

Primena metoda mašinskog učenja pri klasifikaciji satelitskih snimaka

MARTINA T. ČURLIN, Vojnogeografski institut

„General Stevan Bošković“, Beograd

SINIŠA M. DROBNJAK, Vojnogeografski institut

„General Stevan Bošković“, Beograd

IVAN M. POTIĆ, Vojnogeografski institut

„General Stevan Bošković“, Beograd

DEJAN R. ĐORĐEVIĆ, Vojnogeografski institut

„General Stevan Bošković“, Beograd

SASA T. BAKRAČ, Vojnogeografski institut

„General Stevan Bošković“, Beograd

Originalni naučni rad

UDC: 528.8:004.85

DOI: 10.5937/tehnika2402152C

Mašinsko učenje, kao specifična oblast veštačke inteligencije, otvara nove horizonte za teorijska i eksperimentalna istraživanja u domenu daljinske detekcije, posebno u kontekstu klasifikacije satelitskih snimaka. Ovaj rad se fokusira na primenu metoda mašinskog učenja konkretno slučajnih šuma i potpornih vektora, u cilju klasifikacije satelitskih snimaka. Analiza se izvodi koristeći program SAGA GIS na satelitskim snimcima LANDSAT 8 OLI nivoa 2A. Klasifikacija satelitskih snimaka obuhvata dve osnovne grupe kompjuterskih operacija: nenadgledanu (automatsku ili formalnu) i nadgledanu (polu-automatsku ili logičku) klasifikaciju. U ovom istraživanju, praktična klasifikacija satelitskih snimaka izvedena je primenom dve metode mašinskog učenja, slučajnih šuma i potpornih vektora. Rezultati pokazuju da se dobijeni klasifikovani rasteri ne samo poklapaju, već i u potpunosti zamenjuju dosadašnje metode klasifikacije i identifikacije geoprostornih objekata. Ovo istraživanje, stoga, doprinosi značajnom unapređenju u načinu prikupljanja i analize geoprostornih podataka.

Ključne reči: *daljinska detekcija, veštačka inteligencija, klasifikacija satelitskih snimaka, obrada podataka*

1. UVOD

Veštačka inteligencija (VI) predstavlja sofisticiranu disciplinu unutar multidisciplinarne oblasti koja se bavi simulacijom ljudskih kognitivnih funkcija u mašinama. Ove funkcije obuhvataju sposobnost učenja, zaključivanja, rešavanja problema, percepcije i jezičkog razumevanja [1].

VI predstavlja metodologiju donošenja odluka i akcija baziranih na logički izvedenim zaključcima, bez intervencije ljudskog faktora. Ovde, proces razmišljanja i realizacije određenih aktivnosti nije u domenu ljudskih sposobnosti, već je to domen u kojem dominiraju mašine, shvaćene u naj širem značenju tog pojma.

Pod pojmom mašina podrazumevaju se entiteti koji nezavisno izvršavaju kompleksne zadatke. To može uključivati tradicionalno shvaćene mašine koje su opremljene naprednim softverom za vođenje, ali i softverske programe bez materijalne forme vidljive ljudskom oku.

Ove entitete ne karakteriše fiksna fizička lokacija; umesto toga, njihovo delovanje je rasprostranjeno preko mnoštva virtualnih adresa. One imaju sposobnost da simultano obavljaju različite zadatke na mnogobrojnim geografskim lokacijama širom planete, pa čak i u njenoj orbiti [2].

Mašinsko učenje (MU), kao podoblast VI, fokusira se na razvoj algoritama koji omogućavaju mašinama da uče iz podataka i donose odluke ili predviđanja [3]. Iako su prvi teorijski modeli nastali u drugoj polovini 20. veka, MU je doživelo ekspanziju sa pojavom velikih skupova podataka i napretkom računarske moći [4]. MU predstavlja specijalizovanu granu VI koja se fokusira na kreiranje i razvoj adaptivnih računarskih

Adresa autora: Martina Čurlin, Vojnogeografski institut „General Stevan Bošković“, Beograd, Mije Kovačevića 5

e-mail: martinacurlin1999@gmail.com

Rad primljen: 07.12.2023.

Rad prihvaćen: 29.01.2024.

sistema sposobnih za unapređenje sopstvenih performansi kroz obradu i primenu znanja stečenog iz prethodnih iskustava. Ova disciplina se u širem smislu bavi istraživanjem metodologija generalizacije, kao i formulisanjem i analiziranjem algoritama koji omogućavaju ovaj proces. Počeci teorijskog razmatranja MU datiraju iz kasnih 60-ih godina 20. veka, sa radovima naučnika poput Golda, dok su čvrste teorijske osnove počele da se formiraju u toku 80-ih. U domenu MU, neki od najznačajnijih teorijskih pristupa uključuju Goldov model graničnog učenja (eng. learning in the limit), Valiantov model (eng. Probably Approximately Correct - PAC), i široko priznatu statističku teoriju učenja koja se smatra jednom od najkompletnijih [5].

U mnogim praktičnim situacijama, postoji izazov u preciznom definisanju pravila koja bi rešila određeni problem, kao i u karakterizaciji samog rešenja. Uzimimo za primer proces prepoznavanja lica; nije jednostavno jasno definisati šta zapravo čini lice. Pokušaji da se lice opiše kroz karakteristike njegovih pojedinačnih delova neizbežno vode ka složenom nizu pitanja koja se tiču ne samo definicija tih konkretnih delova, već i razumevanja kako se ti delovi međusobno povezuju i interaguju [6]. Iako je očigledno da postoji određena pravilnost u procesima poput prepoznavanja lica, kompleksnost i prisustvo izuzetaka često čine formalno definisanje i precizno opisivanje pravila izuzetno izazovnim. U kontekstu prevođenja rečenica s jednog jezika na drugi, situacija postaje još složenija. Iako jezici neosporno imaju svoje bogate i složene strukture, one su prepune nijansi i izuzetaka koji onemogućavaju jednostavno i precizno formalno definisanje pravila. Dodatni sloj kompleksnosti se pojavljuje kada je potrebno razumeti i adekvatno prevesti strukturu i suštinu jednog jezika na potpuno drugi jezik, što zahteva ne samo razumevanje gramatičkih i sintaksičkih pravila oba jezika, već i duboko razumevanje kulturnih, kontekstualnih i stilističkih nijansi koji čine jezik živim i dinamičnim entitetom [7].

Da bi se problem uspešno rešio, ključno je imati jasnu specifikaciju tog problema. Međutim, u situacijama poput onih opisanih u prethodnim primerima, pružanje eksplicitne i precizne specifikacije može biti izuzetno teško. U takvim okolnostima, specifikacija problema često se definiše implicitno, putem velike količine podataka koji sadrže primere rešenih i nerešenih problema. Ova vrsta specifikacije, iako možda ne pruža formalnu preciznost, može biti izuzetno korisna u praksi. Ona omogućava sistemima, poput onih zasnovanih na MU, da uče i prilagođavaju se na osnovu stvarnih podataka, što često predstavlja najefikasniji pristup kada je formalno definisanje problema previše složeno ili nepraktično [2]. MU se fokusira na adresiranje problema za koje je izazovno kreirati formalnu specifikaciju, ali za koje postoji obilje podataka

koji na praktičan način prikazuju različite aspekte i nivoe problema. U jezgru metoda MU leži zadatak prepoznavanja i ekstrakcije relevantnih obrazaca i pravilnosti iz dostupnih podataka, i njihove transformacije u odgovarajući matematički model. Osnovni cilj metoda MU nije samo kreiranje modela koji precizno opisuje postojeće podatke, već i postizanje visokog stepena generalizacije. To znači da model treba da pokazuje dobre performanse i u slučajevima ili situacijama koje nisu direktno obuhvaćene podacima korišćenim za njegovo treniranje. Suštinski, cilj je razviti model koji će efikasno donositi pouzdane zaključke i pružiti rešenja za buduće instance problema, time demonstrirajući pravu vrednost i primenljivost u praksi [8–10]. Greške u modelovanju podataka su sastavni deo MU, bilo tokom treniranja ili kasnije primene modela. Posebno su bitne greške koje se javljaju prilikom primene modela, jer je glavni cilj MU razviti modele koji efikasno rešavaju buduće probleme. Stoga je važno ne samo identifikovati i korigovati greške tokom treniranja, već i neprestano unapređivati model kako bi precizno funkcionisao u različitim i nepredvidivim situacijama [11].

Podaci korišćeni u MU mogu biti neprecizni ili nepotpuni, a ponekad sadrže i protivrečnosti. Zbog toga tradicionalna logika nije adekvatan okvir za MU, koje se oslanja na induktivne, a ne deduktivne metode, često koristeći verovatnoću i statistiku. Ovaj pristup omogućava robusnost u odnosu na nesavršenosti podataka i omogućava kvantifikaciju pouzdanosti rešenja. Iako algoritmi MU nemaju garanciju potpunosti i ne postoje apsolutne garancije u pogledu performansi algoritama naučenih kroz MU, ova tehnika često postiže izvanredne rezultate u mnogim oblastima, značajno nadmašujući tradicionalne pristupe uprkos inherentnim ograničenjima i mogućim greškama [12].

MU metode omogućavaju brzo donošenje odluka, daleko brže u poređenju sa metodama automatskog rasuđivanja koje, iako precizne, zahtevaju znatnu računsku snagu. Zbog ove efikasnosti, MU se često preferira čak i u situacijama gde je moguće precizno definisati problem, ali gde je prostor pretrage izuzetno velik. Međutim, uprkos brzini i efikasnosti, rešenja zasnovana na MU nisu imuna na greške. Stoga, upotreba metoda MU može biti neprimerena u kritičnim domenima gde greške nisu dopuštene, kao što su automatska kontrola metro sistema, kontrola avio-saobraćaja i slične oblasti gde je potrebna maksimalna pouzdanost [13].

Proces rešavanja problema pomoću MU uključuje sledeće osnovne korake [2]:

- Modelovanje problema;
- Rešavanje problema, odnosno treniranje modela;
- Evaluacija dobijenog rešenja, odnosno modela.

2. VRSTE KLASIFIKACIJE I PRIMENA KLASIFIKACIONIH METODA MAŠINSKOG UČENJA KOJE SE KORISTI U DALJINSKOJ DETEKCIJI

Tokom digitalne obrade satelitskih snimaka jedan od postupaka, pored rektifikacije i korekcije, poboljšavanja slike, povezivanja podataka i uklapanja u GIS, jeste klasifikacija. Definisati željeni tematski sadržaj je cilj interpretacije satelitskog snimka (kao što je: vrste i način korišćenja zemljišta, vrste vegetacije, hidrografski sadržaj i slično). Suština digitalne obrade satelitskog snimka je klasifikacija snimka. Svrha svake klasifikacije leži u identifikovanju i vizualizaciji, odnosno u preciznom kategorizovanju ili sistematskom raspoređivanju piksela u jednu ili više grupa (klasa), čime se pikseli povezuju sa specifičnim prostornim entitetima koje te klase reprezentuju [14].

2.1. Vrste klasifikacije

Postoje dva pristupa u klasifikaciji satelitskih snimaka. To su automatska i poluautomatska klasifikacija. Automatska klasifikacija se zasniva na primeni posebnih kompjuterskih operacija koje upravljaju procesom izdvajanja područja snimka sa različitim osobinama. Klasifikacija se obavlja identifikacijom svakog piksela ponaosob i njihovim pridruživanjem određenoj klasi. Automatska klasifikacija se uglavnom koristi za brzu, preliminarnu analizu snimka. Automatska klasifikacija je razvijena za primenu u okviru tradicionalnih multispektralnih snimaka, ali je značajno da se istakne kako većina softverskih paketa podržava učitavanje neograničenog broja rasterskih slojeva (kanala), čime se proširuje njena primenljivost i fleksibilnost [15].

Poluautomatska klasifikacija se bazira na primeni predefinisanih instrukcija baziranih na specijalno kreiranom „ključu“ koji se koristi za analizu i interpretaciju multispektralnih snimaka omogućavajući time kontrolisanu i ciljanu obradu podataka. Utvrđivanje identiteta i prostornog položaja područja određenih svojstava određuje se terenskim radom, analizom topografskih i drugih karata i avio snimaka, ili na osnovu drugih pouzdanih podataka. Identifikacija regiona od značaja za istraživanje se realizuje putem vizuelnog određivanja njihove lokacije na snimcima. Ova odabrana područja poseduju posebne spektralne karakteristike, što omogućava formiranje određenih tematskih grupa ili klasa koje se sastoje od podataka sa sličnim atributima [16].

Drugi segment poluautomatske klasifikacije podrazumeva detaljnu analizu vrednosti piksela određenih kategorija putem kompjuterske obrade, koristeći se za to različitim statističkim metodama, uključujući analizu srednjih vrednosti, standardnih

devijacija itd. Ova metoda omogućava evaluaciju i kategorizaciju svakog piksela, dodeljujući ga specifičnoj klasi na osnovu njegove pripadnosti, usklađeno sa prethodno postavljenim kriterijumima za klasifikaciju. Postupak se završava procenom kvaliteta rezultata klasifikacije. Ako se utvrdi da rezultati nisu na zadovoljavajućem nivou, preduzimaju se određene korektivne radnje, koje najčešće uključuju finije podešavanje granica između klasa, posebno u situacijama gde dolazi do značajnog preklapanja piksela sa sličnim vrednostima iz različitih klasa [17].

2.2. Primena klasifikacionih metoda i rezultati klasifikacije

Primenjene metode MU za klasifikaciju satelitskih snimaka u ovom radu su metoda slučajnih šuma i metoda potpornih vektora [18] i predstavljaju primer poluautomatske klasifikacije.

Kao ulazni podatak u procesu klasifikacije satelitskih snimaka metodama mašinskog učenja u našem slučaju korišćen je snimak satelitskog sistema za konstantno osmatranje LANDSAT 8 OLI.

Površina koju zahvataju LANDSAT snimci iznosi 120x120 km². Za potrebe klasifikacije ovog satelitskog snimka metodama mašinskog učenja korišćen je već pripremljen satelitski snimak isečen tačno da zahvata područje jednog lista topografske karte razmere 1:25 000 (TK25).

U zavisnosti od potreba korisnika LANDSAT snimci se isporučuju na različitim nivoima obrade (1A, 1V i 2A) [19]. Snimak je dostavljen u formatu nivoa 2A, što implicira da je već izvršena osnovna obrada, uključujući radiometrijsku korekciju za poboljšanje kvaliteta slike, georeferenciranje u UTM WGS 84 koordinatni sistem bez korišćenja kontrolnih tačaka, kao i rektifikaciju snimka uzimajući u obzir srednju visinu terena na sceni. Položajna tačnost scene je 350 m. Zbog toga je prvo urađeno georeferenciranje snimka u Digitalnom modelu terena (eng. Digital Terrain Models - DTM) biranjem karakterističnih tačaka za koje su koordinate očitavane sa topografskih karata razmere 1:25 000 (TK25) i topografskih karata razmere 1:50 000 (TK50). Srednja kvadratna greška je imala vrednosti od 0,21 do 1,28, odnosno 21 tačka je imala vrednost ispod vrednosti 1,0, a osam preko 1,0 po pikselu (srednja kvadratna greška ne bi trebalo da bude veća od vrednosti jednog piksela) [19].

Za potrebe klasifikacije, korišćeno je šest spektralnih kanala prostorne rezolucije 30 m. Iz vidljivog dela spektra je korišćeno tri kanala: plavi, zeleni i crveni (eng. Blue, Green, Red – BGR), jedan kanal iz bliskog infracrvenog dela spektra (eng. Near Infra Red – NIR) i dva kanala iz srednjeg infracrvenog dela spektra (eng. Short Wave Infra Red – SWIR) (tabela 1).

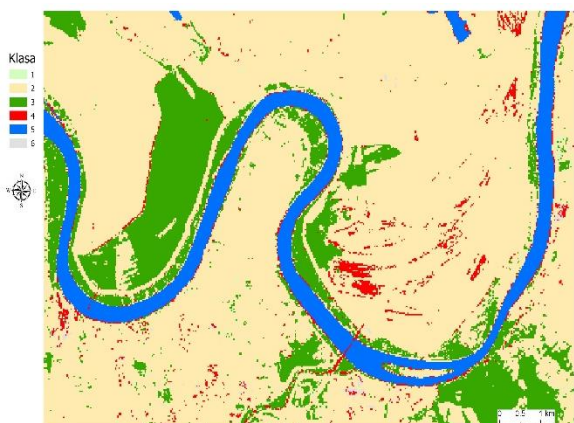
Tabela 1. Kanali korišćeni u klasifikaciji i njihove karakteristike [19].

Broj kanala	Talasne dužine	Područje u spektru
Kanal 2	0,45 - 0,51 μm	Plavo
Kanal 3	0,53 - 0,59 μm	Zeleno
Kanal 4	0,64 - 0,67 μm	Crveno
Kanal 5	0,85 - 0,88 μm	Blisko infracrveno
Kanal 6	1,57 - 1,65 μm	Srednje infracrveno
Kanal 7	2,11 - 2,29 μm	Srednje infracrveno

Praktična klasifikacija predmetnog snimka, metodom mašinskog učenja, urađena je u programu SAGA GIS [20]. Inicijalno je određeno 6 klasa koje će biti klasifikovane: 1-pašnjaci, 2-poljoprivreda, 3-šume, 4-urbana sredina, 5-voda i 6-gola zemlja.

Nakon učitavanja satelitskog snimka u okruženje softvera SAGA GIS, pokrenuta je procedura za prvu metodu mašinskog učenja, metodu slučajnih šuma (eng. Random Forest - RF). U sledećem koraku se, u prozoru za unošenje parametara, biraju metode klasifikacije. U polju Grid system se učitava raster satelitskog snimka, dok se u poljima Training Area i Class Identifier unose se već pripremljeni test uzorci i polje atributa za svih 6 klasa po kojem se vrši klasifikacija korišćenjem izabrane metode klasifikacije mašinskog učenja.

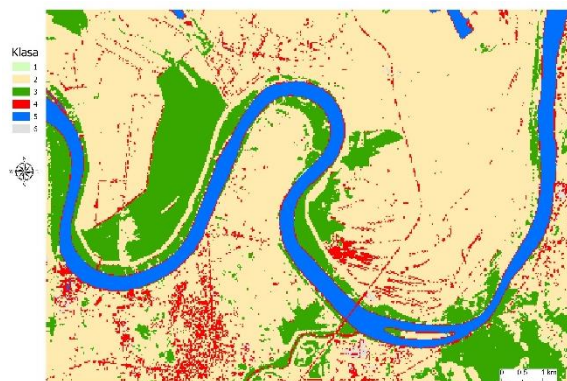
Parametri za izabranu metodu mašinskog učenja se nalaze u poljima Options. Prilikom korišćenja metode stablo odlučivanja koristimo već predefinisane parametre koji se nalaze u prikazanom prozoru. Rezultat procesa je klasifikovan raster koji je prikazan na slici 1.



Slika 1 - Klasifikovan raster korišćenjem metode stabla odlučivanja

Procedura za metodu potpornih vektora (eng. Support Vector Machine - SVM) se u odnosu na RF metodu razlikuje samo u podešavanju parametara za izabranu metodu mašinskog učenja u poljima Options. Prilikom korišćenja SVM metode za kernel funkciju biramo radial basis funkciju. Nakon unetih parametara

za izabranu metodu klasifikacije dobijamo klasifikovan raster koji je prikazan na slici 2.



Slika 2 - Klasifikovan raster korišćenjem metode potpornih vektora

U poređni rezultati klasifikacije su prikazani u tabeli 2, gde se mogu videti površine klasifikovane kroz dva različita modela, RF i SVM u kvadratnim kilometrima. Najveću površinu obuhvata Klasa 2 sa 84.23 km² za RF i 78.30 km² za SVM, dok su Klasa 1 i Klasa 6 najmanje zastupljene. U poređenju dva modela, RF generalno prikazuje veće površine osim za Klasu 3 i Klasu 4 gde SVM pokazuje veće vrednosti.

Tabela 2. U poređni prikaz površina klasifikovanih putem RF i SVM modela u km² i procentima

	Klasa 1	Klasa 2	Klasa 3	Klasa 4	Klasa 5	Klasa 6	Ukupno
RF (km ²)	0.30	84.23	19.38	3.11	10.53	0.28	117.88
RF (%)	0.25	71.46	16.44	2.64	8.93	0.24	100
SVM (km ²)	0.13	78.30	20.87	7.45	10.69	0.44	117.88
SVM (%)	0.11	66.43	17.70	6.32	9.06	0.37	100

2.3. Ocena tačnosti dobijenih rezultata

Proces ocene tačnosti klasifikovanog rastera u ArcGIS PRO verziji 3.2 počinje stvaranjem tačaka za procenu tačnosti (eng. Accuracy Assessment Points), koje su ključne za obezbeđivanje validnosti klasifikacije. Podaci za ocenu tačnosti su definisani nezavisno od podataka koji su učestvovali u procesu treniranja kako bi se obezbedila potpuno nezavisna procedura. Odabir podataka je izvršen na osnovu snimaka veće rezolucije od rezolucije Landsat 8 snimka. Za potrebe rada su kao osnova korišćeni snimci preuzeti sa Sentinel 2 satelitskog sistema prostorne rezolucije 10 m [21] i snimci dostupni na platformi Google Earth [22]. Sledeći korak je izračunavanje matrice konfuzije, koja nam pruža tabelarni prikaz odnosa između stvarnih i klasifikovanih kategorija. Ova matrica je ključna za detaljnu analizu tačnosti klasifikacije [23, 24]. Korisnička tačnost pokazuje verovatnoću da klasifikovana tačka u određenoj kategoriji zaista pripada toj kategoriji, dok proizvođačka tačnost ukazuje

na verovatnoću pravilnog klasifikovanja tačke koja stvarno pripada određenoj kategoriji [24].

U ovoj studiji, primenjene su dve metode za klasifikaciju rasterskih podataka: RF i SVM. U RF klasifikaciji, ukupan broj tačaka korišćenih za ocenu tačnosti bio je 119 (tabela 3), dok je u SVM klasifikaciji taj broj iznosio 116 (tabela 4). Za ocenu tačnosti je korišćeno svih 6 klasa.

Tabela 3. Ocena tačnosti RF klasifikacije

RF Klasifikacija									
Klasa	1	2	3	4	5	6	Total	Korisnička tačnost	Kappa
1	9	0	1	0	0	0	10	0.9	-
2	5	41	8	5	0	5	64	0.64	-
3	0	0	15	0	0	0	15	1	-
4	1	4	1	4	0	0	10	0.4	-
5	0	0	0	0	10	0	10	1	-
6	0	0	0	2	0	8	10	0.8	-
Total	15	45	25	11	10	13	119	0	-
Proizvođačka tačnost	0.6	0.91	0.6	0.36	1	0.62	0	0.73	-
Kappa	-	-	-	-	-	-	-	-	0.63

Kada poredimo podatke iz tabela 3 i 4, uočavamo da SVM klasifikacija uopšteno ima bolje rezultate za korisničku i proizvođačku tačnost u poređenju sa RF klasifikacijom. Na primer, u SVM klasifikaciji, korisnička tačnost se kreće od 70% za klasu 4 do 100% za klasu 5, dok proizvođačka tačnost varira od 53.33% za klasu 1 do 100% za klasu 6. Za razliku od toga, u RF klasifikaciji, korisnička tačnost se kreće od 40% za klasu 4 do 100% za klase 3 i 5, dok proizvođačka tačnost varira od 36.36% za klasu 4 do 100% za klasu 5. Kapa statistika pokazuje ukupnu tačnost klasifikacije u odnosu na slučajnu tačnost, ističući da SVM klasifikacija ima veću saglasnost sa vrednošću Kapa statistike od 75.76%, u poređenju sa 63.44% kod RF klasifikacije. Ovi rezultati sugerišu da SVM klasifikacija generalno postiže veći nivo tačnosti u ovom istraživanju.

Tabela 4. Ocena tačnosti SVM klasifikacije

SVM Klasifikacija									
Klasa	1	2	3	4	5	6	Total	Korisnička tačnost	Kappa
1	8	1	1	0	0	0	10	0.8	-
2	7	48	5	0	0	0	60	0.8	-
3	0	2	14	0	0	0	16	0.88	-
4	0	1	1	7	1	0	10	0.7	-
5	0	0	0	0	10	0	10	1	-
6	0	0	0	1	0	9	10	0.9	-
Total	15	52	21	8	11	9	116	0	-
Proizvođačka tačnost	0.53	0.92	0.67	0.88	0.91	1	0	0.83	-
Kappa	-	-	-	-	-	-	-	-	0.76

Izbor metode klasifikacije treba uvek pažljivo razmotriti u skladu sa specifičnim zahtevima i karakteristikama podataka u svakom istraživanju. Primeenom ovog pristupa omogućava se preciznija i pouzdanija ocena tačnosti klasifikovanih rasterskih podataka, što je ključno za dalje analize i primenu rezultata u oblasti geografskih informacionih sistema.

3. ZAKLJUČAK

Malo je oblasti, disciplina ili metodologija koje su doživele tako ubrzan i dinamičan razvoj kao daljinska detekcija. Danas oblast daljinska detekcija obuhvata širok spektar stručnjaka iz različitih specijalizovanih oblasti. Kontinuirano se unapređuje tehnologija snimanja, oprema, softverski alati, tehnike obrade i područja primene, što rezultira konstantnim širenjem i evolucijom ovog polja. Zbog ovog brzog napretka, literatura u oblasti brzo zastareva, dok temeljni principi i osnovni postulati ostaju nepromenjeni.

Daljinska detekcija sve više dobija na značaju kao ključna metoda za prikupljanje informacija o zemljištu i okolini. Vodeće satelitske misije i programi ne samo da prate i snimaju raznovrsne prirodne i društvene pojave, već takođe koriste tehnike mašinskog učenja i klasifikaciju snimaka pomoću različitih naprednih algoritama, potvrđujući svoju nezamenljivu ulogu u modernoj naučnoj i praktičnoj primeni.

Na osnovu analize i poređenja rezultata iz tabela 3 i 4 može se zaključiti da SVM metoda pokazuje bolje performanse u odnosu na RF. To se posebno ogleda u većim vrednostima korisničke i proizvođačke tačnosti kod SVM klasifikacije, kao i u višoj vrednosti Kapa statistike (75.76% naspram 63.44%).

Dok RF klasifikacija pokazuje varijabilnost u tačnosti, sa nižim minimalnim vrednostima, SVM metoda konzistentno ostvaruje visoke rezultate, što sugeriše njenu veću efikasnost i pouzdanost u klasifikaciji rasterskih podataka. Ovi nalazi ukazuju na prednost SVM metode u konkretnom kontekstu istraživanja, potvrđujući njenu primenljivost u preciznijoj GIS analizi.

Konačno, može se zaključiti da mašinsko učenje i razni algoritmi, koji će u budućnosti biti sve raznovrsniji i perspektivniji, zamenjuju u potpunosti dosadašnje metode klasifikacije i dovode do unapređenja daljinske detekcije odnosno njenog sadržaja.

4. ZAHVALNOST

Ovaj rad je deo projekta 1.23: *Prikupljanje geotopografskih podataka bespilotnim aerofotogrametrijskim sistemima*, koji finansira Ministarstvo odbrane Republike Srbije.

LITERATURA

- [1] Russell S. Norvig, P. Artificial Intelligence: A Modern Approach (Global Edition). *Artif. Intell. A Mod. Approach*, 2021.
- [2] Nikolić M, Zečević A. *Mašinsko učenje*; Univerzitet u Beogradu, Matematički fakultet: Beograd, 2019;
- [3] Taherdoost, H. Machine Learning Algorithms. In *Encyclopedia of Data Science and Machine Learning*; IGI Global, pp. 938–960, 2022.
- [4] Kalita, J.K. *Machine Learning: Theory and Practice*; 1st Edition.; CRC Press: Boca Raton, ISBN 9780367433543, 2023.
- [5] Novaković J. Đ. *Rešavanje klasifikacionih problema mašinskog učenja*; Veljović, A., Ed.; Reinženjering pp.; Fakultet tehničkih nauka u Čačku: Čačak, Vol. 4; ISBN 9788677761578, 2013.
- [6] Orozco M, Castellanos, G. Nearest Feature Rules and Dissimilarity Representations for Face Recognition Problems. In *Face Recognition*; 2007.
- [7] Orbay A, Akarun L. Neural Sign Language Translation by Learning Tokenization. In Proceedings of the Proceedings - 15th *IEEE International Conference on Automatic Face and Gesture Recognition*, FG 2020; 2020.
- [8] Benedicta, E.; Werner, M. Learning Finite State Machine Specifications from Test Cases. doi: 10.53846-/GOEDISS-2462, 2010.
- [9] Aigner S, Körner M. The Importance of Loss Functions for Increasing the Generalization Abilities of a Deep Learning-Based Next Frame Prediction Model for Traffic Scenes. *Mach. Learn. Knowl. Extr. Vol. 2*, Pages 78-98 2020, 2, 78–98, doi: 10.3390/MAKE2020006, 2020.
- [10] Tugay R, Ögüdücü Ş. G. Demand Prediction Using Machine Learning Methods and Stacked Generalization. In Proceedings of the DATA 2017 - Proceedings of the 6th International Conference on Data Science, Technology and Applications; 2017.
- [11] Kamath, C. Intelligent Sampling for Surrogate Modeling, Hyperparameter Optimization, and Data Analysis. *Mach. Learn. with Appl.* 9, 100373, doi: 10.1016/J.MLWA.2022.100373, 2022.
- [12] Milošević B, Milošević D, Obradović, S. Mašinsko Učenje u Obrazovanju. In Proceedings of the Infoteh-jahorina; Jahorina, Vol. 13, pp. 964–968, March 2014.
- [13] Heminger, A. Automated Decision Making and Machine Learning: Regulatory Alternatives for Autonomous Settings, Portland State University, 2021.
- [14] Potić I, Srdić Z, Vakanjac B. Bakrač S, Đorđević D, Banković R, Jovanović, J.M. Improving Forest Detection Using Machine Learning and Remote Sensing: A Case Study in Southeastern Serbia. *Appl. Sci.* 13, 8289, doi: 10.3390/app13148289, 2023.
- [15] Tempa K, Aryal K. R. Semi-Automatic Classification for Rapid Delineation of the Geohazard-Prone Areas Using Sentinel-2 Satellite Imagery. *SN Appl. Sci.* 4, doi: 10.1007/s42452-022-05028-6, 2022.
- [16] Potić I. M, Komazec N. M, Mihajlović, L. M, Milić A. M. & Bakrač S. T. Upravljanje rizikom od neeksplozivnih ubojnih sredstava u Republici Srbiji u funkciji zaštite životne sredine - studija slučaja Borovac. *Vojnotehnički glasnik.* 71(4), 1087-1114. <https://doi.org/10.5937/vojtehg71-44656>, 2023
- [17] Hornstra T. J, Lemmens M. J. P. M, Wright, G.L. Incorporating Intra-Pixel Reflectance Variability in the Multispectral Classification Process of High-Resolution Satellite Imagery of Urbanised Areas. *Cartography*, 28, doi: 10.1080/00690805-1999.9714313, 1999.
- [18] Ramezan C. A, Warner T. A, Maxwell A. E, Price, B.S. Effects of Training Set Size on Supervised Machine-Learning Land-Cover Classification of Large-Area High-Resolution Remotely Sensed Data. *Remote Sens.* 13, doi: 10.3390/rs13030368, 2021.
- [19] Engebretson, C. Landsat 8-9 Operational Land Imager (OLI) - Thermal Infrared Sensor (TIRS) Collection 2 Level 2 (L2) Data Format Control Book (DFCB). Vol. 2. Department of the Interior U.S. Geological Survey, 2020.
- [20] Conrad O, Bechtel B, Bock M, Dietrich H, Fischer E, Gerlitz L, Wehberg J, Wichmann V, Böhner J. System for Automated Geoscientific Analyses (SAGA) v. 2.1.4. *Geosci. Model Dev.* 2015, 8, doi: 10.5194/gmd-8-1991-2015, 2015.
- [21] Copernicus. 2020. Copernicus Open Access Hub [online]. Available at: <https://scihub.copernicus.eu/> [Accessed: 20 November 2023]
- [22] Google Earth. Geospatial Solutions: Google Earth Pro 7.3.3.7786 [online]. Available at: <https://www.google.com/earth/versions/> [Accessed: 20 November 2023], 2020.
- [23] Esri. (n.d.). Accuracy assessment. ArcGIS Pro. Retrieved November 26, from <https://pro.arcgis.com/en/pro-app/latest/help/analysis/image-analyst/accuracy-assessment.htm>, 2023.
- [24] Congalton R. G. & Green K. Assessing the Accuracy of Remotely Sensed Data: Principles and Practices (3rd ed.). CRC Press. 2019. ISBN 9780429052729. doi: 10.1201/9780429052729

SUMMARY

APPLICATION OF MACHINE LEARNING METHODS IN THE CLASSIFICATION OF SATELLITE IMAGES

Machine learning, as a specific domain within artificial intelligence, opens new horizons for both theoretical and experimental research in remote sensing, particularly in satellite imagery classification. This study focuses on applying machine learning methods, specifically decision trees and support vector machines, to classify satellite images. The analysis uses the SAGA GIS software on LANDSAT 8 OLI Level 2A satellite images. Satellite image classification encompasses two primary groups of computer operations: unsupervised (automatic or formal) and supervised (semi-automatic or logical) classification. This research executes the practical classification of satellite images by applying the aforementioned machine learning methods. The results indicate that the obtained classified rasters not only align with but also fully replace existing classification and identification methods of geospatial objects. Consequently, this research contributes to a significant advancement in collecting and analysing geospatial data.

Key Words: *remote sensing, artificial intelligence, classification of satellite images, data processing*