

# Detekcija solarnih panela korištenjem satelitskih snimki i dubokog učenja

Antonio Bilić

Prirodoslovno-matematički fakultet

Sveučilište u Splitu

abilic@pmfst.hr

**Sažetak**—U ovom radu istražuje se primjena dubokog učenja za automatsku detekciju solarnih panela na satelitskim snimkama visoke rezolucije. Porast korištenja solarne energije kao ključnog obnovljivog izvora zahtijeva precizne i skalabilne metode za identifikaciju i praćenje solarnih instalacija. Predložena metoda temelji se na računalnom vidu i konvolucijskim neuronskim mrežama (CNN) te omoguće pouzdano prepoznavanje solarnih panela iz zraka. Model je treniran na označenim skupovima podataka koji obuhvaćaju različite geografske regije, uvjeti osvjetljenja, vremenske prilike i varijacije u kutovima snimanja. Za postizanje bolje generalizacije i preciznosti, korištene su tehničke povećanja podataka (data augmentation), normalizacije i prilagodbe modela (fine-tuning) specifičnim regijama.

Cilj rada je pokazati kako se duboke neuronske mreže mogu učinkovito koristiti za detekciju solarnih instalacija, čime se doprinosi boljem razumijevanju njihovog prostornog rasporeda, podršci u planiranju energetske infrastrukture, tržišnoj analizi te donošenju strateških odluka. Eksperimentalni rezultati ukazuju na visoku ukupnu točnost (preko 90%) te dobre metrike poput F1 mjere, koeficijenta podudarnosti (IoU) i Kappa koeficijenta, čak i u izazovnim scenarijima s djelomično zaklonjenim ili degradiranim snimkama. U radu se raspravlja i o izazovima vezanim uz neuravnoteženost klasa, prikupljanje reprezentativnih podataka te skalabilnost sustava za veća zemljopisna područja.

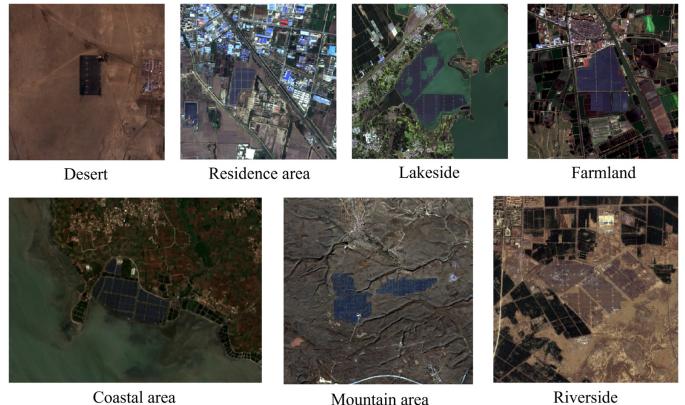
Predloženi pristup pokazuje veliki potencijal za širu primjenu u automatskom mapiranju obnovljivih izvora energije i integraciju u GIS sustave i nacionalne prostorne baze podataka, čime se omogućuje donošenje informiranih i održivih odluka temeljenih na podacima.

**Index Terms**—duboko učenje, satelitske snimke, solarni paneli, računalni vid, detekcija objekata

## I. UVOD

Globalno ubrzano usvajanje obnovljivih izvora energije ključno je u borbi protiv klimatskih promjena i smanjenju ovisnosti o fosilnim gorivima. Među različitim oblicima obnovljive energije, solarna fotonaponska (PV) energija doživjava najbrži rast zbog pada cijena tehnologija, poboljšanja učinkovitosti i političke podrške [1]. Međunarodna agencija za energiju (IEA) procjenjuje da će globalni kapacitet solarne energije rasti po stopi od 12% godišnje, dosežući 2764 TWh do 2030. godine [2].

S povećanjem broja solarnih instalacija, javlja se potreba za učinkovitim metodama praćenja njihove geografske distribucije. Tradicionalni pristup ručnog prikupljanja podataka često je nepotpun, dugotrajan i radno intenzivan [3]. Mnoge zemlje nemaju centralizirane baze prostornih podataka o solarnim elektranama, a postojeće baze poput Wiki-Solar često nisu potpune [1].



Slika 1: Primjeri solarnih instalacija u različitim okruženjima: (a) pustinja, (b) stambeno područje, (c) pored jezera, (d) poljoprivredno zemljишte, (e) obalno područje, (f) planinska regija, (g) riječno područje. Slika sintetizirana prema podacima iz [2].

Automatsko otkrivanje i mapiranje solarnih panela korištenjem satelitskih snimki omogućuje:

- Energetskim plannerima optimizaciju distribucijske mreže
- Investitorima identificiranje potencijalnih lokacija za nove solarne elektrane
- Regulatorima praćenje usklađenosti s energetskom politikom
- Istraživačima procjenu utjecaja solarnih instalacija na okoliš

### A. Klasifikacija solarnih PV sustava

Solarne PV instalacije mogu se klasificirati u dvije glavne kategorije [3]:

- **Komunalne solarne elektrane** (engl. utility-scale) - velike zemaljske PV sustave na otvorenim područjima kojima upravljaju proizvođači energije za opskrbu mreže. Imaju kapacitet od najmanje 1 MW i pokrivaju 2-4 hektara zemljišta.
- **Distribuirani PV sustavi** - manja postrojenja instalirana blizu krajnjeg korisnika za lokalnu potrošnju. Najčešće se postavljaju na krovove ili kao manji zemaljski sustavi povezani s lokalnom mrežom.

## B. Izazovi u detekciji solarnih panela

Automatska detekcija solarnih panela na satelitskim snimkama suočava se s nekoliko izazova:

- **Varijabilnost izgleda:** Solarni paneli mogu znatno varirati po obliku, veličini, orijentaciji i razmještaju.
- **Sličnost s drugim objektima:** U urbanim područjima, solarni paneli mogu imati spektralne karakteristike slične krovovima, parkiralištima ili staklenicima.
- **Uvjeti osvjetljenja i vremenske prilike:** Različito osvjetljenje, kut sunca, oblaci i magle mogu utjecati na izgled solarnih panela.
- **Prostorna rezolucija:** Ograničena rezolucija satelitskih snimki otežava detekciju manjih solarnih instalacija.
- **Neuravnoteženost klase:** Solarni paneli obično zauzimaju mali dio satelitske snimke, što stvara problem neuravnoteženosti za algoritme strojnog učenja.

## II. PREGLED PODRUČJA

### A. Tehnike detekcije solarnih panela

#### 1) Metode temeljene na pikselima

Metode bazirane na pikselima klasificiraju svaki piksel satelitske snimke pojedinačno, na temelju njegovih spektralnih karakteristika [1]:

- **Random Forest (RF):** RF klasifikatori učinkoviti su u detekciji solarnih panela zbog robusnosti prema šumu i sposobnosti rada s velikim brojem varijabli. Plakman i sur. [1] koristili su RF za detekciju solarnih parkova s Sentinel-2 snimki, postigavši visoku preciznost (preko 99%).
- **Support Vector Machine (SVM):** Malof i sur. [4] prvi su predložili korištenje SVM pristupa za lokalizaciju solarnih panela sa satelitskih i zračnih snimki.
- **Naive Bayes:** Dalagan i Principe [3] usporedili su SVM, RF i Naive Bayes klasifikatore, pri čemu je SVM pokazao najbolje rezultate za klasifikaciju temeljenu na pikselima.

#### 2) Objektno orijentirane metode

Objektno-orientirane metode prvo segmentiraju sliku u homogene objekte, a zatim klasificiraju te objekte [1], [3]. Prednosti:

- Smanjenje efekta "soli i papra" koji se često pojavljuje u klasifikaciji temeljenoj na pikselima
- Analiza spektralnih svojstava, oblika, tekture i kontekstualnih informacija
- Bolje očuvanje detalja i kontura solarnih panela

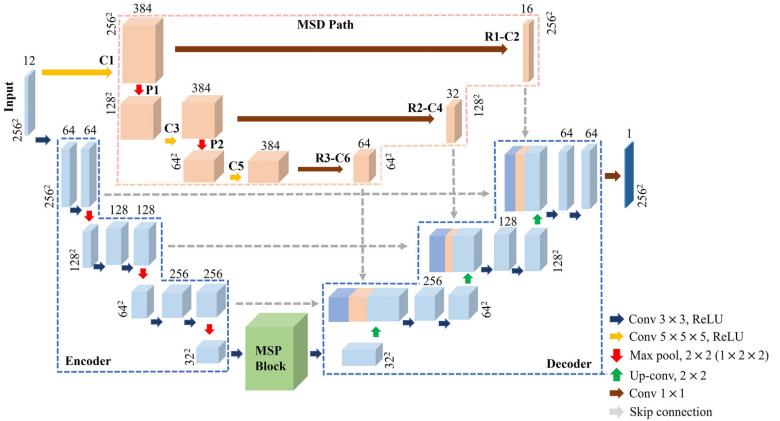
#### 3) Duboko učenje

S napretkom dubokog učenja, konvolucijske neuronske mreže (CNN) pokazale su se izuzetno učinkovitima u detekciji solarnih panela:

- **FCN (Fully Convolutional Networks):** Hou i sur. [5] predložili su SolarNet, okvir dubokog učenja koji kombinira FCN i modul pozornosti za lociranje i procjenu površine solarnih elektrana.

Tablica I: Usporedba točnosti različitih metoda detekcije solarnih panela

Metoda	Ukupna točnost	F1 mjera	IoU	Kappa koef.
RF (POBC) [3]	0.987	0.88	0.90	0.934
E-UNET [2]	0.989	0.869	0.80	0.934
U-Net+ [2]	0.987	0.855	-	0.919
SVM (PBC) [3]	0.986	0.84	0.88	0.910
SegNet [2]	0.973	0.803	-	0.845
RF (pikselski) [1]	0.929	0.85	-	-



Slika 2: Arhitektura E-UNET za detekciju solarnih panela [2]. Arhitektura se sastoji od: modula za kodiranje-dekodiranje značajki, MSP bloka (Multi-Scale Pooling) i MSD (Multi-Spectral 3D) konvolucijskog puta.

- **U-Net:** Dui i sur. [2] predložili su E-UNET, poboljšanu verziju U-Net arhitekture koja koristi multispektralni 3D konvolucijski put (MSD) i blok višeskalnog udruživanja (MSP) za efikasno izvlačenje spektralnih i prostornih značajki.
- **Kombinirani pristupi:** Dalagan i Principe [3] predložili su pristup koji kombinira klasifikaciju temeljenu na pikselima (PBC) i objektima (OBC) za detekciju solarnih panela.

### B. Korištenje multispektralnih podataka

Većina ranih studija o detekciji solarnih panela koristila je samo RGB satelitske snimke. Međutim, novija istraživanja pokazuju da korištenje multispektralnih podataka može značajno poboljšati preciznost detekcije, posebno kada su solarni paneli i pozadina vizualno slični [2].

Sentinel-2 satelitski podaci, s 13 spektralnih kanala prostornih rezolucija od 10, 20 i 60 metara, pokazali su se posebno korisnima. Kruitwagen i sur. [6] koristili su multispektralne podatke iz Sentinel-2 i SPOT-6/7 za globalni pregled komunalnih solarnih instalacija.

Dui i sur. [2] istražili su vrijednost multispektralnih podataka kombinirajući različite kanale Sentinel-2 s indeksima putem NDVI (Normalized Difference Vegetation Index), NDWI (Normalized Difference Water Index) i PVSI (Photovoltaic Spectral Index) za poboljšanje detekcije.

### III. METODOLOGIJA

U ovom poglavlju opisuje se metodologija korištena za detekciju solarnih fotonaponskih (PV) panela iz satelitskih snimki koristeći tehnike dubokog učenja. Metodologija obuhvaća pripremu podataka, pristup segmentaciji, dizajn arhitekture mreže i procedure treniranja. Predloženi UNet++ model unapređuje mogućnosti tradicionalnih mreža za semantičku segmentaciju kroz naprednu encoder-decoder arhitekturu s nested i dense skip vezama.

#### A. Obrada ulaznih i izlaznih podataka

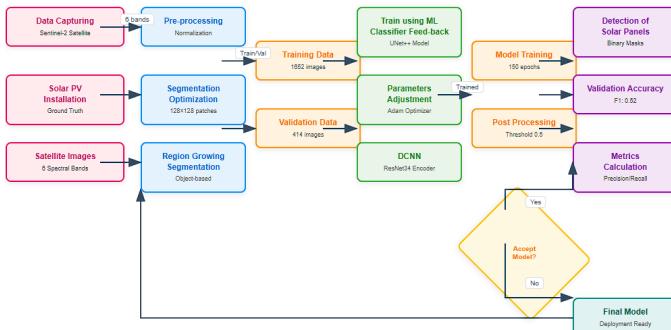
##### 1) Karakteristike ulaznih podataka

Ulagani podaci sastoje se od multispektralnih satelitskih snimki prikupljenih Sentinel-2 satelitima. Za našu metodologiju koristimo specifične spektralne kanale koji su pokazali visoku diskriminativnu sposobnost za detekciju solarnih panela:

- B2 (Blue)**: 490 nm, prostorna rezolucija 10m
- B3 (Green)**: 560 nm, prostorna rezolucija 10m
- B4 (Red)**: 665 nm, prostorna rezolucija 10m
- B8 (NIR)**: 842 nm, prostorna rezolucija 10m
- B11 (SWIR1)**: 1610 nm, prostorna rezolucija 20m
- B12 (SWIR2)**: 2190 nm, prostorna rezolucija 20m

Skup podataka obuhvaća 2066 slike koje sadrže PV postrojenja u različitim okruženjima, omogućujući robusno učenje modela kroz različite geografske i okolišne kontekste.

Solar Panel Detection Data Pipeline



Slika 3: Tok podataka za detekciju solarnih panela, prikazujući progresiju od multispektralnih Sentinel-2 ulaznih podataka kroz faze predobrade do UNet++ modela i konačno do binarne segmentacijske maske.

##### 2) Priprema podataka i augmentacija

Za pripremu podataka za treniranje modela implementirali smo sljedeće korake:

- Podjela skupa podataka**: Ukupni skup od 2066 slika podijeljen je na:

- Trening skup: 1652 slike
- Validacijski skup: 414 slike

- Spektralni kanali**: Korišteno je 6 odabranih spektralnih kanala [B2, B3, B4, B8, B11, B12] koji su se pokazali najinformativnijima za detekciju solarnih panela
- Prostorna rezolucija**: Sve slike imaju dimenzije  $128 \times 128$  piksela
- Augmentacija podataka**: Za poboljšanje robusnosti modela korištene su standardne tehnike augmentacije podataka
- Normalizacija**: Svi spektralni kanali normalizirani su za poboljšanje stabilnosti treniranja i konvergencije

##### 3) Reprezentacija izlaza

Izlaz modela je binarna segmentacijska maska istih prostornih dimenzija kao ulazni patch ( $128 \times 128$  piksela). Svakom pikselu u masci dodijeljena je vrijednost 1 ako predstavlja solarni panel ili 0 u suprotnom. Ove binarne maske služe kao ground truth tijekom treniranja i koriste se za evaluaciju tijekom testiranja.

Tablica II: Specifikacije ulaznih i izlaznih podataka za segmentaciju

Parametar	Ulazni podaci	Izlazni podaci
Dimenzije	$128 \times 128$ piksela	$128 \times 128$ piksela
Broj kanala	6 spektralnih kanala	1 kanal (binarna maska)
Prostorna rezolucija	10–20 m (ovisno o kanalu)	Naslijedena od ulaza
Tip podataka	Float32 (normalizirano)	Binarni (0 ili 1)
Značenje vrijednosti	Reflektancija u više spektralnih kanala	Prisutnost (1) ili odsutnost (0) PV panela
Izvor podataka	Sentinel-2 satelitske snimke	Model segmentacije
Predobrada	Selekcija kanala, normalizacija	Postprocesiranje s pragom (0.5)

#### B. Modalitet segmentacije

Zadatak detekcije solarnih panela formuliran je kao problem semantičke segmentacije, gdje se svaki piksel satelitske snimke klasificira kao pripadajući PV panelu (klasa 1) ili ne (klasa 0). Za razliku od detekcije objekata, koja identificira granične okvire oko objekata, semantička segmentacija pruža klasifikaciju na razini piksela koja čuva precizni oblik i opseg PV instalacija.

##### 1) Binarna semantička segmentacija

U našem pristupu koristimo binaru semantičku segmentaciju, gdje je cilj klasificirati svaki piksel u jednu od dvije klase:

- Klasa 0**: Pozadina (pikseli koji nisu dio PV panela)
- Klasa 1**: Solarni PV paneli

Ovaj binarni pristup pojednostavljuje zadatok dok istovremeno održava visoku preciznost u ocrtavanju granica PV instalacija.

##### 2) Izazovi u segmentaciji PV panela

Segmentacija PV panela iz satelitskih snimki predstavlja nekoliko jedinstvenih izazova:

- Varijabilnost izgleda**: PV paneli značajno variraju u veličini, obliku, orientaciji i uzorcima rasporeda

- **Spektralna sličnost:** U nekim okruženjima, PV paneli mogu imati spektralne karakteristike slične drugim objektima
- **Varijacije u osvjetljenju:** Različiti uvjeti osvjetljenja utječu na izgled PV panela
- **Detekcija malih objekata:** Rezidencijalne PV instalacije mogu zauzimati samo nekoliko piksela na slici
- **Neuravnoteženost klase:** PV paneli tipično zauzimaju mali postotak piksela na slici

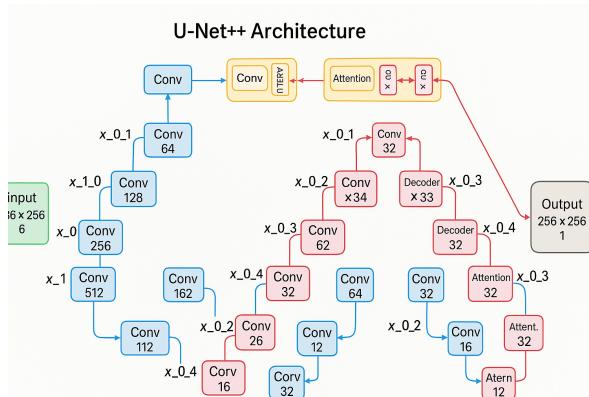
### C. UNet++ arhitektura

UNet++ arhitektura predstavlja značajno unapređenje klasične U-Net arhitekture kroz uvođenje nested i dense skip poveznica koje omogućavaju bolju propagaciju značajki između encoder i decoder dijelova mreže. Ova arhitektura je posebno prikladna za zadatke semantičke segmentacije gdje je potrebna precizna lokalizacija objekata.

#### 1) Ključne karakteristike UNet++ arhitekture

UNet++ se sastoji od sljedećih ključnih komponenti:

- **Redesigned skip pathways:** Umjesto direktnih skip poveznica, UNet++ koristi dense konvolucijske blokove koji postupno spajaju semantičke razine
- **Nested arhitektura:** Encoder i decoder su povezani kroz seriju nested U-Net struktura različitih dubina
- **Deep supervision:** Mogućnost generiranja izlaza na različitim dubinama mreže za poboljšanu optimizaciju
- **Fleksibilnost modela:** Mogućnost rada s modelima različite složenosti ovisno o dostupnim računalnim resursima



Slika 4: Arhitektura UNet++ modela s prikazom nested skip poveznica i dense konvolucijskih putanja

#### 2) Implementacija modela

Model je implementiran koristeći HuggingFace biblioteku s sljedećim specifikacijama:

- **Encoder:** ResNet34 kao backbone mreža
- **Encoder weights:** ImageNet pretrained weights
- **Ulagni kanali:** 6 (odabrani Sentinel-2 kanali)
- **Klase:** 1 (binarna segmentacija)
- **Aktivacija:** None (sigmoid se primjenjuje u loss funkciji)

## CREATE MODEL

```
[ ] model = smp.UnetPlusPlus(
    encoder_name="resnet34",
    encoder_weights="imagenet",
    in_channels=6,
    classes=1,
    activation=None
)
```

Slika 5: Kod za kreiranje UNet++ modela s ResNet34 encoderom

#### 3) Encoder-Decoder struktura

UNet++ koristi asimetričnu encoder-decoder strukturu gdje:  
**Encoder (ResNet34 backbone):**

- Postupno smanjuje prostornu rezoluciju uz povećanje dubine značajki
- Koristi residual blokove za bolje učenje dubokih reprezentacija
- Pretreniran na ImageNet skupu podataka za bolje inicijalne težine

#### Decoder s nested skip vezama:

- Postupno rekonstruira prostornu rezoluciju
- Dense skip veze omogućavaju bolju integraciju multi-scale značajki
- Nested putanje omogućavaju postupno spajanje semantičkih informacija

#### D. Funkcija gubitka i optimizacija

Za treniranje modela korištena je kombinacija funkcija gubitka prilagođena za binarnu segmentaciju:

$$\mathcal{L}_{total} = \mathcal{L}_{BCE} + \alpha \cdot \mathcal{L}_{Dice} \quad (1)$$

gdje je:

- $\mathcal{L}_{BCE}$  - binarna unakrsna entropija
- $\mathcal{L}_{Dice}$  - Dice loss za bolje rukovanje neuravnoteženošću klase
- $\alpha$  - težinski faktor za balansiranje gubitaka

#### 1) Parametri treniranja

Model je treniran s sljedećim hiperparametrima:

- **Optimizator:** Adam
- **Početna stopa učenja:**  $1 \times 10^{-4}$
- **Scheduler:** ReduceLROnPlateau
- **Batch size:** Prilagođen dostupnoj GPU memoriji
- **Epohe:** 150 s early stopping

## IV. REZULTATI

U ovom poglavlju prikazani su rezultati treniranja modela temeljenog na U-Net arhitekturi, razvijenog za segmentaciju solarnih panela na Sentinel-2 satelitskim snimkama niske prostorne rezolucije (22–26 piksela). Model koristi šest odabranih spektralnih kanala te je evaluiran kvantitativno pomoću standardnih metrika i kvalitativno kroz vizualne usporedbe predikcija.

### A. Kvantitativna evaluacija

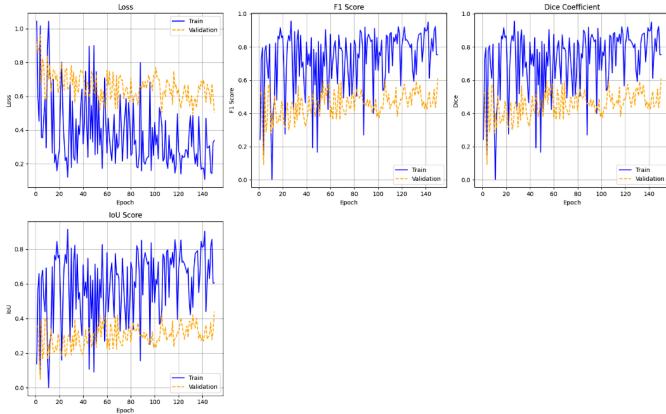
Model je treniran kroz 150 epoha korištenjem Adam optimizatora s početnom stopom učenja  $1 \cdot 10^{-4}$ , uz automatsko smanjenje stope učenja u slučaju stagnacije validacijskog gubitka. Evaluacija je provedena na posebno izdvojenim skupovima za treniranje i validaciju.

U Tablici III prikazane su vrijednosti glavnih metrika. Na trening skupu postignut je F1 score od 0,79, dok je na validacijskom (evaluacijskom) skupu dosegnut rezultat od 0,52, što ukazuje na ograničenu generalizaciju modela. Vrijednosti Dice koeficijenta i IoU također potvrđuju ovu razliku između performansi na trening i evaluacijskom skupu.

Tablica III: Usporedba metrika na trening i evaluacijskom skupu

Metrika	Trening skup	Evaluacijski skup
F1 score	0,7907	0,5207
Dice koeficijent	0,7907	0,5207
IoU score	0,6538	0,4372
Preciznost	0,7666	0,5696
Odziv (Recall)	0,8164	0,5415
Točnost (Accuracy)	0,9341	0,9581

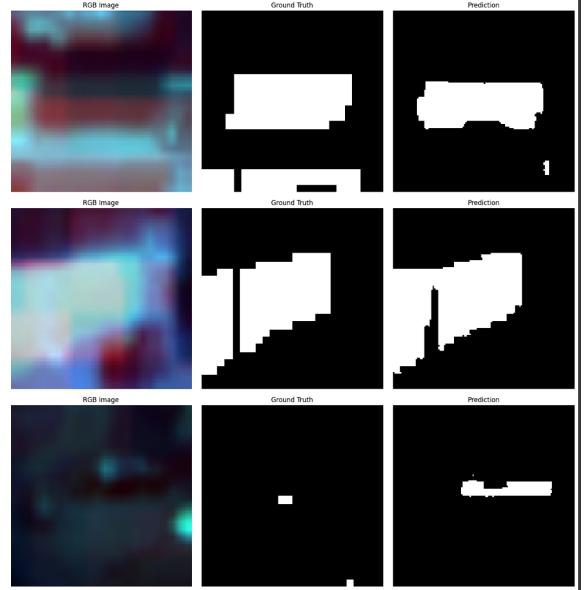
Na slici 6 prikazani su grafovi svih metrika kroz epohe. Jasno je vidljivo da model vrlo brzo uči već u prvima epohama, nakon čega slijedi stabilizacija performansi. Razlika između trening i evaluacijskih rezultata sugerira da je model sklon overfittingu, što je čest problem kod segmentacijskih zadataka s ograničenim podatcima.



Slika 6: Grafovi gubitka i metrika (F1, Dice, IoU, preciznost i odziv) tijekom treniranja U-Net modela

### B. Kvalitativna evaluacija

Pored numeričkih metrika, kvaliteta modela procijenjena je i vizualno. Na slici 7 prikazana je usporedba između ulazne slike, stvarne maske i predikcije modela.



Slika 7: Usporedba ulazne slike, stvarne maske i predikcije U-Net modela

Model pokazuje sposobnost segmentacije područja solarnih panela, iako s varijabilnim uspjehom. Predikcija uglavnom prati oblik i položaj panela, ali s određenim odstupanjima duž rubova i povremenim lažno pozitivnim detekcijama. Niska rezolucija satelitskih snimaka predstavlja značajan izazov za preciznu segmentaciju.

### C. Usporedba s rezultatima natjecanja

Za dodatnu perspektivu, uspoređeni su postignuti rezultati s onima iz stvarnog natjecanja. Na slici 8 prikazan je izvadak iz javne ljestvice rezultata (Leaderboard), gdje su najbolji timovi postigli F1 score do 0,66.

Usporedbom s vodećim rezultatima, model treniran u ovom radu s rezultatom od 0,52 na evaluacijskom skupu ostvaruje umjeren uspjeh. Iako rezultat nije na razini najboljih timova, s obzirom na ograničene resurse i osnovnu implementaciju U-Net arhitekture, ovaj rezultat predstavlja solidnu osnovu za daljnji razvoj.

### D. Zaključak

Dobiveni rezultati potvrđuju da je U-Net arhitektura prikladan izbor za zadatak segmentacije solarnih panela iz Sentinel-2 satelitskih podataka, ali također ukazuju na potrebu za dodatnim optimizacijama. Model ostvaruje dobre performanse na trening skupu, ali pokazuje ograničenu generalizaciju na neviđenim podatcima. Razlika između trening i evaluacijskih rezultata sugerira overfitting, što je čest problem kod kompleksnih segmentacijskih zadataka s ograničenim podatcima.

#	Team name/User name	Best Score	Entries	Last Submit	Solution
1	motokimura	0.6632834	17	1 year ago	
2	KouFu	0.6579446	38	1 year ago	
3	bysk2	0.6507268	47	1 year ago	
4	kisuke	0.6477581	79	1 year ago	
5	kumalab	0.6463335	40	1 year ago	
6	SalutekKsw	0.638887	65	1 year ago	
7	elsh	0.6387431	62	1 year ago	
8	Michal	0.6377874	38	1 year ago	
9	Anil45	0.625784	45	1 year ago	
10	bitGuber	0.6170287	92	1 year ago	

Slika 8: Dio ljestvice rezultata (Leaderboard) natjecanja

U budućem radu, dodatne nadogradnje modela, kao što su napredne tehnike augmentacije podataka, regularizacijske tehnike, balansiranje klase ili zamjena gubitne funkcije (npr. Focal loss), moglo bi značajno poboljšati konačne rezultate i generalizaciju modela. .

#### LITERATURA

- [1] V. Plakman, J. Rosier, and J. van Vliet, "Solar park detection from publicly available satellite imagery," *GIScience Remote Sens.*, vol. 59, no. 1, pp. 462–481, 2022, doi: 10.1080/15481603.2022.2036056.
- [2] Z. Dui, Y. Huang, J. Jin, and Q. Gu, "Automatic detection of photovoltaic facilities from Sentinel-2 observations by the enhanced U-Net method," *J. Appl. Remote Sens.*, vol. 17, no. 1, pp. 1-14, 2023, doi: 10.1117/1.JRS.17.014516.
- [3] A. G. Dalagan and J. A. Principe, "Spatial inventory of solar photovoltaic (PV) installations using remote sensing and machine learning: Case of Pampanga Province, Philippines," *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.*, vol. XLVIII-4/W9-2024, pp. 115–123, 2024, doi: 10.5194/isprs-archives-XLVIII-4-W9-2024-115-2024.
- [4] J. M. Malof, R. Hou, L. M. Collins, K. Bradbury, and R. Newell, "Automatic solar photovoltaic panel detection in satellite imagery," in International Conference on Renewable Energy Research and Applications (ICRERA), 2015, pp. 1428–1431.
- [5] X. Hou, B. Wang, W. Hu, L. Yin, and H. Wu, "SolarNet: A deep learning framework to map solar power plants in China from satellite imagery," *arXiv Prepr. arXiv1912.03685*, 2019.
- [6] L. Kruitwagen, K. Story, J. Friedrich, L. Byers, S. Skillman, and C. Hepburn, "A global inventory of photovoltaic solar energy generating units," *Nature*, vol. 598, no. 7882, pp. 604–610, 2021, doi: 10.1038/s41586-021-03957-7.