljinsko zaznavanje, obdelava satelitskih slik in prostorsko modeliranje v okviru geografskih informacijskih sistemov [14].

Arhitektura U-Net, uporabljena pri predlagani metodi, sprejme 3-kanalni vhodni tenzor, kjer vsak kanal predstavlja sloj prostorskih podatkov površja. Ti skupaj tvorijo vhodne informacije, na podlagi katerih model predvidi senčne vzorce. Poleg prostorskih podatkov arhitektura vključuje tudi pomožni vhod z relativno sferično pozicijo Sonca. Vrednosti se obdelajo in vgradijo v globlje sloje mreže, kar omogoča boljše prilagoditev napovedi osvetlitvenim pogojem. Mreža ima simetrično zgradbo s padajočim delom (kodirnik) in naraščajočim delom (dekodirnik), ki sestavi izhodno sliko. Končni rezultat je 1-kanalni izhod, ki za vsak piksel poda verjetnost pripadnosti senci. Za binarno klasifikacijo vrednosti pragovno razmejimo pri 0,5, pri čemer vrednost 1 predstavlja osvetljeno, vrednost 0 pa senčeno območje.

2.3 Postopek učenja in izgubna funkcija

Postopek učenja modela temelji na postopnem prilagajanju uteži z metodo nadzorovanega učenja, kjer se model uči napovedovati binarne maske senčenja na podlagi znanih vhodno-izhodnih parov. V fazi učenja poseben poudarek namenjamo izbiri in uporabi ustrezne izgubne funkcije. Uporabili smo binarno navzkrižno-entropijsko izgubno funkcijo (angl. binary cross-entropy) [15]. Izguba se izračuna kot:

$$\mathcal{L}(s, t) = \log(1 + e^{-|s|}) + \max(0, s) - s \cdot t, \quad (5)$$

kjer s predstavlja napoved modela, $t \in \{0, 1\}$ pa ciljno vrednost, ali gre za senčno ali osvetljeno območje.

V vsakem učnem koraku model izvede naprej usmerjen prehod ter iz vhodnih podatkov napove senčno masko, ki jo primerja s ciljno. Napaka se izračuna z izgubno funkcijo, na podlagi katere gradienti posodobijo uteži. Optimizacija je izvedena z algoritmom Adam [16] z vnaprej določeno hitrostjo učenja. Po vsakem prehodu skozi učni nabor model ocenimo na validacijski množici z metrikama F1 in razmerjem preseka skozi unijo (angl. Intersection over Union, IoU). IoU meri prekrivanje med napovedano in referenčno binarno masko. Metrika F1 pa združuje informaciji o natančnosti in občutljivosti, da omogoča uravnoteženo oceno modela tudi v primerih neuravnoteženih podatkov. Za preprečevanje prenašenosti se uporabi zgodnje ustavljanje, ki prekine učenje ob stagnaciji metrik. Končno vrednotenje poteka na testni množici, ločeni od učenja in validacije. S tem zagotovimo objektivno oceno sposobnosti modela za generalizacijo. Hiperparametri, izgube in metrike so prikazane v naslednjem poglavju.

3 Rezultati

V eksperimentih smo ocenili točnost naučenega napovednega modela v različnih svetlobnih pogojih za vsako uro v letu. Pri tem smo za vhod uporabili podatke LiDAR dela mesta Maribor, prostorske velikosti $750\text{ m} \times 1000\text{ m}$, pri čemer smo nastavili $G_{res} = 1\text{ m}^2$. Eksperimenti so bili izvedeni na sistemu z 13th Gen Intel Core i9-13900K 24-jedrnim procesorjem, grafično kartico NVIDIA GeForce RTX 5080, s 16 GB pomnilnika in 128 GB delovnega pomnilnika.

V tabeli 1 so predstavljeni ključni hiperparametri učenja ter časovna zahtevnost simulacije in učenja pri različnih velikostih ploščic, skupaj s povprečnim časom napovedi.

Parameter	Vrednost
Povprečno število epoh:	15
Velikost serije (angl. batch size):	32
Hitrost učenja (angl. learning rate):	10^{-4}
Optimizacijski algoritem:	Adam
Povprečen čas simulacije [ms]:	300
Povprečen čas napovedi (128 x 128) [ms]:	238
Povprečen čas napovedi (256 x 256) [ms]:	237
Povprečen čas napovedi (512 x 512) [ms]:	243
Povprečen čas učenja (128 x 128) [h]:	5
Povprečen čas učenja (256 x 256) [h]:	4
Povprečen čas učenja (512 x 512) [h]:	75

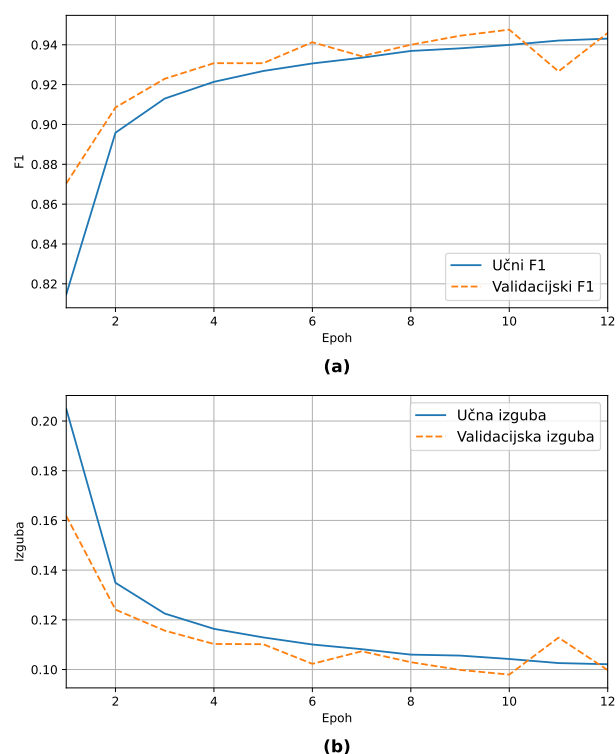
Tabela 1: Učni parametri in časovna zahtevnost napovednega modela.

Rezultati točnosti modela, pridobljeni z uporabo metrik IoU in F1, so prikazani v tabeli 2 in prikazujejo visoko skladnost med napovedanimi in referenčnimi senčnimi maskami. Za primer velikosti ploščic 256×256 je na validacijski (IoU: 90,1%, F1: 95,4%) in testni množici (IoU: 90,4%, F1: 95,9%) model dosegel visoke vrednosti, kar potrjuje njegovo zmožnost generalizacije. V tabeli so prav tako prikazani rezultati točnosti za ploščice velikosti 128×128 in 512×512 .

Množica:	IoU [%]:	F1 [%]:
Validacijska (128 x 128):	91,8	95,7
Validacijska (256 x 256):	90,1	95,4
Validacijska (512 x 512):	84,8	91,7
Testna (128 x 128):	92,1	96,2
Testna (256 x 256):	90,4	95,9
Testna (512 x 512):	84,2	92,3

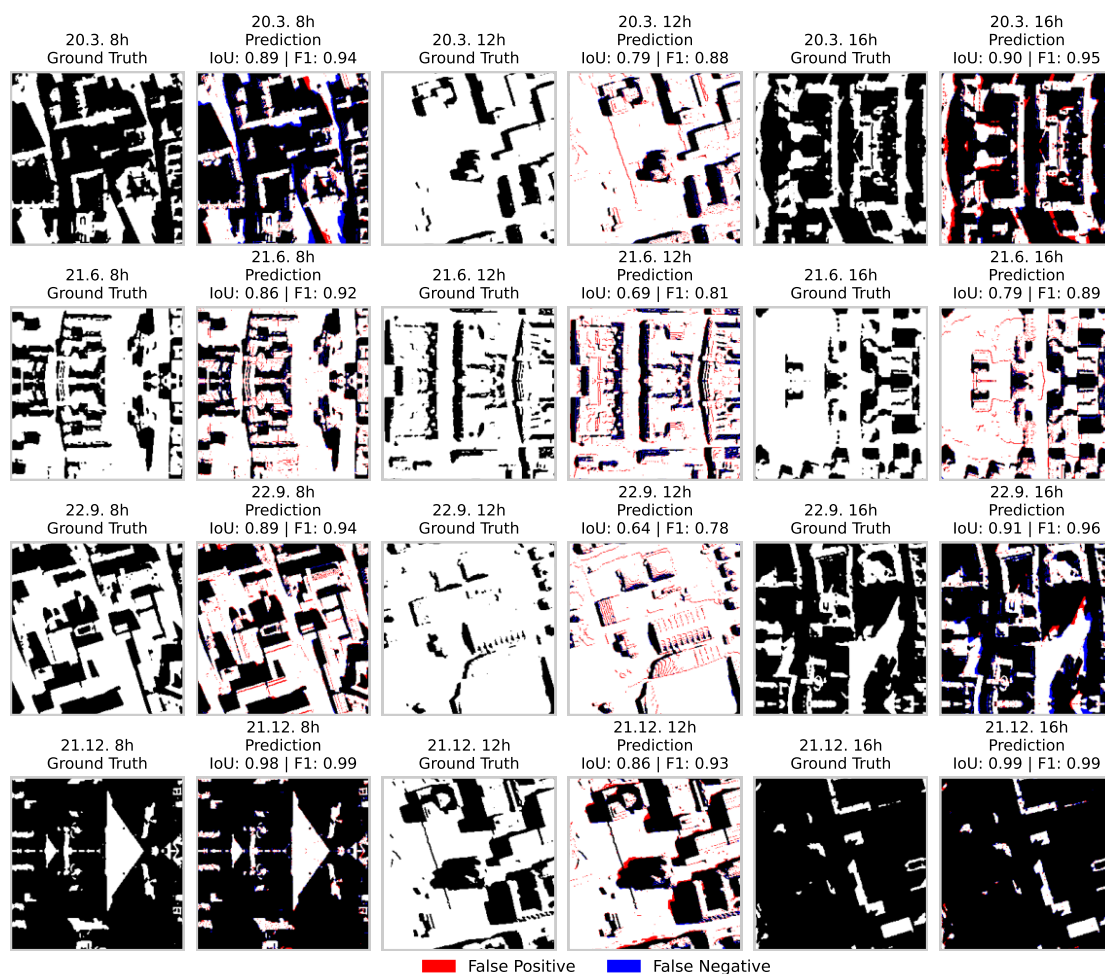
Tabela 2: Primerjava metrik uspešnosti na vseh množicah.

Na sliki 2 je na grafu (a) prikazan rezultat metrike F1 v odvisnosti od števila epoh za učno in validacijsko množico pri velikosti ploščic 256×256 . F1 hitro naraste in po peti epohi preseže rezultat 0.9. Krivulji ostajata ves čas blizu, visoka končna vrednost validacijskega F1 (95,4 %) pa nakazuje robustnost modela pri detekciji senčnih območij. Na grafu (b) je prikazan potek izgube za učno in validacijsko množico za identično velikost. Obe krivulji se enakomerno zmanjšujeta brez znakov divergence.



Slika 2: Metrika F1 ter padec izgube skozi epohe pri velikosti ploščic 256×256 .

Za namen vizualizacije zmogljivosti modela smo izbrali reprezentativne časovne točke, ki predstavljajo pomembne mejnike v Sončnem ciklu in sicer pomladni in jesenski ekvinokcij (20. marec in 22. september) ter poletni in zimski solsticij (21. junij in 21. december). Vizualna ponazoritev teh ocen je predstavljena na sliki 3.



Slika 3: Primerjava med dejanskimi in napovedanimi senčnimi maskami za štiri dni v letu ob različnih urah pri velikosti ploščic 256 x 256.

Ob ekvinokciju je trajanje dneva in noči izenačeno, Sonce pa se nahaja neposredno nad ekvatorjem. Poletni solsticij zaznamuje najdaljši dan v letu, ko Sonce doseže najvišjo točko na nebu. Zimski solsticij pa prinaša najnižji položaj Sonca v letu. Z izbiro teh dni smo omogočili uravnoteženo oceno robustnosti modela v raznolikih sezonskih pogojih, ki vplivajo na geometrijo osvetlitve ter hkrati na kompleksnost zaznavanja senc.

Model dosega najvišjo natančnost v zimskem času ter v jutranjih in popoldanskih urah, ko so sence izrazitejšje zaradi nizkega položaja Sonca. Najboljši rezultati so bili doseženi 21. decembra ob 8. in 16. uri, kjer je bila skladnost med napovedjo in dejanskimi podatki skoraj popolna (IoU 98%, F1 99%). Visoko natančnost model ohrani tudi v jesenskem in spomladanskem ekvinokciju ob 16. uri, z vrednostmi F1 nad 94%. Nasprotno pa model izkazuje slabšo zmogljivost opoldne, zlasti 22. septembra ob 12. uri (IoU 64%, F1 78%) in 21. junija ob 12. uri (IoU 69%, F1 81%), kar sovпада s krajšimi in manj kontrastnimi sencami zaradi visokega Sončevega položaja. Rezultati kažejo, da ima geometrija osvetlitve ključno vlogo pri natančnosti napovedi, saj model dosega najboljše rezultate v razmerah, kjer so sence daljše in izrazitejše.

4 Zaključek

V članku smo predstavili pristop za napoved senčnih območij, ki temelji na vhodnih slikah površja in informacijah o sončni legi. Uporabili smo pristop globokega učenja arhitekture U-Net, ki smo jo razširili z dodatnim slojem za vnos globalnih podatkov o azimutu in višini sonca. Model smo učili na podatkovnem naboru urbanega okolja. Rezultati potrjujejo učinkovitost predlaganega modela pri zaznavi senčnih območij.

Na testnem naboru je model dosegel F1 rezultat 94,9% in IoU 90,4 %, s čimer potrjujemo njegovo sposobnost prilagajanja modela za dano lokacijo. Analiza učnega procesa je potrdila, da model konvergira stabilno, kar je razvidno iz enakomernega padca funkcije izgube ter hitrega izboljšanja metrike F1 v začetnih epohah. Vizualni rezultati napovedi dodatno kažejo, da model zanesljivo zaznava vzorce osenčenosti v različnih svetlobnih pogojih.

Kljub temu smo v eksperimentih izvedli delitev podatkov po času, ne pa strogo po prostoru, kar pomeni, da se lahko iste prostorske ploščice pojavijo v različnih sklopih ob različnih urah in dnevih. Ta nastavek pokaže, da se model dobro posploši na različne

osvetlitvene pogoje za isto lokacijo. Vendar to pomeni tudi možnost prostorskega uhajanja informacij (angl. spatial leakage).

V nadaljnjem delu načrtujemo razširitev metode z vključitvijo fizikalno informiranih nevronske mreže (angl. physics-informed neural networks). Metodo nameravamo dodatno preizkusiti na širšem naboru raznovrstnih urbanih območij, da bi celoviteje ocenili njeno robustnost pri generalizaciji na različne tipologije urbanega površja in svetlobnih pogojev. Poleg tega bomo rezultate sistematično primerjali z obstoječo literaturo na referenčnih metodah ter validacijo razširili z ujemanjem napovedanih senc z ortofoto posnetki, ki sence že vključujejo, kar nam bi omogočilo neposredno časovno usklajeno primerjavo.

Zahvala

Raziskovalno delo je bilo sofinancirano s strani Javne agencije za znanstvenoraziskovalno in inovacijsko dejavnost Republike Slovenije, v okviru temeljnega projekta št. J7-50095 in raziskovalnega programa št. P2-0041. Za podatke LiDAR se zahvaljujemo Agenciji Republike Slovenije za okolje.

Literatura

- [1] Q. Wang, L. Yan, Anisotropic Scattering Shadow Compensation Method for Remote Sensing Image with Consideration of Terrain, *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (2016) 401–405doi:10.5194/ISPRS-ARCHIVES-XLI-B1-401-2016.
- [2] R. Sundberg, The Impact of Cloud Shadows on Subpixel Target Detection, 2018, pp. 1–5. doi:10.1109/WHISPERS.2018.8747206.
- [3] X. Dong, J. Cao, W. Z. and, A review of research on remote sensing images shadow detection and application to building extraction, *European Journal of Remote Sensing* 57 (1) (2024) 2293163. arXiv:https://doi.org/10.1080/22797254.2023.2293163, doi:10.1080/22797254.2023.2293163.
- [4] N. Lukač, D. Žlaus, S. Seme, B. Žalik, G. Štumberger, Rating of roofs' surfaces regarding their solar potential and suitability for PV systems, based on LiDAR data, *Applied Energy* 102 (2013) 803–812. doi:https://doi.org/10.1016/j.apenergy.2012.08.042.
- [5] N. Lukač, D. Mongus, B. Žalik, G. Štumberger, M. Bizjak, Novel gpu-accelerated high-resolution solar potential estimation in urban areas by using a modified diffuse irradiance model, *Applied Energy* 353 (2024) 122129. doi:https://doi.org/10.1016/j.apenergy.2023.122129.
- [6] M. Žalik, D. Mongus, N. Lukač, High-resolution spatiotemporal assessment of solar potential from remote sensing data using deep learning, *Renewable Energy* 222 (2024) 119868.
- [7] M. Bizjak, B. Žalik, G. Štumberger, N. Lukač, Large-scale estimation of buildings' thermal load using lidar data, *Energy and Buildings* 231 (2021) 110626.
- [8] I. Reda, A. Andreas, Corrigendum to "Solar position algorithm for solar radiation applications" [*Solar Energy* 76 (2004) 577–589], *Solar Energy* 81 (6) (2007) 838. doi:10.1016/j.solener.2007.01.003.
- [9] S. Musleh, M. Sarfraz, H. Raafat, Shadow detection of the satellite images of earth using ratio image pixels, *Earth Science Informatics* 14 (2021) 1–16. doi:10.1007/s12145-020-00537-7.
- [10] T. Zhou, H. Fu, C. Sun, S. Wang, Shadow detection and compensation from remote sensing images under complex urban conditions, *Remote Sensing* 13 (4) (2021). doi:10.3390/rs13040699.
- [11] Y. Xie, D. Feng, X. Shen, Y. Liu, J. Zhu, T. Hussain, S. W. Baik, Clustering feature constraint multiscale attention network for shadow extraction from remote sensing images, *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022) 1–14. doi:10.1109/TGRS.2022.3151901.
- [12] Q. Zhang, H. Guo, G. Li, T. Zhang, Q. Jiao, Deep unsupervised shadow detection with curriculum learning and self-training, *Computer Vision and Image Understanding* 248 (2024) 104124. doi:https://doi.org/10.1016/j.cviu.2024.104124.
- [13] O. Ronneberger, P. Fischer, T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation, in: N. Navab, J. Hornegger, W. M. Wells, A. F. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, Vol. 9351, Springer International Publishing, Cham, 2015, pp. 234–241, series Title: Lecture Notes in Computer Science.
- [14] X. He, Y. Zhou, J. Zhao, D. Zhang, R. Yao, Y. Xue, Swin Transformer Embedding UNet for Remote Sensing Image Semantic Segmentation, *IEEE Transactions on Geoscience and Remote Sensing* 60 (2022) 1–15. doi:10.1109/TGRS.2022.3144165.
- [15] A. Painsky, G. Wornell, On the Universality of the Logistic Loss Function, in: 2018 IEEE International Symposium on Information Theory (ISIT), 2018, pp. 936–940. doi:10.1109/ISIT.2018.8437786.
- [16] R. N. Singarimbun, E. B. Nababan, O. S. Sitompul, Adaptive moment estimation to minimize square error in backpropagation algorithm, in: 2019 International Conference of Computer Science and Information Technology (ICoSNiKOM), 2019, pp. 1–7. doi:10.1109/ICoSNiKOM48755.2019.9111563.