

# Detekcija snježnih površina u Alpama iz multispektralnih satelitskih snimaka korištenjem strojnog učenja u programskom jeziku R

---

Subotić, Josipa

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Mining, Geology and Petroleum Engineering / Sveučilište u Zagrebu, Rudarsko-geološko-naftni fakultet**

Permanent link / Trajna poveznica: <https://urn.nsk.hr/urn:nbn:hr:169:726499>

Rights / Prava: [In copyright](#) / [Zaštićeno autorskim pravom](#).

Download date / Datum preuzimanja: **2025-04-03**



Repository / Repozitorij:

[Faculty of Mining, Geology and Petroleum Engineering Repository, University of Zagreb](#)



SVEUČILIŠTE U ZAGREBU  
RUDARSKO-GEOLOŠKO-NAFTNI FAKULTET  
Diplomski studij geološkog inženjerstva

**DETEKCIJA SNJEŽNIH POVRŠINA U ALPAMA IZ MULTISPEKTRALNIH  
SATELITSKIH SNIMAKA KORIŠTENJEM STROJNOG UČENJA U  
PROGRAMSKOM JEZIKU R**

Diplomski rad

Josipa Subotić

G2181

Zagreb, 2024.



KLASA: 602-04/24-01/116  
URBROJ: 251-70-14-24-1  
U Zagrebu, 19.06.2024.

Josipa Subotić, studentica

## RJEŠENJE O ODOBRENJU TEME

Na temelju vašeg zahtjeva primljenog pod KLASOM 602-04/24-01/116, URBROJ: 251-70-14-24-1 od 26.06.2024. priopćujemo vam temu diplomskog rada koja glasi:

### DETEKCIJA SNJEŽNIH POVRŠINA U ALPAMA IZ MULTISPEKTRALNIH SATELITSKIH SNIMAKA KORIŠTENJEM STROJNOG UČENJA U PROGRAMSKOM JEZIKU R

Za mentora ovog diplomskog rada imenuje se u smislu Pravilnika o izradi i obrani diplomskog rada Doc.dr.sc. Ivan Medved nastavnik Rudarsko-geološko-naftnog fakulteta Sveučilišta u Zagrebu.

Mentor:



(potpis)

Doc.dr.sc. Ivan Medved

(titula, ime i prezime)

Predsjednica povjerenstva za  
završne i diplomske ispite:



(potpis)

Izv.prof.dr.sc. Ana Maričić

(titula, ime i prezime)

Prodekan za nastavu i studente:



(potpis)

Izv.prof.dr.sc. Borivoje  
Pašić

(titula, ime i prezime)

**DETEKCIJA SNJEŽNIH POVRŠINA U ALPAMA IZ MULTISPEKTRALNIH  
SATELITSKIH SNIMAKA KORIŠTENJEM STROJNOG UČENJA U  
PROGRAMSKOM JEZIKU R**

Josipa Subotić

Rad izrađen: Sveučilište u Zagrebu  
Rudarsko-geološko-naftni fakultet  
Zavod za geofizička istraživanja i rudarska mjerenja  
Pierottijeva 6, 10 000 Zagreb

Sažetak

Svrha ovog rada je razvoj i evaluacija modela za detekciju snježnog pokrivača korištenjem algoritama strojnog učenja. Korištenjem multispektralnih satelitskih snimaka iz misije Sentinel-2 i naprednih algoritama poput randomForest, Support Vector Machine i Neural Network, cilj je točno klasificirati snježni pokrivač i druge klase poput stijena, urbanih područja, vode i vegetacije. Analizirani su spektralni potpisi snježnog pokrivača u različitim godišnjim razdobljima, čime je omogućeno stvaranje pouzdanih klasificiranih rastera za vizualizaciju i procjenu učinkovitosti modela.

Ključne riječi: Sentinel-2, QGIS, R, RStudio, RF, SVM, NNET, ALPE, ITALIJA, AUSTRIJA, snijeg, urbano područje, vegetacija, stijene, voda.

Završni rad sadrži: 62 stranice, 16 tablica, 40 slika, 0 priloga, i 33 reference.

Jezik izvornika: Hrvatski

Pohrana rada: Knjižnica Rudarsko-geološko-naftnog fakulteta, Pierottijeva 6, Zagreb

Mentori: Dr. sc. Ivan Medved, docent RGNF

Ocjenjivači: Dr. sc. Ivan Medved, docent RGNF  
Dr. sc. Jasna Orešković, profesor RGNF  
Dr. sc. Tomislav Korman, izvanredni profesor RGNF

DETECTION OF SNOW SURFACES IN THE ALPS FROM MULTISPECTRAL SATELLITE IMAGES  
USING MACHINE LEARNING IN R PROGRAMMING LANGUAGE

Josipa Subotić

Thesis completed at: University of Zagreb

Faculty of Mining, Geology and Petroleum Engineering  
Department of Geophysical Exploration and Mine Surveying  
Pierottijeva 6, 10 000 Zagreb

Abstract

The purpose of this work is to develop and evaluate a model for snow cover detection using machine learning algorithms. Using multispectral satellite images from the Sentinel-2 mission and advanced algorithms such as randomForest, Support Vector Machine and Neural Network, the goal is to accurately classify snow cover and other classes such as rocks, urban areas, water and vegetation. The spectral signatures of the snow cover in different periods of the year were analyzed, which enabled the creation of reliable classified rasters for visualization and evaluation of the model's effectiveness.

Keywords: Sentinel-2, QGIS, R, RStudio, RF, SVM, NNET, ALPS, ITALY, AUSTRIA, snow, urban area, vegetation, rocks, water.

Thesis contains: 62 pages, 16 tables, 40 figures, 0 appendixes, i 33 references.

Original in: Croatian

Archived in: Library of Faculty of Mining, Geology and Petroleum Engineering, Pierottijeva 6, Zagreb

Supervisors: PhD. Ivan Medved, Assistant Professor RGNF

Reviewers: PhD. Ivan Medved, Assistant Professor RGNF  
PhD Jasna Orešković, Professor RGNF  
PhD Tomislav Korman, Associate Professor RGNF

## SADRŽAJ

1. UVOD .....	1
2. CILJEVI RADA .....	2
3. PODRUČJE ISTRAŽIVANJA .....	3
3.1. Prethodna istraživanja .....	4
4. MATERIJAL I METODE.....	6
4.1. Copernicus (ESA) i Sentinel-2 .....	6
4.2. QGIS .....	8
4.3. R programski jezik .....	9
4.3.1. R paketi.....	10
4.4. Metoda strojnog učenja za klasifikaciju.....	11
5. ANALIZA SPEKTRALNOG POTPISA SATELITSKE MISIJE SENTINEL-2	14
6. OBRADA PODATAKA METODAMA STROJNOG UČENJA.....	16
6.1. Korištenje metode <i>randomForest (RF)</i> iz paketa CARET.....	19
6.2. Korištene metode <i>Support Vector Machine (SVM)</i> iz paketa CARET .....	24
6.3. Korištenje metode <i>Neural Network (NNET)</i> iz paketa CARET .....	27
6.4. Usporedba dobivenih modela .....	31
7. REZULTAT RASPRAVA KORIŠTENJEM SENTINEL-2.....	34
7.1. Potvrde <i>modela rf</i> starosti stijena sa kartama ljetnog i zimskog razdoblja .....	40
8. ZAKLJUČAK.....	48
9. ZAHVALE .....	50
10. LITERATURA .....	51

## POPIS SLIKA

Slika 2.1 Prikaz rada algoritma.....	2
Slika 3.1 Područje istraživanja Italije i Austrije. ....	3
Slika 4.1 Copernicus sentinel-2 (Esri, 2023).....	6
Slika 4.2 Copernicus- dobivanje i preuzimanje podataka (Copernicus-b, 2024) .....	7
Slika 4.3 Područje istraživanja u programu QGIS. ....	9
Slika 4.4 R studio. ....	10
Slika 4.5 Prikaz rada <i>RF</i> -a (Javatpoint-b, 2021.). ....	12
Slika 4.6 Prikaz rada <i>SVM</i> -a (Datacamp, 2018.). ....	12
Slika 4.7 Prikaz rada <i>Neural Network</i> (Geeksforgeeks, 2020.).....	13
Slika 5.1 Spektralni potpis zimskog razdoblja Sentinel-2.....	15
Slika 5.2 Spektralni potpis ljetnog razdoblja Sentinel-2. ....	15
Slika 6.1 Prikaz histograma multispektralnih snimaka za zimsko razdoblje. ....	17
Slika 6.2 Prikaz histograma multispektralnih snimaka za ljetno razdoblje.....	18
Slika 6.3 Grafički prikaz predikcije <i>RF</i> -a iz razdoblja zime .....	20
Slika 6.4 Grafički prikaz predikcije <i>RF</i> -a iz razdoblja ljeta.....	21
Slika 6.5 Prikaz <i>heatmap</i> 1 razdoblje zima. ....	22
Slika 6.6 Prikaz <i>heatmap</i> 1 razdoblje ljeto.....	23
Slika 6.7 Slika <i>heatmap</i> 2 razdoblje zime. ....	23
Slika 6.8 Slika <i>heatmap</i> 2 razdoblje ljeto.....	24
Slika 6.9 Grafički prikaz modela <i>SVM</i> -a razdoblje zime. ....	25
Slika 6.10 Grafički prikaz modela <i>SVM</i> -a razdoblje ljeta. ....	26
Slika 6.11 Grafički prikaz modela <i>Neural Network</i> za razdoblje zime.....	28
Slika 6.12 Grafički prikaz modela <i>Neural Network</i> za razdoblje ljeta.....	28
Slika 7.1 <i>Predict_nnet</i> zimsko razdoblje.....	35
Slika 7.2 <i>Predict_nnet</i> ljetno razdoblje. ....	36
Slika 7.3 <i>Predict_rf</i> zimsko razdoblje. ....	37
Slika 7.4 <i>Predict_rf</i> ljetno razdoblje.....	38
Slika 7.5 <i>Predict_svm</i> zimsko razdoblje. ....	39
Slika 7.6 <i>Predict_svm</i> ljetno razdoblje. ....	40
Slika 7.7 Približena slika snježnog pokrova u Alpama tijekom zimskog razdoblja.....	41
Slika 7.8 Područje Padanske nizine iz zimskog razdoblja modela <i>nnet</i> . ....	41
Slika 7.9 Detekcija vodenog pokrova zimskog razdoblja modela <i>nnet</i> .....	42

Slika 7.10. Problematika u sjenovitim područjima modela <i>nnet</i> .....	42
Slika 7.11 Problematika u detektiranju snijega modela <i>nnet</i> . ....	43
Slika 7.12 Pogrešna detekcija urbanog područja modela <i>nnet</i> . ....	43
Slika 7.13 Približena slika snježnog pokrova u Alpama tijekom ljetnog razdoblja modela <i>rf</i> . .....	44
Slika 7.14 Približeno područje planinskog područja modela <i>rf</i> .....	45
Slika 7.15 Područje Padanske nizine iz ljetnog razdoblja modela <i>rf</i> .....	45
Slika 7.16 Detekcija vodenog pokrova ljetnog razdoblja modela <i>rf</i> . ....	46
Slika 7.17 Preciznost detekcije modela <i>rf</i> . ....	47



## POPIS TABLICA

Tablica 4.1 Sentinel-2 kanali.....	8
Tablica 4.2 Tablica paketa u R-u.....	10
Tablica 6.1 Prikaz vrijednost dobivene tablice za zimsko razdoblje.....	16
Tablica 6.2 Prikaz vrijednost dobivene tablice za ljetno razdoblje .....	17
Tablica 6.3 Količina treninga podataka klasa iz zimskog razdoblja. ....	18
Tablica 6.4 Količina testa podataka klasa iz zimskog razdoblja .....	19
Tablica 6.5 Količina treninga podataka iz ljetnog razdoblja. ....	19
Tablica 6.6 Količina testa podataka iz ljetnog razdoblja.....	19
Tablica 6.7 Statistički prikaz <i>random forest</i> klasifikacije za razdoblje zime.....	21
Tablica 6.8 Statistički prikaz <i>random forest</i> klasifikacije za razdoblje ljeta.....	22
Tablica 6.9 Statistički prikaz <i>SVM</i> klasifikacije razdoblje zime. ....	26
Tablica 6.10 Statistički prikaz <i>SVM</i> klasifikacije razdoblje ljeta. ....	26
Tablica 6.11 Statistički prikaz <i>nnet</i> klasifikacije iz razdoblja zime. ....	29
Tablica 6.12 Statistički prikaz <i>nnet</i> klasifikacije iz razdoblja ljeto. ....	29
Tablica 6.13 Prikaz neuronske mreže za razdoblje zime.....	30
Tablica 6.14 Prikaz neuronske mreže za razdoblje ljeta. ....	30

## **POPIS KORIŠTENIH OZNAKA I JEDINICA**

ALOS - Advanced Land Observing Satellite

ArcGIS - Geografski informacijski sustav (GIS) softver razvijen od strane tvrtke Esri

DMR - digitalni model reljefa

DOF - digitalna ortofoto karta

GADM - Global Administrative Areas

GIS - geografski informacijski sustav

IDE - Integrated development environment

NIPP – Nacionalna infrastruktura prostornih podataka

NNET – neuronska mreža (Neural Network)

OLI - Operation Land Imager

ORS2A - Ortho Ready Standard

PRISM - Panchromatic Remote-sensing Instrument for Stereo Mapping

QGIS – geografski informacijski sustav otvorenog koda (Quantum Geographic Information System)

R – programski jezik

RF – slučajna šuma (Random Forest)

S2A/B - Sentinel-2A i Sentinel-2B

S2MSI2A - MultiSpectral Instrument (MSI) na Sentinel-2A

SRTM - Shuttle Radar Topography Mission

SVM - Support Vector Machine

TIFF – format datoteke za pohranu slika (Tagged Image File Format)

TIRS - Thermal Infrared Sensor

USGS – Američki geološki zavod (United States Geological Survey)

WFS – Mrežni servis za vektorske podatke (Web Feature Service)

WMS-C - Web Map Service-Cached

WMS – Mrežni kartografski servis (Web Map Service)

WMTS - Web Map Tile Service

WV2 - WorldView-2

## 1. UVOD

Klasifikacija snježnog pokrivača ima ključnu ulogu u razumijevanju klimatskih promjena, upravljanju vodnim resursima i prevenciji prirodnih katastrofa poput lavina. Snježni pokrivač, osobito u planinskim područjima, predstavlja važan izvor vode te igra presudnu ulogu u ekosustavima i ljudskim aktivnostima. Međutim, točna i pravovremena detekcija snježnog pokrivača na velikim geografskim područjima predstavlja izazov, posebno u uvjetima miješanja sa sličnim klasama poput vegetacije ili stijena. Satelitske snimke, poput onih iz misije Sentinel-2 (Europska svemirska agencija, Copernicus), nude visoku rezoluciju i mogućnost precizne analize površinskih pokrivača. Korištenjem ovih podataka, u kombinaciji s naprednim algoritmima strojnog učenja, moguće je automatski i učinkovito klasificirati snježni pokrivač.

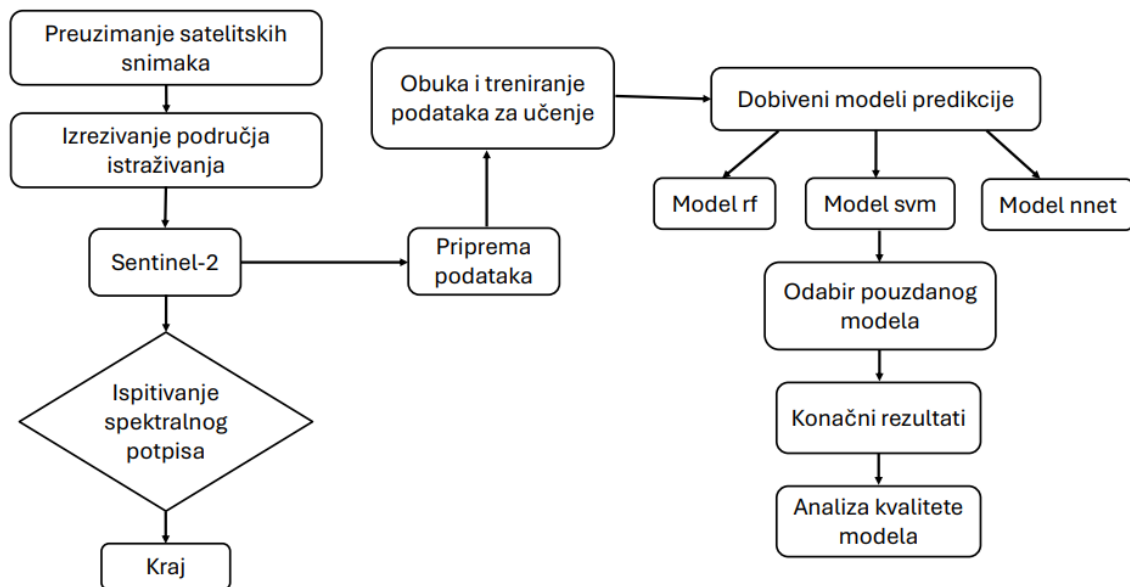
Glavni cilj ovog rada je razvoj i evaluacija modela za detekciju snježnog pokrivača koristeći algoritme strojnog učenja. Koristeći aplikacije i programske jezike otvorenog koda, poput QGIS-a i R-a, kreiran je model koji koristi multispektralne satelitske snimke za klasifikaciju snijega te drugih klasa poput stijena, urbanog područja, vode i vegetacije. Nadzirana klasifikacija primijenjena je kroz algoritme poput *randomForest* (RF), *Support Vector Machine* (SVM) i *Neural Network* (nnet) iz paketa CARET, pri čemu su ovi algoritmi obučeni na temelju poligona koji su označavali različite klase. Tijekom ovog istraživanja, provedene su detaljne analize spektralnih potpisa snježnog pokrivača u zimskim i ljetnim razdobljima, koristeći podatke dobivene iz deset različitih spektralnih kanala. Dobiveni rezultati omogućili su stvaranje klasificiranih rastera koji su potom vizualizirani i analizirani kako bi se odredila pouzdanost svakog modela u detekciji snijega.

## 2. CILJEVI RADA

Cilj ovog rada bio je izraditi algoritam u programskom jeziku R za detekciju snježnog pokrova u području Alpa. U tu svrhu koriste se multispektralne satelitske snimke s europske svemirske agencije Copernicus (Sentinel-2). Uzimaju se multispektralne satelitske snimke područja Talijanske Republike i Republike Austrije na kojima će se provesti istraživanje.

Na temelju odabranih lokacija, koristeći program QGIS, izrađeni su poligoni koji služe kao podaci za uvježbavanje sustava. Algoritam u R-u implementira nadziranu klasifikaciju koja, temeljem sličnosti danih poligona, pronalazi nove lokacije sa snježnim pokrovom.

Glavni cilj je razviti algoritam koji se može primijeniti na bilo koje područje za istraživanje i detekciju snježnog pokrova, kako bi se omogućilo nadziranje dinamika promjene površine snijega tijekom godine. Tijek izrade algoritma prikazan je na slici 2.1.



Slika 2.1 Prikaz rada algoritma.

### 3. PODRUČJE ISTRAŽIVANJA

Za područje istraživanja odabrane su Istočne Alpe, točnije područje Republike Italije i Austrijske Republike. Satelitska snimka prikazuje područje Zapadne Austrije, sjeveroistočne Italije i Padanske nizine (Slika 3.1). Zapadna Austrija na satelitskoj snimci dijeli se na saveznu državu Tirol i saveznu državu Korušku. Detaljnije na njihove kotare: Spittal an der Drau i Hermagor u Koruškoj i Lienz u Tirolu.

Talijanski dio satelitske snimke uključuje sjeveroistočni dio Italije koji obuhvaća regije Furlanija-Julijska krajina, Trentino-Južni Tirol i Veneto. Na jugoistočnom dijelu snimke nalazi se područje Padanske nizine

Regija Furlanija-Julijska Krajina sastoji se od općina Porderone i Udine. Veneto regija pokriva općinu Belluno koja gotovo u cijelosti pokriva površinu satelitske snimke te općinu Treviso. Regija Trentino-Južni Tirol sadrži općine Bolzano i Trento. Općine Belluno, Treviso, Porderone i Udine djelomično pokrivaju područje Padanske nizine.



Slika 3.1 Područje istraživanja Italije i Austrije.

### 3.1. Prethodna istraživanja

Pregledom dostupne znanstvene literature iz područja detekcije snježnih površina u Alpama pomoću multispektralnih satelitskih snimaka i strojnog učenja u programskom jeziku R, izdvajaju se sljedeći rezultati.

Istraživanje Chanussota i suradnika (2022) predstavlja pristup koji koristi kombinaciju različitih metoda obrade slike i klasifikacije, uključujući *random forest*, *SVM* i *neuralne mreže* za detekciju snježnih linija na ledenjacima u Alpama. Rezultati pokazuju visoku pouzdanost metoda, s Pearsonovom korelacijom od 79% do 96%, što ukazuje na visoku točnost u usporedbi s ručno označenim podacima. Glavna prednost predloženog pristupa je što ne zahtijeva kalibraciju za svaki pojedini ledenjak.

Zhai i suradnici (2022) istražuju korištenje Sentinel-2 satelitskih snimaka za kartiranje snježnog pokrivača pomoću metoda strojnog učenja, kao što su *random forest* i *U-Net* duboko učenje. Rezultati pokazuju da *U-Net* model, koji kombinira spektralne i prostorne informacije, nadmašuje tradicionalne metode, omogućavajući preciznije kartiranje snježnog pokrivača. Eksperimentalni rezultati pokazuju da kombinacija spektralnih traka B2, B11, B4 i B9 daje najbolje rezultate za *U-Net* model.

Iako je fokus istraživanja na Himalaji, metode korištene u istraživanju Noi i Kappasa (2017) su relevantne i za Alpe. Istraživanje uspoređuje različite klasifikatore, uključujući *random forest*, k-najbliže susjede (*KNN*) i *SVM*, za klasifikaciju korištenja zemljišta i detekciju snježnog pokrivača koristeći Sentinel-2 snimke. Rezultati pokazuju da *random forest* daje visoku točnost klasifikacije, što naglašava važnost odabira odgovarajućih metoda za precizno mapiranje.

Ovo istraživanje razvija metode za segmentaciju oblaka koristeći *konvolucijske neuronske mreže (CNN)*, s posebnim fokusom na razlikovanje između oblaka i snijega. Iako se fokusira na oblake, metodologija se može primijeniti na probleme detekcije snijega, pružajući koristan okvir za razvoj sličnih modela u programu R (Morale et al., 2024.).

He i suradnici (2021) koriste poboljšani *U-Net* model za automatsko kartiranje glacijalnih jezera u Himalajama. Ova metodologija, koja se temelji na ekstrakciji karakteristika iz satelitskih snimaka, može se prilagoditi za detekciju snježnog pokrivača u Alpama, pružajući dodatne uvide u primjenu dubokog učenja za analizu multispektralnih snimaka (.

Ova istraživanja pružaju sveobuhvatan pregled različitih metoda i tehnika koje se koriste za detekciju snijega i srodne primjene, naglašavajući važnost strojnog učenja i odabira odgovarajućih spektralnih podataka za postizanje preciznih rezultata. Korištenjem ovih

informacija, moguće je razviti i implementirati vlastite modele za detekciju snježnih površina u Alpama koristeći R.

## 4. MATERIJAL I METODE

U okviru ovog istraživanja bit će predstavljeni postupci i metode za prikupljanje i obradu podataka. Za prikupljanje podataka korišteni su multispektralni satelitski snimci dobiveni misijom Copernicus (Sentinel-2). Obrada ovih podataka provedena je koristeći program QGIS i programski jezik R. Prednosti korištenih materijala i metoda uključuju njihovu besplatnost i jednostavnu dostupnost.

### 4.1. Copernicus (ESA) i Sentinel-2

Copernicus je program Europske unije za promatranje Zemlje, koji pruža širok raspon usluga primjenjivih u različitim sektorima kao što su urbani razvoj, zaštita okoliša, poljoprivreda, šumarstvo, ribarstvo, zdravstvo, infrastruktura, turizam i drugi (Slika 4.1). Ove usluge su besplatne i dostupne putem web stranice Copernicus Open Access Hub (Slika 4.2).

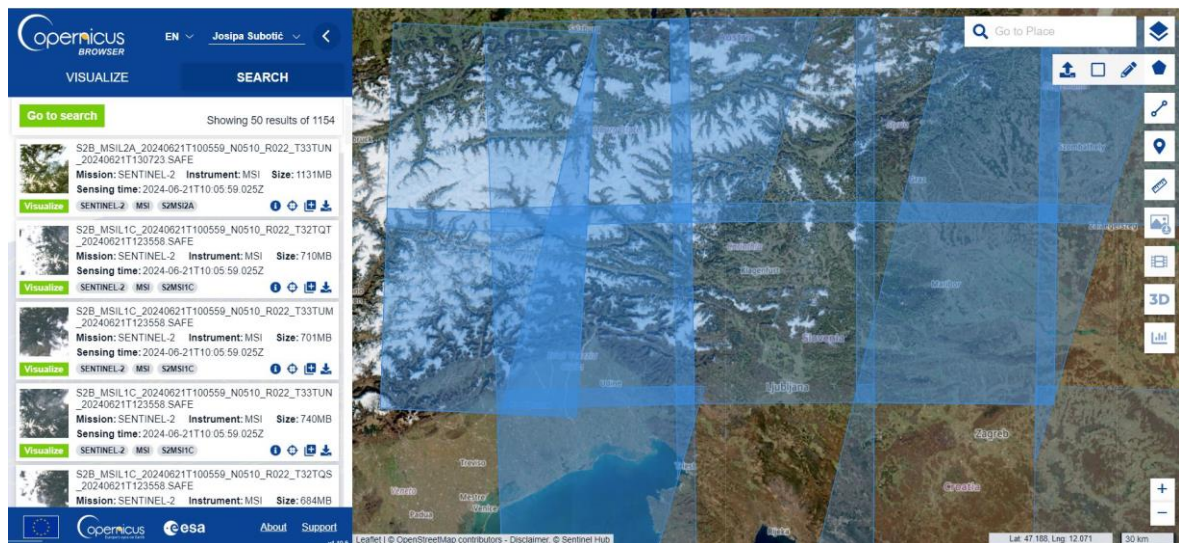


Slika 4.1 Copernicus sentinel-2 (Esri, 2023)

Sentinel-2 je europska satelitska misija koja pruža širokopojasne, visoko rezolutne multispektralne slike. Ovaj sustav je opremljen optičkim instrumentom koji nudi 13 spektralnih kanala: četiri kanala s prostornom rezolucijom od 10 metara, šest kanala s rezolucijom od 20 metara i tri kanala s rezolucijom od 60 metara. Misiju čine dva satelita, Sentinel-2A i Sentinel-2B, koji rade u sinkroniziranoj sunčevoj orbiti na srednjoj visini od 786 km, postavljeni pod kutom od  $180^\circ$  jedan prema drugome. Glavni cilj Sentinel-2 misije je podržati program Copernicus u praćenju klimatskih promjena, nadzoru zemljišta, te upravljanju hitnim situacijama i sigurnošću. Misija omogućava visoko kvalitetne podatke koji se koriste za klasifikaciju zemljišnog pokrova, korekciju atmosferskih utjecaja te smanjenje efekata oblaka i snijega. Multispektralni satelitski snimci mogu se preuzeti odabirom područja interesa, vremenskog perioda, postotka oblaka i proizvoda S2MSI2A,



nakon čega sustav generira slike s rezolucijama od 10 m, 20 m i 60 m, prilagođene specificiranim kriterijima. (Sentinel Online, 2023.)



Slika 4.2 Copernicus- dobivanje i preuzimanje podataka (Copernicus-b, 2024)

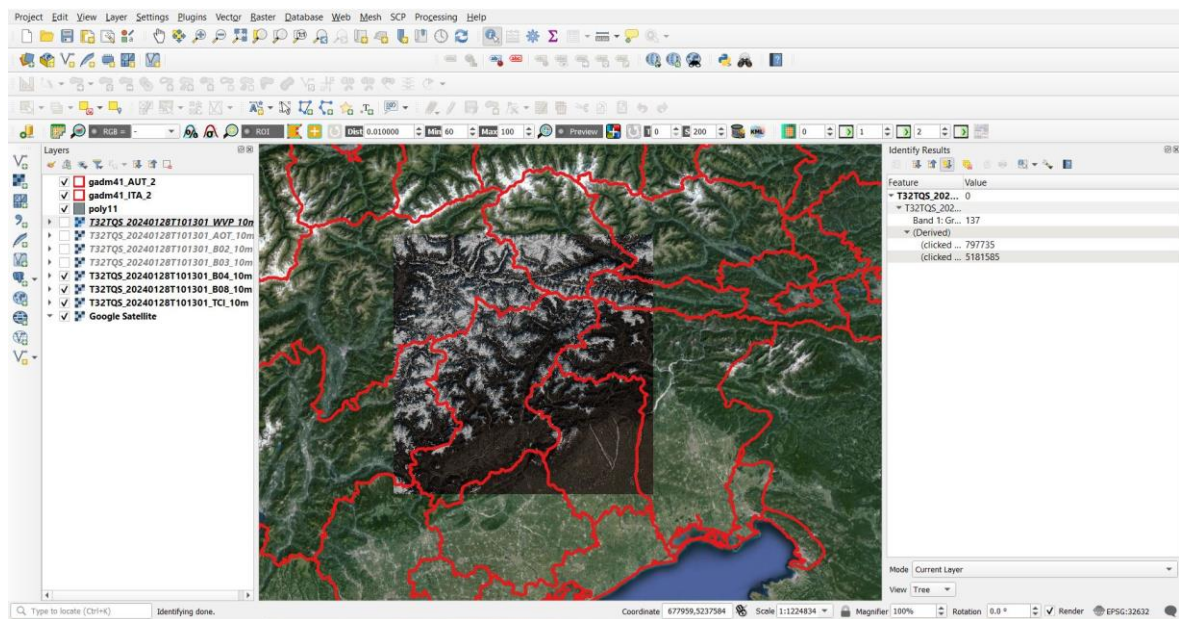
Multispektralna satelitska snimka sastoji se od 13 kanala (Tablica 4.1 Sentinel-2 kanali.), od kojih svaki ima različitu prostornu rezoluciju: 10 m (kanali B2 – plavi, B3 – zeleni, B4 – crveni, B8 – bliski infracrveni), 20 m (kanali B5, B6, B7, B8a – VNIR; B11, B12 – SWIR) i 60 m (kanali B1 – obalni aerosol, B10 – infracrveni SWIR). Kombiniranjem ovih kanala moguće je istaknuti određene značajke na snimci, kao što su geološke strukture, poljoprivredni usjevi ili vegetacijski obrasci. U ovom radu koristi se deset kanala: B2 (plavi), B3 (zeleni), B4 (crveni), B5, B6, B7, B8, B8a (VNIR), B11 i B12 (SWIR). Prikaz boja na snimci omogućava vizualizaciju sličnu ljudskom vidu, gdje je vegetacija prikazana zelenom bojom, urbana područja bijelim i sivim nijansama, a vodena područja tamnoplavom bojom, ovisno o čistoći vode (GISGeography, 2022.).

**Tablica 4.1** Sentinel-2 kanali.

Kanali	Rezolucija	Valna duljina	Opis
B1	60m	443 nm	Ultra plava
B2	10m	490 nm	Plava
B3	10m	560 nm	Zelena
B4	10m	665 nm	Crvena
B5	20m	705 nm	Vidljivi i bliski infracrveni (VNIR)
B6	20m	740 nm	Vidljivi i bliski infracrveni (VNIR)
B7	20m	783 nm	Vidljivi i bliski infracrveni (VNIR)
B8	10m	842 nm	Vidljivi i bliski infracrveni (VNIR)
B8a	10m	865 nm	Vidljivi i bliski infracrveni (VNIR)
B9	60m	940 nm	Kratkovalno infracrveni (SWIR)
B10	60m	1357 nm	Kratkovalno infracrveni (SWIR)
B11	20m	1610 nm	Kratkovalno infracrveni (SWIR)
B12	20m	2190 nm	Kratkovalno infracrveni (SWIR)

## 4.2. QGIS

QGIS (Quantum GIS) je besplatan geografski informacijski sustav otvorenog koda koji omogućuje rad s vektorskim i rasterskim podacima te bazama podataka. QGIS podržava pregled, uređivanje, kreiranje, upravljanje i izvoz vektorskih i rasterskih podataka u raznim formatima, uključujući karte, te nudi nadogradnju osnovnih GIS funkcija (QGIS-b,2023., QGIS-a, 2023.). Karte dostupne na internetu mogu se pretraživati i uvesti u QGIS pomoću servisa poput WMS, WMTS, WMS-C, WFS i drugih. Instalacijom UMN MapServer-a ili GeoServer-a zajedno s web serverom, moguće je podatke objaviti na internetu.



Slika 4.3 Područje istraživanja u programu QGIS.

U QGIS-u, koristeći multispektralne satelitske snimke, određene su granice područja istraživanja. Izrađuju se poligoni za pet vrsti klasifikacija (snijeg, voda, stijene, urbano područje, vegetacija) koji se kasnije koriste za daljnje istraživanje (Slika 4.3.).

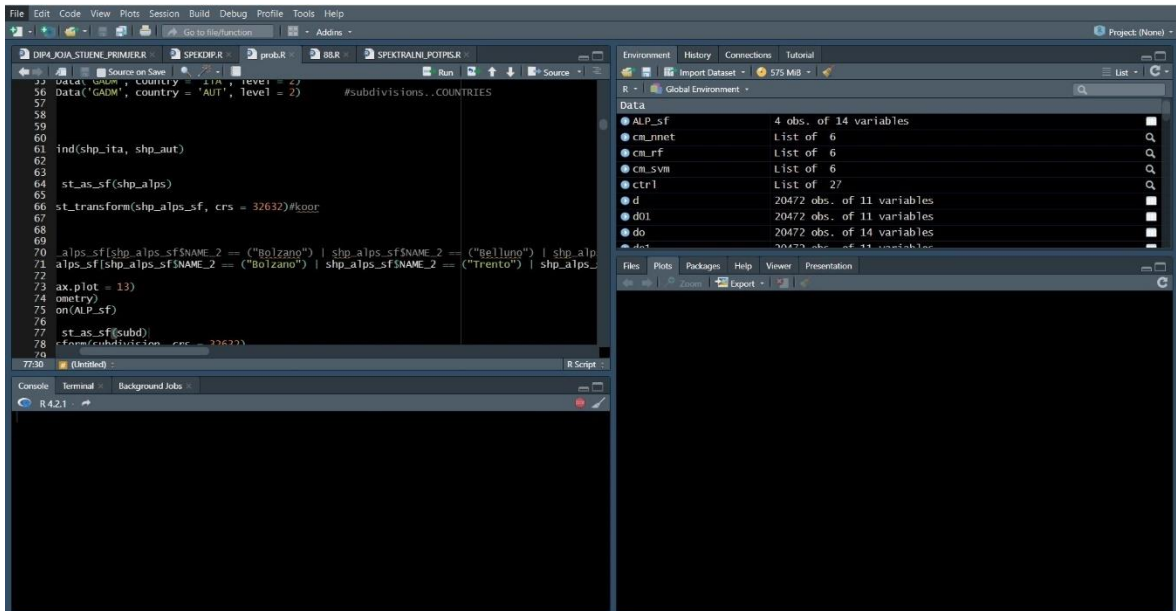
### 4.3. R programski jezik

R je besplatni softver otvorenog koda namijenjen statističkoj analizi. RStudio je integrirano razvojno okruženje (IDE) za R koje omogućuje statističko računanje i izradu grafika. U RStudio-u možete uvoziti datoteke bez potrebe za pisanjem koda, a također podržava integraciju koda i tekstualne dokumentacije za istraživačko i pisano programiranje. RStudio izvorno podržava RMarkdown sučelje, što omogućuje prikaz rezultata u HTML, PDF ili Word formatima. Također omogućuje dijeljenje projekata putem povezivanja s GitHub-om i drugim sustavima za kontrolu verzija (Giorgi et al., 2022). Dostupno je više od 10.000 R paketa koji se mogu instalirati pomoću naredbe `install.packages("naziv paketa")`.

Sučelje RStudio-a podijeljeno je na četiri glavna prozora (Slika 4.4):

1. R Skripta (gore lijevo): Namijenjen je pisanju naredbi koje se mogu spremi, ponovno koristiti ili mijenjati.
2. Konzola (dolje lijevo): Naredbe se izvršavaju izravno u konzoli, a rezultati se prikazuju odmah.
3. Okruženje (gore desno): Ovdje se nalaze kartice Environment, koja prikazuje sve varijable stvorene u R-u, i History, koja bilježi sve prethodno unesene naredbe.

2. Parcele (dolje desno): Kartica Plot prikazuje grafove i dijagrame, kartica Packages prikazuje sve dostupne pakete, dok kartica Help definira funkcije i omogućuje provjeru njihovog korištenja. (R Learning Infrastructure Home Web Page, 2023., Patwal M., 2020.)



Slika 4.4 R studio.

### 4.3.1. R paketi

R paketi predstavljaju skup funkcija, koji uključuju kodove, podatke i dokumentaciju u R programskom jeziku. Instalacija određenog paketa može se pokrenuti koristeći funkciju *install.packages()*, pri čemu se u zagradu unosi ime željenog paketa. Instalirani R paketi pohranjuju se u direktorij nazvan *library()*. Kako bi se funkcije, kodovi i podaci iz paketa mogli koristiti u R Studio-u, potrebno je pokrenuti naredbu *library()*. Tablica 4.2 Tablica paketa u R-u. prikazuje R pakete koji su korišteni u ovom istraživanju (Tutorialspoint, 2024., Posit Support, 2024., Javatpoint-a, 2021., Spatial Data Science, 2023., CRAN, 2024.)

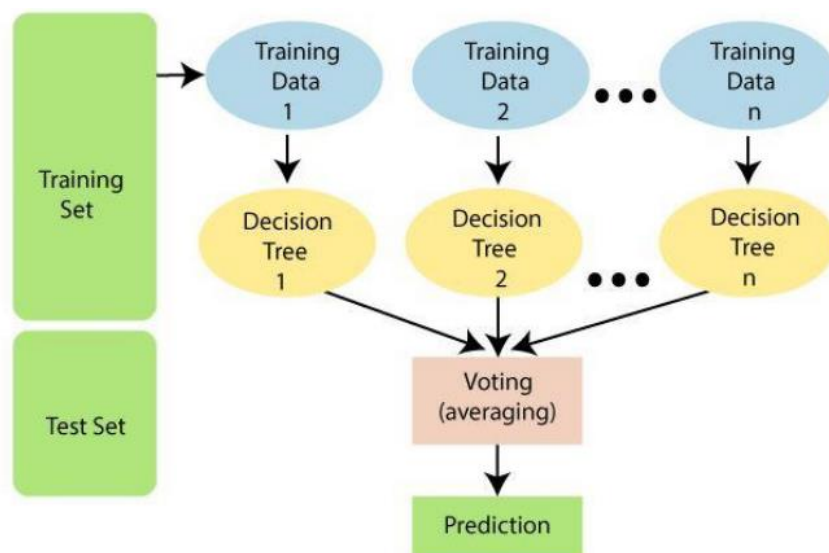
Tablica 4.2 Tablica paketa u R-u.

Naziv R paketa	Svrha paketa
<b>terra</b>	Paket se koristi za stvaranje, čitanje, manipuliranje i pisanje rasterskih podataka korištenih za razvoj specifičnih funkcija.
<b>sf</b>	Paket kodiranjem pojednostavljuje prikaz prostornog vektorskog podatka.
<b>dplyr</b>	Paket stvoren za brzu manipulaciju podataka gdje služi kao prečac za podskupove, sažimanje, preuređivanje i spajanje skupova podataka.
<b>ggplot2</b>	Paket služi za kreiranje kvalitetnih, elegantnih grafikona i slojevitih crteža.

<b>data.table</b>	Paket koji pri velikim podacima koristi alternativni način organiziranja skupova za brze operacije.
<b>caret</b>	Paket omogućuje obavljanje klasifikacije podataka.
<b>randomForest</b>	Paket za modeliranje podataka iz strojnog učenja.
<b>MLmetrics</b>	Paket procjene gubitka, rezultata i korisnosti koji mjeri regresiju, klasifikaciju i rangiranje podataka.
<b>plotly</b>	Paket pruža kvalitetne grafikone i online interakciju.
<b>LiblineaR</b>	Paket koji rješava velike linearne klasifikacije. Sadrži više klasnu klasifikaciju, unakrsnu provjeru valjanosti pri odabiru modela i brzu procjenu vjerojatnosti modela .
<b>grDevice</b>	Paket koji pruža podršku osnovnoj i mrežnoj grafici.
<b>RColorBrewer</b>	Paket nudi različite harmonije boja za karte i druge grafike.
<b>NeutralNetTools</b>	Paket s alatima vizualizacije i analize koji pomažu u interpretaciji modela s funkcijama iscertavanja, kvantificiranja važnosti varijable, provođenja analiza osjetljivosti i dobivanje jednostavnog popisa modela.
<b>mapview</b>	Paket koji brzo izrađuje prostorne karte sa ili bez pozadinskih karata, obuhvaća metode za vizualizaciju rasterskih slika u pravim i lažnim bojama i graničnih okvira.
<b>leafsync</b>	Paket kreira višestruke web-karte s brošurama koje imaju sinkroniziranu kontrolu pomicanja, povećavanja, usporedbe atributa s različitom geometrijom i mogućnost sinkroniziranja bilo koje kombinacije karata.

#### 4.4. Metoda strojnog učenja za klasifikaciju

*Random Forest* je nadzirani algoritam namijenjen klasifikaciji i regresiji. Ovaj algoritam integrira predikcije više stabala odlučivanja kako bi riješio složene probleme i poboljšao performanse modela. Ključne prednosti *Random Forest* algoritma uključuju veću točnost, mogućnost procjene važnosti varijabli, pouzdanu obradu podataka s nedostacima, brzu obuku, te prilagodljivost različitim vrstama podataka. R pruža jednostavnu i praktičnu implementaciju *Random Forest* algoritma. Postupak obuhvaća učitavanje skupa podataka, razdvajanje na skupove za treniranje i testiranje, te primjenu *Random Forest* algoritma za predikciju na testnom skupu (Slika 4.5).. Izvedba modela procjenjuje se koristeći različite metrike, uključujući preciznost i točnost (Simplilearn, 2024., Javatpoint-b, 2021.).



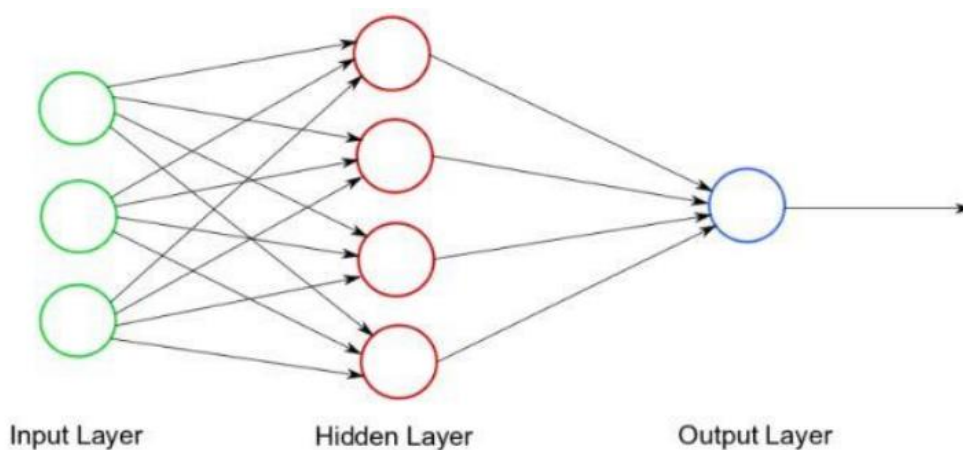
**Slika 4.5** Prikaz rada *RF*-a (Javatpoint-b, 2021.).

*Support Vector Machine (SVM)* je algoritam koji se koristi za rješavanje klasifikacijskih i regresijskih problema. *SVM* pronalazi hiper-ravninu koja najbolje razdvaja podatke u prostoru, omogućujući klasifikaciju temeljem razgraničenja između različitih skupina podataka (Slika 4.6). Prednosti *SVM*-a uključuju sposobnost rada s velikim dimenzijama, fleksibilnost u modeliranju različitih oblika klasifikacija te prilagodljivost za razne probleme. U RStudio-u, *SVM* algoritam se lako implementira. Postupak uključuje učitavanje skupa podataka, razdvajanje na skupove za treniranje i testiranje, te obuku *SVM* modela za generiranje predikcija na testnom skupu. Izvedba modela procjenjuje se koristeći mjere poput preciznosti i točnosti (UC Bussines Analytics R Programming Guide, 2023., Datacamp-b, 2018., Geeksforgeeks-a, 2021.).



**Slika 4.6** Prikaz rada *SVM*-a (Datacamp, 2018.).

*Neural Network* (neuronska mreža) je algoritam inspiriran radom neurona u ljudskom mozgu, osmišljen za rješavanje različitih problema. Oponaša način na koji neuroni međusobno komuniciraju kako bi donosili odluke (Slika 4.7). Mreža se sastoji od ulaznog sloja (koji prima neobrađene podatke), skrivenog sloja (koji obrađuje podatke) i izlaznog sloja (koji generira rezultate). U R-u, neuronske mreže se implementiraju pomoću paketa `nnet`, koji omogućuje obuku modela i analizu rezultata. Prednosti neuronskih mreža uključuju njihovu učinkovitost u rješavanju regresijskih i klasifikacijskih problema, sposobnost modeliranja nelinearnih odnosa s velikim brojem ulaza, visoku brzinu predviđanja, te fleksibilnost u treniranju s različitim brojevima ulaza i slojeva. Neuronske mreže su najefikasnije kada se koriste s velikim količinama podataka (Datacamp, 2023., Geeksforgeeks, 2020., Packt, 2024.).



**Slika 4.7** Prikaz rada *Neural Network* (Geeksforgeeks, 2020.).

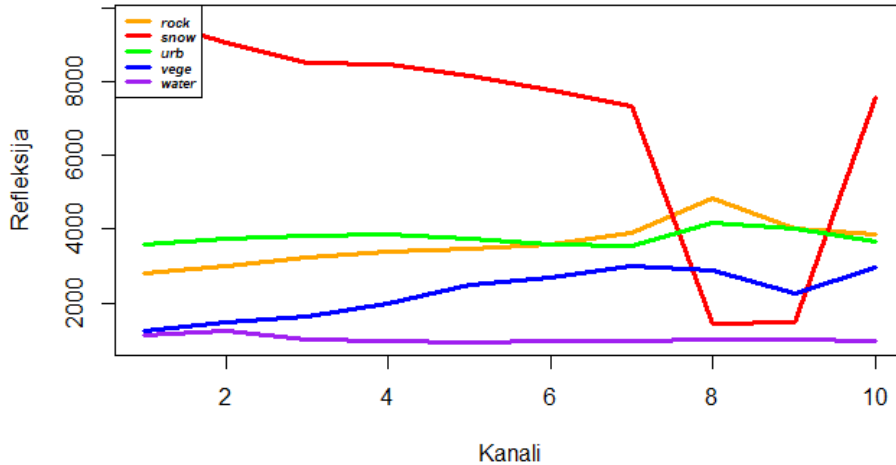
## 5. ANALIZA SPEKTRALNOG POTPISA SATELITSKE MISIJE SENTINEL-2

Grafički prikaz spektralnog potpisa za multispektralnu snimku Sentinel-2 iz zimskog i ljetnog razdoblja (28.01.2024. i 11.08.2023.) prikazuje refleksiju svake klase za pojedini kanal. Izrađene su dvije detaljne analize spektralnog potpisa pomoću satelitske misije Sentinel-2. Obje satelitske misije su usmjerene na određivanje pet klasa (snijeg, vegetacija, urbano područje, vodeni pokrov i stijene). U obje analize korišteni su podaci dobiveni kroz deset različitih kanala. Svaki kanal predstavlja određeni dio elektromagnetskog spektra i nosi informacije o različitim aspektima promatranog područja. Tablice s prikupljenim podacima učitane su i pretvorene u matrice koje sadrže odgovarajuće klase. Ove matrice su ključne za stvaranje grafičkih prikaza krivulja, čime se olakšava interpretacija spektralnih potpisa. Klase su definirane različitom bojom kako bi se jasno vizualizirali rezultati analize. Korišteni su isti kodovi za dobivanje spektralnih potpisa Sentinel-2 za pet klasa u ljetnom i zimskom razdoblju. U nastavku slijedi kod:

```
set.seed(555)
p <- st_sample( sampl, 1000, method= "regular")
p_sf <- st_sf(p)
p_at <- st_join(p_sf, sampl)
#df <- terra::extract(as(sentinel, "Raster"), p_at)
df <- terra::extract(as(s1, "Raster"), p_at)
df <- as.data.frame(df)
df$polygonID <- p_at$polygonID
df$polygon <- p_at$polygon
ms <- aggregate(df, list(p_at$polygon), mean, na.rm=TRUE)
#mjenjanje brojeva
ms <- ms[-12]
rownames(ms) <- ms[,1]
ms <- ms[,-1]
mycolor <- c('white', 'red', 'green', 'blue', 'purple', 'black')
ms <- as.matrix(ms)
plot(0, ylim=c(min(ms),max(ms)), xlim = c(1,10), type='n', xlab="Kanali", ylab = "Refleksija")
for (i in 1:nrow(ms)){lines(ms[i,], type = "l", lwd = 3, lty = 1, col = mycolor[i])}
title(main="Spektralni potpis klasifikacija na karti Sentinel-2-ZIMA", font.main = 2)
legend("topleft", rownames(ms), cex=0.55, text.font=4, col=mycolor, lty = 1, lwd =3, bty = "y")
```

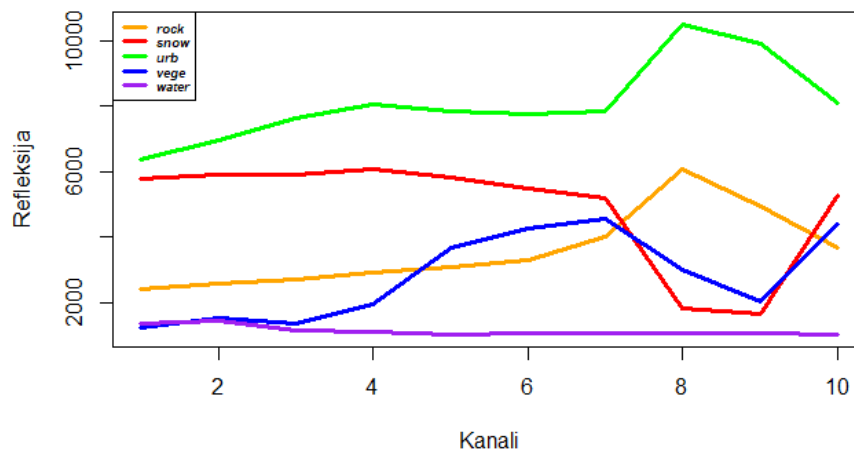
Dobiveni grafički prikaz spektralnog potpisa pet klasa u zimskom razdoblju za multispektralne snimke Sentinel-2 (Slika 5.1.). Sve krivulje imaju 5 klasa definirane različitim bojama, snijeg (crveno), stijena (narančasto), urbano područje (zeleno), vegetacija (plavo), voda (ljubičasto). Svaka krivulja ima karakterističan oblik i postižu različite trendove u pojedinim kanalima. Najizraženija je krivulja snijega (crvena) zbog značajnog pada u odnosu na većinu krivulja koje imaju sličan trend. Krivulja snijega ima značajan pad kod osmog kanala te nagli rast kod desetog kanala. Ostale klase nemaju izraženo povećanje i padanje trendova.





Slika 5.1 Spektralni potpis zimskog razdoblja Sentinel-2.

Spektralni potpis ljetnog razdoblja snimke Sentinel-2 prikazan je također za pet vrsta klasifikacije (Slika 5.2). Sve krivulje su definirane različitim bojama: snijeg (crveno), stijena (narančasto), urbano područje (zeleno), vegetacija (plavo) i voda (ljubičasto). Sve krivulje imaju karakterističan oblik. Za razliku od spektralnog potpisa iz zimskog razdoblja Sentinel-2, krivulje iz ljetnog razdoblja pokazuju različite trendove. Krivulja snijega ima trend opadanja kod sedmog kanala te ponovnog rasta kod desetog kanala. Krivulja stijene pokazuje lagani rast od sedmog kanala te ponovno lagani pad od osmog kanala. Krivulja vegetacije ima mali porast, pad od sedmog kanala te lagani porast od devetog kanala. Najizraženija je krivulja urbanog područja koja ima najveću vrijednost refleksije i rast trenda od sedmog kanala te lagani pad od osmog kanala. Krivulja vode nema trend rasta ni pada u zimskom i ljetnom spektralnog potpisu Sentinel-2.



Slika 5.2 Spektralni potpis ljetnog razdoblja Sentinel-2.

## 6. OBRADA PODATAKA METODAMA STROJNOG UČENJA

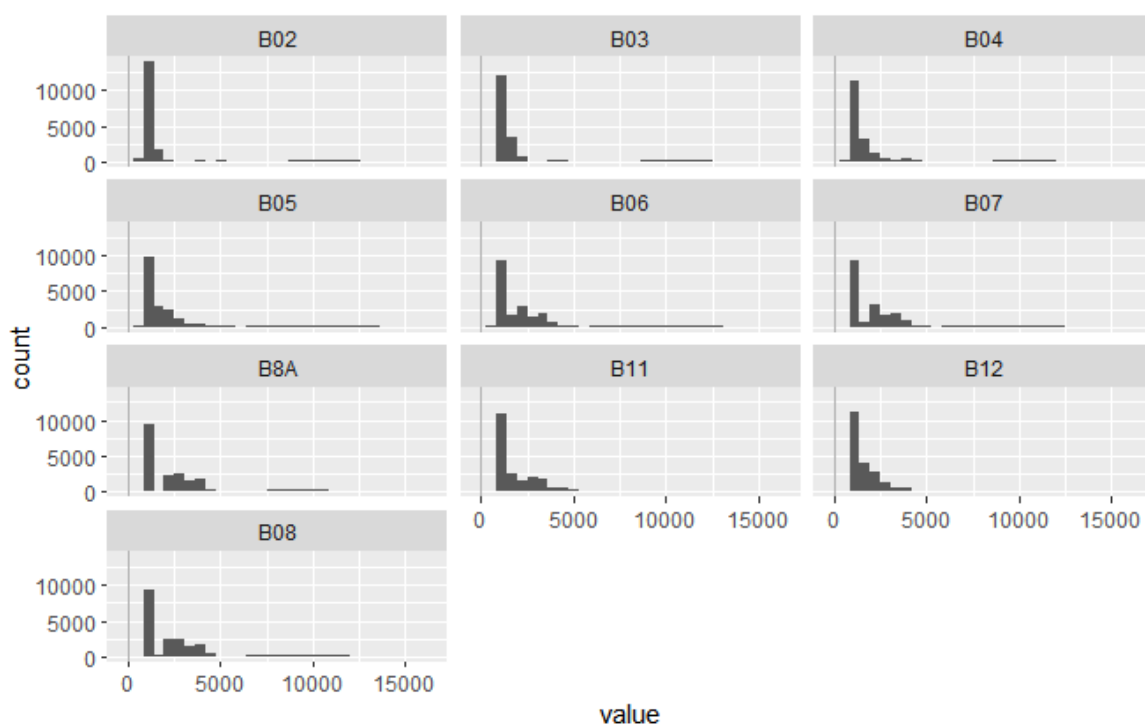
Primjenom Copernicus Open Access Hub-a preuzimaju se multispektralne snimke satelita Sentinel-2 za odabrano područje istraživanja. Definirano je 10 kanala satelitskih snimaka. Korištenjem GADM sustava, koji nudi karte i prostorne podatke svih zemalja, učitane su granice općina (razina 2) u Italiji i Austriji iz baze podataka. Izdvojena su područja istraživanja koja uključuju općine u Italiji (Bolzano, Trento, Belluno, Treviso, Pordenone, Udine) i kotari u Austriji (Lienz, Spittal an der Drau, Hermagor) iz većeg skupa geoprostornih podataka. Izdvojeno područje na temelju zadanih granica koristi se za maskiranje vrijednosti *Raster* objekta izvan područja istraživanja. Ovo područje preklapa se s vektorskim objektom koji predstavlja poligone različitih klasa (snijeg, vegetacija, urbano područje, vode, stijena). Korištenjem funkcija za izdvajanje pojedinačnih informacija o vrijednostima piksela, stvorena je tablica koja sadrži vrijednosti svih kanala i pripadajuću klasu za svaki piksel. Zbog velike količine numeričkih podataka dobivenih za poligone pet klasa na obje satelitske slike, prikazani su samo poligoni snijega i vegetacije. Prikaz nekoliko elemenata i struktura dobivene tablice nalazi se u tablici 6.1., koja sadrži ukupno 20472 elementa, te u tablici 6.2., koja sadrži ukupno 13630 elemenata. Te tablice služe kao ulazni podaci za korištenje algoritama strojnog učenja u ovom radu. Na slikama 6.1. i 6.2. prikazani su histogrami za svaki kanal multispektralnih snimaka za ljeto i zimu.

**Tablica 6.1** Prikaz vrijednost dobivene tablice za zimsko razdoblje.

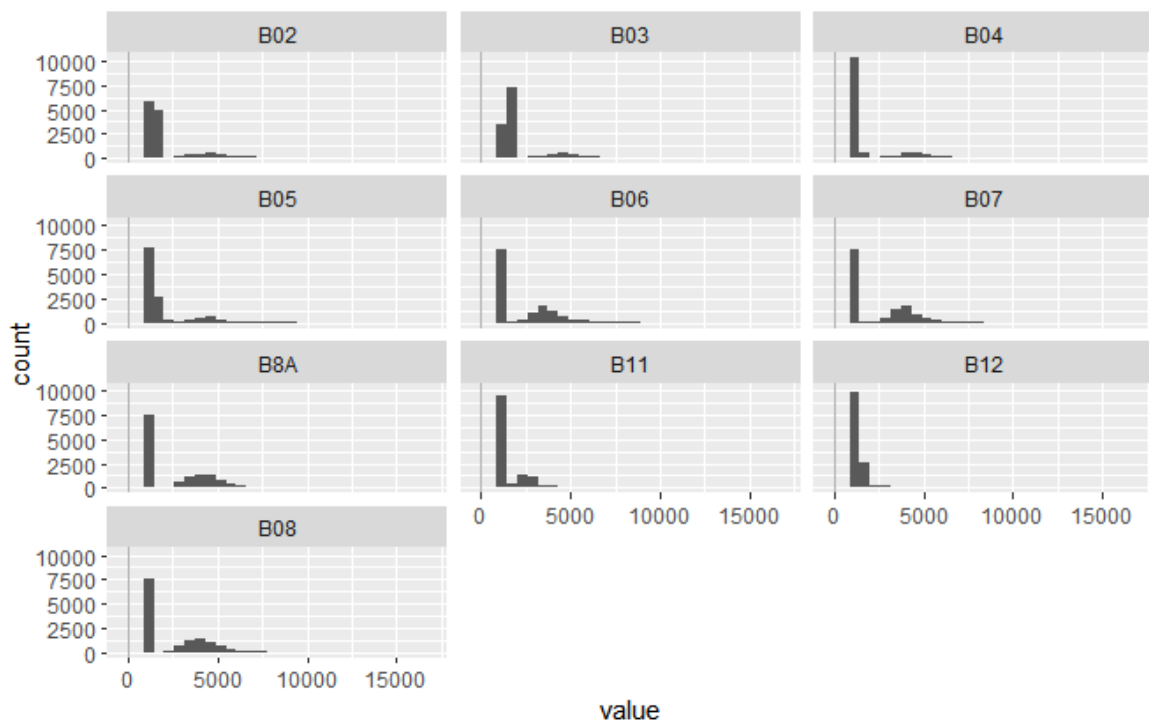
	B02	B03	B04	B05	B06	B07	B8A	B11	B12	B08	Klasa
1:	10036	10316	10413	10646	10117	9824	9473	2139	2127	10217	snow
2:	9373	9556	9829	10186	9900	9710	9400	2005	2011	9823	snow
3:	11466	11735	11973	12383	12008	11370	10960	2115	2140	11609	snow
4:	11075	11270	11507	11673	11444	10842	10516	2075	2126	11119	snow
5:	10709	10991	11131	11336	11011	10723	10379	2008	2061	10766	snow
-----											
20468:	1644	1948	2229	2693	3303	3461	3839	3540	2657	3769	vege
20469:	1697	1986	2323	2800	3372	3567	3922	3776	2815	4017	vege
20470:	1696	1999	2317	2834	3363	3590	3909	3815	2861	3873	vege
20471:	1712	2037	2352	2828	3370	3605	3885	3838	2892	4000	vege
20472:	1698	2009	2295	2756	3317	3521	3838	3713	2822	3797	vege

**Tablica 6.2** Prikaz vrijednost dobivene tablice za ljetno razdoblje

	B02	B03	B04	B05	B06	B07	B8A	B11	B12	B08	Klasa
1:	4249	4881	5539	5816	5825	5963	6024	6957	5578	6192	snoww
2:	4329	4888	5459	5685	5758	5827	5906	6792	5391	6015	snow
3:	4656	5171	5715	5922	6056	6124	6289	7082	5659	6404	snow
4:	4613	5093	5582	5909	5930	6064	6235	6966	5521	5999	snow
5:	4429	4902	5341	5556	5590	5710	5845	6702	5348	5696	snow
-----											
13626:	1250	1641	1324	2123	4661	5213	5585	2911	1940	5837	vege
13627:	1201	1630	1248	2091	5008	5775	6244	2923	1842	6108	vege
13628:	1223	1639	1250	2095	5106	5933	6217	2869	1835	6004	vege
13629:	1232	1639	1255	2137	5105	6002	6182	2899	1843	6028	vege
13630:	1231	1639	1269	2127	5070	5954	6131	2870	1833	6037	vege



**Slika 6.1** Prikaz histograma multispektralnih snimaka za zimsko razdoblje.



**Slika 6.2** Prikaz histograma multispektralnih snimaka za ljetno razdoblje.

Korištenjem funkcije „*createDataPartition()*“ iz paketa *caret*, početni skup podataka dijeli se na skup za treniranje i skup za testiranje. Skup za treniranje „*dt\_train*“ sastoji se od odabranih redaka izvornog skupa podataka koji se koriste za treniranje modela (Tablica 6.3. i Tablica 6.5.), dok se u skupu za testiranje „*dt\_test*“ koriste preostali redci izvornog skupa podataka za testiranje modela (Tablica 6.4. i Tablica 6.6.). U nastavku slijedi kod:

```
set.seed(321)
idx_train <- createDataPartition(dt$class,
                                p = 0.7, # percentage of data as training
                                list = FALSE)
dt_train <- dt[idx_train]
dt_test <- dt[-idx_train]

table(dt_train$class)
table(dt_test$class)
```

**Tablica 6.3** Količina treninga podataka klasa iz zimskog razdoblja.

Rock	Snow	Urb	Vege	Water
165	2459	452	4694	6563

**Tablica 6.4** Količina testa podataka klasa iz zimskog razdoblja

Rock	Snow	Urb	Vege	Water
70	1053	193	2011	2812

**Tablica 6.5** Količina treninga podataka iz ljetnog razdoblja.

Rock	Snow	Urb	Vege	Water
30	2089	26	2142	5257

**Tablica 6.6** Količina testa podataka iz ljetnog razdoblja.

Rock	Snow	Urb	Vege	Water
12	894	11	917	2252

Stvaraju se preklopni skupovi za slučajni generator koji se koriste tijekom treniranja i evaluacije modela. Koristeći naredbu „*n\_folds*“, definira se broj preklopnih skupova i generira se popis za svaki od njih. Funkcija „*trainControl()*“ omogućuje upravljanje procesom učenja i procjene modela. Ova funkcija primjenjuje strategiju unakrsne potvrde „*cv*“ za ocjenu pouzdanosti i izvedbe modela, određuje broj presavijanja („*number*“), način pretrage hiperparametara „*search*“ te način spremanja predikcija „*savePredictions*“. Za izračunavanje vjerojatnosti klasa koristi se „*classProbs*“, dok se „*seeds*“ koriste za kontrolu nasumičnosti procesa. U nastavku slijedi kod:

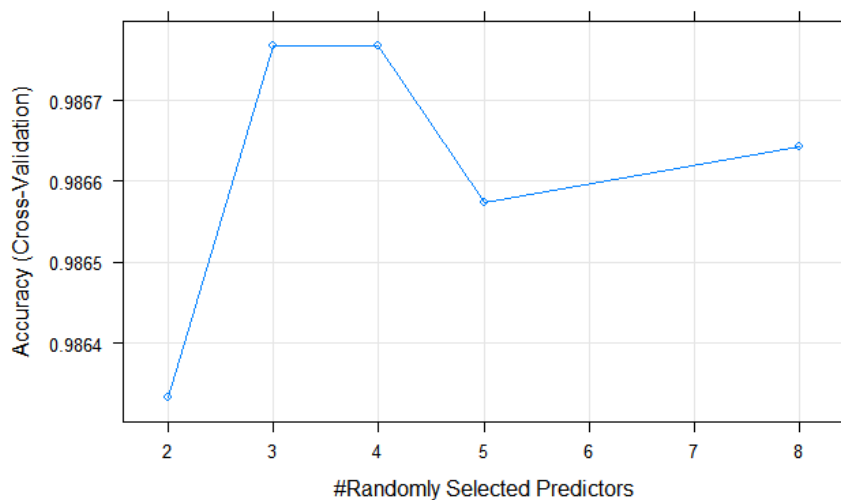
```
n_folds <- 10
set.seed(321)
folds <- createFolds(1:nrow(dt_train), k = n_folds)
seeds <- vector(mode = "list", length = n_folds ) # +1 for the final model nnet
for(i in 1:n_folds) seeds[[i]] <- sample.int(1000, n_folds)
seeds[n_folds + 1] <- sample.int(1000, 1) # seed for the final model
ctrl <- trainControl(summaryFunction = multiClassSummary,
                     method = "cv",
                     number = n_folds,
                     search = "grid",
                     classProbs = TRUE, # not implemented for SVM; will just get a warning
                     savePredictions = TRUE,
                     index = folds,
                     seeds = seeds)
```

## 6.1. Korištenje metode *randomForest* (RF) iz paketa CARET

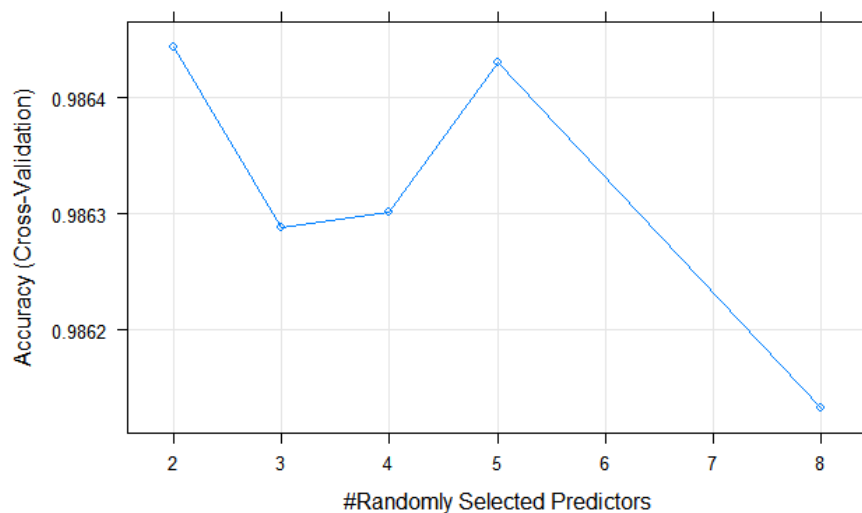
Primjenjuje se metoda *randomForest* (RF). Varijabla „*model\_rf*“ koristi funkciju „*train()*“ za obučavanje modela slučajne šume, optimizira hiperparametre modela putem unakrsne provjere i procjenjuje važnost prediktorskih varijabli. Nakon završetka treniranja,

funkcija vraća model zajedno s njegovom procjenom izvedbe. Funkcija „`model_rf$times$everything`“ prikazuje vrijeme potrebno za svaku iteraciju procesa preklapanja, uključujući prilagodbu i predviđanje modela. Rezultat je grafički prikaz koji prikazuje kako točnost opada s povećanjem broja nasumičnih predikcija (Slika 6.3. i 6.4.).

```
model_rf <- caret::train(class ~ . , method = "rf", data = dt_train,
  importance = TRUE, # passed to randomForest()
  # run CV process in parallel;
  # see https://stackoverflow.com/a/44774591/5193830
  #allowParallel = TRUE,
  tuneGrid = data.frame(mtry = c(2, 3, 4, 5, 8)),
  trControl = ctrl)
saveRDS(model_rf, file = "~/model_rf3.rds")
model_rf
model_rf$times$everything
plot(model_rf)
```



Slika 6.3 Grafički prikaz predikcije RF-a iz razdoblja zime



Slika 6.4 Grafički prikaz predikcije *RF*-a iz razdoblja ljeta.

Funkcija „*confusionMatrix()*“ za varijablu *cm\_rf* generira matricu zabune na temelju modela slučajne šume. Ova matrica prikazuje broj točnih i netočnih predikcija za svaku klasu ciljne varijable (Tablica 6.7 i Tablica 6.8). Objekt *model\_rf\$finalModel* predstavlja obučeni model slučajne šume i koristi se za predviđanje ciljne varijable za nove podatke pomoću naredbe „*predict()*“. U nastavku slijedi kod:

```
cm_rf <- confusionMatrix(data = predict(model_rf, newdata = dt_test), as.factor(dt_test$class))
cm_rf
model_rf$finalModel
```

Tablica 6.7 Statistički prikaz *random forest* klasifikacije za razdoblje zime.

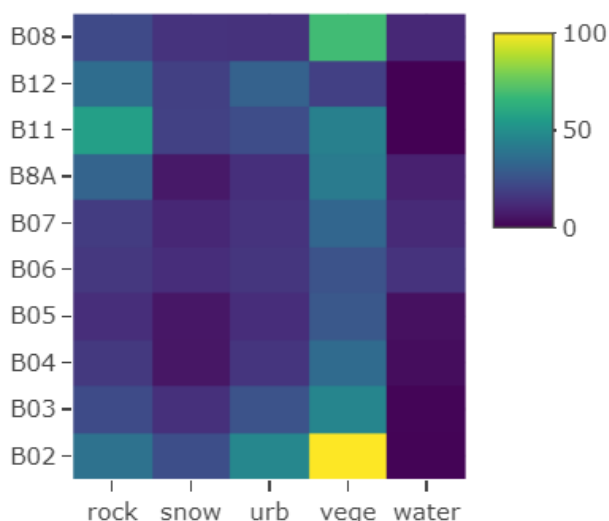
Statistics by Class:	Class: rock	Class: snow	Class: urb	Class: vege	Class: water
Sensitivity	0.95714	1.0000	0.83938	0.9990	0.9957
Specificity	0.99967	0.9998	0.99899	0.9910	0.9994
Pos Pred Value	0.97101	0.9991	0.96429	0.9819	0.9993
Neg Pred Value	0.99951	1.0000	0.99481	0.9995	0.9964
Prevalence	0.01140	0.1715	0.03144	0.3276	0.4581
Detection Rate	0.01091	0.1715	0.02639	0.3273	0.4561
Detection Prevalence	0.01124	0.1717	0.02737	0.3333	0.4564
Balanced Accuracy	0.97841	0.9999	0.91918	0.9950	0.9976

Tablica 6.8 Statistički prikaz *random forest* klasifikacije za razdoblje ljeta.

Statistics by Class:	Class: rock	Class: snow	Class: urb	Class: vege	Class: water
Sensitivity	0.666667	0.9888	0.727273	0.9935	0.9991
Specificity	0.999755	0.9969	0.999755	0.9959	1.0000
Pos Pred Value	0.888889	0.9888	0.888889	0.9859	1.0000
Neg Pred Value	0.999019	0.9969	0.999264	0.9981	0.9989
Prevalence	0.002937	0.2188	0.002692	0.2244	0.5512
Detection Rate	0.001958	0.2163	0.001958	0.2230	0.5507
Detection Prevalence	0.002203	0.2188	0.002203	0.2261	0.5507
Balanced Accuracy	0.833211	0.9928	0.863514	0.9947	0.9996

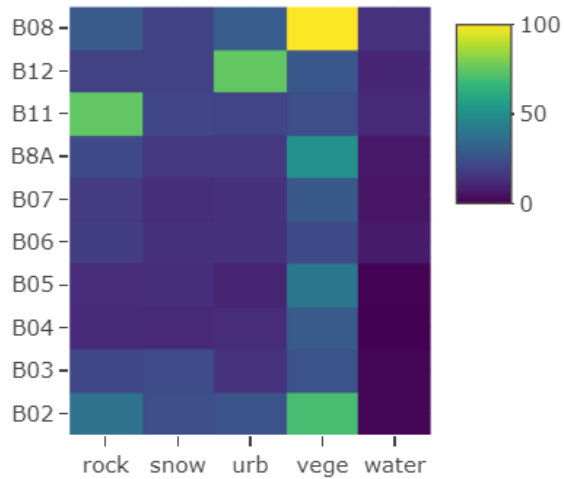
Funkcija „*varImp()*“ pretvara izračunate vrijednosti u matricu. Rezultat je *heatmap* grafikon koji prikazuje značajnost svake komponente za klasifikaciju ciljne varijable (Slika 6.5. i Slika 6.6.). Tamnija boja (plava) označava veću značajnost za određeni kanal, dok svjetlija boja (žuta) označava manju značajnost.

```
caret::varImp(model_rf)$importance %>%
  as.matrix %>%
  plot_ly(x = colnames(.), y = rownames(.), z = ., type = "heatmap",
          width = 350, height = 300)
```



Slika 6.5 Prikaz *heatmap* 1 razdoblje zima.





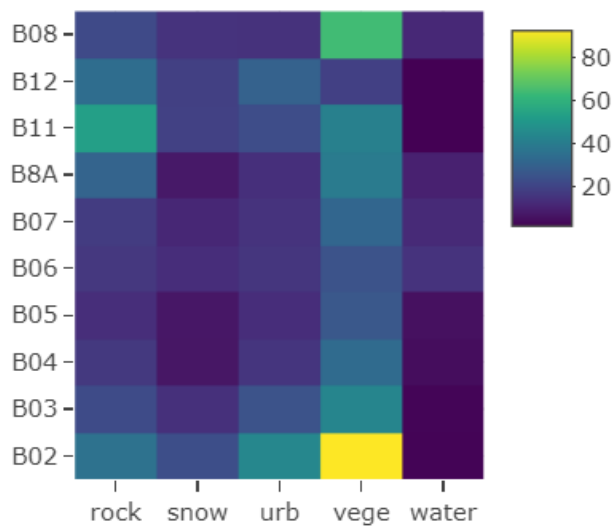
Slika 6.6 Prikaz heatmap 1 razdoblje ljeto.

Izračunavanje važnosti značajki u modelu slučajne šume provodi se pomoću funkcije „*importance()*“ iz paketa *randomForest*. Izračunate vrijednosti pretvaraju se u matricu, a zatim se stvara heatmap grafikon koristeći naredbu „*plot\_ly()*“ iz paketa *plotly* (Slika 6.7. i Slika 6.8.). U nastavku slijedi kod:

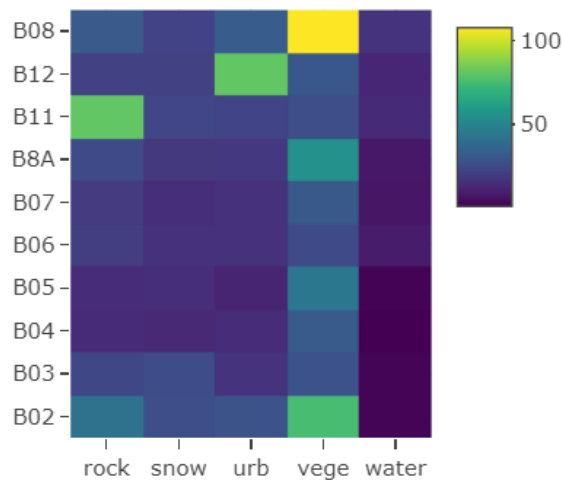
```

randomForest::importance(model_rf$finalModel) %>%
  .[, - which(colnames(.) %in% c("MeanDecreaseAccuracy",
    "MeanDecreaseGini"))] %>%
  plot_ly(x = colnames(.), y = rownames(.), z = ., type =
    "heatmap",
    width = 350, height = 300)

```



Slika 6.7 Slika heatmap 2 razdoblje zime.



Slika 6.8 Slika *heatmap* 2 razdoblje ljeto.

## 6.2. Korištene metode *Support Vector Machine (SVM)* iz paketa *CARET*

Sljedeća metoda koja se koristi uključuje stvaranje mreže parametara za algoritam poznat kao *Support Vector Machine (SVM)*. Ova metoda pomaže u pronalaženju optimalnih vrijednosti parametara, čime poboljšava rezultate analize podataka. Korištenjem naredbe „*expand.grid()*“, kombiniraju se tri vrijednosti za parametar „*cost*“ (0.2, 0.5, 1) s dvije vrijednosti za parametar „*Loss*“ ("L1" i "L2"), čime se stvara mreža od šest kombinacija. Model se obučava koristeći skup podataka i linearni *SVM* algoritam s ciljem pronalaženja najboljih vrijednosti parametara za poboljšanje prilagodbe modela podacima. Za podešavanje hiperparametara koristi se mreža „*svm\_grid*“, za kontrolu izvedbe modela koristi se kontrola treniranja „*trControl*“, a za obuku modela koristi se funkcija „*caret::train()*“. Podskup podataka sadrži informacije o vremenu izvođenja svakog koraka u procesu obuke *SVM* modela. Naredba „*plot(model\_svm)*“ prikazuje grafički prikaz izvedbe modela obučenog pomoću objekta „*model\_svm*“ (Slika 6.9. i Slika 6.10.). Izračunava se matrica zabune koja prikazuje koliko je model ispravno i pogrešno predvidio svaku klasu, na temelju predviđanja modela na testnom skupu podataka i stvarnih vrijednosti ciljne varijable „*dt\_test\$class*“ (Tablica 6.9 i Tablica 6.10). U nastavku slijedi kod:

```

svm_grid <- expand.grid(cost = c(0.2, 0.5, 1),
                      Loss = c("L1", "L2"))
model_svm <- caret::train(class ~ . , method = "svmLinear3",
data = dt_train,
                        allowParallel = TRUE,
                        tuneGrid = svm_grid,
                        trControl = ctrl)
saveRDS(model_svm, file = "~/model_svm3.rds")

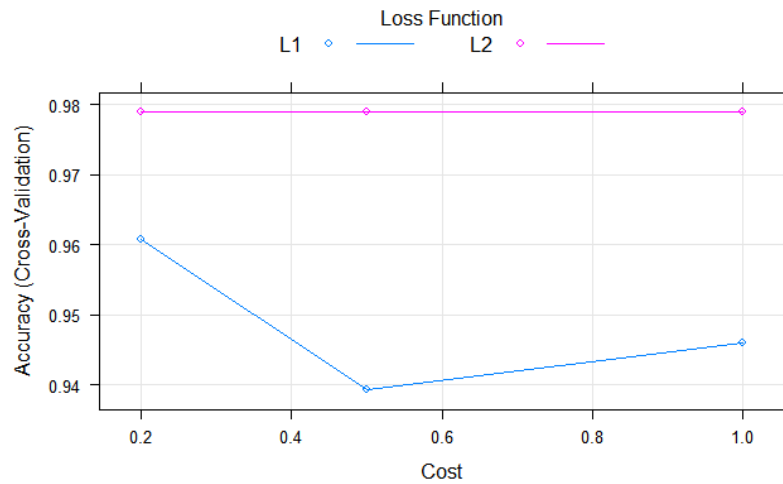
```

```

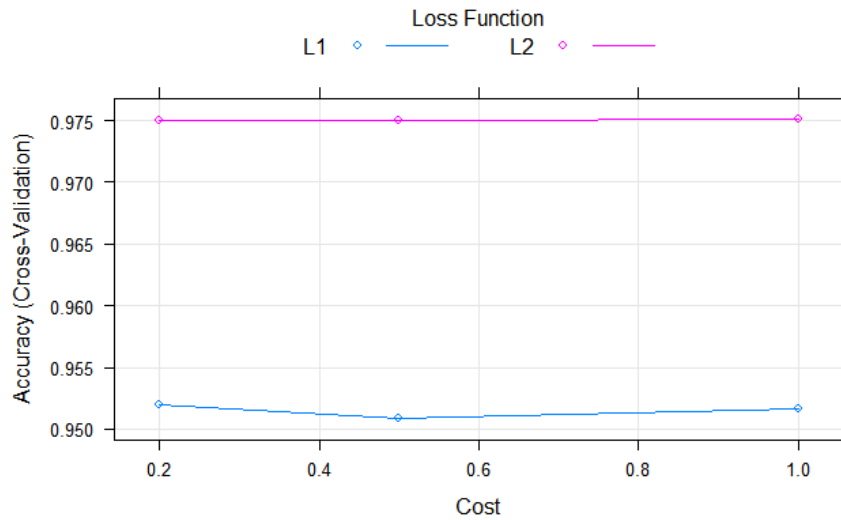
model_svm$times$everything
plot(model_svm)

m_svm <- confusionMatrix(data = predict(model_svm, newdata =
dt_test), as.factor(dt_test$class))
cm_svm

```



Slika 6.9 Grafički prikaz modela SVM-a razdoblje zime.



Slika 6.10 Grafički prikaz modela SVM-a razdoblje ljeta.

Tablica 6.9 Statistički prikaz SVM klasifikacije razdoblje zime.

Statisticsby Class:	Class: rock	Class: snow	Class: urb	Class: vege	Class: water
Sensitivity	0.757143	1.0000	0.53368	0.9995	0.9950
Specificity	0.997858	1.0000	0.99849	0.9813	0.9931
Pos Pred Value	0.803030	1.0000	0.91964	0.9631	0.9918
Neg Pred Value	0.997201	1.0000	0.98507	0.9998	0.9958
Prevalence	0.011403	0.1715	0.03144	0.3276	0.4581
Detection Rate	0.008633	0.1715	0.01678	0.3274	0.4558
Detection Prevalence	0.010751	0.1715	0.01824	0.3400	0.4595
Balanced Accuracy	0.877500	1.0000	0.76608	0.9904	0.9941

Tablica 6.10 Statistički prikaz SVM klasifikacije razdoblje ljeta.

Statisticsby Class:	Class: rock	Class: snow	Class: urb	Class: vege	Class: water
Sensitivity	0.500000	0.9318	0.0909091	0.9738	0.9987
Specificity	0.999264	0.9890	1.0000000	0.9867	0.9869
Pos Pred Value	0.666667	0.9597	1.0000000	0.9551	0.9894
Neg Pred Value	0.998528	0.9810	0.9975520	0.9924	0.9983
Prevalence	0.002937	0.2188	0.0026921	0.2244	0.5512
Detection Rate	0.001468	0.2039	0.0002447	0.2186	0.5504

Detection Prevalence	0.002203	0.2124	0.0002447	0.2288	0.5563
Balanced Accuracy	0.749632	0.9604	0.5454545	0.9803	0.9928

### 6.3. Korištenje metode *Neural Network (NNET)* iz paketa CARET

Posljednja metoda koja se koristi je *Neural Network (NNET)*. Ova metoda omogućuje prikupljanje informacija o treniranom modelu kako bi se procijenila njegova izvedba. Stvara se mreža svih kombinacija vrijednosti parametara temeljenih na definiranim argumentima. Tri veličine sloja (*size*) koje predstavljaju broj neurona u sloju (5, 10 ili 15) kombiniraju se s tri stope degradacije *decay* koje određuju brzinu smanjenja težina mreže tijekom treniranja (0.001, 0.01 ili 0.1). Model se trenira koristeći sve dostupne značajke iz varijable „*class*“. Parametri modela koje će se ispitati postavljeni su varijablom *nnet\_grid*. Tijekom treniranja i evaluacije izvedbe modela primjenjuje se unakrsna validacija. Naredba „*plot(model\_nnet)*“ generira različite dijagrame koji prikazuju izvedbu modela (Slika 6.11. i Slika 6.12.). Nakon toga, ponovno se izračunavaju matrice zabune na temelju predviđanja modela i stvarnih vrijednosti iz testnog skupa (Tablica 6.8. i Tablica 6.9.).

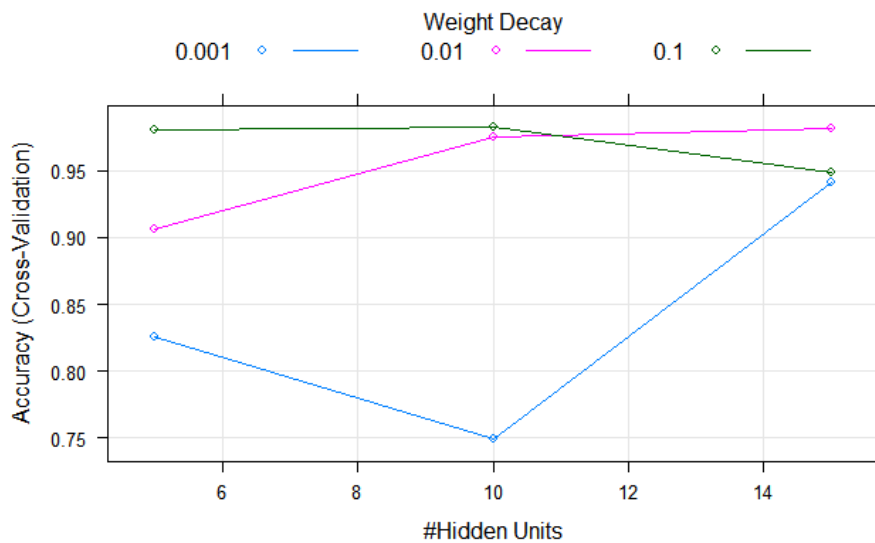
```
nnet_grid <- expand.grid(size = c(5, 10, 15),
                        decay = c(0.001, 0.01, 0.1))

model_nnet <- train(class ~ ., method = 'nnet', data = dt_train,
                   importance = TRUE,
                   maxit = 1000, # set high enough so to be sure that it converges
                   allowParallel = TRUE,
                   tuneGrid = nnet_grid,
                   trControl = ctrl)

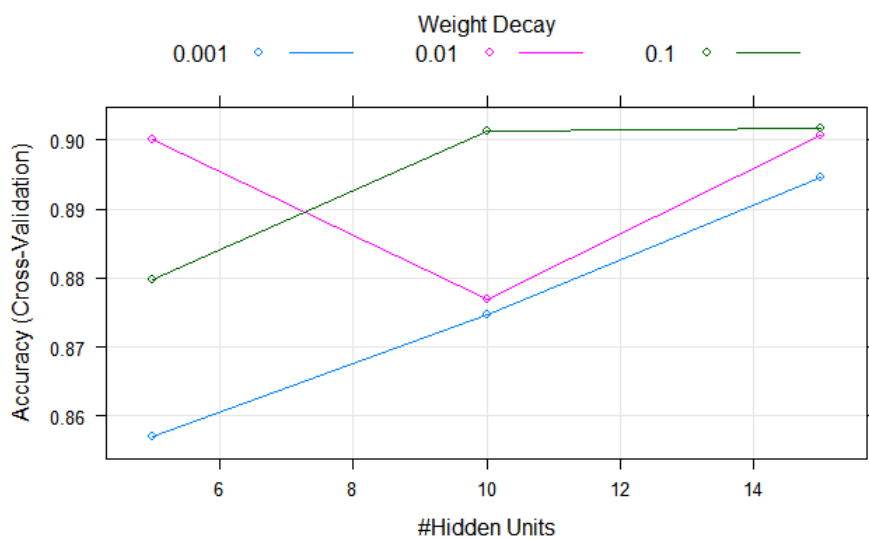
saveRDS(model_nnet, file = "~/model_nnetALPE.rds")

model_nnet$times$everything # total computation time
plot(model_nnet) # tuning results

cm_nnet <- confusionMatrix(data = predict(model_nnet,
newdata = dt_test), as.factor(dt_test$class))
cm_nnet
```



Slika 6.11 Grafički prikaz modela *Neural Network* za razdoblje zime.



Slika 6.12 Grafički prikaz modela *Neural Network* za razdoblje ljeta.

**Tablica 6.11** Statistički prikaz *nnet* klasifikacije iz razdoblja zime.

<b>Statisticsby Class:</b>	<b>Class: rock</b>	<b>Class: snow</b>	<b>Class: urb</b>	<b>Class: vege</b>	<b>Class: water</b>
Sensitivity	0.92857	1.0000	0.83938	0.9960	0.9968
Specificity	0.99984	1.0000	0.99832	0.9925	0.9967
Pos Pred Value	0.98485	1.0000	0.94186	0.9848	0.9961
Neg Pred Value	0.99918	1.0000	0.99480	0.9981	0.9973
Prevalence	0.01140	0.1715	0.03144	0.3276	0.4581
Detection Rate	0.01059	0.1715	0.02639	0.3263	0.4566
Detection Prevalence	0.01075	0.1715	0.02802	0.3313	0.4584
Balanced Accuracy	0.96420	1.0000	0.91885	0.9943	0.9967

**Tablica 6.12** Statistički prikaz *nnet* klasifikacije iz razdoblja ljeto.

<b>Statisticsby Class:</b>	<b>Class: rock</b>	<b>Class: snow</b>	<b>Class: urb</b>	<b>Class: vege</b>	<b>Class: water</b>
Sensitivity	0.583333	0.9586	0.000000	0.9782	0.9982
Specificity	0.999018	0.9897	1.000000	0.9880	0.9989
Pos Pred Value	0.636364	0.9629	NaN	0.9594	0.9991
Neg Pred Value	0.998773	0.9884	0.997308	0.9937	0.9978
Prevalence	0.002937	0.2188	0.002692	0.2244	0.5512
Detection Rate	0.001713	0.2097	0.000000	0.2195	0.5502
Detection Prevalence	0.002692	0.2178	0.000000	0.2288	0.5507
Balanced Accuracy	0.791176	0.9741	0.500000	0.9831	0.9986

Funkcija za interpolaciju boja `grDevices::colorRampPalette` koristi naredbu „`brewer.pal`“ iz paketa `RColorBrewer` za kreiranje palete od 10 različitih boja odabranih za vizualizaciju podataka i generiranje grafova. Važnost značajki u neuronskoj mreži računa se pomoću funkcije „`garson(model_nnet)`“, koja je obučena na skupu podataka. Naredbe „`scale_y_continuous`“ i „`scale_fill_gradientn`“ koriste se za prilagodbu osi y i gradijenta boja u izlaznom grafu. Obučeni `model_nnet` omogućuje prikaz neuronske mreže (Slika 6.13. i Slika 6.14). Boje koje odgovaraju funkciji „`cols`“ koriste se za prikaz dijagrama neuronske mreže, pri čemu je poredak sortiran prema važnosti neurona (Slika 6.15. i Slika 6.16). Za obučeni model, kanali imaju značajan utjecaj na ciljani rezultat, pri čemu se pouzdanost smanjuje s desna na lijevo. Poveznice između neuronske mreže i

dijagrama prikazane su bojom kanala, gdje tamnije boje označavaju pouzdanije kanale u usporedbi sa svjetlijim bojama. U nastavku slijedi kod:

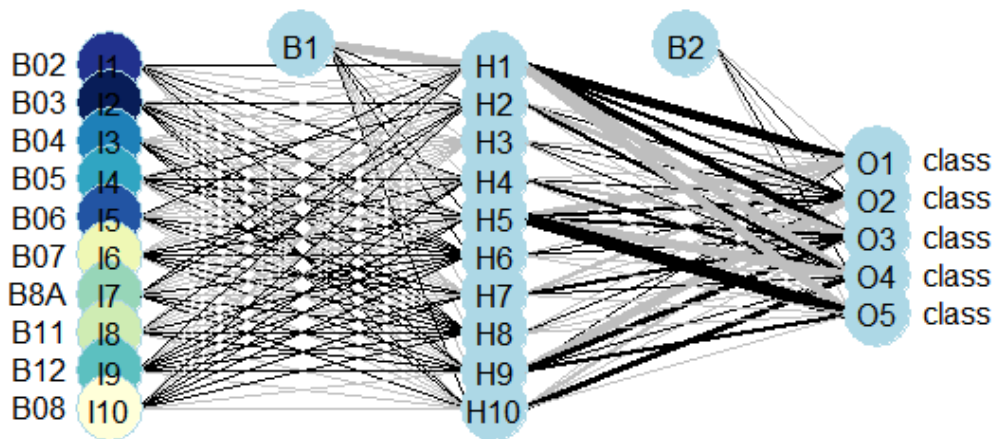
```

cols <- grDevices::colorRampPalette(colors = brewer.pal(n = 9,
name = "YlGnBu"))(10)

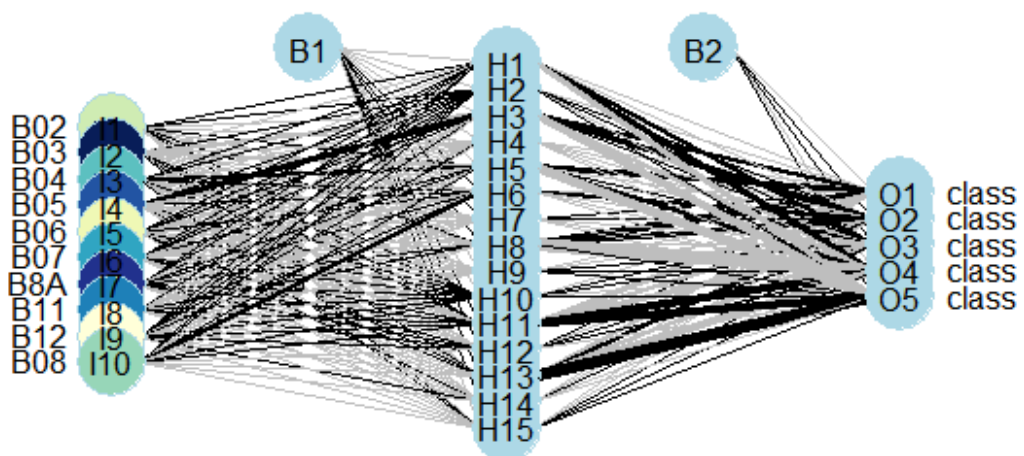
garson(model_nnet) +
  scale_y_continuous('Rel. Importance') +
  scale_fill_gradientn(colours = cols)

cols_rank_import <- cols[rank(garson(model_nnet, bar_plot =
FALSE)$rel_imp)]
plotnet(model_nnet, circle_col = list(cols_rank_import,
'lightblue'))

```

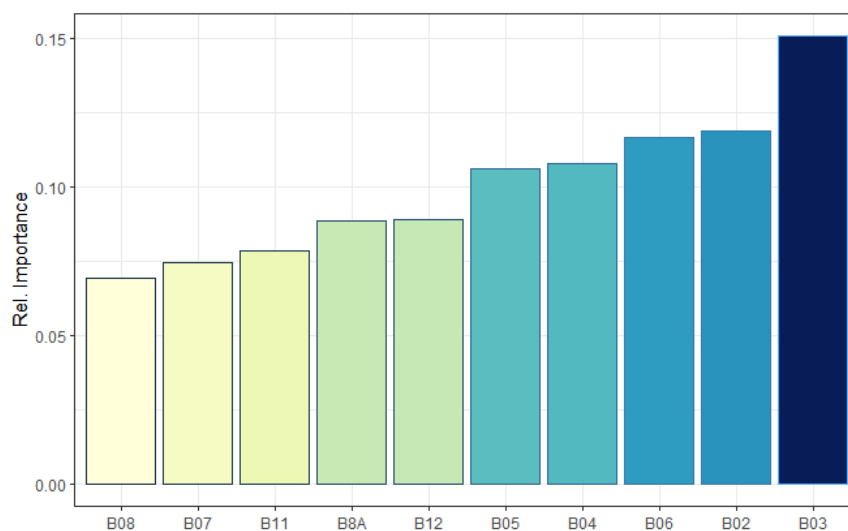


Tablica 6.13 Prikaz neuronske mreže za razdoblje zime.

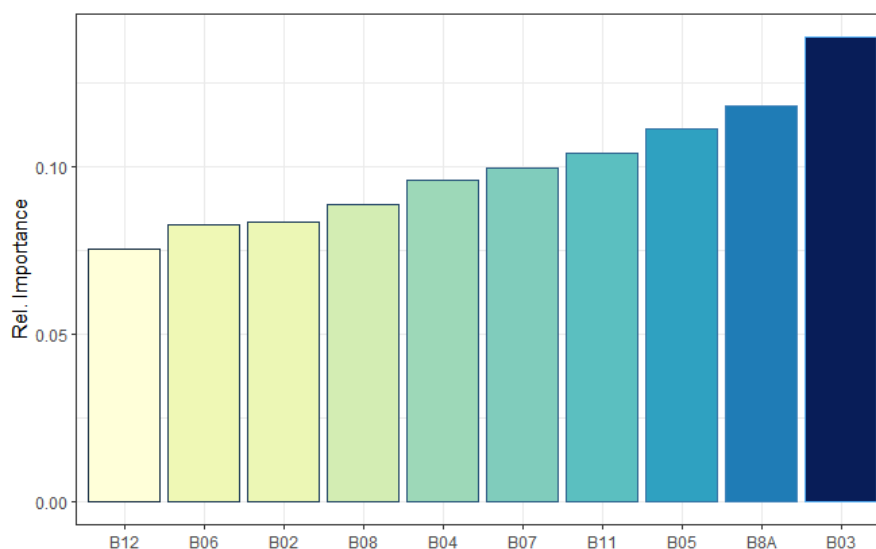


Tablica 6.14 Prikaz neuronske mreže za razdoblje ljeta.





Slika 6.15 Dijagram neuronske mreže zimskog razdoblja.



Slika 6.16 Dijagram neuronske mreže ljetnog razdoblja.

#### 6.4. Usporedba dobivenih modela

Stvara se lista modela za usporedbu izvedbe koja uključuje tri različita modela: „*rf*“ (*Random Forest*), „*svm*“ (*SVM*) i „*nnet*“ (*Neural Network*). Funkcijom *caret::resamples* provodi se postupak unakrsne validacije za svaki model u listi (Slika 6.17. i Slika 6.18.). Naredba „*bwplot()*“ generira (*box-and-whisker*) grafove za svaki model u objektu *resamples*. Funkcija *resamples\$metrics* vraća tablicu koja sadrži različite vrijednosti za svaki model. Razlike između svih kombinacija modela u listi izračunavaju se pomoću funkcije „*diff*“ iz paketa *caretEnsemble*, dok se izračunavanje *p*-vrijednosti i konstrukcija tablice sažetka provode naredbom „*summary*“. Rezultati su pouzdaniji što su bliži

vrijednosti jedan za određene funkcije. Model „*nnet*“ postiže najbolje rezultate za funkcije *recall* i *sensitivity*, dok su rezultati ostalih modela na sličnoj razini. Kod:

```

model_list <- list(rf = model_rf, svm = model_svm, nnet =
model_nnet)
model_list <- list(rf = model_rf, nnet = model_nnet)

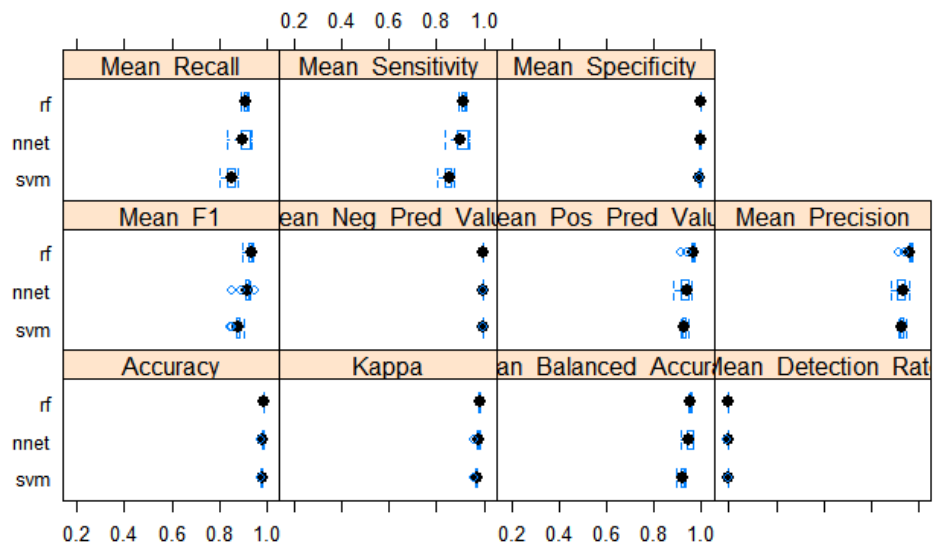
resamples <- caret::resamples(model_list)

bwplot(resamples)

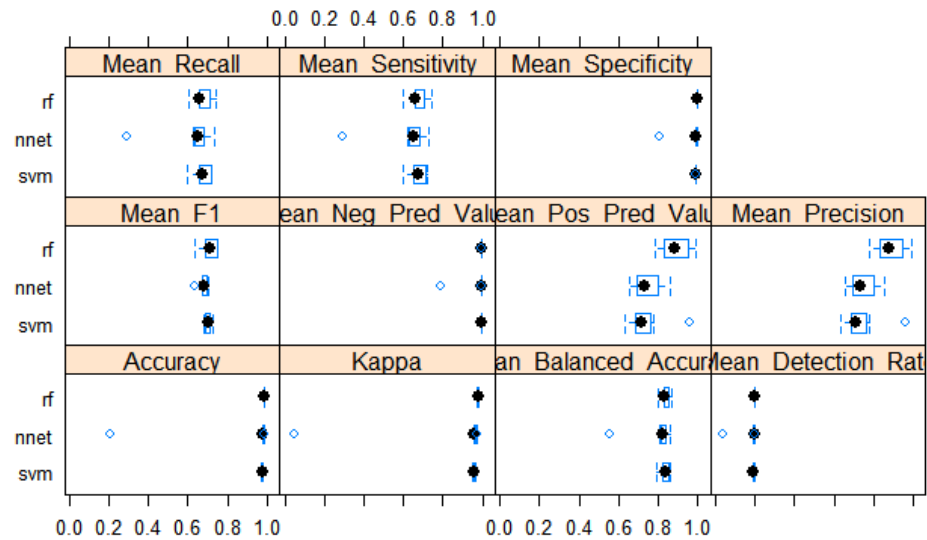
resamples$metrics

t_tests <- resamples %>%
  diff(metric = "Accuracy") %>%
  summary
t_tests

```



Slika 6.17 Prikaz unakrsne validacije razdoblje zime



Slika 6.18 Prikaz unakrsne validacije razdoblje ljeta.

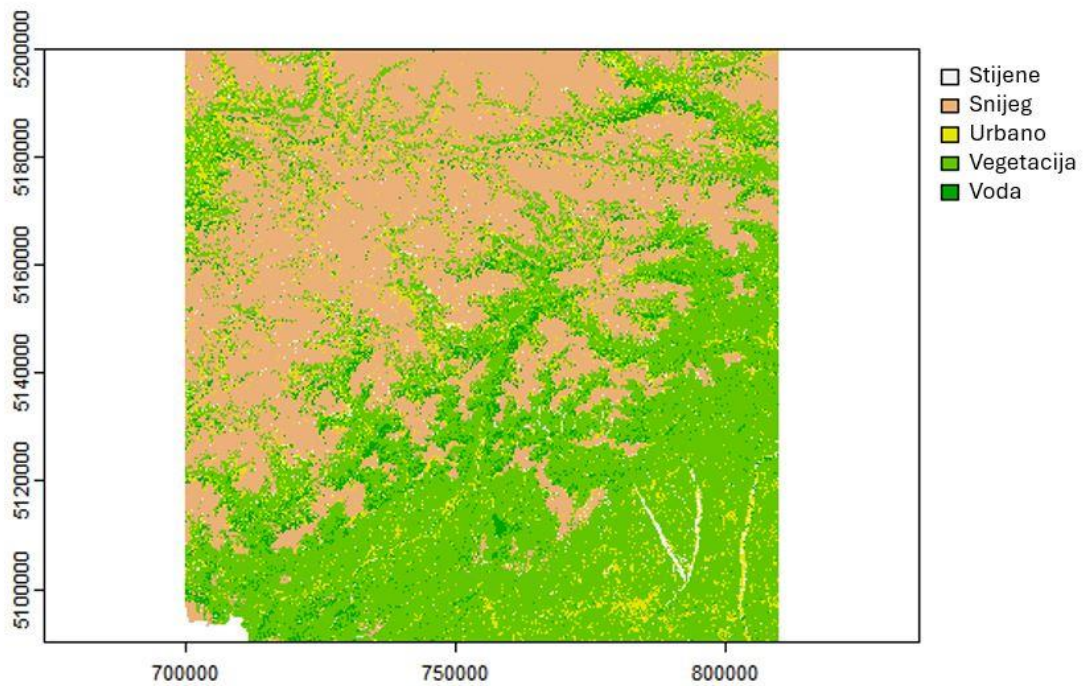
Vrijeme izvršavanja tri modela, koji su trenirani i pohranjeni kao *model\_rf*, *model\_svm* i *model\_nnet*, mjeri se i prikazuje pomoću funkcije „*system.time()*“. Primjena modela na piksele iz testnog skupa podataka obavlja se naredbom *raster::predict*. Na kraju se generira grafički prikaz za sva tri modela, a rezultatni TIFF datoteke pohranjuju se za daljnju primjenu i istraživanje.

```
system.time({
  predict_rf <- raster::predict(object = s1,
                                model = model_rf, type = 'raw',
                                na.rm=TRUE)
  predict_svm <- raster::predict(object = s1,
                                 model = model_svm, type = 'raw',
                                 na.rm=TRUE)
  predict_nnet <- raster::predict(object = s1,
                                  model = model_nnet, type = 'raw',
                                  na.rm=TRUE)})
```

## 7. REZULTAT RASPRAVA KORIŠTENJEM SENTINEL-2

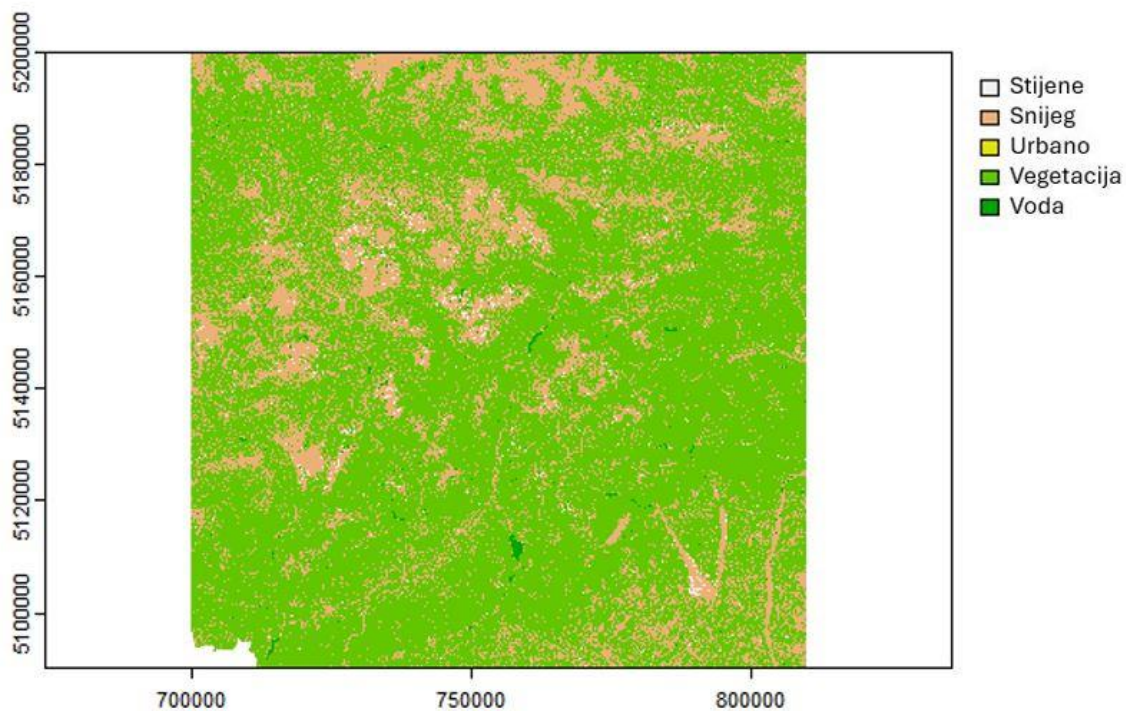
Grafički prikazi tri modela, *model\_rf*, *model\_svm* i *model\_nnet*, prikazani su na slikama dobivenim iz programskog jezika R, kako bi se pregledala klasifikacija područja Italije i Austrije. Konkretno, obuhvaćene su općine u Italiji (Bolzano, Trento, Belluno, Treviso, Pordenone, Udine) i kotari u Austriji (Lienz, Spittal an der Drau, Hermagor). Za to područje klasificirano je pet numeriranih klasa koje imaju svoje boje: stijene - bijela, snijeg- krem, urbano područje- žuta, vegetacija - svijetlozelena, voda - tamnozeleno. Svaka metoda primijenjena je na satelitske snimke iz zimskog i ljetnog razdoblja. Tijekom zime, u planinskim područjima prevladava veći snježni pokrivač, dok su glacijalna jezera i rijeke često prekriveni ledom. Vegetacija, poput šuma, travnjaka i poljoprivrednih zemljišta, nije zelene boje, već pretežno smeđe ili tamnozeleno nijanse. Zbog bijele boje, stijene na satelitskim snimkama mogu biti zamijenjene za snježni pokrivač, što može dovesti do pogrešne interpretacije. U ljetnom razdoblju, količina snijega je manja, a zelene površine su značajnije zastupljene. Jezera i potoci više nemaju pokrov bijelog ili svijetloplavog leda, već se njihova boja kreće od plave do zelene. U oba godišnja doba dolazi do poteškoća u pravilnoj detekciji stijena. Na satelitskim snimkama, to su vrhovi planina i planinska područja koja su u kontaktu sa snijegom i niskom vegetacijom. U nizinskim područjima, stijenskoj klasifikaciji pridružuju se bijeli do bijelo-žuti aluvijalni nanosi te sivi izdanci stijena iz zelenog pokrova. Dodatni problem u prepoznavanju tipova klasifikacije javlja se zbog prisutnosti manjih i većih kamenih fragmenata raspršenih po zelenim i smeđim travnatim površinama. Nijedan model nije inherentno bolji ili lošiji od drugog, već njihova učinkovitost ovisi o svojstvima podataka i rezultatu koji se želi postići. Svaki model pruža određeni tip podataka, specifičan za njegove karakteristike i naredbe.

Model koji se pokazao najprikladnijim za zimsko razdoblje je model *nnet*. Ovaj model može prepoznati kompleksne i nepravilne oblike, što ga čini snažnim alatom za klasifikaciju, no u ovom slučaju nije dao zadovoljavajuće rezultate. Iako je koristan za klasifikacije s velikim brojem dimenzija te može prilagoditi bilo koji oblik za zamišljenu klasifikaciju, u zimskom razdoblju prepoznaje veliku količinu snježnog pokrivača (Slika 7.1.). Međutim, model je pogrešno klasificirao vegetacijsko područje kao urbano, što je vjerojatno posljedica sličnosti zimskih boja između trave i šuma (smeđa) te krovova kuća (narančasto-smeđa). Uz snježni pokrivač, pojavljuju se i točkasta područja stijena.



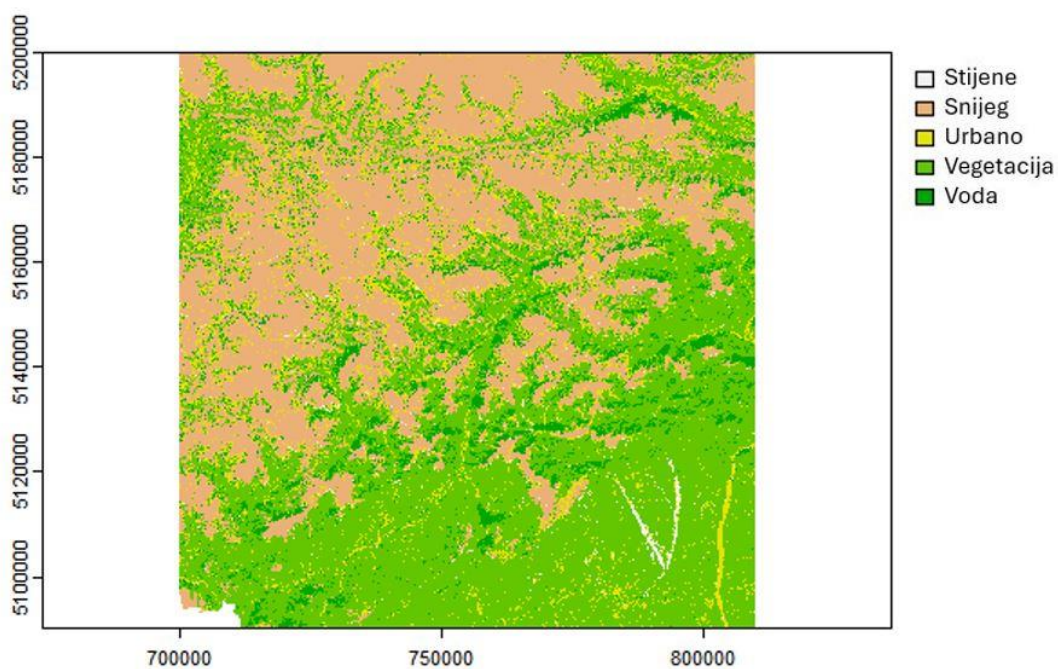
**Slika 7.1** *Predict\_nnet* zimsko razdoblje.

U ljetnom razdoblju, kada na planinskim područjima i vrhovima ima manje količine snijega, model *nnet* često zamjenjuje stijene i urbana područja za snijeg. Vodene površine su, međutim, točno klasificirane (Slika 7.2.).



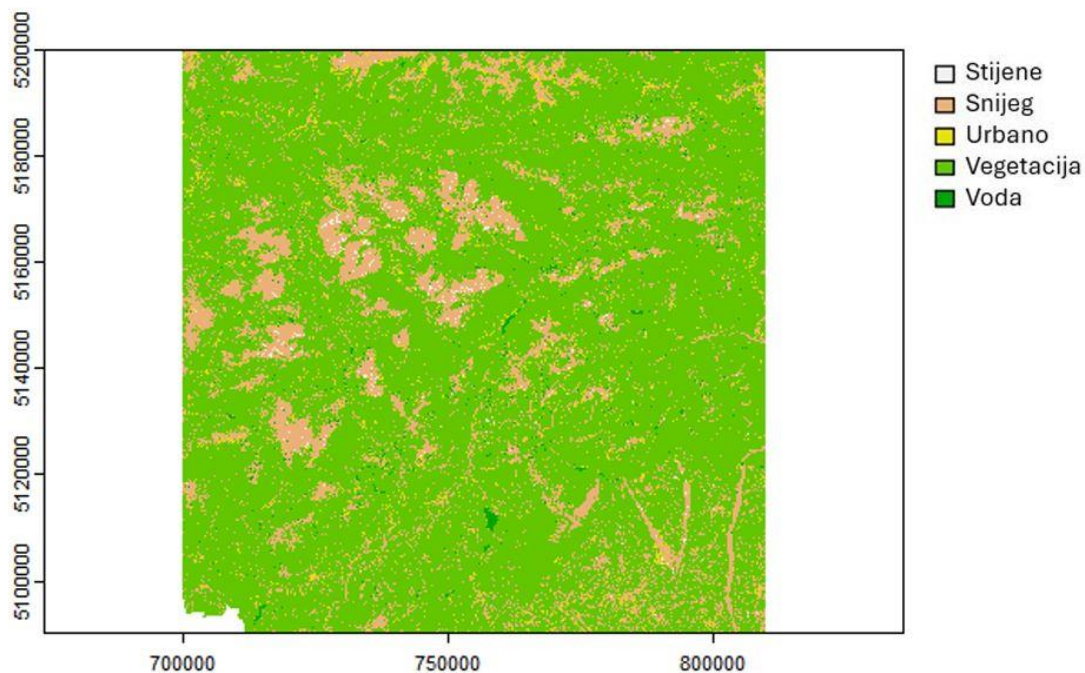
**Slika 7.2** *Predict\_nnet* ljetno razdoblje.

Model *rf* odabran je za daljnja istraživanja u ljetnom razdoblju jer nudi najjasniju podjelu između različitih klasa. Ovaj model se obično koristi za klasifikaciju kada postoji mnogo šuma i nejasnoća u podacima. Međutim, u zimskom razdoblju (Slika 7.3.) nije dao zadovoljavajuće rezultate: izjednačava klase snijega s klasama urbanog područja, a vodene površine su prikazane u prekomjernim količinama u planinskim dijelovima. Također, riječna i aluvijalna područja zamijenjena su za urbana područja.



**Slika 7.3** *Predict\_rf* zimsko razdoblje.

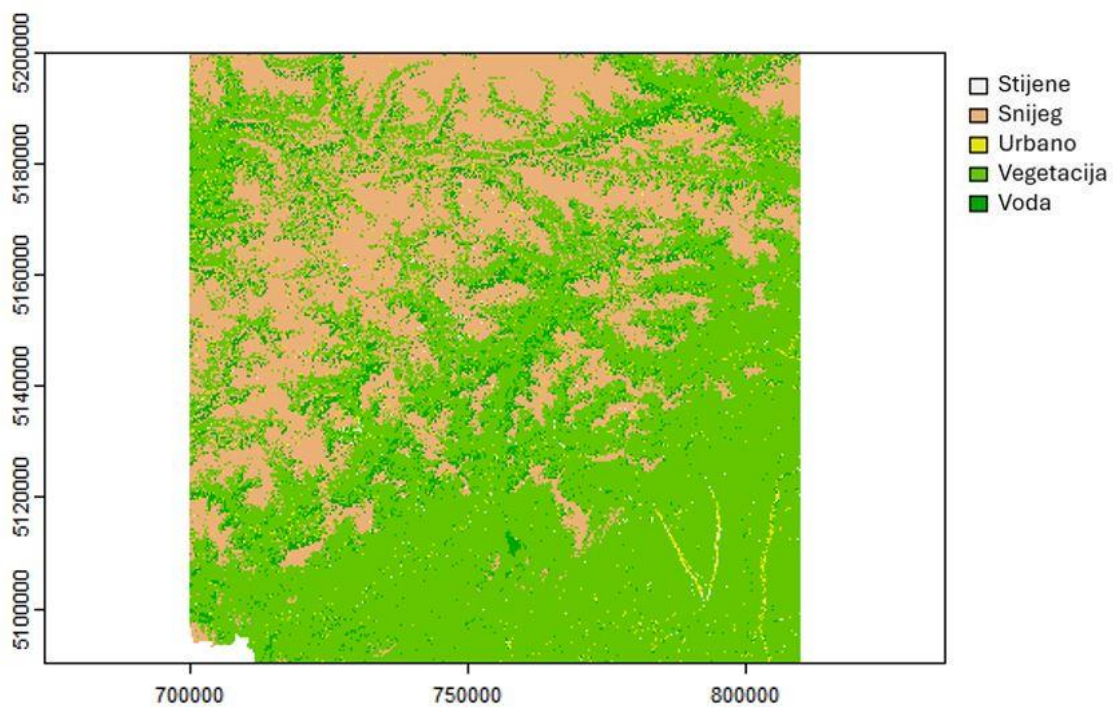
U modelu za ljetno razdoblje (Slika 7.4.), smanjena je zamjena klase urbanog područja za klasu snijega. Planinska područja, iako imaju manji broj klasificiranih stijena, ne miješaju se s vodenim površinama. Međutim, aluvijalna područja i dalje su klasificirana kao snježni pokrov.



**Slika 7.4** *Predict\_rf* ljetno razdoblje.

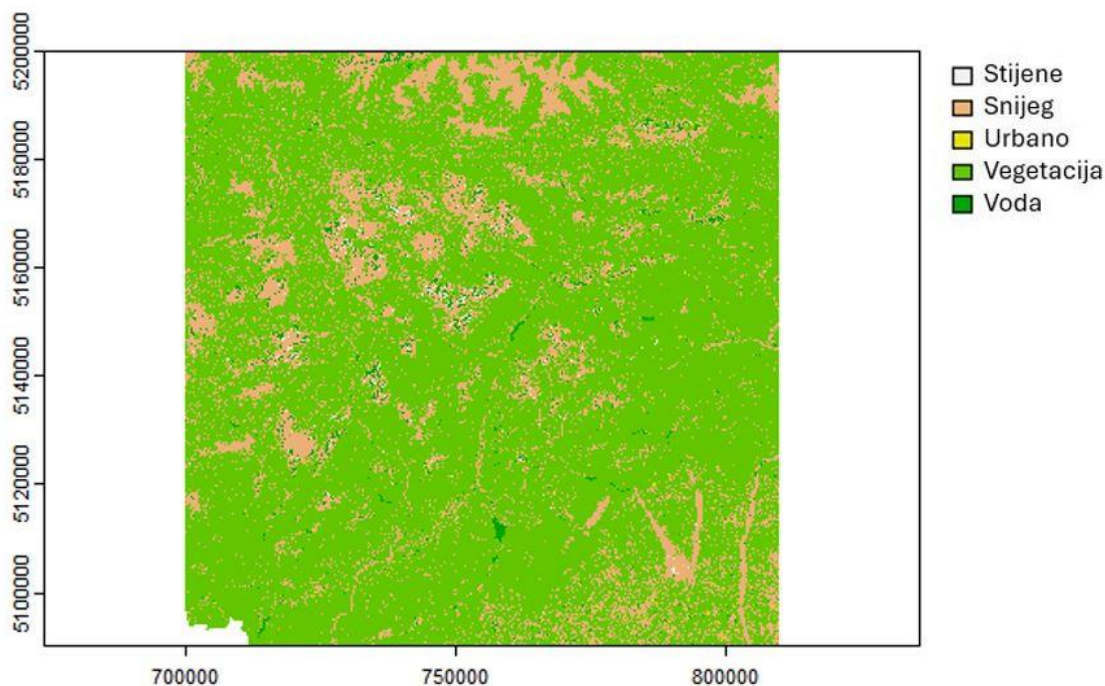
Model *SVM* koristi se za rad s velikim dimenzionalnostima, oponašanje i prilagodbu različitih oblika za svaku zamišljenu klasifikaciju. Međutim, zahtijeva više vremena i parametara za trening nego model *NNET*. Rezultati nisu zadovoljavajući za zadani problem jer miješa većinu klasa. U modelu za zimsko razdoblje (Slika 7.5.), *SVM* ne razlikuje dobro snijeg od vodenih površina, koje se na planinskom području prikazuju u tamnozelenim bojama u jarugama. Aluvijalna područja pogrešno su klasificirana kao urbano područje.





**Slika 7.5** *Predict\_svm* zimsko razdoblje.

U ljetnom razdoblju model također zamjenjuje planinske vrhove s vodenim površinama, prikazujući ih tamnozelenim točkicama, dok se aluvijalna područja klasificiraju kao snježna klasa (Slika 7.6.).

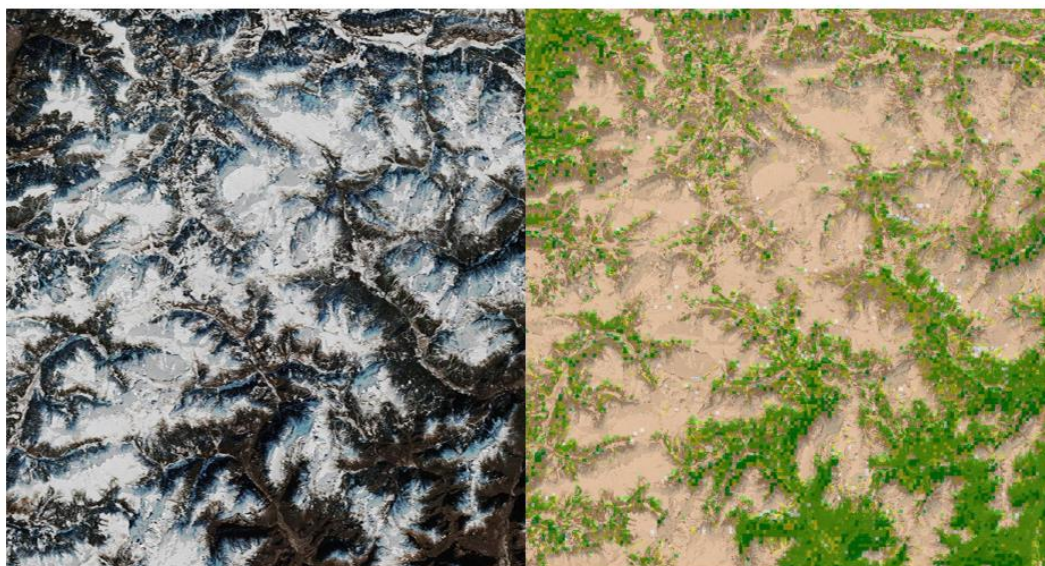


Slika 7.6 *Predict\_svm* ljetno razdoblje.

### 7.1. Potvrde *modela rf* starosti stijena sa kartama ljetnog i zimskog razdoblja

Primjenom vizualnog pregleda u QGIS-u dobiva se uvid u klasifikaciju pet klasa. Za jasniju usporedbu, preklapile su se slike odabranih modela preko satelitskih snimaka. Na preklapljenim slikama, snijeg je prikazan krem bojom, stijene su bijele, vegetacijsko područje svijetlozeleno, vodene površine tamnozeleno, dok je urbano područje označeno žutom bojom.

Tijekom zime, u planinskim područjima dominira veća količina snijega bijele boje, koji se može miješati s vegetacijom, poput šuma prekrivenih snijegom. Vegetacija, uključujući šume, travnjake i poljoprivredna zemljišta, nije zelena, već uglavnom smeđih ili tamnozelenih nijansi. Potoci i jezera su većinom plave boje, ali zaleđeni dijelovi, osobito uz rubove, mogu poprimiti bijelu boju, sličnu snježnom pokrovu. Stijene obuhvaćaju aluvijalna područja blijedožute i bijele boje te pojedine sive izdanke. Urbana područja ostaju nepromijenjena tijekom ljetne i zimske sezone, uključujući kuće, zgrade, ceste, betonirane površine i slično. Na Slika 7.7 jasno su vidljive razlike u prepoznatim klasifikacijama na većem prostornom području modela *nnet* zimskog razdoblja. Površina snijega označena je krem bojom, dok je vegetacija prikazana svijetlozelenom bojom.



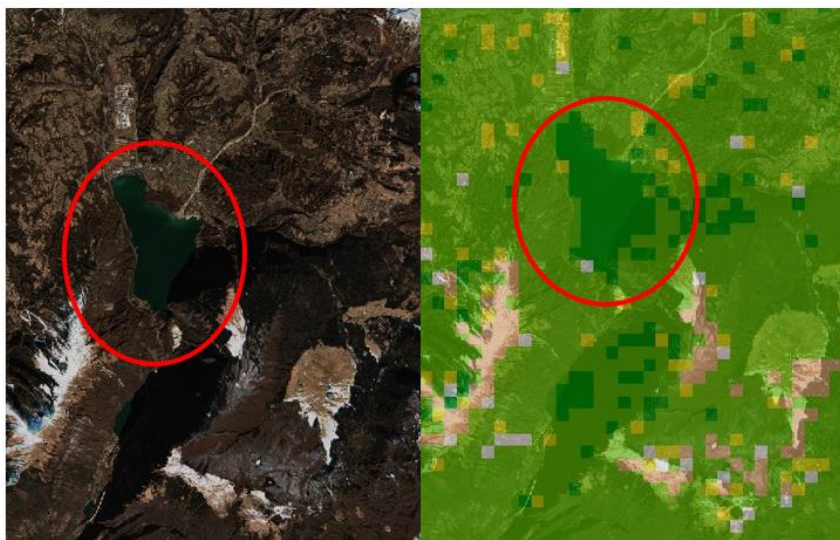
**Slika 7.7** Približna slika snježnog pokrova u Alpama tijekom zimskog razdoblja.

Na Slika 7.8 na području Padanske nizine modela *nnet* zimskog razdoblja, jasno su vidljiva dva aluvijalna područja s malo vidljive vodene površine. Jedno područje ima oblik slova "V", dok je drugo, smješteno desno od njega, duguljaste strukture. Oba područja su klasificirana kao stijene. Područje u obliku slova "V" točno je klasificirano bijelom bojom, dok je drugo aluvijalno područje pogrešno klasificirano kao stjenovito i urbano područje. Osim toga, ostali dijelovi slike su dobro detektirani, poput urbanih područja (crveni krugovi) označenih žutom bojom te poljoprivredne površine svijetlozelene boje.



**Slika 7.8** Područje Padanske nizine iz zimskog razdoblja modela *nnet*.

Vodeni pokrov, jezero, na Slika 7.9 je označeno tamnozelenom bojom (crveni kružić) , no i ostatak slike koji nije pokriven vodenim područjem je označen je u tamnozelenim pikselima.



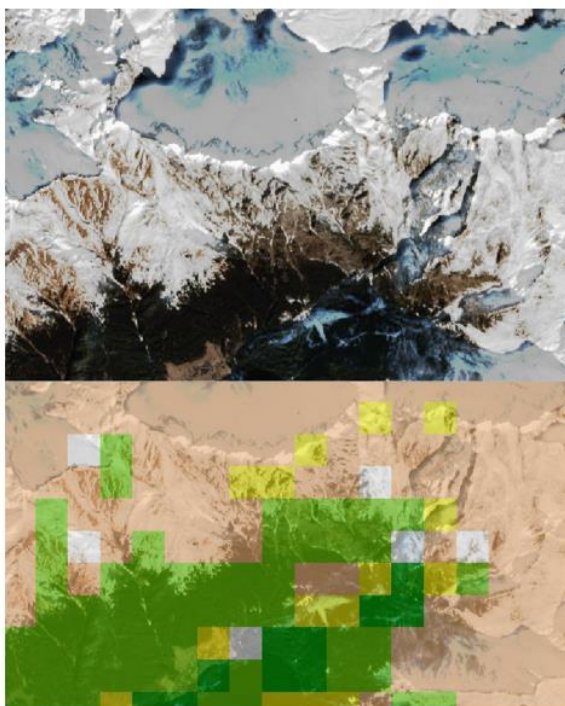
**Slika 7.9** Detekcija vodenog pokrova zimskog razdoblja modela *nnet*.

Na Slika 7.10 se jasno vidi da u području podno planinskog prostora gdje se javlja sijena, dolazi do pogrešne detekcije (crveni kružić). Umjesto pokrivanja tog područja kao vegetacijski ili stjenoviti prostor, zbog raznih valnih duljina, sustav to stavlja u klasu vodenog područje . S obzirom na vizualnu detekciju mogu se lakše detektirati greške.



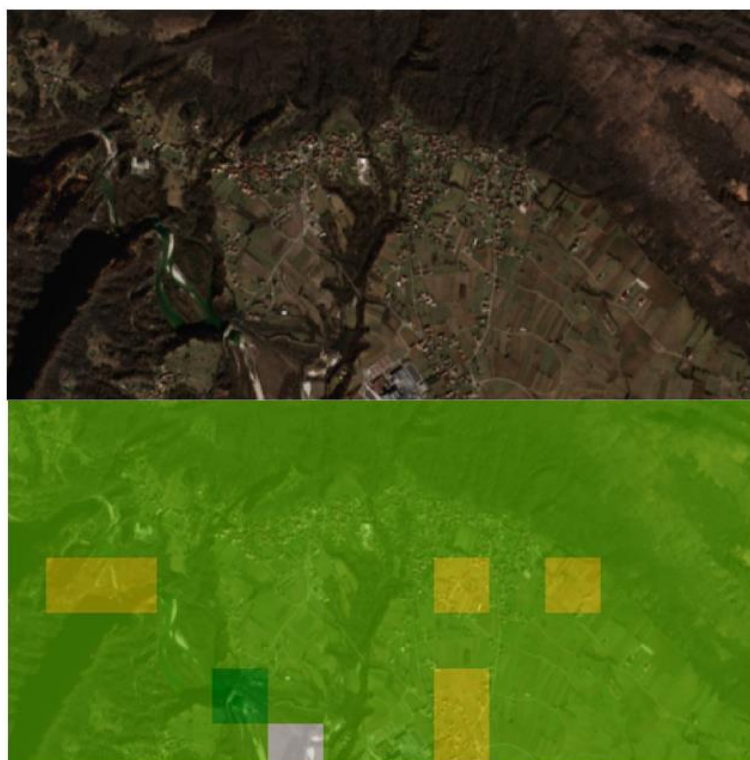
**Slika 7.10.** Problematika u sjenovitim područjima modela *nnet*.

Poneka područja koja nemaju niti opravdanu krivu klasifikaciju . Na Slika 7.11 se na području planinskog vrha (krem boja) i vegetacije (svijetlo zelena boja) javljaju žuti pikseli kao naznaka detektirane urbane klasifikacije.



**Slika 7.11** Problematika u detektiranju snijega modela *nnet*.

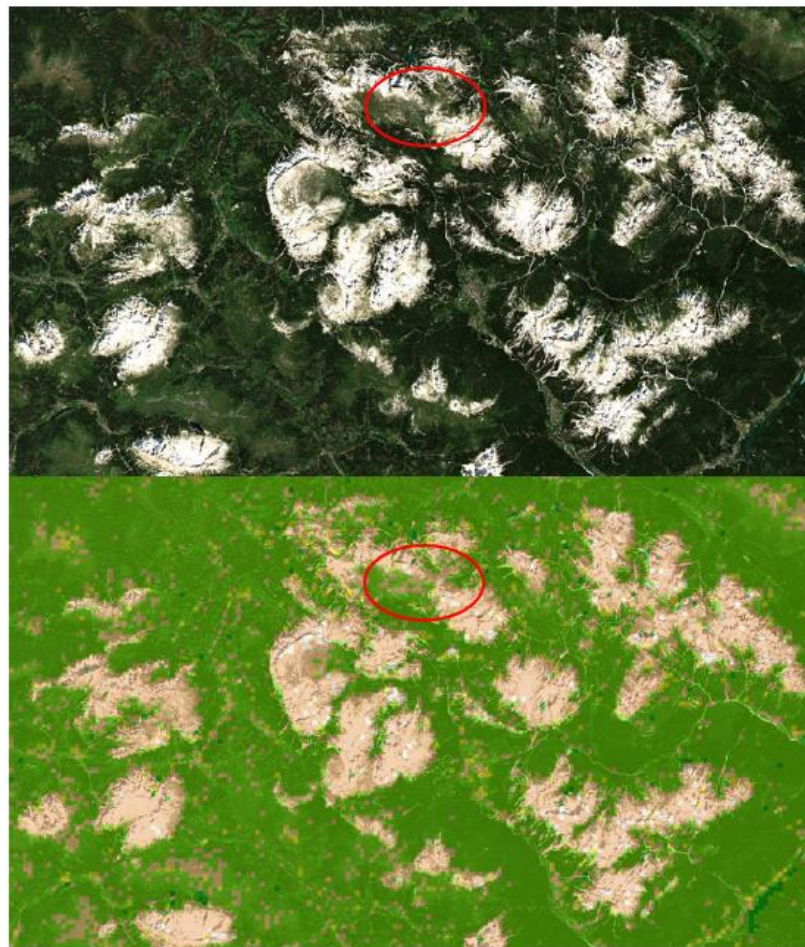
Slika 7.12 prikazuje kako je veliki dio urbane sredine detektiran kao područje vegetacije (svijetlozelena boja). Na Ponekom prostoru urbane razine se vide odgovarajući pikseli (žuta boja). No većina je krivo klasificirana.



**Slika 7.12** Pogrešna detekcija urbanog područja modela *nnet*.

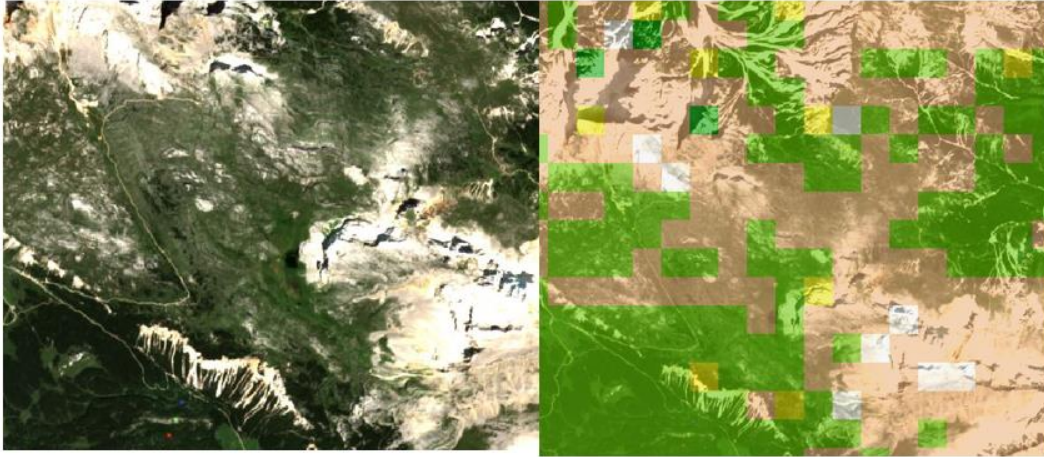
U ljetnom razdoblju, količina snijega je manja, a zelene površine su značajnije zastupljene. Jezera i potoci su plave do zelene boje. Na satelitskim snimkama, vrhovi planina i planinska područja su zastupljena snijegom bijele boje, no može doći do kontakta sa stijenama i niskom vegetacijom. Dodatni problem u prepoznavanju tipova klasifikacije javlja se zbog prisutnosti manjih i većih kamenih fragmenata raspršenih po zelenim travnatim površinama. U nizinskim područjima, stijenskoj klasifikaciji pridružuju se bijeli do bijelo-žuti aluvijalni nanosi te sivi izdanci stijena iz zelenog pokrova.

Na Slika 7.13 jasno su vidljive razlike u prepoznatim klasifikacijama na većem prostornom području modela *rf* ljetnog razdoblja. Površina snijega označena je krem bojom, dok je vegetacija prikazana svijetlozelenom bojom. Približavanjem područja označenim crvenim kružićem dobivamo jasniji uvid u točnost detekcije.



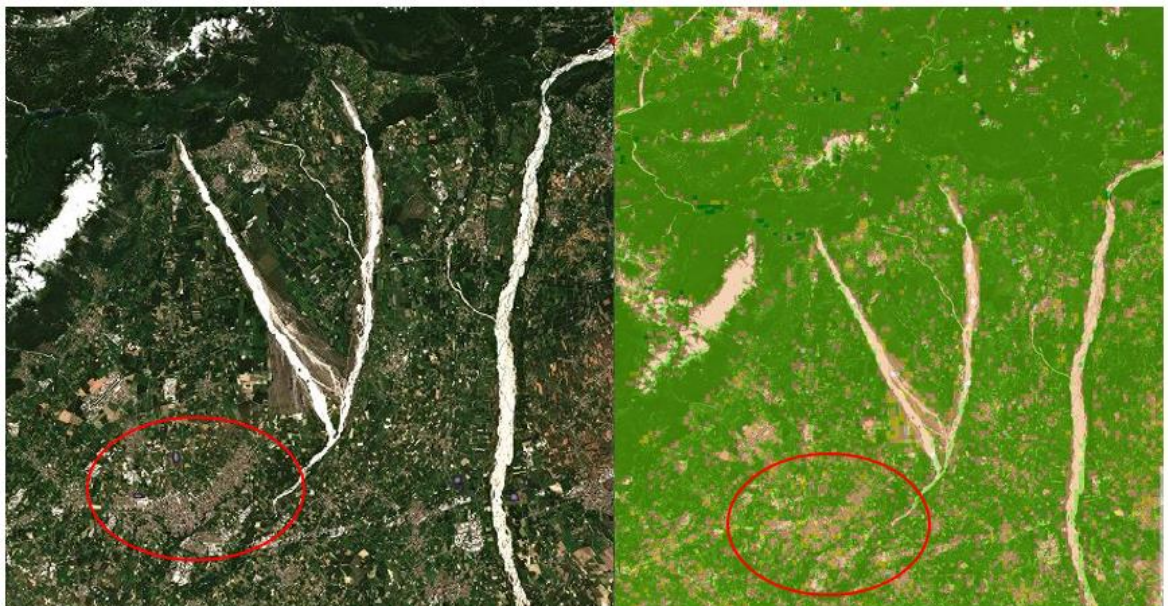
**Slika 7.13** Približena slika snježnog pokrova u Alpama tijekom ljetnog razdoblja modela *rf*.

Slika 7.14. detaljnije pokazuje približno područje. Gdje je vidljivo kako je stjenovito područje (bijeli pikseli) pogrešno klasificirano kao snijeg (krem pikseli). Razlog je tome što zbog lošije rezolucije sustav nije mogao razlikovati klase snijega i klase stijene.



**Slika 7.14** Približno područje planinskog područja modela *rf*.

Na Slika 7.15, na području Padanske nizine iz modela *rf* ljetnog razdoblja, jasno su vidljiva dva aluvijalna područja s malo vidljive vodene površine. Jedno područje ima oblik slova "V", dok je drugo, smješteno desno od njega, duguljaste strukture. Oba stjenovita područja su pogrešno klasificirana kao snijeg (krem boja). Osim toga, ostali dijelovi slike, poput urbanih područja (crveni krugovi) detektirani također kao sniježi pokrov krem bojom.



**Slika 7.15** Područje Padanske nizine iz ljetnog razdoblja modela *rf*.

Vodeni pokrov, odnosno jezero, na Slika 7.16 označen je tamnozelenom bojom (crveni kružići). Na ostatku slike vidljivo je prisustvo krem piksela u području vegetacije i urbanog područja, posebno u gusto urbaniziranom dijelu označenom žutim kružićem. Područje označeno plavim kružićem prikazuje oblake koji su klasificirani kao snijeg.



Slika 7.16 Detekcija vodenog pokrova ljetnog razdoblja modela *rf*.

Na **Error! Reference source not found.** crvenim kružićem označena je stijenska površina koja je točno detektirana bijelim pikselima. Pojava pomaka piksela rezultat je nepodudarnosti referentnog koordinatnog sustava i prostorne neusklađenosti. Zbog toga je potrebno pretražiti okolni prostor oko tih piksela kako bi se osigurala ispravna klasifikacija.





**Slika 7.17** Preciznost detekcije modela rf.

Unatoč greškama u detekciji, oba modela uspješno su identificirala snijeg na širem prostornom području. Iako su se pojavili slučajevi pogrešne detekcije i miješanja klasifikacija, ovi modeli pokazuju da će s daljnjim istraživanjem i razvojem tehnologija detekcija snijega, kao i drugih klasifikacija na različitim geografskim područjima, postati preciznija.

## 8. ZAKLJUČAK

Korištenjem slobodno dostupnih satelitskih snimaka iz misije Sentinel-2 (Europske svemirske agencije Copernicus) te primjenom aplikacija i programskih jezika otvorenog koda (QGIS i R), razvijen je novi model za detekciju snježnog pokrivača. Model za predikciju snijega stvoren je uz pomoć algoritama strojnog učenja.

Poligoni pet različitih klasa (snijeg, stijene, urbano područje, voda i vegetacija) izrađeni su u aplikaciji QGIS i služili su kao podaci za treniranje modela. Algoritmi nadzirane klasifikacije, uključujući *randomForest*, *Support Vector Machine* i *Neural Network* iz paketa CARET, korišteni su za izradu novih modela za detekciju snijega. Ovi modeli su također omogućili razlikovanje klasa stijena, urbanih područja, vode i vegetacije.

Provedene su dvije detaljne spektralne analize snimaka misije Sentinel-2, jedna za zimsko, a druga za ljetno razdoblje, koristeći podatke iz deset različitih kanala. Novorazvijeni modeli primijenjeni su na multispektralne satelitske snimke, generirajući šest klasificiranih rastera. Grafički prikaz ovih rastera u QGIS-u omogućio je uvid u pouzdanost modela u predikciji snijega i drugih klasa.

Za daljnja istraživanja odabrani su modeli *Neural Network* za zimsko razdoblje i *randomForest* za ljetno razdoblje, jer su se pokazali kao najpouzdaniji za postizanje ciljanih rezultata u klasifikaciji snijega. Vizualna analiza u QGIS-u omogućila je bolji uvid u klasifikaciju snijega i ostalih klasa.

Analizirane su slike dijelova satelitskih snimaka misije Sentinel-2, uključujući usporedbu s preklapljenim izbljedjelim rasterima radi lakše procjene točnosti detekcije. Rezultati pokazuju da su u široj slici najzastupljenije klase bolje detektirane, s naglaskom na klasu

snijega u zimskom razdoblju i klasu vegetacije u ljetnom razdoblju. Međutim, primijećeno je miješanje ili višak pojedinih klasa u zimskim i ljetnim rasterima.

Pogrešna detekcija na određenim područjima rezultat je miješanja klasa (snijeg i vegetacija, snijeg i stijene, stijene i vegetacija) te sličnih frekvencija koje sustav grupira u jednu klasu. Na ovom radu pokazuje se da je metodama strojnog učenja moguće razviti model za klasifikaciju snijega, ali i drugih klasa poput vegetacije, stijena, vode i urbanih područja.

Za poboljšanje učinkovitosti modela potrebni su podaci veće prostorne rezolucije kako bi se povećala pouzdanost detekcije. Osim toga, trebali bi se koristiti i satelitski radarski snimci iz misije Sentinel-1. Daljnja istraživanja trebala bi se usmjeriti na pronalaženje naprednijih tehnologija za klasifikaciju snijega u Alpama na različitim područjima, uključujući informacije o stalnom snijegu na planinskim područjima te druge ključne podatke koji bi omogućili precizniju i pouzdaniju analizu.

## **9. ZAHVALE**

Zahvaljujem svojem mentoru, doc. dr. sc. prof. Ivanu Medvedu, na velikoj podršci, strpljenju i korisnim savjetima tijekom cijelog procesa izrade ovog rada. Bez vaše pomoći i smjernica, ovaj rad ne bih uspjela ostvariti.

Također, iskreno zahvaljujem svojoj obitelji i prijateljima koji su uvijek vjerovali u mene i pružali mi podršku kada mi je bila najpotrebnija. Bez vas, ovaj rad ne bi bio moguć.

## 10. LITERATURA

Giorgi, F.M., Ceraolo, C., Mercatelli, D. 2022. The R Language: An Engine for Bioinformatics and Data Science. *Life*, 12, 648. [https:// doi.org/10.3390/life12050648](https://doi.org/10.3390/life12050648) (5.3.2024.)

Morale, G., Ramirez, A., Telles, J. End-to-end Cloud Segmentation in High-Resolution Multispectral Satellite Imagery Using Deep Learning. (2024) URL: <https://arxiv.labs.arxiv.org/html/1904.12743>

Prieur, C., Rabatel, A., Thomas, J-B., Farup, I. & Chanussot, J. (2022). Machine Learning Approaches to Automatically Detect Glacier Snow Lines on Multi-Spectral Satellite Images. *Remote Sens.* 2022, 14(16), 3868; <https://doi.org/10.3390/rs14163868> (7.3.2024)

Saini, R. & Singh, S. (2023). Land Use Land Cover Mapping and Snow Cover Detection in Himalayan Region Using Machine Learning and Multispectral Sentinel-2 Satellite Imagery. *nt. j. inf. tecnol.* (February 2024) 16(2):675–686 <https://doi.org/10.1007/s41870-023-01673-1>

Sharma, A., Prakash, C. Glacial lakes mapping using satellite images and deep learning algorithms in Northwestern Indian Himalayas. *Model. Earth Syst. Environ.* 10, 2063–2077 (2024). <https://doi.org/10.1007/s40808-023-01885-1>

Wang, Y., Su, J., Zhai, X., Meng, F. & Lu, C. (2022). Snow Coverage Mapping by Learning from Sentinel-2 Satellite Multispectral Images via Machine Learning Algorithms. *Remote Sens.* 2022, 14(3), 782; <https://doi.org/10.3390/rs14030782>

## WEB IZVORI:

CRAN-a. *Leafsync: Small Multiples for Leaflet Web Maps*. URL: <https://cran.r-project.org/web/packages/leafsync/index.html> (01.04.2024.)

CRAN-b. *MapView: Interactive Viewing of Spatial Data in R*. URL: <https://cran.r-project.org/web/packages/mapview/index.html> (01.04.2024.)

CRAN-c. *NeuralNetTools: Visualisation and Analysis Tools for Neural Networks*. URL: <https://cran.r-project.org/web/packages/NeuralNetTools/index.html>. (01.04.2024.)

CRAN-d. *RColorBrewer: ColorBrewer Palettes*, URL: <https://cran.r-project.org/web/packages/RColorBrewer/index.html> (01.04.2024.)

CRAN-e. *LiblineaR: Linear Predictive Models Based on the LIBLINEAR C/C++ Library*. URL: <https://cran.r-project.org/web/packages/LiblineaR/index.html> (01.04.2024.)

CRAN-f. *Terra: Spatial Data Analysis*. URL: <https://cran.r-project.org/web/packages/terra/index.html> (01.04.2024.)

Copernicus-b, 2024. *Copernicus Browser*. LINK: <https://browser.dataspace.copernicus.eu/> . (25.02.2024.)

Datacamp-a, 2023. *Building Neural Network (NN) Models in R*. URL: <https://www.datacamp.com/tutorial/neural-network-models-r> (6.4.2024.)

Datacamp-b, 2018. *Support Vector Machines in R*. URL: <https://www.datacamp.com/tutorial/support-vector-machines-r> (7.4.2024.)

Esri, 2017. *Sentinel-2 operations* URL: [https://www.esa.int/Enabling\\_Support/Operations/Sentinel-2\\_operations](https://www.esa.int/Enabling_Support/Operations/Sentinel-2_operations) (23.1.2024.)

Geeksforgeeks-a, 2021. *Classifying data using Support Vector Machines (SVMs) in R*. URL: <https://www.geeksforgeeks.org/classifying-data-using-support-vector-machines-in-r/> (5.3.2024.)

Geeksforgeeks-b, 2020. *How Neural Networks are used for Classification in R Programming*. URL: <https://www.geeksforgeeks.org/how-neural-networks-are-used-for-classification-in-r-programming/> (5.3.2024.)

GISGeography, 2022. *Sentinel 2 Bands and Combinations*. URL: [Sentinel 2 Bands and Combinations - GIS Geography](#) (5.3.2024.)

GitHub, 2023. *MLmetrics*. URL: <https://github.com/yanyachen/MLmetrics> (5.3.2024.)

Javatpoint-a, 2021. *List of R packages*. URL: <https://www.javatpoint.com/list-of-r-packages> (8.3.2024)

Javatpoint-b, 2021. *Random Forest Algorithm*. URL: <https://www.javatpoint.com/machine-learning-random-forest-algorithm> (8.3.2024)

QGIS-a, 2024. *Features*. URL: [https://docs.qgis.org/3.28/en/docs/user\\_manual/preamble/features.html](https://docs.qgis.org/3.28/en/docs/user_manual/preamble/features.html) (1.3.2024.)

QGIS-b. *About QGIS*. URL: <https://www.qgis.org/en/site/about/index.html> (1.3.2024.)

Packt, *Pros and cons od neural network*. URL: <https://subscription.packtpub.com/book/big-data-&-business-intelligence/9781788397872/1/ch01lv11sec27/pros-and-cons-of-neural-networks> (1.3.2024.)

Patwal, M., (2020.): *A Short Introduction to R Programming*. URL: <https://bookdown.org/manishpatwal/bookdown-demo/> (13.3.2024.)

Posit Support, 2024. *Recommended Packages*. URL: <https://support.posit.co/hc/en-us/articles/201057987-Quick-list-of-useful-R-packages> (25.3.2024.)

R Learning Infrastructure Home Web Page. *Introduction to R and RStudio*. URL: [http://mercury.webster.edu/aleshunna/R\\_learning\\_infrastructure/Introduction\\_to\\_R\\_and\\_R\\_Studio.html](http://mercury.webster.edu/aleshunna/R_learning_infrastructure/Introduction_to_R_and_R_Studio.html) (25.3.2024.)

Sentinel Online, 2023. SENTINEL-2 MISSION GUIDE. URL: <https://sentinels.copernicus.eu/web/sentinel/missions/sentinel-2> (25.3.2024.)

Simplilearn, 2024. *Introduction to Random Forest in R*. URL: <https://www.simplilearn.com/tutorials/data-science-tutorial/random-forest-in-r> (25.3.2024.)

Spatial Dana Science, 2023. *The terra package*. URL: <https://rspatial.org/pkg/1-introduction.html#the-terra-package> (25.3.2024.)

Tutorialspoint, 2024. *R-Packages*. URL: [https://www.tutorialspoint.com/r/r\\_packages.htm](https://www.tutorialspoint.com/r/r_packages.htm) (25.3.2024.)

UC Bussines Analytics R Programming Guide, *Support Vector Machine*. URL: <https://uc-r.github.io/svm> (25.3.2024.)