

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

PROJEKT

**Utjecaj sintetičkih podataka u segmentaciji
površinskih objekata korištenjem WASR-T mreže**

Maja Magdalenić

Voditelj: *Nikola Mišković*

Zagreb, Veljača 2023

Uvod

Ovaj rad izrađen u sklopu predmeta Projekt prolazi kroz proces treniranja modela za segmentaciju površinskih objekata koristeći WASR-T mrežu. Nakon što je ukratko opisana korištena mreža i setovi podataka za treniranje, prikazani su i komentirani dobiveni rezultati za sve modele.

1. WASR

WASR mreža koristi se za segmentaciju pomorskih objekata. Predtrenirana mreža sliku odvaja u 3 klase (water, sky i obstacle). Mreža je specifična po tome što je u potpunosti prilagođena pomorskom okruženju, a prilagođena je na način da je uveden novi gubitak razdvajanja vode od prepreka koji se izračunava u zadnjem stadiju koda i pomaže učenju značajki različitih pojava vode kako bi te pojave odvojio od obilježja koja odgovaraju preprekama. WASR -T ili WASR-temporal mreža koja je korištena u ovom projektu je proširenje WASR mreže vremenskim kontekstom. Mreža je treniran na setu podataka koji sadrži kontekstualne frameove. Efektivno to izgleda ovako : uzima se primjerice 5 frameova, prvo 1 „pravi“ frame i onda još 4 framea snimljena u kratkom vremenu (između kojih je mala razlika) koji dolaze nakon tog prvog framea - time dobivamo taj vremenski kontekst. Poanta je u tome da u tih 4 kontekstualna framea očekujemo različit odsjaj (jer voda nikad nije skroz mirna) dok objekt ostaje relativno isti, i ovo se koristi za poboljšanje performansi mreže na odsjajima u vodi. WASR-T je trenutno jedna od segmentacijskih mreža koja daje najbolje rezultate u ovakvim uvjetima.

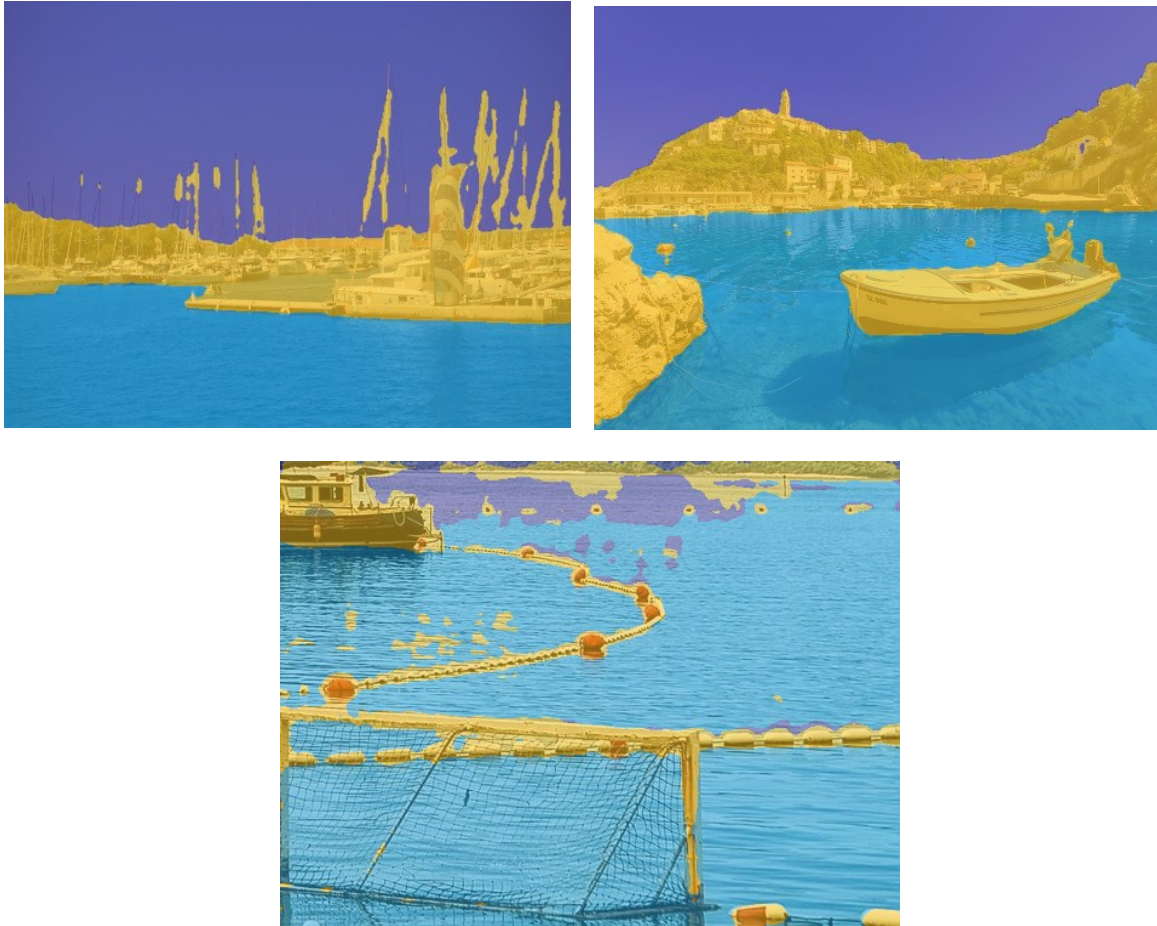
2. Treniranje na realnim podacima

Set podataka korišten za treniranje mreže na realnim podacima jest MaSTR1478 (Maritime Semantic Segmentation Training Dataset). Dataset sadrži 1325 slika snimljenih USV-om tijekom razdoblja od 2 godine, i za svaku sliku 5 kontekstualnih frameova. Uključuje niz različitih uvjeta koje nalazimo u vodenom okruženju. Sve slike su označene za semantičku segmentaciju po svakom pikselu. Rezultati treniranog modela prikazani su u tablici ispod.

Tablica 1 Rezultati modela treniranog na realnim podacima

| Dataset | IoU obstacle | IoU sky | IoU water |
|-----------|--------------|---------|-----------|
| MaSTR1478 | 0,9961 | 0,9994 | 0,9997 |

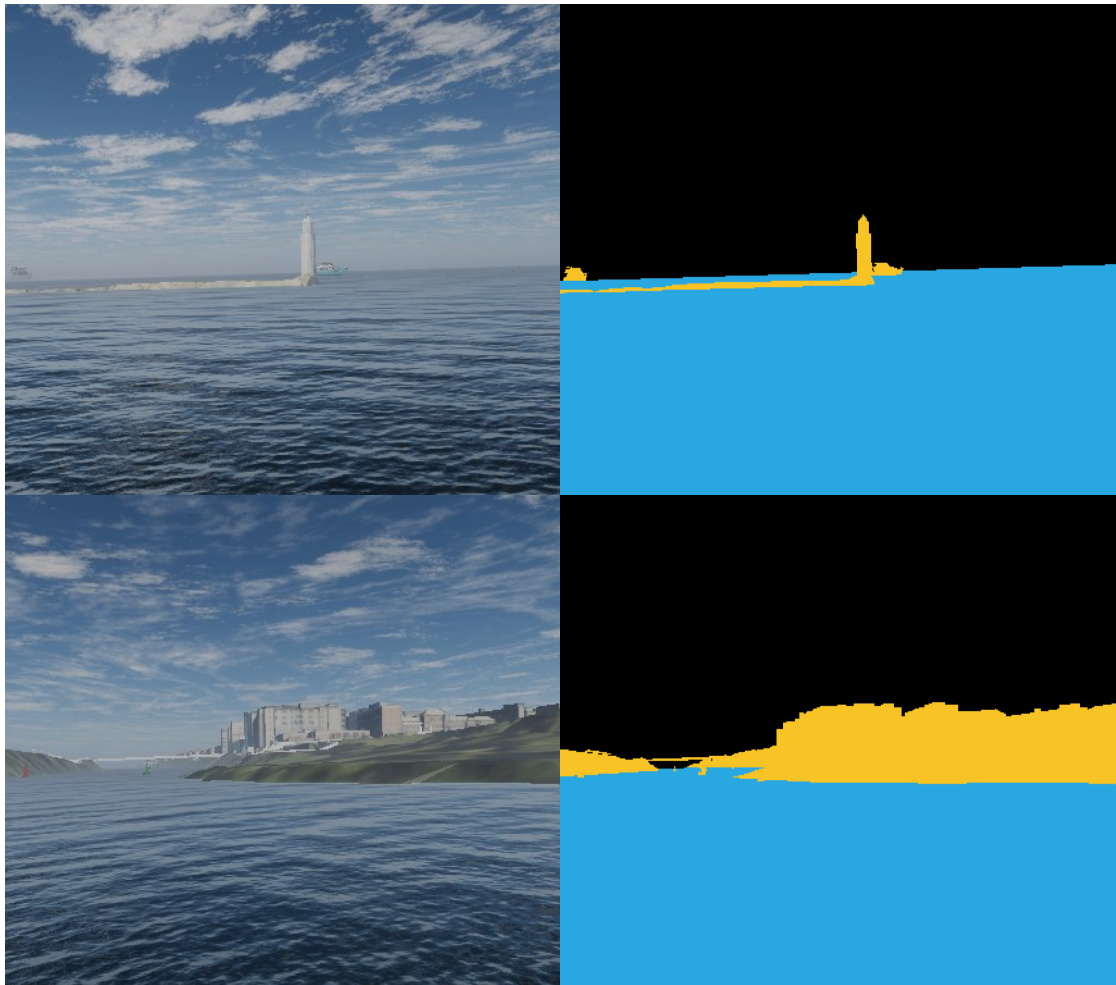
Dobivamo vrlo visoku točnost u segmentaciji, no bitno je spomenuti da su ovo rezultati testiranja na slikama vrlo sličnima onima na kojima je model treniran i validiran. Na slikama ispod prikazani su rezultati segmentacije na slikama koje nisu iz MaSTR1478 dataseta.



3. Treniranje na sintetičkim podacima

Nakon treniranja na realnim podacima potrebni su nam sintetički podaci. Sintetički podaci skupljeni su u Unity programskom okruženju. Najprije je napravljena scena koja sadrži pomorsko okruženje te su dodani pomorski markeri i brodovi kako bi imali dodatne objekte za segmentaciju. Prikupljanje podataka napravljeno je uz pomoć Unity Perception paketa. Ovaj paket sadrži mnogo alata koji služe upravo za generiranje velikih skupova sintetičkih podataka za korištenje u strojnom i dubokom učenju. Neki od njih su Instance segmentation, Detection, Semantic segmentation itd. Za svaki objekt u sceni potrebno je označiti kojoj klasi pripada, u suprotnome će biti klasificiran kao „*undefined*“. Na kameru u sceni dodana je Semantic Segmentation skripta koja je spremala parove slika rgb – label određenom frekvencijom. Frekvenciju postavljamo sami. Primjeri slika dobiveni u

simulatoru prikazani su ispod. Iz ovih labela potrebno je samo napraviti maske za što je napisana kratka skripta u pythonu i mogu se koristiti za treniranje.



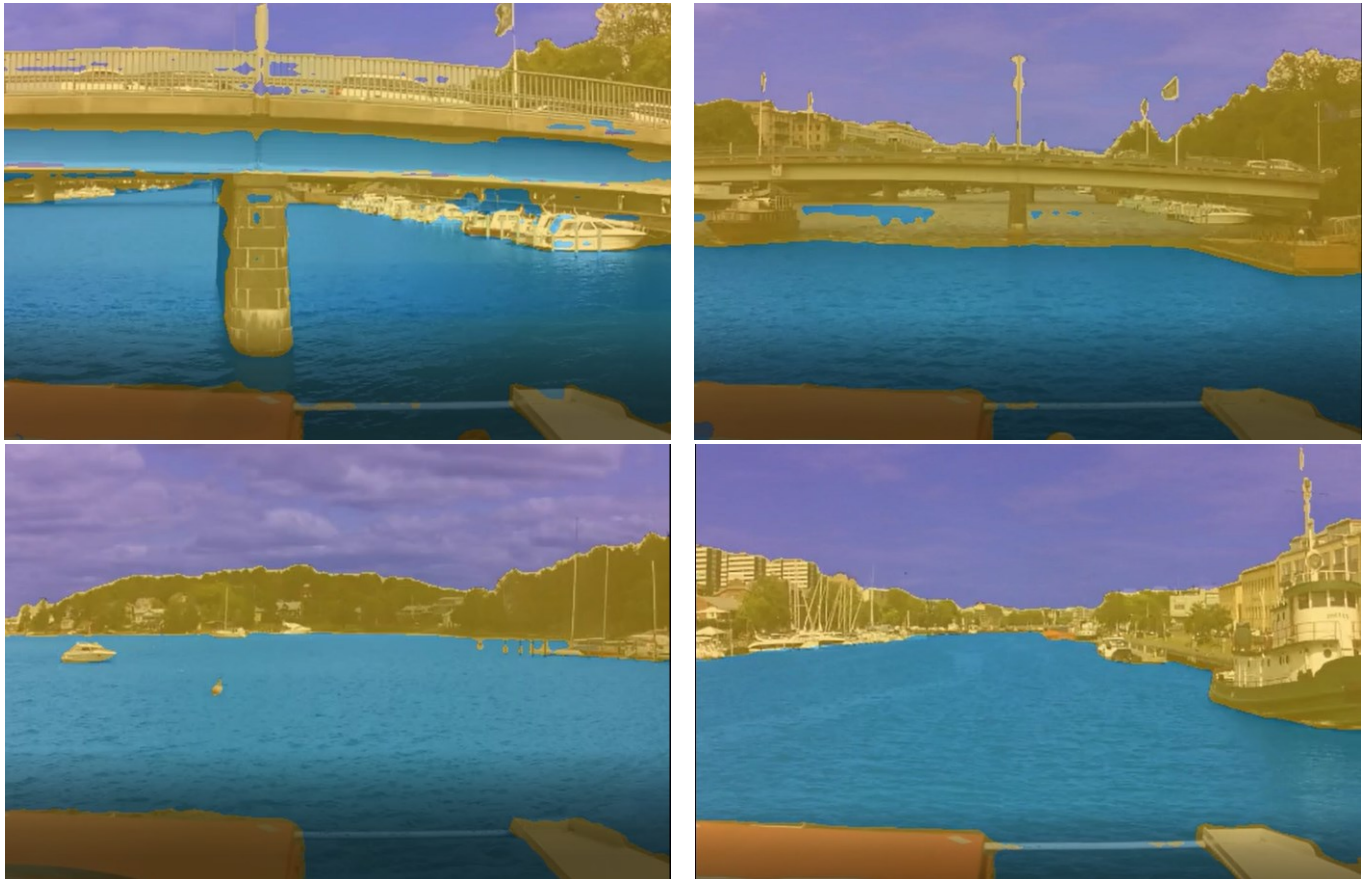
Na ovaj je način prikupljeno 8000 slika (sa kontekstualnim framovima), otprilike toliko ih imamo u realnom MaSTr datasetu. WASR-T mreža najprije je trenirana na setu podataka koji sadrži $\frac{1}{4}$ slika iz realnog dataseta i $\frac{3}{4}$ iz simulacijskog dataseta, a nakon toga na setu podataka koji sadrži samo $\frac{1}{4}$ realnog dataseta (2000 slika), kako bi mogli vidjeti utjecaj simulacijskih podataka. Rezultati treniranja za sva 3 modela prikazani su u Tablica 2.

Tablica 2 Rezultati za sva 3 trenirana modela

| Dataset | IoU obstacle | IoU sky | IoU water |
|---|--------------|---------|-----------|
| MaSTr1478 | 0,9961 | 0,9994 | 0,9997 |
| $\frac{1}{4}$ MaSTr1478 + $\frac{3}{4}$ simulation | 0,9909 | 0,9987 | 0,9991 |
| $\frac{1}{4}$ MaSTr1478 | 0,9937 | 0,9992 | 0,9993 |

Kao što se moglo i očekivati, najbolje rezultate dobivamo kod treniranja samo na realnom skupu podataka. Kod preostala 2 skupa podataka na kojima je trenirano: mješoviti skup od 8000 slika, radi lošije nego da smo imali samo 2000 realnih slika. Dakle, sintetički podaci su pogoršali performanse modela i prema tome bolje bi nam bilo imati jako malo podataka, ali da su realni. No moramo imati na umu da ovo nije prava slika stanja jer su sva 3 modela validirana na slikama iz realnog skupa podataka - što je logično zato što želimo da model što bolje radi u stvarnom svijetu, ali je problem što su te slike za validaciju onda jako slične realnim slikama na kojima je trenirano i time bi i mali udio sintetičkih podataka samo pokvario model. Bilo bi bolje da smo validirali model na 3. nezavisnom skupu realnih podataka. No u tom slučaju, trebali bi imati nezavisni treći anotirani set podataka.

Na kraju je najbolji model testiran na slikama iz dataseta brodova, budući da ove slike imaju vrlo sličnu okolinu. Na slikama ispod prikazani su rezultati.



Model ne radi toliko savršeno kao na slikama prije, što je bilo za očekivati budući da su one slike bile vrlo slične onima koje je vidio. Vizualno možemo analizirati ove slike i vidjeti da model posebno ima poteškoće pri prepoznavanju mostova, ali dataset opet nije anotiran pa u brojkama ne možemo predložiti točnost.

4. Poteškoće u radu – kriva anotacija

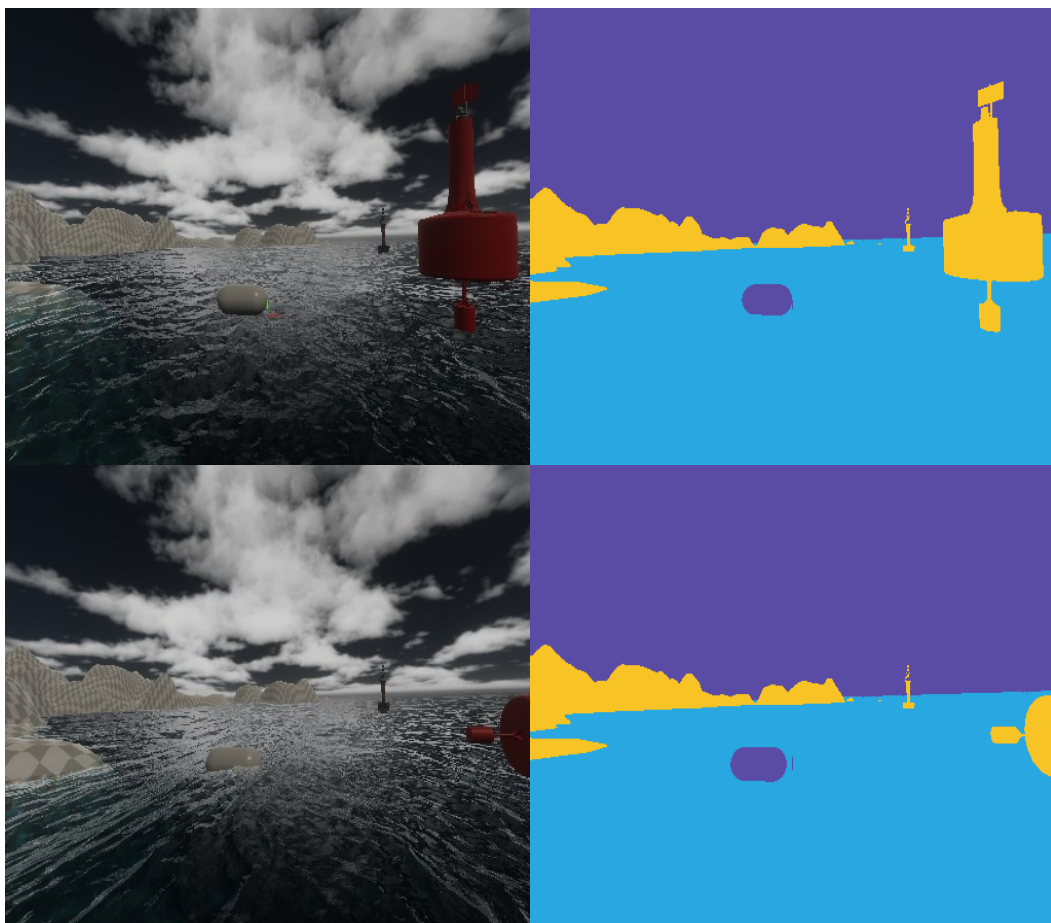
Jedna od poteškoća s kojima smo se susreli u radu je kriva anotacija oceana.

U Unity programskom okruženju ocean je podijeljen na dijelove koji se spajaju. Na spojevima tih dijelova ne postoji ništa i ovo je vidljivo u obliku isprekidanih crtica koje su anotirane kao objekt.



Problem je riješen dodavanjem prozirne plohe na ocean koja je onda označena kao more.

Drugi problem s anotacijom je bio taj što Semantic Segmentation skripta ne vidi valove, već vidi samo površinu oceana na nultoj razini. Ako na donjim slikama pogledamo donji lijevi dio kopna, na realnoj slici on je više ili manje prekriven, dok je na segmentiranoj označen isto.



Ovo bi bila značajna pogreška kada bismo dodali neke velike valove ili ekstremne vremenske uvjete. Kako bi izbjegli krivu anotaciju, u našem slučaju zadržali smo se na mirnijem moru.