Segmentacija satelitskih snimaka šume za klasifikaciju tipa šume

Sičić, Sara

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Zagreb, Faculty of Electrical Engineering and Computing / Sveučilište u Zagrebu, Fakultet elektrotehnike i računarstva**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:168:158788

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2025-03-23**

Repository / Repozitorij:

FER Repository - University of Zagreb Faculty of Electrical Engineering and Computing repozitory

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 466

SEGMENTACIJA SATELITSKIH SNIMAKA ŠUME ZA KLASIFIKACIJU TIPA ŠUME

Sara Sičić

Zagreb, lipanj 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 466

SEGMENTACIJA SATELITSKIH SNIMAKA ŠUME ZA KLASIFIKACIJU TIPA ŠUME

Sara Sičić

Zagreb, lipanj 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Zagreb, 4. ožujka 2024.

DIPLOMSKI ZADATAK br. 466

Pristupnica:	**Sara Sičić (0036515679)**
Studij:	Računarstvo
Profil:	Znanost o podacima
Mentor:	prof. dr. sc. Marko Subašić

Zadatak: **Segmentacija satelitskih snimaka šume za klasifikaciju tipa šume**

Opis zadatka:

Ispitati metode segmentacije slika. Posebnu pažnju obratiti na postupke segmentacije slika vrlo visoke rezolucije. Pripremiti bazu satelitskih snimaka šumskog područja i odgovarajućih ručno segmentiranih mapa. Implementirati metodu za segmentaciju prema tipu šume. Ispitati točnost na pripremljenoj bazi satelitskih snimaka. Ispitati mogućnosti augmentacije ulazne baze slika s ciljem poboljšanja performansi. Ispitati optimalan odabir ulaznih kanala multispektralne satelitske snimke. Za implementaciju koristiti programski jezik Python.

Rok za predaju rada: 28. lipnja 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

DIPLOMSKI RAD br. 466

SEGMENTACIJA SATELITSKIH SNIMAKA ŠUME ZA KLASIFIKACIJU TIPA ŠUME

Sara Sičić

Zagreb, rujan, 2024.

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

Zagreb, 4. ožujka 2024.

DIPLOMSKI ZADATAK br. 466

Pristupnica:	**Sara Sičić (0036515679)**
Studij:	Računarstvo
Profil:	Znanost o podacima
Mentor:	prof. dr. sc. Marko Subašić

Zadatak: **Segmentacija satelitskih snimaka šume za klasifikaciju tipa šume**

Opis zadatka:

Ispitati metode segmentacije slika. Posebnu pažnju obratiti na postupke segmentacije slika vrlo visoke rezolucije. Pripremiti bazu satelitskih snimaka šumskog područja i odgovarajućih ručno segmentiranih mapa. Implementirati metodu za segmentaciju prema tipu šume. Ispitati točnost na pripremljenoj bazi satelitskih snimaka. Ispitati mogućnosti augmentacije ulazne baze slika s ciljem poboljšanja performansi. Ispitati optimalan odabir ulaznih kanala multispektralne satelitske snimke. Za implementaciju koristiti programski jezik Python.

Rok za predaju rada: 28. lipnja 2024.

Zahvaljujem mentoru, roditeljima, prijateljima i šumama.

Sadržaj

1. Uvod **3**

2. Duboko učenje **5**

2.1. Struktura duboke neuronske mreže 5

2.1.1. Umjetni neuron 6

2.1.2. Neuronska mreža 6

2.1.3. Slojevi CNN mreže 7

2.1.4. Podatci 8

2.1.5. Funkcija gubitka 8

2.1.6. Propagacija pogreške unazad 9

2.2. Primjene u analizi slika 9

3. Modeli računalnog vida **10**

3.1. U-Net Model 10

3.1.1. Arhitektura U-Net modela 10

3.1.2. Primjena u segmentaciji satelitskih snimaka 11

3.2. VGG16 Model 11

3.2.1. Arhitektura VGG16 modela 12

3.2.2. Primjena u segmentaciji satelitskih snimaka 12

4. Priprema podataka **14**

4.1. Izvor podataka 14

4.2. Priprema podataka 16

4.3. Analiza distribucije podataka 17

4.4. Podjela podataka 21

5. Eksperimentalni postav **22**
5.1. Implementacijski detalji 22
5.2. Generator podataka 22
5.2.1. Augmentacija podataka 22
5.3. U-net model 23
5.3.1. U-net model, 3 ulaza, 5 izlaza 23
5.3.2. U-net model, 8 ulaza, 5 izlaza 24
5.4. VGG16 model 24
5.4.1. VGG16 model, 3 ulaza, 5 izlaza 24
5.4.2. VGG16 model, 3 ulaza, 6 izlaza 25
5.5. *Callback* funkcije 25
5.6. Metrike 26
5.6.1. Mean Squared Error (MSE) 26
5.6.2. Root Mean Square Error (RMSE) 26
5.6.3. R-squared 27

6. Rezultati i rasprava **28**
6.1. Analiza rezultata modela 28
6.1.1. U-Net model, 3 ulaza, 5 izlaza 28
6.1.2. U-Net model, 8 ulaza, 5 izlaza 32
6.1.3. VGG16 model, 3 ulaza, 5 izlaza 36
6.1.4. VGG16 model, 3 ulaza, 6 izlaza 40
6.2. Ograničenja i izazovi 44

7. Zaključak **45**

Literatura **46**

Sažetak **47**

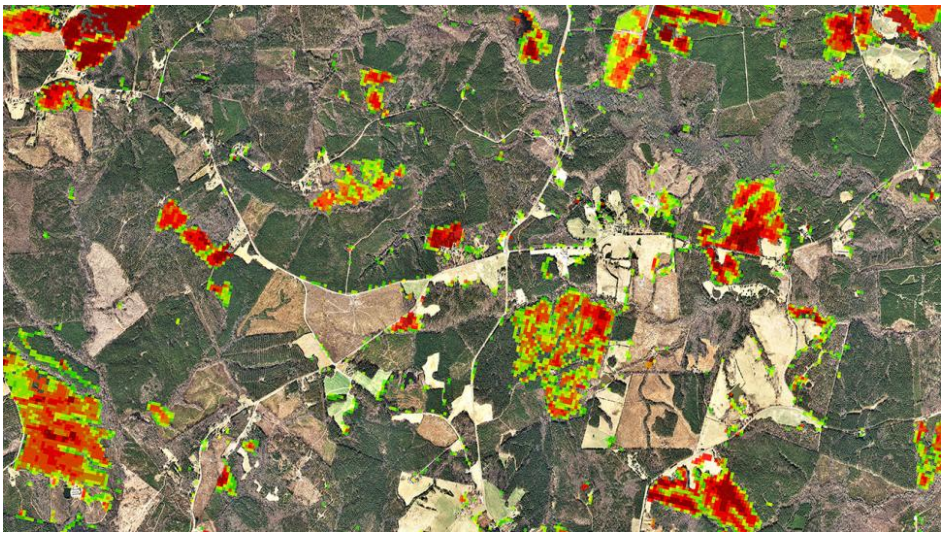
Abstract **48**

1. Uvod

Tehnološki napredak otvorio je nove mogućnosti za istraživanje i razumijevanje našeg okoliša. Jedno od ključnih područja istraživanja čini analiza i upravljanje šumskim ekosustavima, koji su ključni za održavanje biološke raznolikosti, regulaciju klimatskih promjena te očuvanje prirodnih resursa. Segmentacija satelitskih snimaka šuma tehnika je kojom možemo pratiti stanje šuma, procjenjivati njihovu površinu i zdravlje te putem koje otkrivamo promjene uzrokovane prirodnim i antropogenim faktorima.

Semantičkom segmentacijom precizno možemo klasificirati piksele na satelitskoj slici u odgovarajuću kategoriju, čime se dobiva jasna i detaljna mapa šumskog pokrova. Ova tehnika koristi se i u svrhu planiranja korištenja zemljišta, prevencije šumskih požara te za implementaciju strategija očuvanja prirode. Uz pomoć algoritama strojnog učenja i dubokog učenja, semantičkom segmentacijom možemo efikasno analizirati puno veću količinu satelitskih podataka nego što je to prije bilo moguće.

Ovaj diplomski rad usredotočit će se na primjenu dubokog učenja za segmentaciju satelitskih snimaka šuma, s ciljem razvoja pouzdanog modela koji može klasificirati različite tipove šuma. Proučavat će se različiti modeli konvolucijskih neuronskih mreža, posebice one koje su se pokazale uspješnima u području obrade slika. Također, rad će obuhvatiti i analizu učinkovitosti različitih modela za segmentaciju, uključujući evaluaciju njihove preciznosti i sposobnosti generalizacije različitih tipova šume.



Slika 1.1. SouthFACT aplikacija za detekciju promjene šuma

2. Duboko učenje

Duboko učenje područje je strojnog učenja koje se temelji na višeslojnim neuronskim mrežama sposobnim za učenje složenih reprezentacija podataka. Ovaj pristup omogućuje obradu velikih skupova podataka te postiže odlične rezultate u područjima kao što su prepoznavanje uzoraka, obrada prirodnog jezika i računalni vid. Segmentacija slika šuma predstavlja ključnu primjenu dubokog učenja u analizi okoliša, a koristeći složene modele poput modela U-Net, moguće je automatski prepoznati i segmentirati područja šuma na temelju optičkih i multispektralnih podataka. Pomoću ove tehnologije brzo možemo identificirati područja šuma i njihovu klasifikaciju prema vrsti i starosti, što je od velike važnosti u očuvanju okoliša i donošenju odluka o upravljanju prirodnim resursima [1].

2.1. Struktura duboke neuronske mreže

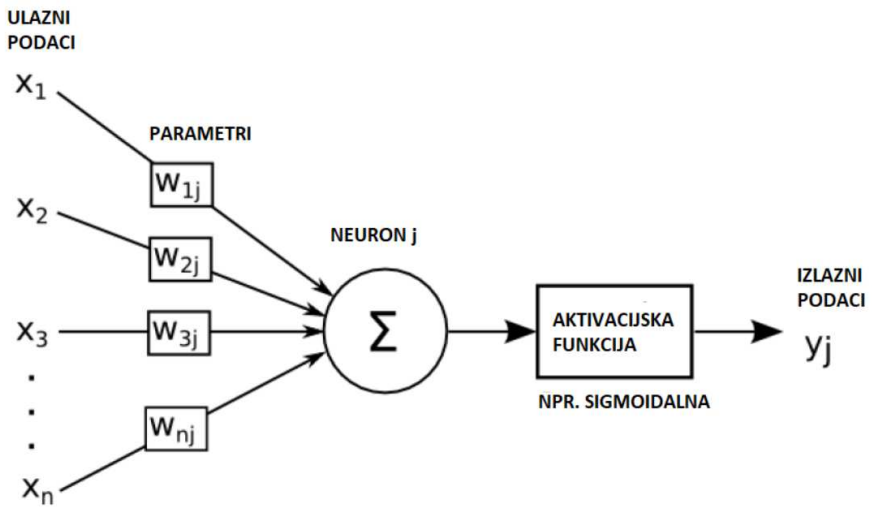
Duboke neuronske mreže čine temelj dubokog učenja i modelima omogućuju da uče složene reprezentacije iz podataka. Osnovna ideja dubokog učenja je imitirati način na koji ljudski mozak obrađuje informacije. Neuronska mreža sastoji se od slojeva neurona koji su međusobno povezani, gdje svaki neuron prima ulaz, vrši određene izračune i prenosi rezultat dalje kroz mrežu.

Mreže su obično organizirane u slojeve: ulazni sloj prima sirove podatke, skriveni slojevi obrađuju te podatke, a izlazni sloj generira konačna predviđanja. Skriveni slojevi mogu biti brojni i svaki sloj može imati stotine ili tisuće neurona, ovisno o složenosti zadatka i veličini mreže [2].

2.1.1. Umjetni neuron

Umjetni neuron prikazan na slici 2.1., osnovna je jedinica duboke neuronske mreže. Svaki neuron prima jedan ili više ulaza kojima su pridijeljene težine koje određuju važnost svakog ulaza. Neuron zatim zbraja ulaze i rezultat prolazi kroz aktivacijsku funkciju koja uvodi nelinearnost u model. Najčešće korištene aktivacijske funkcije su ReLU (Rectified Linear Unit), sigmoidna funkcija i tangens hiperbolična funkcija.

Aktivacijske funkcije omogućuju mreži da uči složene uzorke u podatcima. Bez nelinearnosti, mreža bi bila poput jednostavne linearne regresije, što bi ograničilo njezinu sposobnost učenja složenih reprezentacija [3].

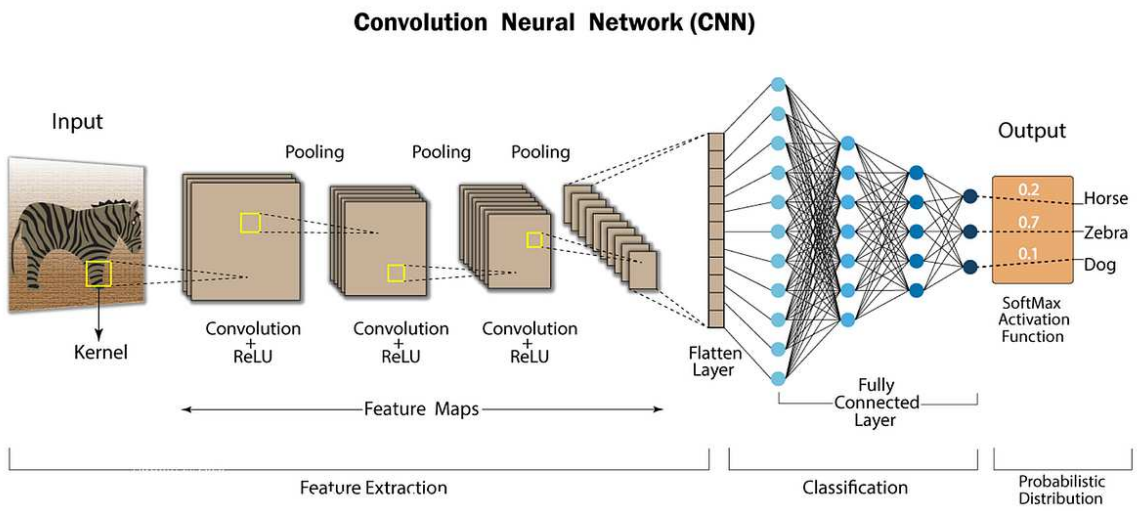


Slika 2.1. Umjetni neuron

2.1.2. Neuronska mreža

Neuronska mreža skup je međusobno povezanih umjetnih neurona organiziranih u slojeve. Postoji nekoliko vrsta neuronskih mreža, ali najčešće korištene u analizi slika su konvolucijske neuronske mreže. CNN mreže specijalizirane su za obradu podataka s rešetkastom topologijom, kao što su slike, koristeći konvolucijske slojeve koji automatski otkrivaju prostorne hijerarhije značajki.

Svaki sloj mreže pretvara ulazne podatke u oblik koji je lakše obraditi u sljedećem sloju. Na primjer, prvi sloj može detektirati osnovne oblike kao što su rubovi, dok kasniji slojevi detektiraju složenije uzorke poput tekstura ili objekata [4].



Slika 2.2. CNN mreža

2.1.3. Slojevi CNN mreže

CNN arhitektura prikazana na slici 2.2. sastoji se od nekoliko slojeva, od kojih svaki ima specifičnu funkciju:

1. **Konvolucijski sloj:** Sastoji se od skupa konvolucijskih filtera (jezgara) koji obrađuju ulazne slike kako bi generirali izlazne mape značajki. Jezgre imaju inicijalne težine koje se prilagođavaju tijekom procesa učenja kako bi izdvojile ključne značajke iz slike. Operacija konvolucije uključuje pomicanje jezgre preko slike, pri čemu se izračunava umnožak između slike i jezgre, stvarajući tako značajke izlazne mape.

2. **Sloj za uzorkovanje (Pooling sloj):** Ovaj sloj smanjuje dimenzije mapa značajki, čime se smanjuje složenost modela dok se zadržava većina bitnih informacija. Postoji nekoliko metoda uzorkovanja, uključujući max pooling i average pooling, koje se najčešće koriste. Iako pooling sloj pomaže u sažimanju informacija, može dovesti do gubitka važnih podataka.

3. **Aktivacijska funkcija:** Aktivacijska funkcija mapira ulazne vrijednosti na izlazne i odlučuje hoće li određeni neuron biti aktiviran. Najčešće korištene funkcije su sigmoidna, hiperbolička tangens funkcija i ReLU, pri čemu je ReLU najčešće korištena zbog svoje jednostavnosti i niskog računalnog opterećenja. Aktivacijske funkcije omogućuju CNN-u da modelira složene nelinearne odnose u podatcima.

4. **Potpuno povezani sloj (Fully Connected Layer):** Ovaj sloj se nalazi na kraju CNN arhitekture i povezuje sve neurone prethodnog sloja. On funkcionira kao klasifikator mreže, a njegov izlaz predstavlja konačnu klasifikaciju. Ulaz u ovaj sloj dolazi iz zadnjeg pooling ili konvolucijskog sloja, nakon što su značajke spljoštene u vektor.

5. **Funkcije gubitka (Loss Functions):** Funkcije gubitka koriste se u izlaznom sloju za izračunavanje razlike između stvarnog i predviđenog izlaza. Ova razlika koristi se za optimizaciju težina mreže tijekom procesa učenja, čime se poboljšava preciznost modela [5].

2.1.4. Podatci

Kvalitetni podatci temelj su uspješne primjene dubokog učenja. U kontekstu segmentacije satelitskih snimaka, podatci često dolaze iz različitih izvora poput satelita, zračnih snimaka ili dronova. Ovi podatci moraju biti točno označeni kako bi model mogao naučii razliku između različitih klasa.

Priprema podataka uključuje normalizaciju, augmentaciju i podjelu na skup za treniranje, validaciju i testiranje. Normalizacija pomaže u standardizaciji podataka, augmentacija povećava raznolikost podataka kroz transformacije kao što su rotacije i skaliranja, dok podjela podataka omogućava validaciju modela na neviđenim podatcima [6].

2.1.5. Funkcija gubitka

Većina algoritama dubokog učenja koristi neku vrstu optimizacije. Optimizacija se odnosi na zadatak minimiziranja ili maksimiziranja neke funkcije f(x). Funkcija koju želimo minimizirati ili maksimizirati naziva se ciljana funkcija ili kriterij. Kada je minimiziramo, možemo je također nazvati funkcijom troška, funkcijom gubitka ili funkcijom pogreške. Funkcija gubitka mjeri koliko su predviđanja modela udaljena od stvarnih vrijednosti. Za zadatke semantičke segmentacije često se koristi unakrsna entropija, koja mjeri razliku između predviđene distribucije i stvarne distribucije piksela.

Optimizacija se vrši minimizacijom funkcije gubitka korištenjem algoritama gradijentnog spusta, koji ažuriraju težine mreže na temelju gradijenta funkcije gubitka [7].

2.1.6. Propagacija pogreške unazad

Propagacija pogreške unazad ili *back-propagation* je algoritam za treniranje višeslojnih neuronskih mreža. Računa se gradijent funkcije gubitka s obzirom na težine mreže pomoću pravila lanca, omogućujući iterativno ažuriranje težina kako bi se smanjila pogreška. Svaka iteracija treniranja uključuje prolazak unaprijed kroz mrežu, izračun funkcije gubitka, te povratni prolazak kroz mrežu kako bi se ažurirale težine na temelju gradijenata [8].

2.2. Primjene u analizi slika

Duboko učenje je značajno unaprijedilo analizu slika, omogućujući preciznu klasifikaciju, detekciju i segmentaciju slika.

Prednosti korištenja konvolucijskih neuronskih mreža (CNN) u odnosu na tradicionalne neuronske mreže u kontekstu računalnog vida su sljedeće:

1. Značajka dijeljenja težina, koja smanjuje broj parametara koje mreža treba naučiti, što zauzvrat pomaže mreži da poboljša generalizaciju i izbjegne prekomjerno prilagođavanje (overfitting).

2. Istovremeno učenje slojeva za ekstrakciju značajki i sloja za klasifikaciju uzrokuje da izlaz modela bude visoko organiziran i izrazito ovisan o izdvojenim značajkama.

3. Implementacija velikih mreža mnogo je lakša s CNN-om nego s drugim vrstama neuronskih mreža [5].

3. Modeli računalnog vida

3.1. U-Net Model

U-Net je duboka konvolucijska neuronska mreža koja je prvotno predstavljena 2015. godine u radu Ronnebergera, Fischera i Broxa, te se ističe svojom sposobnošću da precizno segmentira slike na razini piksela.

3.1.1. Arhitektura U-Net modela

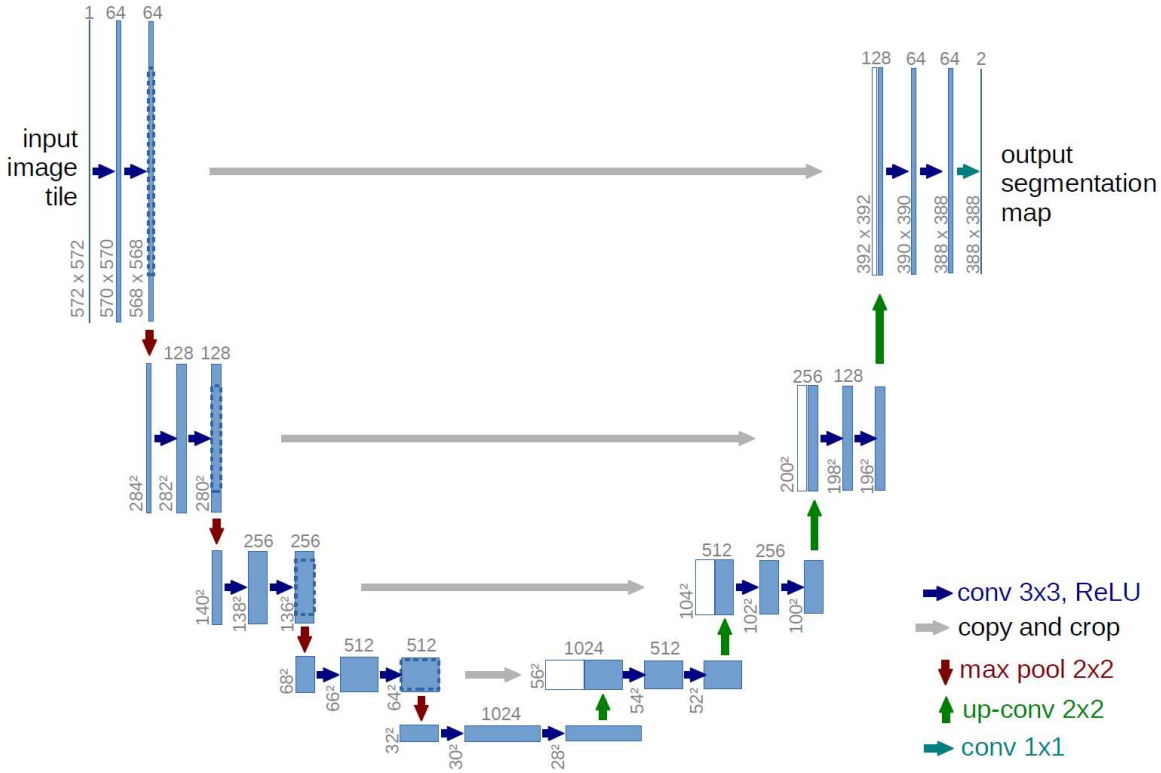
U-Net arhitektura ima oblik slova "U" te se sastoji od simetričnih sažimajućih i proširujućih staza. Sažimajuća staza (eng. *encoder*) koristi konvolucijske slojeve, ReLU aktivaciju i maksimalne pooling slojeve za ekstrakciju značajki i progresivno smanjenje dimenzionalnosti podataka, što omogućava modelu da uči sve apstraktnije značajke na višim razinama hijerarhije.

Proširujuća staza (eng. *decoder*) koristi dekonvolucijske slojeve, često implementirane kao transponirane konvolucije, za postupno povećanje dimenzionalnosti značajki i rekonstrukciju slike u visokoj rezoluciji. Ključna komponenta ove staze je sposobnost rekonstrukcije izvornog prostora, pri čemu se pažljivo integriraju značajke koje su bile izgubljene tijekom sažimanja.

Jedna od ključnih značajki U-Net-a je prisutnost *skip* veza, koje izravno prenose značajke iz odgovarajućih slojeva sažimajuće staze u proširujuću stazu. Ove veze omogućuju modelu da zadrži prostorne informacije, kao što su rubovi i konture objekata, koje bi se inače mogle izgubiti tijekom procesa sažimanja, čime se postiže preciznija segmentacija.

3.1.2. Primjena u segmentaciji satelitskih snimaka

U-Net se pokazao izuzetno učinkovit za segmentaciju satelitskih snimaka, gdje je preciznost na razini piksela ključna. Model može identificirati različite tipove zemljišta, vegetaciju, vodena tijela i druge važne značajke u satelitskim slikama. Popularan je izbor u području segmentacije slika zbog svoje viisoke točnosti segmentacije i relativno jednostavna implementacije.



Slika 3.1. Arhitektura U-Net mreže

3.2. VGG16 Model

VGG16 je duboka konvolucijska neuronska mreža razvijena od strane istraživačkog tima Visual Geometry Group (VGG) na Sveučilištu Oxford. Predstavljen je 2014. godine kao dio rada pod naslovom *"Very Deep Convolutional Networks for Large-Scale Image Recognition"*. Ovaj model postao je jedan od najpoznatijih i najkorištenijih arhitektura u po-

području računalnog vida, posebno za zadatke klasifikacije i segmentacije slika.

3.2.1. Arhitektura VGG16 modela

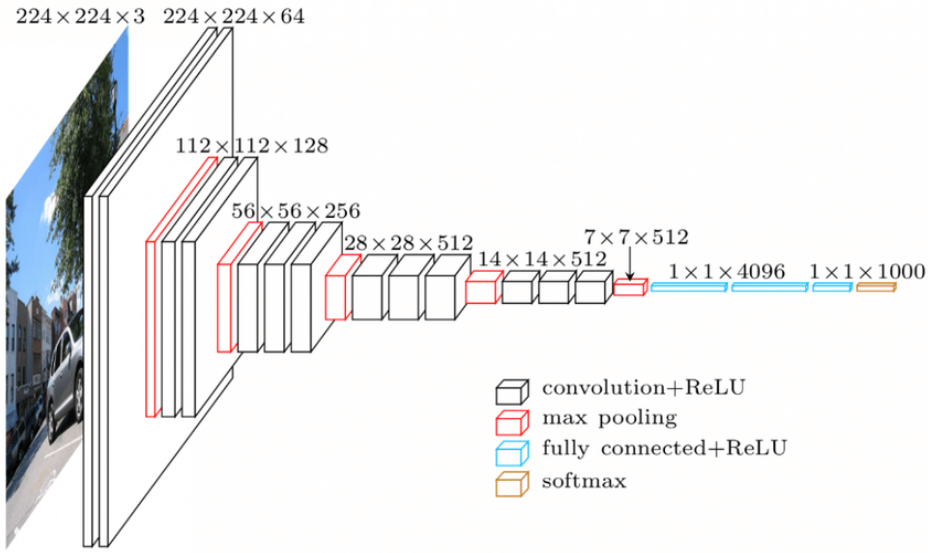
VGG16 se sastoji od 16 slojeva s težinama, uključujući 13 konvolucijskih slojeva, 3 potpuno povezana sloja, te završni sloj za klasifikaciju koji koristi softmax funkciju. Karakteristična značajka VGG16 je uporaba malih (3x3) konvolucijskih filtera, koji omogućuju povećanje dubine mreže bez povećanja broja parametara u značajnoj mjeri. Ova konfiguracija omogućuje mreži da uči složene značajke uz održavanje relativno male veličine modela.

Slojevi u VGG16 su raspoređeni u pet blokova, od kojih svaki završava maksimalnim pooling slojem. Maksimalni pooling (engl. max pooling) slojevi koriste se za smanjenje dimenzionalnosti podataka, pri čemu se zadržavaju ključne značajke, čime se poboljšava prostorna invarijantnost modela.

3.2.2. Primjena u segmentaciji satelitskih snimaka

Iako je VGG16 prvenstveno razvijen za zadatke klasifikacije slika, njegova modularna i duboka arhitektura često se koristi kao temelj za složenije modele segmentacije, kao što su FCN (Fully Convolutional Networks) i DeepLab. Ovi modeli transformiraju završne potpuno povezane slojeve VGG16 u konvolucijske slojeve, čime omogućuju da mreža vrši segmentaciju slike na razini piksela [9].

VGG16 se često kombinira s tehnikama poput transfernog učenja, gdje se unaprijed naučene težine iz velikih skupova podataka kao što je ImageNet prilagođavaju za specifične zadatke poput segmentacije satelitskih snimaka. Ova kombinacija omogućuje brzu i učinkovito treniranje modela na manjim skupovima podataka, uz zadržavanje visoke razine točnosti.



Slika 3.2. Arhitektura VGG16 mreže

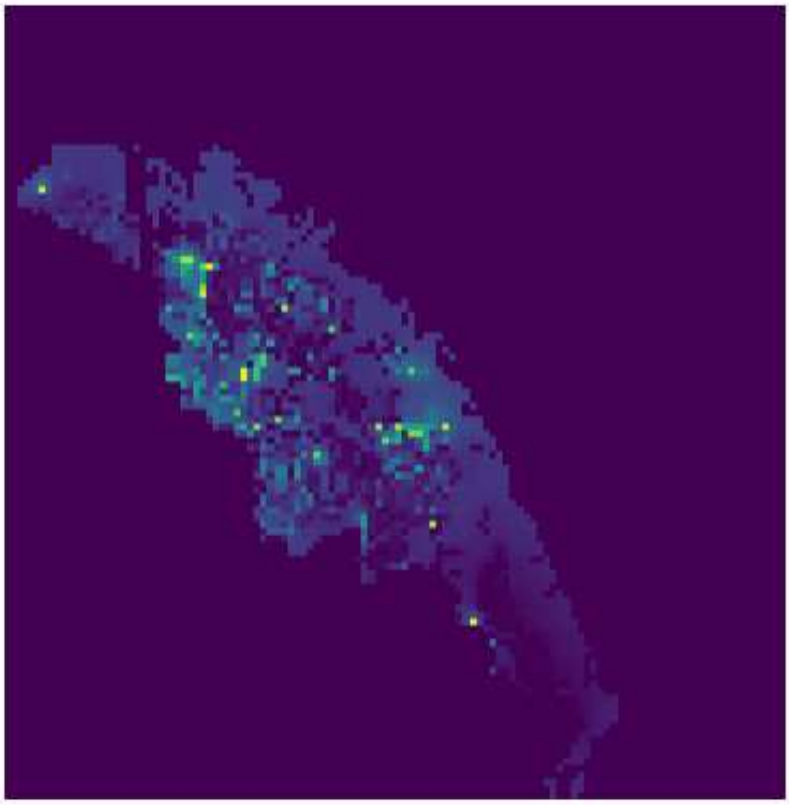
4. Priprema podataka

4.1. Izvor podataka

Ulazne satelitske slike šuma korištene u radu preuzete su od WorldView-2 satelita u vrlo visokoj rezoluciji [10], a referentne oznake prikupljene su od djelatnika Hrvatskog šumskog instituta. Podatci se sastoje od osam satelitskih snimaka istog područja šume na različitim valnim duljinama koje čine ulazne slike te od pet referentnih oznaka za pet tipova šume čija se prisutnost evaluira u radu. Pet tipova šuma koje se procijenjuju u radu su grab, jasen, hrast, joha i vrba.

Referentne oznake koje predstavljaju stvarne podatke o tipu šume su prikupljene od strane djelatnika Hrvatskog šumskog instituta. Oni su za svaku točku u rasteru, za svakih 100 metara procijenili dominantnu vrstu drveća u neposrednoj blizini, a područja između označenih točaka interpolirana su kako bi se dobio kontinuirani prikaz prostorne raspodjele vrsta šuma.

Slika 4.1. Ulazna satelitska snimka šume, kanal 705-745 nm



Slika 4.2. Referentna oznaka šume hrasta

Ulazne satelitske snimke podijeljene su u sljedeće kanale:

- 400-450 nm (coastal blue)
- 450-510 nm (blue)
- 510-580 nm (green)
- 585-625 nm (yellow)
- 630-690 nm (red)
- 705-745 nm (red edge)
- 770-895 nm (NIR1)
- 860-1040 nm (NIR2)

4.2. Priprema podataka

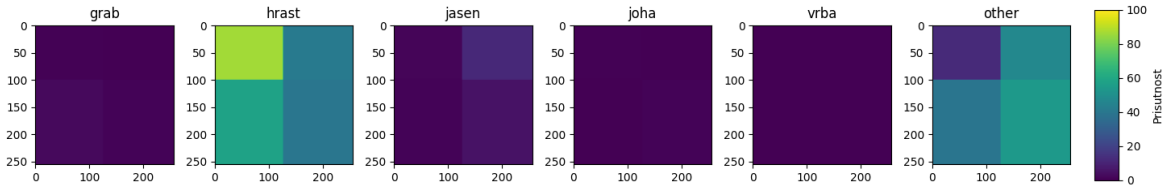
Prije samog učenja na podatcima, podatke je potrebno pripremiti kako bi međusobno pasali i odgovarali ulazu u model. Referentne oznake je bilo potrebno skalirati na razinu ulaznih satelitskih slika šume, za što je bilo potrebno povećati ih 200 puta.

Nakon skaliranja, satelitske slike i referentne oznake bilo je potrebno odrezati na segmente veličine 256×256 piksela kako bi odgovarali ulazu standardnih modela računalnog vida. Potom su odbačene sličice satelitskih šuma i referentnih oznaka koje su u sebi sadržavale piksele pozadine, sličice koje su bile manjih dimenzija jer su na rubovima slike i sličice koje imaju više od 50% piksela vrijednosti nula. Pri odbacivanju gledao se isti segment šume, te ako bi za istu površinu bila odbačena satelitska sličica broj 3, sličica bi bila odbačena i u drugim ulaznim slikama i u referentnim oznakama šume zbog formata podataka koji je potreban za daljnje treniranje modela.

Kod referentnih oznaka dodatno su odbačene sličice sa manje od 10% piksela vrijednosti većih od 1. Ovim pristupom osigurano je da bude što manje neinformativnih referentnih oznaka koje bi umanjile učinkovitost modela.

S obzirom na velik broj nedostajućih vrijednosti između pet klasa šuma u podatcima,

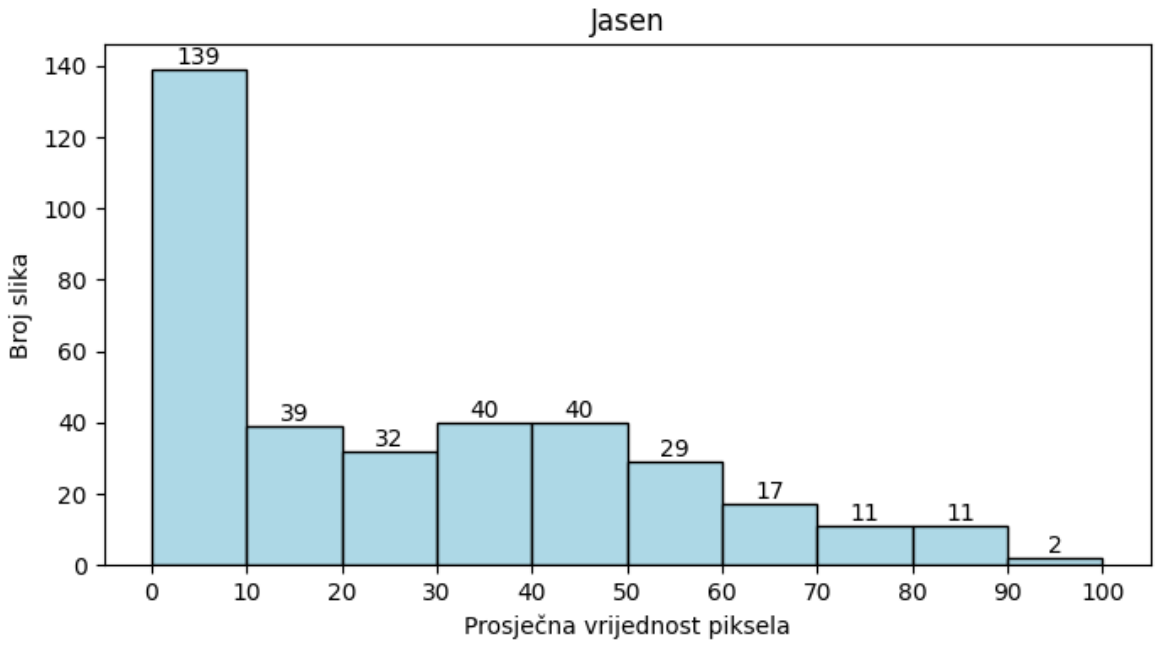
dodana je i šesta klasa *other* koja se koristi u nekim od ekspirementalnih postava u daljnjem radu. Ova klasa nastala je tako što je za svaki piksel pet referentnih oznaka uzeta vrijednost koja nedostaje pikselima ostalih oznaka šuma do zajedničkog zbroja 100. Na slici 4.3. prikazano je kako izgleda šest referentnih oznaka za isti segment šume.



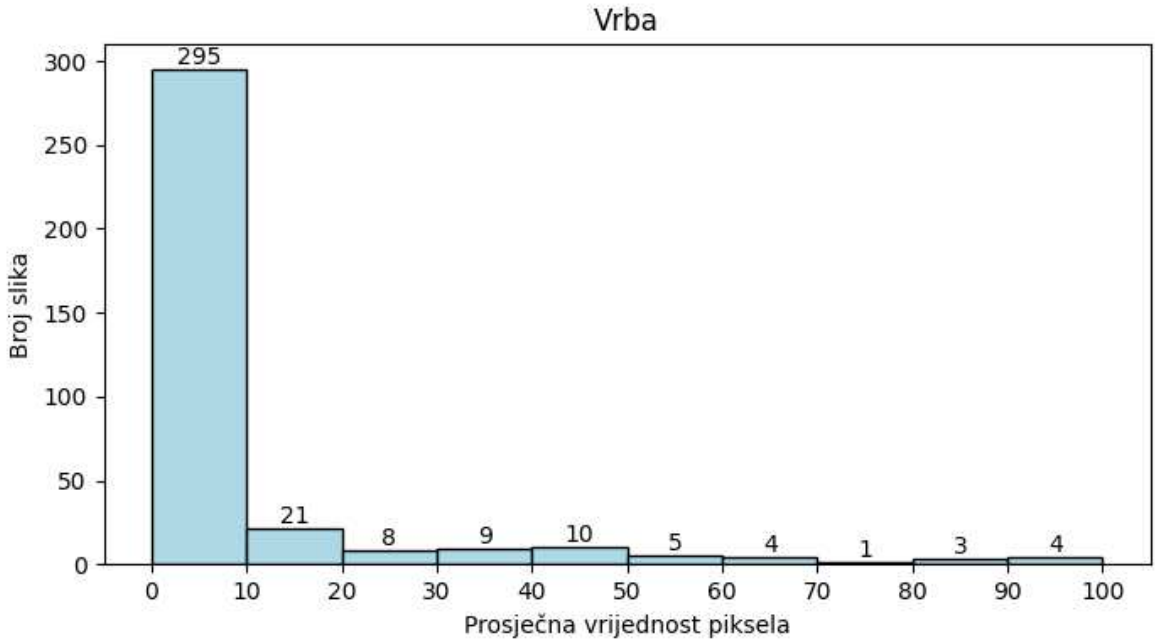
Slika 4.3. Referentne oznake veličine 256×256

4.3. Analiza distribucije podataka

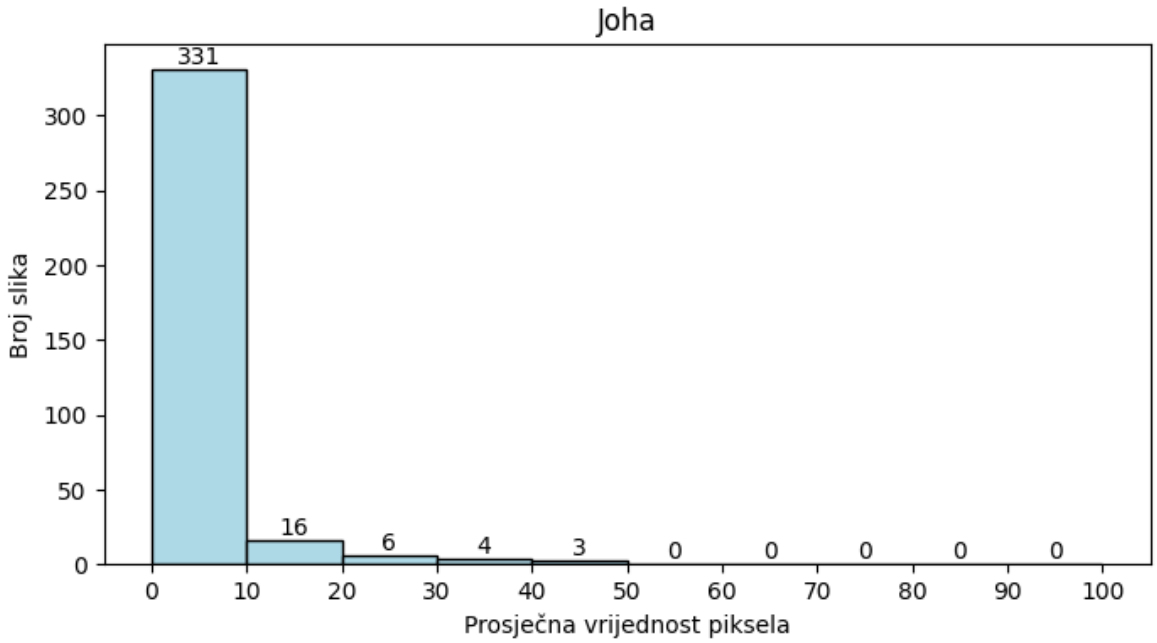
Prije odabira modela, važno je provjeriti podatke kako bi se osigurala uravnoteženost klasa i na temelju toga odabrati odgovarajući pristup treniranju modela.



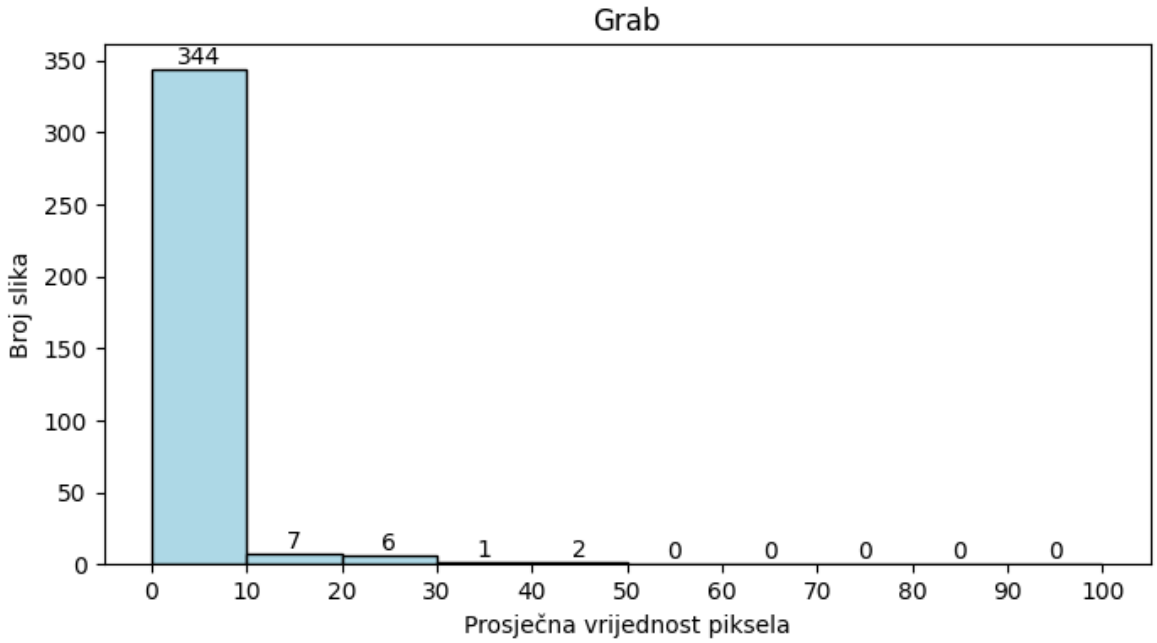
Slika 4.4. Distribucija vrijednosti za klasu *jasen*



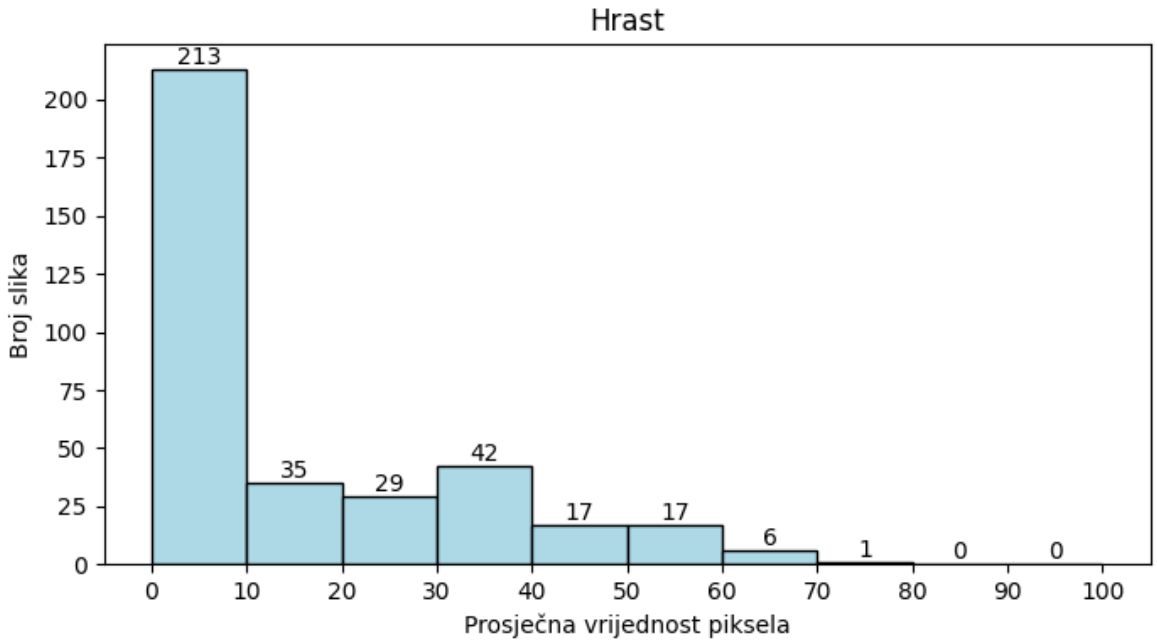
Slika 4.5. Distribucija vrijednosti za klasu *vrba*



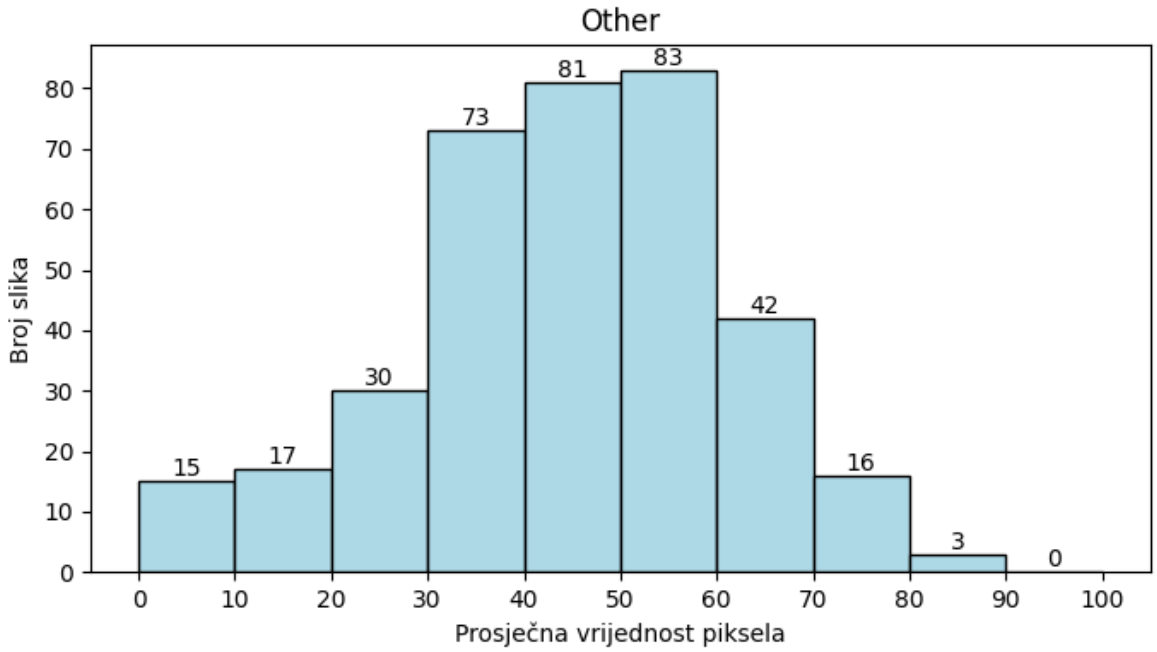
Slika 4.6. Distribucija vrijednosti za klasu *joha*



Slika 4.7. Distribucija vrijednosti za klasu *grab*



Slika 4.8. Distribucija vrijednosti za klasu *hrast*



Slika 4.9. Distribucija vrijednosti za klasu *other*

Na histogramima 4.4. do 4.9. prikazana je uprosječena vrijednost piksela po sličici za svaku klasu. Iz grafova se vidi da podatci referentnih oznaka nisu balansirani s obzirom da klase *joha* i *grab* imaju znatno niže vrijednosti oznaka od ostalih klasa. Klasa *other* ima najviše *"informativnih"* piksela na svojim referentnim oznaka, no s obzirom da je ova klasa dobivena oduzimanjem vrijednosti ostalih klasa na istom pikselu slike od broja 100, zapravo upućuje na to da je u većini slika zbroj piksela ostalih klasa manji od 100. Ovo je problematično jer ukazuje na to da bi referentne oznake mogle biti neinformativne pri treniranju modela.

Problem s manjkom informativnih oznaka može se rješiti generiranjem dodatnih podataka to jest povećavanjem uzorka klasa s manje primjera ili smanjivanjem uzorka klasa koje imaju više primjera. U ovom radu odbačene su referentne oznake koje su u manje od 10% piksela sadržavale vrijednosti manje od jedan i tako su zadržane najinformativnije oznake. Kriterij nije bio stroži od ovoga jer bi u suprotnom ostalo nedovoljno podataka za treniranje modela.

4.4. Podjela podataka

Podatci su ravnomjerno raspoređeni u tri skupa: 60% podataka dodijeljeno je skupu za treniranje, 20% skupu za validaciju, a preostalih 20% skupu za testiranje.

- Skup za treniranje (60%) koristi se za učenje modela, odnosno za prilagodbu parametara modela kako bi što bolje predviđao željene izlaze.
- Skup za validaciju (20%) koristi se za podešavanje hiperparametara modela i praćenje performansi tijekom treniranja, kako bi se spriječilo prenaučavanje i osigurala dobra generalizacija modela.
- Skup za testiranje (20%) koristi se za konačnu evaluaciju modela nakon treniranja, kako bi se procijenile njegove performanse na neviđenim podatcima.

Kako bi se osigurala ravnomjerna distribucija u svakom skupu, sličice su podijeljene u klase ovisno o najzastupljenijoj klasi prikazanog područja, sortirane unutar svake klase te potom podijeljene u skupove za treniranje, validaciju i testiranje u kružnom redoslijedu. Ovaj način raspodjele osigurao je da svi skupovi imaju približno jednako informativne podatke za svaku klasu.

5. Eksperimentalni postav

5.1. Implementacijski detalji

Eksperimenti i treniranje modela provedeni su u Jupyter Notebook okruženju na prijenosnom računalu s macOS operativnim sustavom. Implementacija je izvršena u Pythonu verzije 3.11.5, koristeći TensorFlow verziju 2.15.1 zajedno s Keras API-jem za izgradnju i treniranje dubokih neuronskih mreža.

5.2. Generator podataka

Generator podataka koristi satelitske slike koje obuhvaćaju ili svih osam spektralnih kanala ili tri spektralna kanala, 2, 5 i 6, ovisno o postavu. Kanali 2, 5 i 6 su odabrani jer su tijekom treniranja modela pružali maksimalnu količinu korisnih informacija za segmentaciju. Podatci su označeni s ukupno šest različitih referentnih oznaka koje predstavljaju različite klase unutar šume. Ove klase su *joha*, *jasen*, *hrast*, *grab*, *vrba* i klasa *other*. U eksperimentalnim postavima s 5 izlaza korišten je isti generator ali bez klase *other*.

Sve ulazne slike i referentne oznake poravnate su tako da odgovaraju istim segmentima šumskog područja. Ovaj postupak usklađivanja osigurava da svaki piksel u svim ulaznim slikama i referentnim oznakama predstavlja isti fizički dio šume. Vrijednosti referentnih oznaka skalirane su na raspon [0, 1] zbog kompatibilnosti s arhitekturama korištenih modela.

5.2.1. Augmentacija podataka

Tehnika augmentacije podataka korištena je u nekim od eksperimentalnih postava kako bi se povećala raznolikost skupa za treniranjep, čime se modelu omogućava bolja generalizacija na neviđene podatke. Augmentacija uključuje rotaciju slika unutar raspona od

10 stupnjeva, horizontalno i vertikalno pomicanje do 10% širine i visine slike te horizontalno obrtanje slika.

Za očuvanje kvalitete slika nakon primjene ovih transformacija, korištena je opcija reflektiranja rubova koja osigurava da rubovi slika, koji bi inače bili popunjeni crnim pikselima uslijed transformacija, budu popunjeni reflektiranjem susjednih piksela. Ovaj pristup smanjuje rizik od uvođenja artefakata na rubovima slika, što može negativno utjecati na učenje modela.

5.3. U-net model

Prvi eksperimentalni postav usmjeren na problem segmentacije satelitskih slika šuma izveden je koristeći konvolucijsku neuronsku mrežu U-Net.

5.3.1. U-net model, 3 ulaza, 5 izlaza

Podatci

- **Ulazne slike**: Satelitske slike u rezoluciji 256×256 piksela s tri kanala: 2, 5 i 6.
- **Izlazne slike**: Referentne oznake koje predstavljaju pokrivenost pet različitih tipova šuma, pri čemu su vrijednosti u rasponu od 0 do 1.

Arhitektura modela

- **U-Net arhitektura**: Model koristi modificiranu U-Net arhitekturu, koja se sastoji od kontrahirajućeg puta (enkoder) i ekspanzivnog puta (dekoder). Enkoder reducira dimenzionalnost slike i ekstraktira značajke, dok dekoder povećava dimenzionalnost i rekonstruira sliku.
- **Konvolucijski blokovi**: Svaki konvolucijski blok sastoji se od dva konvolucijska sloja s ReLU aktivacijom i *same* paddingom. Koristi se i batch normalizacija za stabilizaciju treniranja.
- **Downsampling (enkoder)**: Enkoder koristi MaxPooling2D slojeve za smanjenje dimenzija slike, čime se omogućava modelu da uči apstraktne značajke.

- **Upsampling (dekoder)**: Dekoder koristi Conv2DTranspose slojeve za povećanje dimenzija slike, uz korištenje *skip* veza kako bi zadržao informacije iz enkodera.

Izlaz modela

Zadnji sloj modela koristi sigmoidnu aktivacijsku funkciju za svaku od pet klasa, što omogućava predviđanje vrijednosti između 0 i 1 za svaki piksel.

Konfiguracija modela

- **Optimizator**: Adam optimizator sa stopom učenja od 0.0002.
- **Funkcija gubitka**: Srednja kvadratna pogreška (`mean_squared_error`)
- **Metričke funkcije**: Koeficijent determinacije (R-squared), korijen srednje kvadratne pogreške (RMSE) i srednja kvadratna pogreška (MSE) na razini svake klase

5.3.2. U-net model, 8 ulaza, 5 izlaza

Ova konfiguracija sadržava sve elemente prethodnog modela, uz jedinu razliku u ulaznim podatcima. Sada model koristi svih osam satelitskih slika kao ulazne podatke, dok na izlazu ostaje pet klasa šuma.

5.4. VGG16 model

5.4.1. VGG16 model, 3 ulaza, 5 izlaza

Podatci

- **Ulazne slike**: Satelitske slike u rezoluciji 256×256 piksela s tri kanala: 2, 5 i 6.
- **Izlazne slike**: Referentne oznake koje predstavljaju pokrivenost pet različitih tipova šuma, pri čemu su vrijednosti u rasponu od 0 do 1.

Arhitektura modela

- **VGG16 baza**: Model koristi VGG16 kao bazu, pri čemu su težine prethodno trenirane na ImageNet skupu podataka. VGG16 služi kao enkoder koji ekstraktira

značajke iz ulaznih slika, a svi slojevi su inicijalno zamrznuti kako bi se spriječilo njihovo treniranje.

- **Konvolucijski slojevi i upsampling**: Nakon enkodera, model uključuje dodatne konvolucijske slojeve i slojeve za povećanje dimenzionalnosti slike (Conv2DTranspose). Svaki konvolucijski sloj koristi `LeakyReLU` aktivaciju i `BatchNormalization` za stabilizaciju treniranja, dok se `Dropout` slojevi koriste za sprječavanje pretreniranja.

- **Upsampling put**: Dekoder koristi više slojeva `Conv2DTranspose` za postupno povećanje dimenzija slike natrag na originalnu rezoluciju 256×256 piksela, čime se omogućava modelu da generira segmentirane slike.

Izlaz modela

Zadnji sloj modela koristi sigmoidnu aktivacijsku funkciju za svaku od pet klasa.

Konfiguracija modela

- **Optimizator**: Adam optimizator sa stopom učenja od 0.0001

- **Funkcija gubitka**: Srednja kvadratna pogreška (`mean_squared_error`)

- **Metričke funkcije**: Koeficijent determinacije (R-squared), korijen srednje kvadratne pogreške (RMSE) i srednja kvadratna pogreška (MSE) na razini svake klase

5.4.2. VGG16 model, 3 ulaza, 6 izlaza

Arhitektura ovog VGG16 modela identična je prethodnoj, osim što ima 6 izlaznih referentnih oznaka koje uključuju 5 tipova šuma i klasu *other*. Također, ovaj model na izlazu koristi softmax aktivacijsku funkciju kako bi predvidio vjerojatnost za svaku od šest klasa. Za optimizator je opet izabran Adam, ali sa stopom učenja od 0.00005.

5.5. *Callback* funkcije

Pri treniranju modela korištene su *callback* funkcije koje omogućuju kontrolu i prilagodbu ponašanja modela tijekom učenja, poput zaustavljanja treniranja, promjene stope

učenja ili spremanja najboljih rezultata. U ovom radu koristilo se nekoliko *callback* funkcija kako bi se osigurali bolji rezultati:

- **ModelCheckpoint**: koristi se za spremanje modela samo kada postigne najbolji rezultat na skupu podataka za validaciju.
- **EarlyStopping**: zaustavlja treniranje ako se `val_loss` ne poboljša nakon pet iteracija, čime se sprječava prekomjerno treniranje modela.
- **ReduceLROnPlateau**: smanjuje stopu učenja za pola kada se `val_loss` ne poboljša nakon tri iteracije, što omogućuje modelu da bolje uči kada se približi minimalnom gubitku.

5.6. Metrike

5.6.1. Mean Squared Error (MSE)

MSE (*srednja kvadratna pogreška*) je prosječna kvadratna razlika između stvarnih vrijednosti i predviđanja modela. MSE daje veću težinu većim pogreškama zbog kvadriranja razlika, što može pomoći u identificiranju modela koji ima velika odstupanja u predviđanjima. MSE se računa kao:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2$$

gdje su y_i stvarne vrijednosti, $\hat{y}_i$ predviđene vrijednosti, a n broj podataka.

5.6.2. Root Mean Square Error (RMSE)

RMSE (*korijen srednje kvadratne pogreške*) je mjera prosječne magnitude pogreške između predviđenih i stvarnih vrijednosti. RMSE uzima kvadratnu razliku između predviđene i stvarne vrijednosti, računa prosjek tih kvadratnih razlika i zatim uzima kvadratni korijen tog prosjeka. Manja vrijednost RMSE označava bolje performanse modela. RMSE se računa kao:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}$$

gdje su y_i stvarne vrijednosti, $\hat{y}_i$ predviđene vrijednosti, a n broj podataka.

5.6.3. R-squared

R-squared (*koeficijent determinacije*) je statistička mjera koja pokazuje udio varijance u zavisnoj varijabli koji je objašnjen nekom nezavisnom varijablom ili modelom. Vrijednost R-squared varira između 0 i 1, gdje vrijednost bliža 1 znači da model bolje objašnjava varijaciju podataka. R-squared se može izraziti kao:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y}_i)^2}{\sum (y_i - \bar{y})^2}$$

gdje su y_i stvarne vrijednosti, $\hat{y}_i$ predviđene vrijednosti, a $\bar{y}$ srednja vrijednost stvarnih vrijednosti.

6. Rezultati i rasprava

U ovom dijelu rada predstavljeni su rezultati koji su postignuti treniranjem modela na zadatku segmentacije satelitskih slika šuma. Analizirani su učinci različitih arhitektura i augmentacije podataka na konačne performanse modela. Glavni cilj bio je procijeniti sposobnost modela da precizno segmentira slike, identificira različite tipove šuma, te ispitati koliko dobro model radi na neviđenim podatcima.

Rezultati su prikazani u obliku metrika koeficijenta determinacije (R-squared), korijena srednje kvadratne pogreške (RMSE) i srednje kvadratne pogreške (MSE) na razini svake klase.

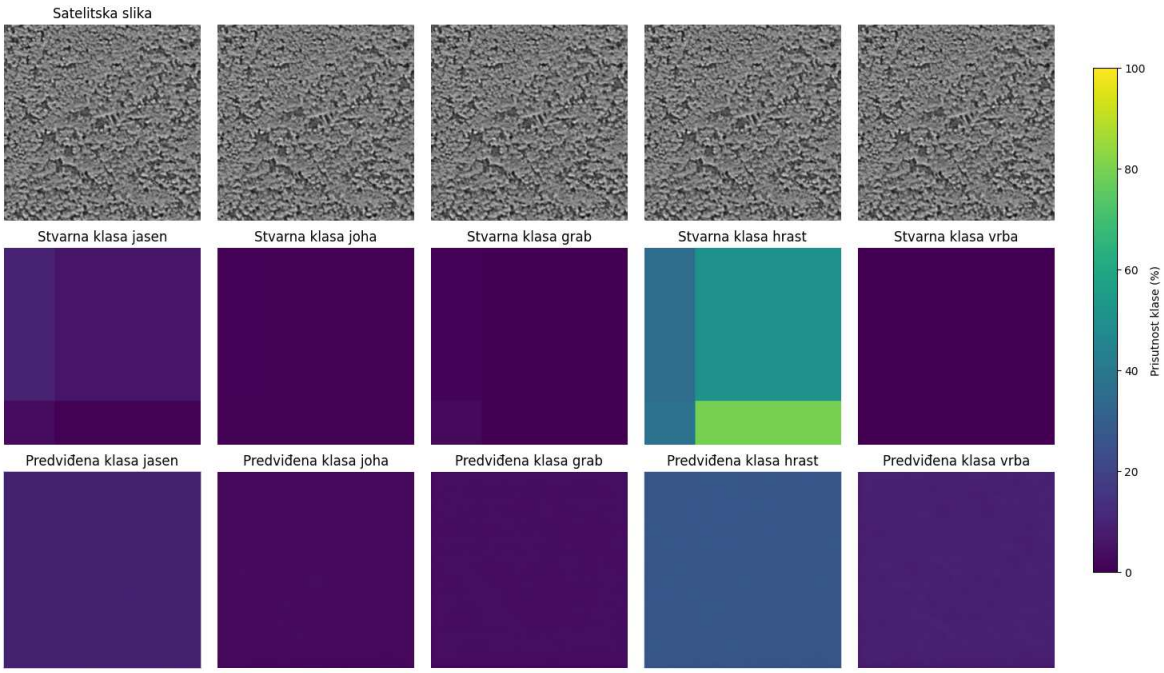
6.1. Analiza rezultata modela

6.1.1. U-Net model, 3 ulaza, 5 izlaza

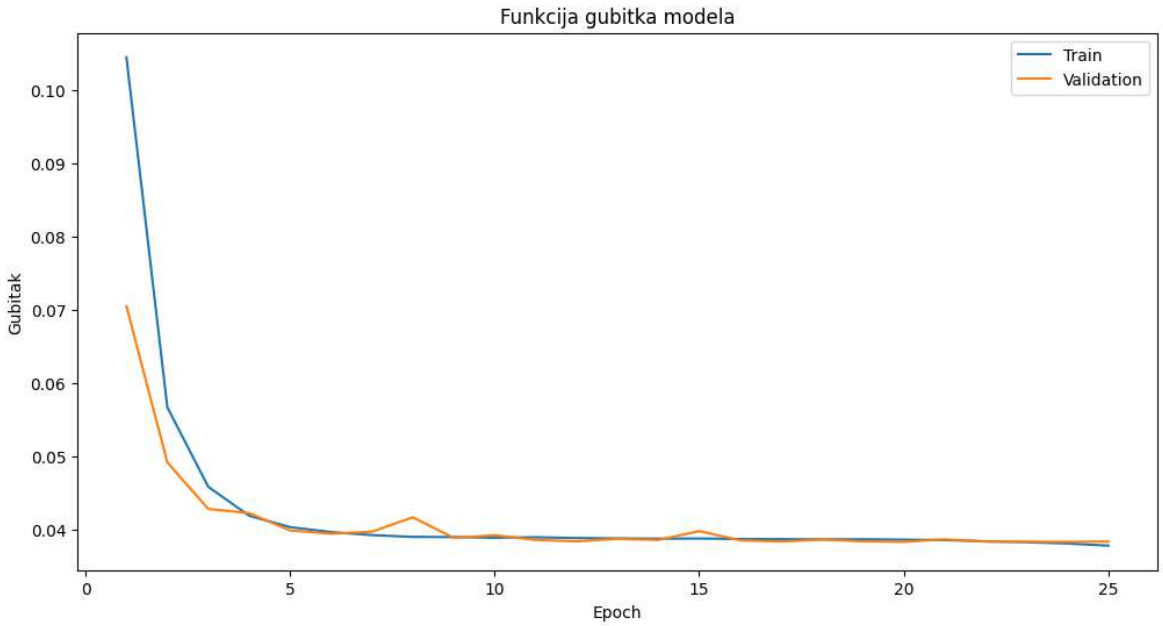
Bez augmentacije podataka

Gubitak modela na skupu za treniranje bio je relativno nizak (0,0215), što sugerira da model uspješno minimizira razliku između predviđenih i stvarnih vrijednosti. Međutim, R-squared vrijednost od 0.2483 pokazuje da model objašnjava samo oko 24% varijance u podatcima, što sugerira da model možda ne uspijeva prepoznati sve obrasce u podatcima.

RMSE od 0,1261 ukazuje na umjerenu razinu pogreške. MSE po klasama pokazao je da model najviše griješi kod klase *jasen*, dok su klase *joha* i *grab* bile bolje predviđene s nižim MSE vrijednostima.



Slika 6.1. Rezultati predviđanja U-Net modela



Slika 6.2. Prikaz funkcije gubitka tijekom treniranja U-Net modela

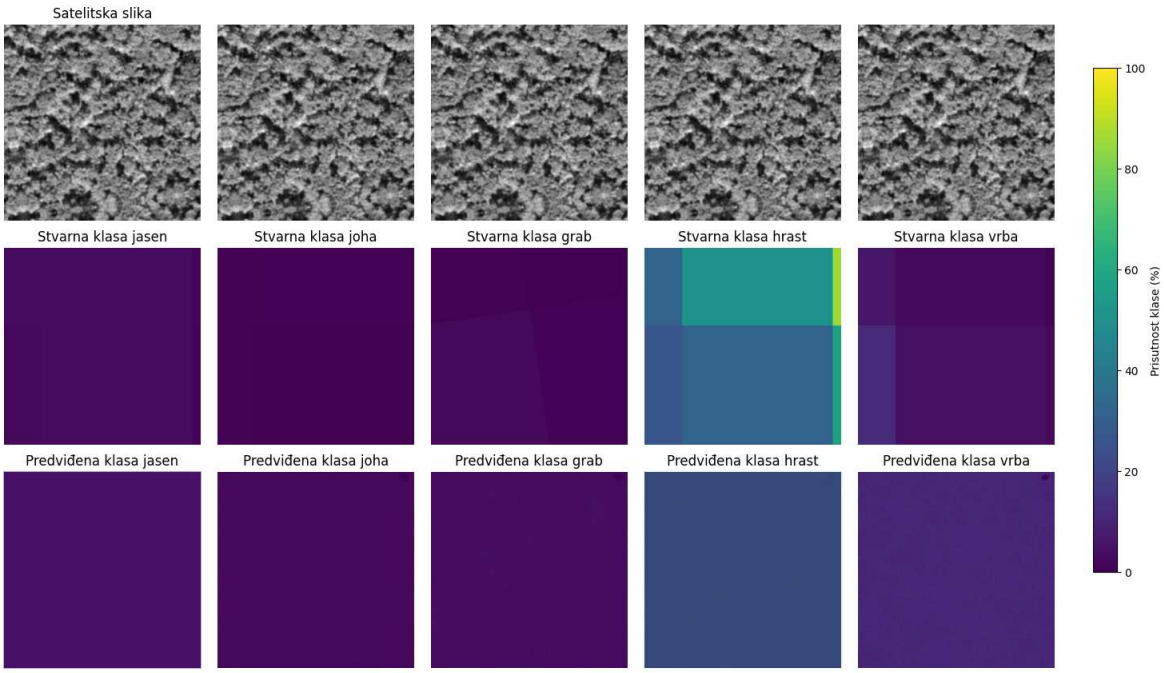
Metrika	**Treniranje**	**Validacija**
Loss	0.0215	0.0302
R-Squared	0.2483	0.1925
RMSE	0.1261	0.1420
MSE klasa *jasen*	0.0497	0.0663
MSE klasa *joha*	0.0013	0.0012
MSE klasa *grab*	0.0012	0.0037
MSE klasa *hrast*	0.0242	0.0364
MSE klasa *vrba*	0.0310	0.0433
Stopa učenja	0.0002	
Broj epoha	25	

Tablica 6.1. Metrike treniranja U-Net modela

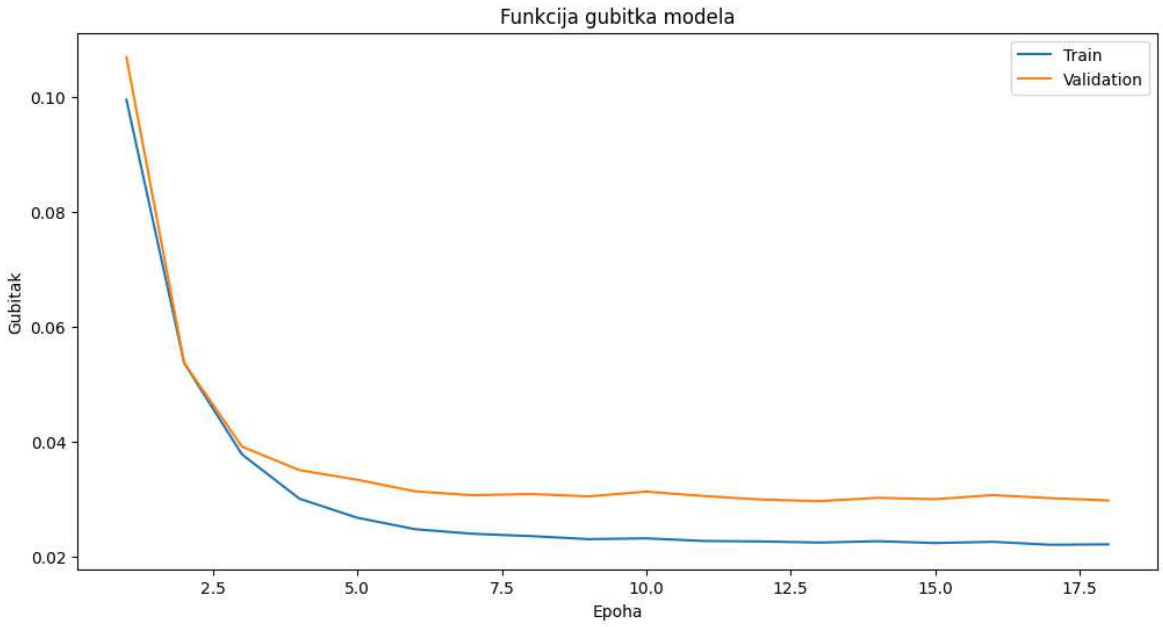
S augmentacijom podataka

Rezultati treniranja iz tablice 6.2. pokazuju kako je model postigao relativno nizak gubitak na skupu podataka za treniranje od 0,0221, što ukazuje na uspješno minimiziranje razlike između predviđenih i stvarnih vrijednosti. R-squared vrijednost od 0,2874 pokazuje da model objašnjava oko 28,74% varijance u podatcima, što sugerira da postoji prostor za poboljšanje modela kako bi se bolje prepoznali obrasci u podatcima. RMSE od 0,1276 ukazuje na nisku prosječnu pogrešku u predviđanjima.

MSE po klasama ukazuje na to da model najviše griješi kod klase *jasen*, dok ostale klase, iako imaju niže vrijednosti, također pokazuju prostor za optimizaciju.



Slika 6.3. Rezultati predviđanja U-Net modela



Slika 6.4. Prikaz funkcije gubitka tijekom treniranja U-Net modela

Metrika	**Treniranje**	**Validacija**
Loss	0.0221	0.0297
R-Squared	0.2874	0.2937
RMSE	0.1276	0.1299
MSE klasa *jasen*	0.0537	0.0645
MSE klasa *joha*	0.0012	0.0010
MSE klasa *grab*	0.0012	0.0035
MSE klasa *hrast*	0.0238	0.0358
MSE klasa *vrba*	0.0306	0.0439
Stopa učenja	0.0002	
Broj epoha	18	

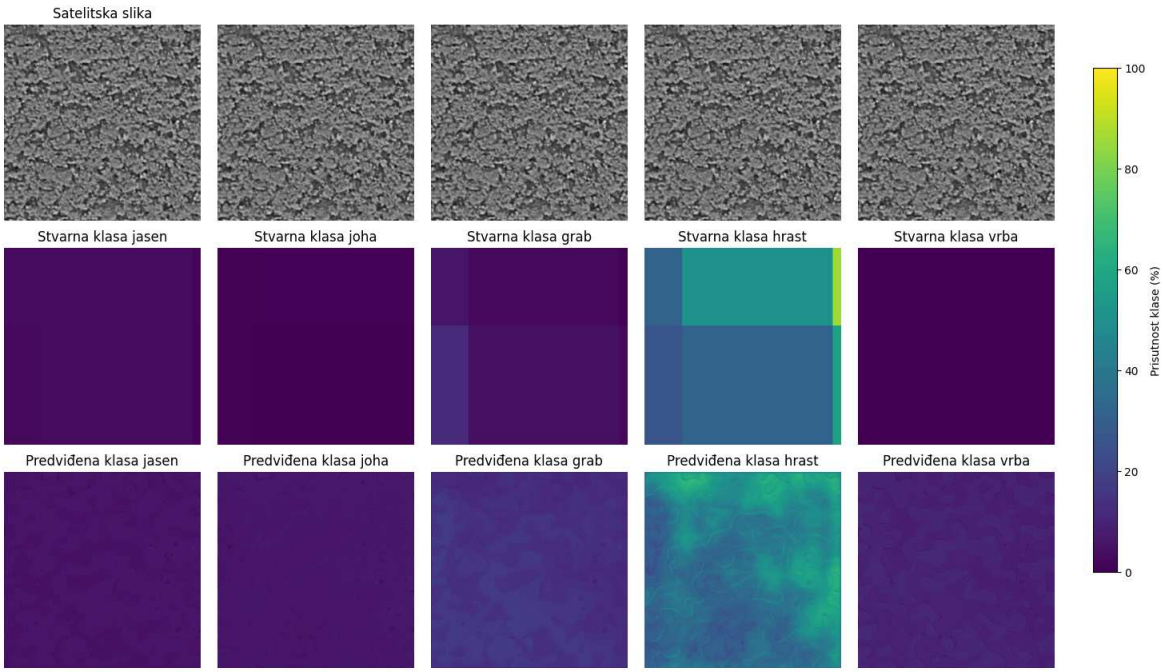
Tablica 6.2. Metrike treniranja U-Net modela

6.1.2. U-Net model, 8 ulaza, 5 izlaza

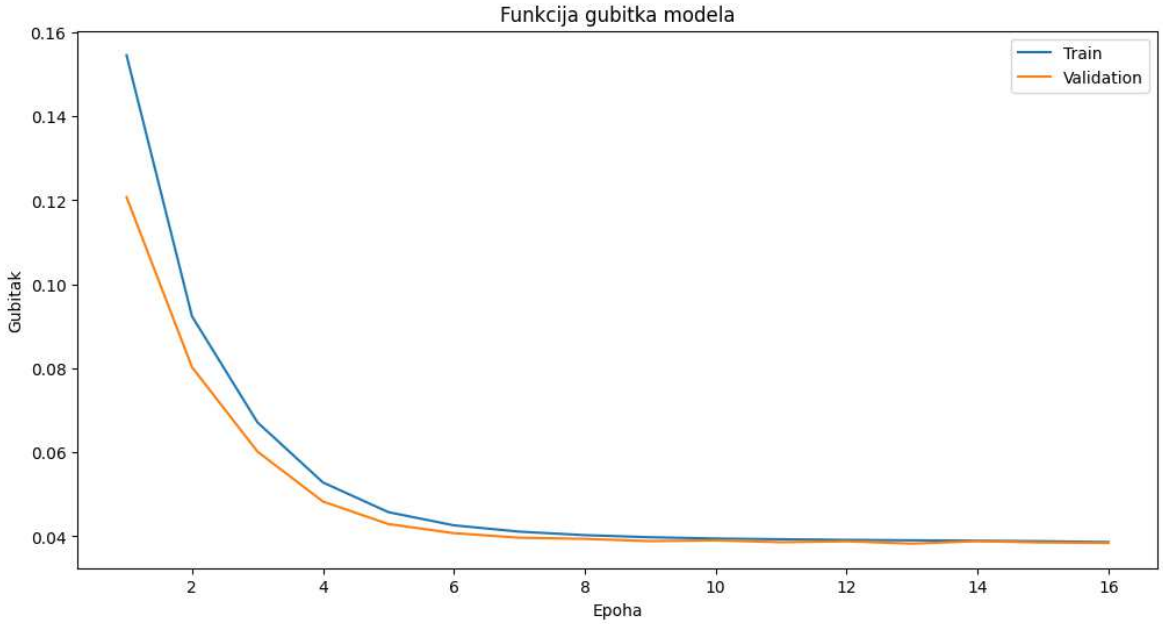
Bez augmentacije podataka

Rezultati pokazuju da model ima solidne performanse, s gubitkom od 0.0386 na skupu za treniranje i sličnim gubitkom od 0.0384 na validacijskom skupu. R-squared vrijednost od 0.2648 na skupu za treniranje i 0.2609 na validacijskom skupu ukazuje da model objašnjava oko 26% varijance u podatcima, što je umjeren rezultat, ali također sugerira da postoji prostor za daljnje poboljšanje modela.

RMSE je približno 0.1883 na skupu za treniranje i 0.1863 na skupu za validaciju, što je konzistentno i sugerira da model ima sličan nivo greške u oba skupa podataka. MSE po klasama pokazuje da je model najmanje grešaka napravio u klasama *joha* i *grab*.



Slika 6.5. Rezultati predviđanja U-Net modela



Slika 6.6. Prikaz funkcije gubitka tijekom treniranja U-Net modela

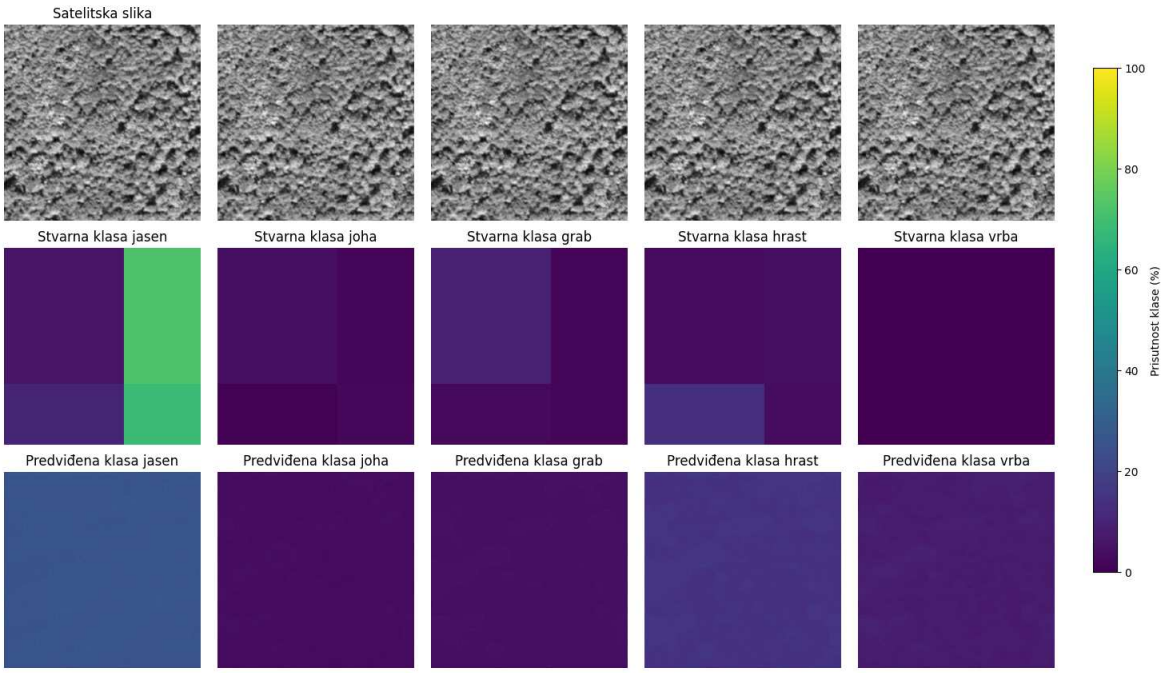
Metrika	**Treniranje**	**Validacija**
Loss	0.0386	0.0384
R-Squared	0.2648	0.2609
RMSE	0.1883	0.1863
MSE klasa *jasen*	0.0903	0.0978
MSE klasa *joha*	0.0085	0.0117
MSE klasa *grab*	0.0066	0.0039
MSE klasa *hrast*	0.0454	0.0484
MSE klasa *vrba*	0.0423	0.0304
Stopa učenja	0.0002	
Broj epoha	16	

Tablica 6.3. Metrike treniranja U-Net modela

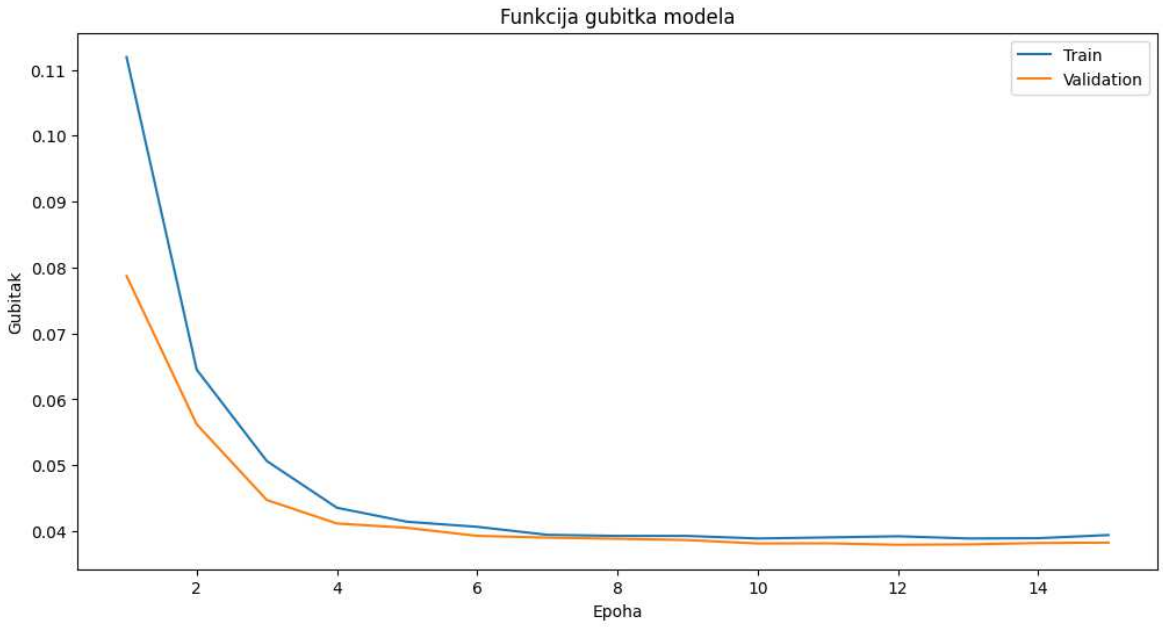
S augmentacijom podataka

Model postiže umjerene performanse, s gubitkom od 0.0394 na skupu za treniranje i nešto manjim gubitkom od 0.0383 na skupu za validaciju, što sugerira da model relativno dobro generalizira na neviđene podatke. Međutim, R-squared vrijednost od 0.2056 na treniranju i 0.1955 na validaciji ukazuje da model objašnjava samo oko 20% varijance u podatcima, što nije visoko i sugerira da model još uvijek ne prepoznaje sve obrasce u podatcima, ostavljajući značajan prostor za poboljšanje.

RMSE vrijednosti od 0.1895 na skupu za treniranje i 0.1851 na validacijskom skupu su konzistentne i ukazuju na ujednačenu razinu greške između treniranja i validacije. MSE po klasama pokazuje da je model najviše grešaka napravio u klasi *jasen*, dok su ostale klase pokazale nešto niže greške, ali s prostorom za poboljšanje.



Slika 6.7. Rezultati predviđanja U-Net modela



Slika 6.8. Prikaz funkcije gubitka tijekom treniranja U-Net modela

Metrika	Treniranje	Validacija
Loss	0.0394	0.0383
R-Squared	0.2056	0.1955
RMSE	0.1895	0.1851
MSE klasa *jasen*	0.0925	0.0974
MSE klasa *joha*	0.0084	0.0102
MSE klasa *grab*	0.0072	0.0044
MSE klasa *hrast*	0.0454	0.0492
MSE klasa *vrba*	0.0435	0.0301
Stopa učenja	0.0002	
Broj epoha	15	

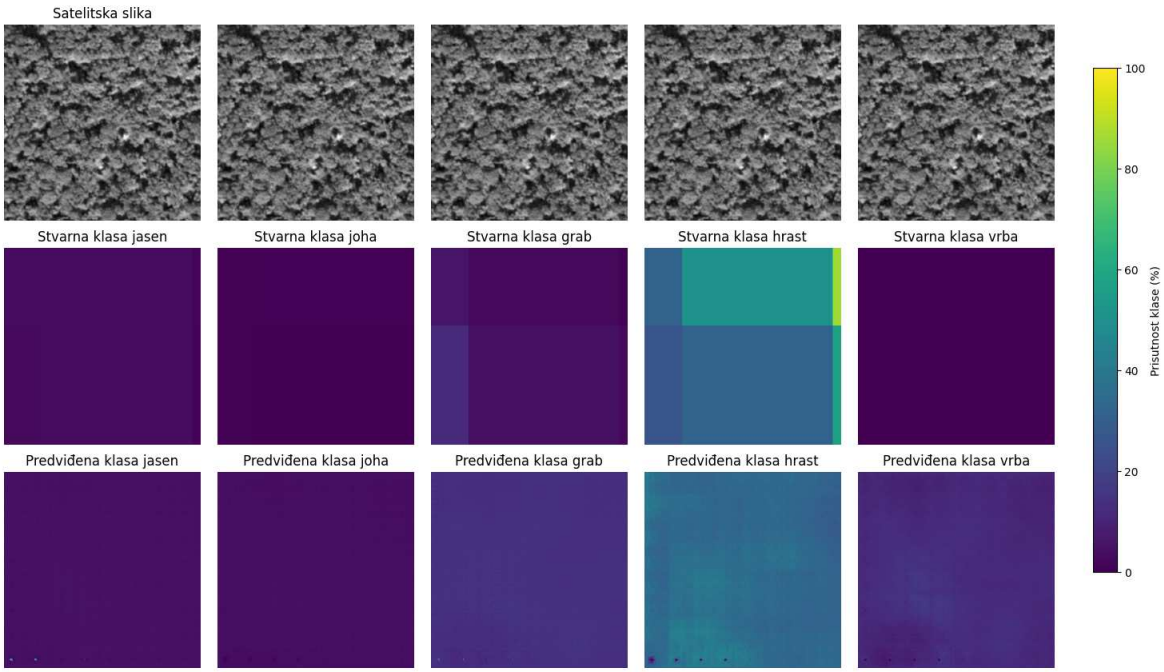
Tablica 6.4. Metrike treniranja U-Net modela

6.1.3. VGG16 model, 3 ulaza, 5 izlaza

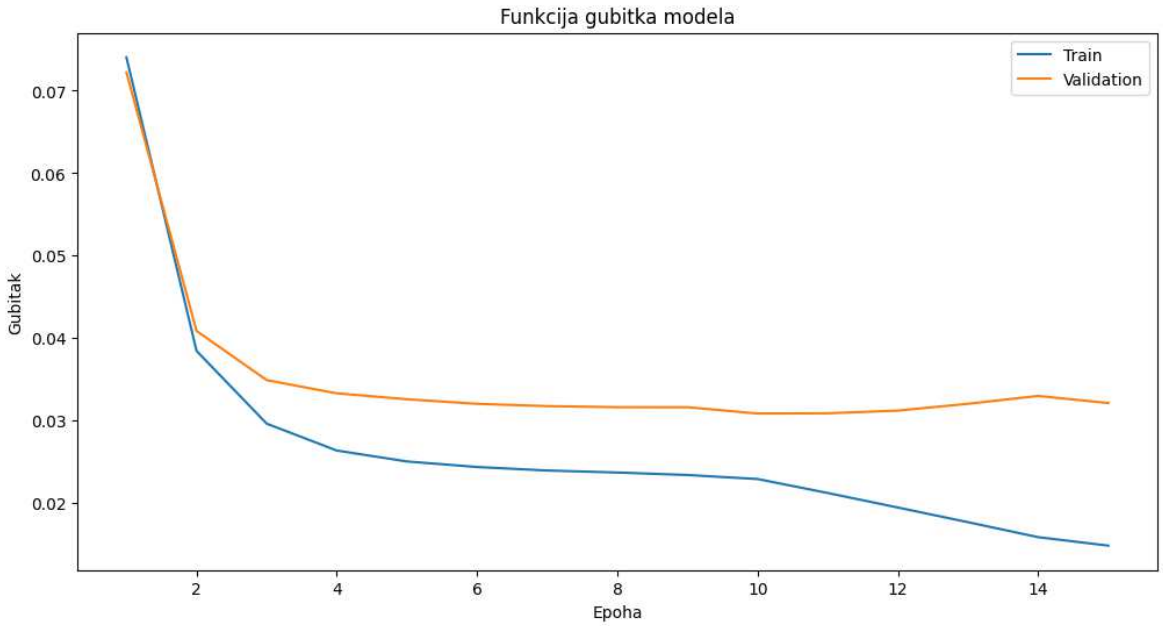
Bez augmentacije podataka

Tijekom treniranja, model je postigao gubitak od 0.0148, uz R-squared vrijednost od 0.4979, što ukazuje da model uspijeva objasniti gotovo 50% varijance u podatcima za treniranje. RMSE iznosi 0.1044, što sugerira relativno nisku prosječnu pogrešku. Gledajući srednje kvadratne pogreške po klasama, najveća pogreška je kod klase *vrba* (0.0282), dok su ostale klase unutar nižih granica pogreške, što pokazuje da model uspješno predviđa većinu klasa s malim pogreškama.

Gubitak na validaciji iznosi 0.0321, što je značajno više u usporedbi s treniranjem, uz manju R-squared vrijednost 0.3183, što sugerira da model ne generalizira dobro na nove podatke. RMSE na validaciji je također viši (0.1598), a MSE vrijednosti po klasama, posebno za klasu *jasen* (0.0744) i klasu *vrba* (0.0439), ukazuju na povećane pogreške u predikciji tih klasa na validacijskom skupu.



Slika 6.9. Rezultati predviđanja VGG16 modela



Slika 6.10. Prikaz funkcije gubitka tijekom treniranja VGG16 modela

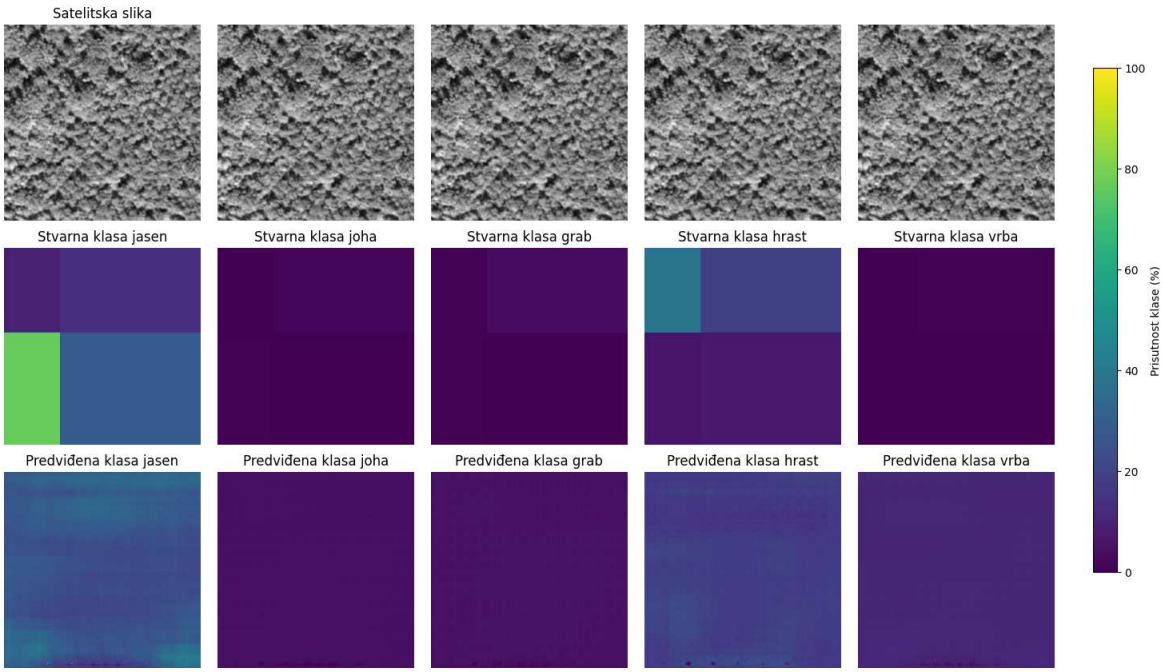
Tablica 6.5. Metrike treniranja VGG16 modela

Metrika	**Treniranje**	**Validacija**
Loss	0.0148	0.0321
R-squared	0.4979	0.3183
RMSE	0.1044	0.1598
MSE klasa *jasen*	0.0219	0.0744
MSE klasa *joha*	0.0019	0.0017
MSE klasa *grab*	0.0020	0.0042
MSE klasa *hrast*	0.0202	0.0365
MSE klasa *vrba*	0.0282	0.0439
Stopa učenja	0.0001	
Broj epoha	15	

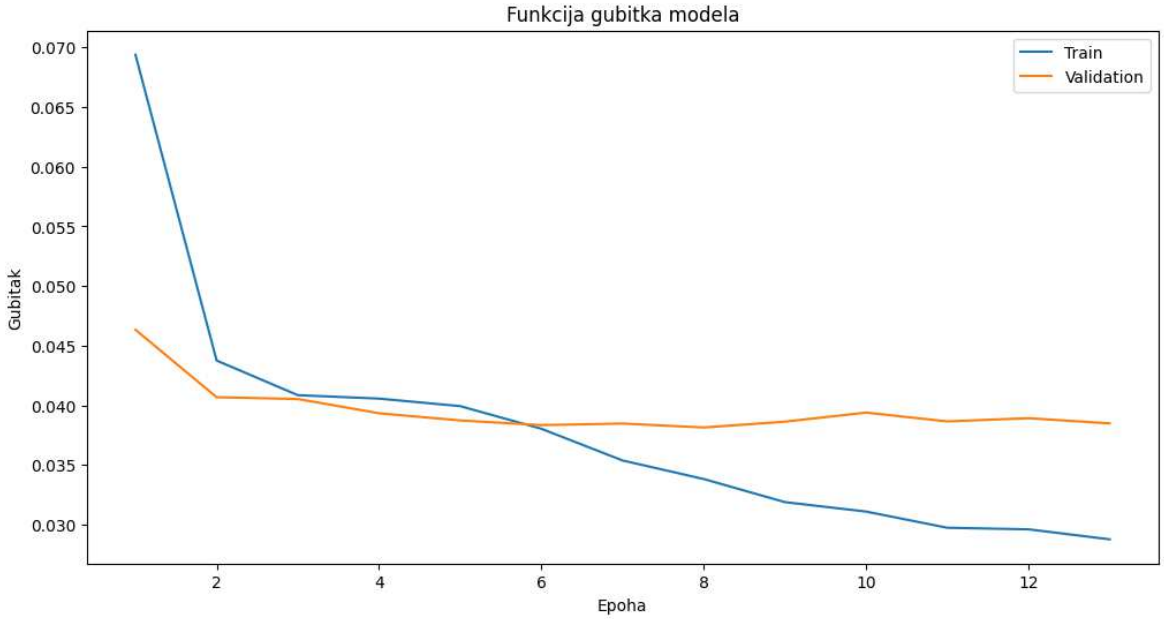
S augmentacijom podataka

Rezultati treniranja pokazuju da je model postigao nešto bolji gubitak na skupu za treniranju, sa vrijednošću od 0.0288, u usporedbi s 0.0385 na validacijskom skupu, što ukazuje na dobru prilagodbu modela podatcima. Međutim, vrijednost R-squared za treniranje je 0.3155, a za validaciju je 0.2684, što sugerira da model ne objašnjava dobro varijaciju podataka u validacijskom skupu. Razlika u RMSE-u između treniranja (0.1573) i validacije (0.1870) također ukazuje na blagi pad performansi na validacijskom skupu.

MSE po klasama ukazuje na neujednačene performanse modela među različitim klasama. Model najbolje predviđa klasu *grab*, s MSE vrijednostima od 0.0064 na treniranju i 0.0040 na validaciji. Najslabije performanse su kod klasa *vrba* i *jasen*, gdje su MSE vrijednosti više.



Slika 6.11. Rezultati predviđanja VGG16 modela



Slika 6.12. Prikaz funkcije gubitka tijekom treniranja VGG16 modela

Tablica 6.6. Metrike treniranja VGG16 modela

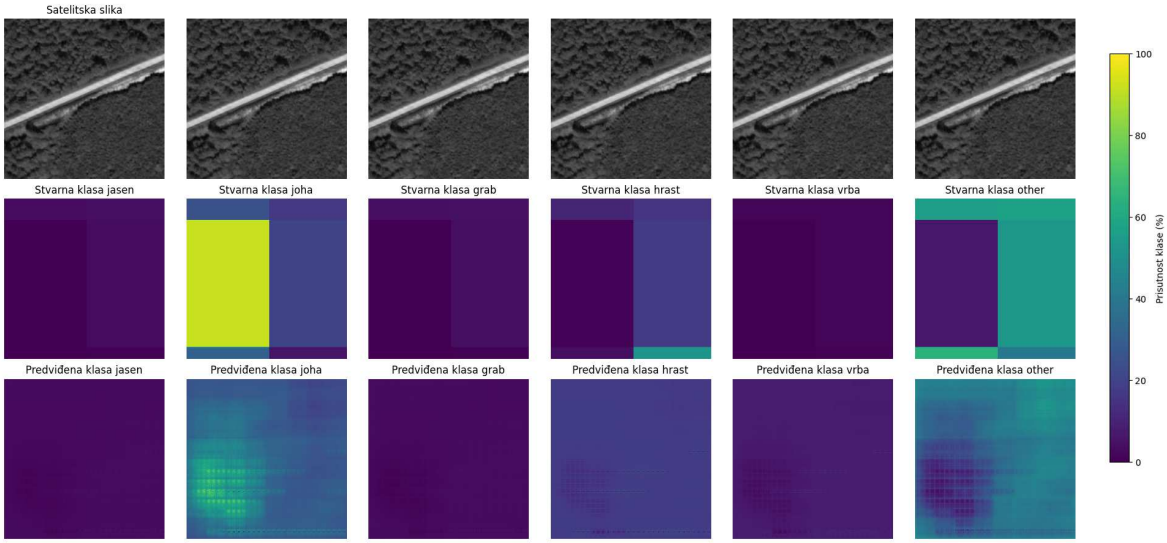
Metrika	**Treniranje**	**Validacija**
Loss	0.0288	0.0385
R-squared	0.3155	0.2684
RMSE	0.1573	0.1870
MSE klasa *jasen*	0.0607	0.0975
MSE klasa *joha*	0.0080	0.0113
MSE klasa *grab*	0.0064	0.0040
MSE klasa *hrast*	0.0291	0.0472
MSE klasa *vrba*	0.0396	0.0324
Stopa učenja	0.0001	
Broj epoha	13	

6.1.4. VGG16 model, 3 ulaza, 6 izlaza

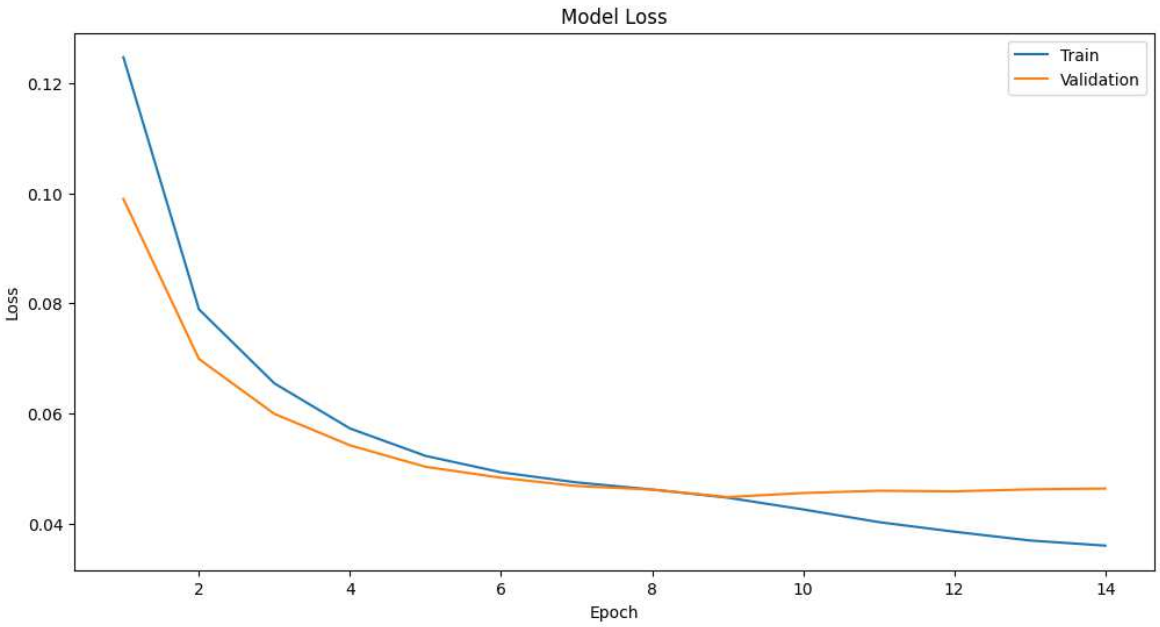
Bez augmentacije podataka

Rezultati prikazani u tablici 6.7. pokazuju da model postiže solidne rezultate na skupu za treniranje, s gubitkom od 0.0359 i R-squared vrijednošću od 0.4690, što sugerira da model uspijeva objasniti približno 47% varijance u podatcima za treniranje. RMSE iznosi 0.1814, što ukazuje na umjeren nivo pogreške. MSE po klasama pokazuje da model ima najviše problema s predviđanjem klasa *jasen* i *hrast*, dok su klase poput *joha* i *grab* bolje predviđene.

Na validacijskom skupu su performanse modela nešto lošije, što je vidljivo iz povećanja gubitka na 0.0463 i smanjenja R-squared vrijednosti na 0.3338. RMSE je također viši na validaciji (0.2058), što sugerira da model ima problema s generalizacijom na nove, neviđene podatke. Najveća pogreška u predviđanjima zabilježena je za klasu *jasen* (0.1039) i klasu *other* (0.0761),



Slika 6.13. Rezultati predviđanja VGG16 modela



Slika 6.14. Prikaz funkcije gubitka tijekom treniranja VGG16 modela

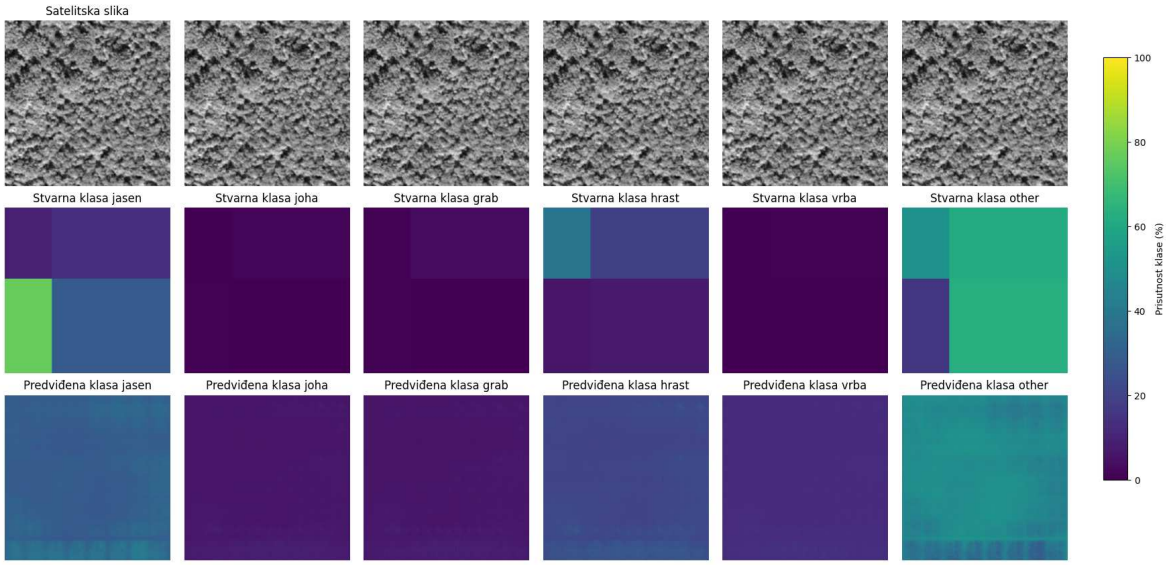
Tablica 6.7. Metrike treniranja VGG16 modela

Metrika	**Treniranje**	**Validacija**
Loss	0.0359	0.0463
R-squared	0.4690	0.3338
RMSE	0.1814	0.2058
MSE klasa *jasen*	0.0752	0.1039
MSE klasa *joha*	0.0104	0.0122
MSE klasa *grab*	0.0083	0.0055
MSE klasa *hrast*	0.0445	0.0468
MSE klasa *vrba*	0.0435	0.0335
MSE klasa *other*	0.0337	0.0761
Stopa učenja	0.00005	
Broj epoha	14	

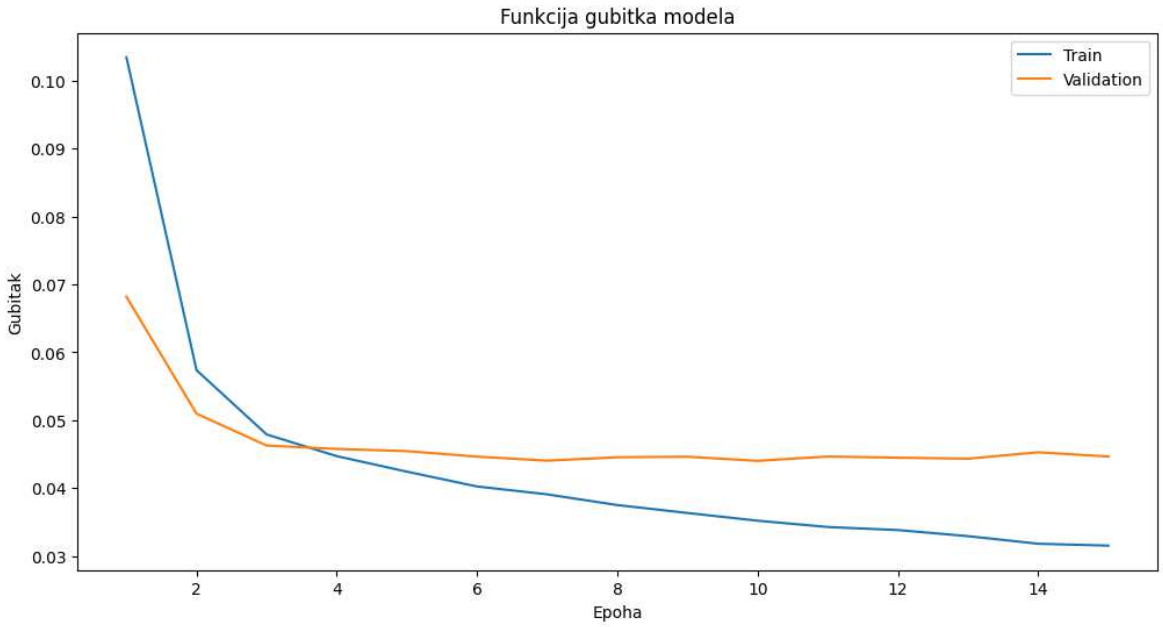
S augmentacijom podataka

Rezultati iz tablice 6.8. prikazuju solidne performanse modela VGG16 tijekom treniranja i validacije. Gubitak na skupu za treniranje iznosi 0.0315, što sugerira dobru sposobnost modela da se prilagodi treniranim podatcima. S druge strane, gubitak na validacijskom skupu je nešto veći (0.0447), što je očekivano jer se model suočava s podatcima koje ranije nije vidio. Vrijednost R-squared, koja iznosi 0.5306 za treniranje i 0.3601 za validaciju, pokazuje da model uspijeva objasniti 53.06% varijacije u treniranju i 36.01% u validaciji.

RMSE vrijednosti, koje su 0.1694 za treniranje i 0.2017 za validaciju, dodatno potvrđuju prihvatljivu točnost modela, iako postoji blagi porast pogreške na validacijskom skupu. Model najlošije predviđa klasu *other*, dok najbolje predviđa klasu *joha* .



Slika 6.15. Rezultati predviđanja VGG16 modela



Slika 6.16. Prikaz funkcije gubitka tijekom treniranja VGG16 modela

Tablica 6.8. Metrike treniranja VGG16 modela

Metrika	**Treniranje**	**Validacija**
Loss	0.0315	0.0447
R-squared	0.5306	0.3601
RMSE	0.1694	0.2017
MSE klasa *jasen*	0.0609	0.0966
MSE klasa *joha*	0.0086	0.0121
MSE klasa *grab*	0.0063	0.0050
MSE klasa *hrast*	0.0326	0.0518
MSE klasa *vrba*	0.0341	0.0341
MSE klasa *other*	0.0466	0.0684
Stopa učenja	0.00005	
Broj epoha	15	

6.2. Ograničenja i izazovi

Jedan od glavnih izazova u ovom radu bio je to što su referentne oznake morale biti znatno uvećane kako bi odgovarale satelitskim slikama u visokoj rezoluciji. Ovo je potencijalni problem jer uvećane oznake možda ne odgovaraju savršeno stvarnim satelitskim podatcima te nesavršenost u usklađivanju može rezultirati manjom preciznošću modela. Osim toga, referentne oznake su često imale male vrijednosti, a podatci su bili nebalansirani i ostalo ih je malo nakon odbacivanja neinformativnih dijelova, što dodatno otežava učenje modela i može dovesti do pristranosti u predviđanju.

Dodatan izazov predstavlja klasa *other*, koja je dodana kako bi popunila nedostajuće podatke. Ova klasa, budući da predstavlja samo ostatak ostalih referentnih oznaka, nije informativna, te iako model najbolje predviđa ovu klasu, to ne pridonosi stvarnoj korisnosti predviđanja.

Jedno od mogućih rješenja za ove probleme je generiranje dodatnih podataka kako bi se uravnotežile klase i poboljšala preciznost modela. Uz to, dodatna augmentacija podataka može pomoći modelu da generalizira bolje na neviđene podatke, smanjujući preprilagođavanje na nerazmjerno male vrijednosti.

7. Zaključak

U ovom radu istražena je primjena naprednih konvolucijskih neuronskih mreža U-Net i VGG16 za segmentaciju satelitskih slika šuma. Cilj je bio razviti modele koji mogu precizno klasificirati tipove šuma na temelju visokorezolucijskih satelitskih podataka. Postignuti rezultati pokazuju da odabrane arhitekture mogu do neke razine uspješno segmentirati različite tipove šuma, no da uspješnost ovih modela ovisi i o samoj kvaliteti podataka.

Usporedba performansi modela pokazala je prednosti korištenja različitih arhitektura, pri čemu je VGG16 pokazao veću sposobnost generalizacije neviđenih podataka od U-Net modela. Augmentirani podatci u ovom radu nisu činili veliku razliku u uspješnosti modela. Model U-Net funkcionirao je bolje s tri satelitske slike na ulazu, a VGG16 model je postizao bolje rezultate sa 6 klasa na izlazu.

Unatoč dobrim početnim rezultatima, vidim potrebu za daljnjim unapređenjima, kao što su finija kalibracija hiperparametara ili korištenje dodatnih podataka za treniranje.

Budući rad mogao bi se usmjeriti na integraciju dodatnih izvora podataka poput težina modela treniranim na satelitskim slikama, generiranjem dodatnih podataka i dodatnom augmentacijom postojećih podataka.

Zaključno, ovaj rad demonstrira potencijal dubokog učenja u području šumskog okoliša i pruža osnovu za daljnje istraživanje i primjenu konvolucijskih neuronskih mreža u segmentaciji satelitskih slika.

Literatura

[1] . H. G. LeCun Y., Bengio Y., “Deep learning”, *Nature*, str. 436-437, 2015.

[2] C. A. Goodfellow I., Bengio Y., *Deep Learning*. MIT, str. 4-6, 2016.

[3] B. C. M., *Pattern Recognition and Machine Learning*. Springer, str. 227-229, 2006.

[4] B. C. M., *Pattern Recognition and Machine Learning*. Springer, str. 329-330, 2006.

[5] M. A.-A. . L. F. M. A. Fadhel, “Review of deep learning: concepts, cnn architectures, challenges, applications, future directions”, *Journal of Big Data*, str. 15–19, 2021.

[6] C. A. Goodfellow I., Bengio Y., *Deep Learning*. MIT, str. 240-241, 2016.

[7] C. A. Goodfellow I., Bengio Y., *Deep Learning*. MIT, str. 82-83, 2016.

[8] C. A. Goodfellow I., Bengio Y., *Deep Learning*. MIT, str. 204, 2016.

[9] K. S. . A. Zisserman, “Very deep convolutional networks for large-scale image recognition”, *arXiv*, 2015.

[10] European Space Agency, https://earth.esa.int/eogateway/missions/worldview-2, [mrežno; stranica posjećena: prosinac 2023.].

Sažetak

Segmentacija satelitskih snimaka šume za klasifikaciju tipa šume

Sara Sičić

U ovom radu istražene su mogućnosti semantičke segmentacije slika šuma s ciljem klasifikacije tipova šuma na temelju označenih podataka pet različitih tipova šuma. Za problem segmentacije s ograničenom količinom podataka ispitane su dvije različite arhitekture: U-Net i VGG16, primijenjene u nekoliko različitih konfiguracija. Testirano je treniranje s augmentiranim podatcima, kao i uspješnost modela u odnosu na broj satelitskih snimaka korištenih kao ulaz u model te broj klasa koje model predviđa. Najuspješnijim se pokazao model VGG16. Daljnja poboljšanja mogla bi uključivati generiranje dodatnih podataka ili korištenje kvalitetnijeg skupa podataka.

Ključne riječi: računalni vid; segmentacija slika; satelitske slike; klasifikacija; duboko učenje; neuronske mreže; VGG16; U-Net

Abstract

Segmentation of satellite forest images for forest type classification

Sara Sičić

This paper explores the potential of semantic segmentation of forest images for the classification of forest types based on labeled data for five different forest types. For the segmentation problem with a limited amount of data, two different architectures were examined: U-Net and VGG16, applied in several different configurations. The training with augmented data was tested, along with the model's performance in relation to the number of satellite images used as input to the model and the number of classes the model predicts. The VGG16 model proved to be the most successful. Further improvements could include generating additional data or using a higher quality dataset.

Keywords: computer vision; image segmentation; satellite pictures; classification; deep learning; neural networks; VGG16; U-Net