

**SVEUČILIŠTE U SPLITU
FAKULTET ELEKTROTEHNIKE, STROJARSTVA I
BRODOGRADNJE**

**POSLIJEDIPLOMSKI DOKTORSKI STUDIJ
ELEKTROTEHNIKE I INFORMACIJSKE TEHNOLOGIJE**

KVALIFIKACIJSKI ISPIT

**DETEKCIJA OBJEKATA NA I POD MOREM
ANALIZOM SLIKA BESPILOTNE LETJELICE**

Ivan Biliškov

Split, rujan 2022.

SADRŽAJ:

1. UVOD	1
2. MORSKI OTPAD	2
3. NEURONSKE MREŽE	4
3.1. Duboke neuronske mreže	4
3.2. Konvolucijske neuronske mreže	5
3.2.1. Konvolucijski slojevi	6
3.2.2. Zašto koristiti CNN za prepoznavanje slika	9
3.2.3. Klasifikatori	10
3.2.4. Detektori čiji se modeli sastoje od dvije faze	14
3.2.5. Detektori čiji se modeli sastoje od jedne faze	17
4. DALJINSKO ISTRAŽIVANJE	21
5. PREGLED PODRUČJA	23
5.1. Radovi u kojima se objekt promatranja nalazi na površini mora	23
5.2. Radovi u kojima se objekt promatranja nalazi pod morem	29
5.3. Ostali relevantni radovi	36
5.4. Baze podataka	40
6. ZAKLJUČAK	44
7. LITERATURA	45

1. UVOD

Jedan od globalnih problema današnjice je morski otpad koji utječe na sve oceane svijeta. Loša praksa gospodarenja čvrstim otpadom, nedostatak infrastrukture, neadekvatno ponašanje i aktivnosti ljudi su doveli do ekoloških, ekonomskih, zdravstvenih i estetskih problema. Spora stopa razgradnje je ono što karakterizira većinu morskog otpada i dovodi do visoke i postupne akumulacije kako u morskom tako i u obalnom okolišu [1]. Nadalje dio morskog otpada se raspada na manje fragmente i nastaju mikročestice. Jedan od najopasnijih predstavnika mikročestica današnjice je mikroplastika koja se nalazi u svim ekosustavima [2]. Ne postoji način na koji bi se mikročestice uklonile iz okoliša, a i da postoji, količina novog smeća koji dospije u okoliš je prevelika i nastaje silovitom brzinom. Diljem svijeta su donesene mjere i zakoni sa svrhom smanjivanja onečišćenja. Neke od mjera su: zabrana proizvodnje i korištenja određenih proizvoda, naknada za pronađeni izgubljeni ribolovni alat ili za boce, maksimalna dnevna granica emisije kemikalija u slivove itd. [3]. Redukcija količine otpada će smanjiti pritisak na morski okoliš, međutim, moramo misliti i na otpad koji se trenutno nalazi u njemu te na koji način ga ukloniti kako ne bi došlo do akumulacije i fragmentacije istog.

Detekcija morskog smeća je prvi korak u otklanjanju otpada iz morskog okoliša. Najnovija otkrića u području računalnog vida (*eng. Computer Vision*), pogotovo napredak konvolucijskih neuronskih mreža (*eng. Convolutional Neural Network; CNN*) unaprijedila su područje mobilne robotike i razvoj aplikacija za autonomnu vožnju, nadzor dronom, strojeva za žetvu, strojeva za muženje... [4].

Cilj ovog rada je istražiti algoritme za detekciju objekata u moru i na obali koristeći tehnologije računalnog vida i konvolucijskih neuronskih mreža. Da bi se izgradio takav algoritam potrebno je osigurati veliki skup podataka na takav način da se što manje narušava morski ekosustav. Najneinvazivniji pristup pokazao se proces prikupljanja podataka iz daljine te je namjera koristiti bepilotnu letjelicu - dron. Ovakav alat bi služio za: detekciju morskog otpada, detekciju morske flore i faune, praćenja stanja zagađenja ili promjene okoliša u obalnim mikrosistemima te istraživanje ekološki osjetljivih područja nedostupnih ronionicima. Dugoročni cilj je učiniti more čistim te smanjiti pritisak na morski okoliš i ojačati podvodni život. Ono što predstavlja izazov je činjenica da podvodna domena nije toliko istraživana koliko terestrička te su metode vođene računalnim vidom ograničene i nailaze na poteškoće.

2. MORSKI OTPAD

Morski otpad se definira kao svaki postojani, obrađeni ili proizvedeni čvrsti materijal koji je odbačen, napušten ili odložen u morski i obalni okoliš. Postupci ljudi, namjerni ili slučajni, izvor su onečišćenja. Otpad pronađen u oceanima potječe s trgovačkih, putničkih, ribarskih, vojnih i istraživačkih brodova, postrojenja za akvakulturu, naftnih / plinskih platformi i bušotina. Morske struje i mijene, vjetar te topografija područja i morskog dna utječu na taloženje i fragmentaciju otpada. Izvori zagađenja i onečišćenja s kopna su: pristaništa, luke, marine, obalni turizam, oborinske vode, komunalne otpadne vode, komunalna odlagališta, riječni prijevoz, industrijski objekti i medicinski otpad. Uragani, tornada, tsunamiji i poplave nanose veliku količinu materijala s kopna dok jaki olujni vjetrovi i valovi vrše degradaciju obale. Velika količina takvog materijala završi u morskome okolišu. Morski otpad nalazimo ne samo u područjima guste naseljenosti već i na područjima koji su daleko od očitih izvora onečišćenja i ljudskog kontakta [1].

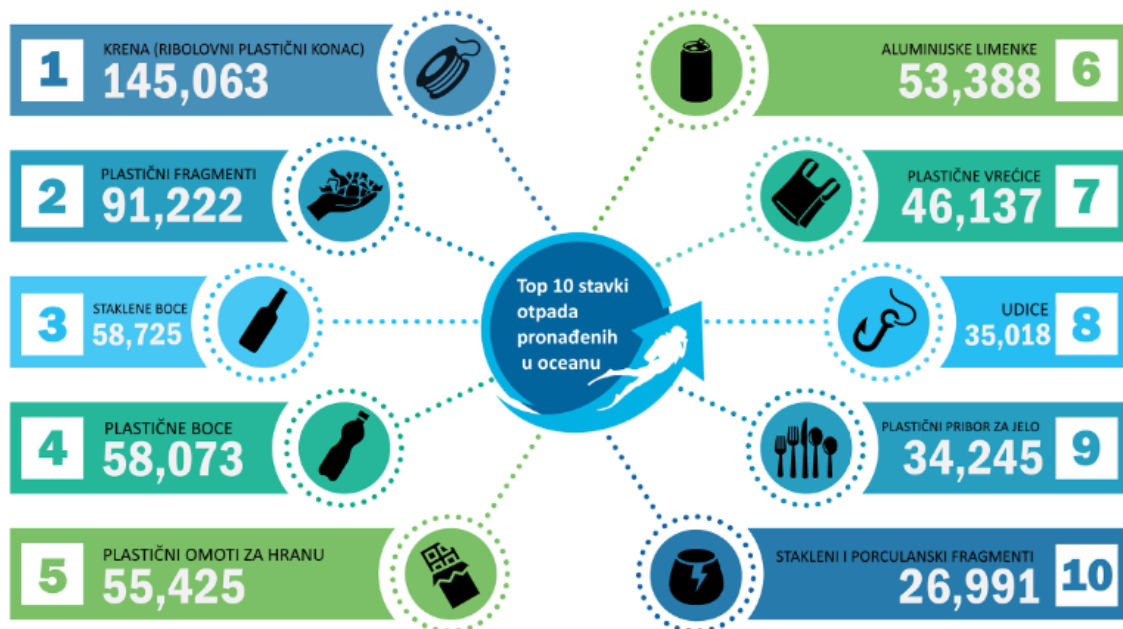
Velika i direktna opasnost po život divljih životinja je gutanje otpada ili zaplitanje u isti. Potonuli morski otpad guši morsko dno i remeti bentičke vrste dok njegovi fragmenti služe kao prijevoz za invazivne vrste. Opasnosti koje prijete ljudima mogu se razvrstati u više kategorija: one koje prijete izravno zdravlju (tjelesne ozljede i bolesti), imovini (uništavanje brodova i opreme), poslovanju (ribarstvo) i sigurnosni rizik na moru (zaplitanje propelera, prepreke u obliku velikih krhotina) [1].

Morski otpad možemo podijeliti u više kategorija:

- Prema mjestu pronalaska: plaža, površina mora, morsko dno, sediment, u/na organizmima (gutanje ili zaplitanje) [5]
- Prema veličini čestica: mega (> 1 m), makro (1 m - 2.5 cm), mezo (2.5 cm - 5 mm), mikro (5 mm - 0.33 mm), nano (< 0.33 mm) [6]
- Prema materijalu: plastika, guma, metal, staklo/keramika, tekstil/prirodna vlakna, drvo, papir/karton, ostalo (određeno/neodređeno) [5]

Top 10 stavki otpada pronađenih u oceanu

Više od milijun stavki uklonjeno i prijavljeno kroz program "Dive against debris"



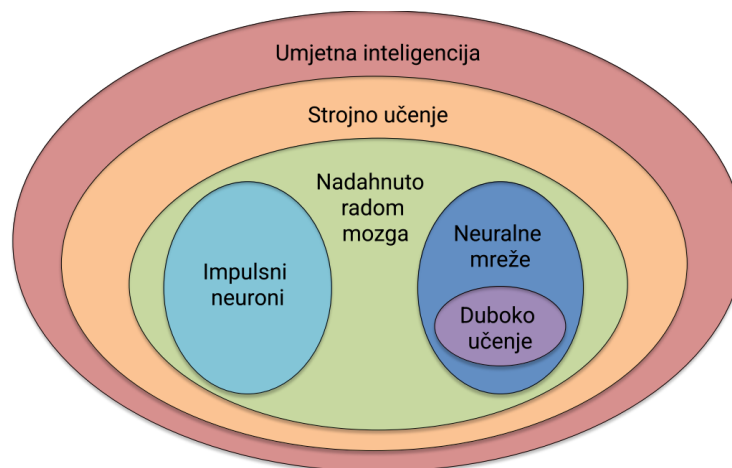
PROJECTAWARE.ORG
Where Conservation Meets Adventure™

* Podaci iz ožujka 2018.

Slika 2.1. Prikaz deset najčešćih stavki pronađenih u oceanu [7]

3. NEURONSKE MREŽE

Kreiranje posebnog, prilagođenog programa koji će rješavati problem po problem je mukotrpan i dug proces. Poddomena umjetne inteligencije (*eng. Artificial Intelligence; AI*) zvana strojno učenje (*eng. Machine Learning; ML*) nudi jednostavnije rješenje. Potrebno je naučiti jedan algoritam koji će rješavati svaki novi problem putem procesa koji se zove treniranje ili obuka. Takvi algoritmi oponašaju radnju koje bi moglo izvršiti ljudsko biće, međutim, nedostatak matematičkih alata za modeliranje na primjenjiv način je nešto što ljudima nedostaje za izvršavanje te radnje (Slika 3.1. “Domena umjetne inteligencije”). Jedna od takvih poddomena strojnog učenja koja je nadahnuta radom ljudskog mozga zove se neuronska mreža (*eng. Neural Network; NN*) [8].

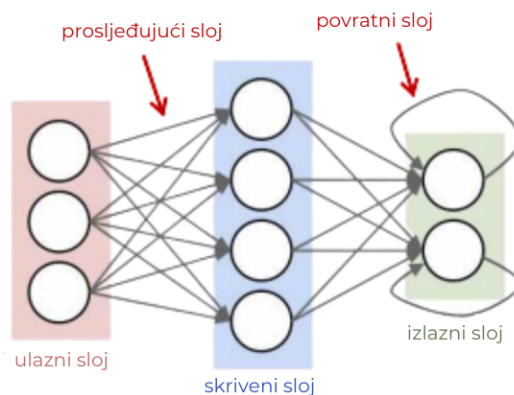


Slika 3.1. Domena umjetne inteligencije

3.1. Duboke neuronske mreže

Duboke neuronske mreže (*eng. Deep Neural Network; DNN*) su podmreže neuronskih mreža koje imaju više od tri sloja, odnosno više od jednog skrivenog sloja. Moderne duboke mreže imaju od pet do više od tisuću slojeva. Duboke mreže za razliku od plićih mreža su sposobne naučiti značajke visoke razine koje su apstraktnije i složenije. Ovisno o primjeni, DNN-ovi dolaze u raznim oblicima i veličinama te se brzo razvijaju kako bi poboljšali učinkovitost i točnost. Pikseli slike, numerički prikaz stanja nekog sustava i amplitude valne duljine zvučnog

zapisa su najčešći ulaz u DNN mreže. Svaki ulaz je skup vrijednosti koji predstavlja informacije koje će mreža analizirati [8]. Neuron u gornjem sloju predstavlja koncept koji se treba protumačiti i on je apstraktan tj. ne možemo ga vidjeti. Suprotno tome vidljiva, ulazna domena (slika, tekst...) je onaj dio koji se može tumačiti [9]. Na slici 3.2. “Prosljeđujući i povratni sloj” možemo vidjeti dva glavna sloja, ulazni - prosljeđujući i izlazni - povratni [8].

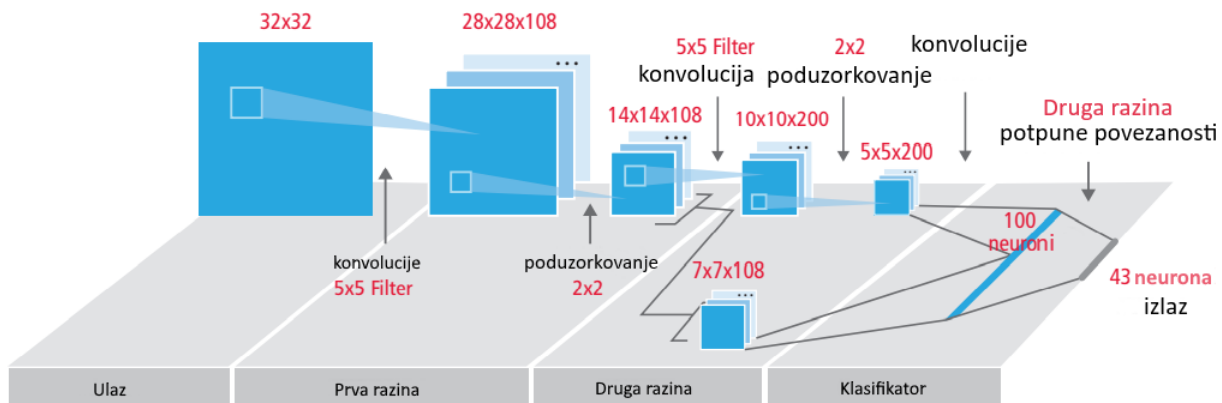


Slika 3.2. Prosljeđujući i povratni sloj

3.2. Konvolucijske neuronske mreže

Konvolucijska neuronska mreža (*eng. ConvNet / CNN*) je poseban slučaj gore opisane neuronske mreže. CNN je vrsta dubokog učenja koji može obraditi ulaznu sliku, dodijeliti karakteristike različitim aspektima / objektima na slici i biti u mogućnosti razlikovati jedan od drugog. Prethodna obrada koja je potrebna u CNN-u mnogo je jednostavnija u usporedbi s ostalim algoritmima za klasifikaciju. Dok se u primitivnim metodama filteri izrađuju ručno, uz dovoljno obuke, CNN ima mogućnost učenja filtera i karakteristika [9]. CNN se sastoji od jednog ili više konvolucijskih slojeva, često s podslojevima, koji slijede jedan ili više potpuno povezanih slojeva baš kao i u standardnoj neuronskoj mreži. Tipični CNN za klasifikaciju predmeta prikazan je na slici 3.3. “Tipični dijagram CNN bloka”. Svaka značajka sloja prima ulaze iz skupa značajki smještenih u malom susjedstvu u prethodnom sloju koje se naziva lokalno prijemno polje. Pomoću lokalnih receptivnih polja značajke mogu izdvojiti elementarne vizualne značajke, poput orijentiranih rubova, krajnjih točaka, kutova itd., koje potom kombiniraju viši slojevi. U tradicionalnom modelu prepoznavanja uzoraka ili slike, ručno dizajnirani ekstraktor značajki prikuplja relevantne podatke iz ulaza i uklanja nebitne varijable. Nakon ekstraktora

slijedi podesiv klasifikator, standardna neuronska mreža koja klasificira značajke vektora u klase. U CNN-u, konvolucijski slojevi imaju ulogu ekstraktora značajki, ali nisu ručno dizajnirane (kao do sad gdje su se za specifične primjene koristile specifične jezgre (*eng. kernel*), algoritmi detektora kutova i slično). Vrijednosti *kernel* filtera se odrede tokom procesa treniranja [9].



Slika 3.3. Tipični dijagram CNN bloka

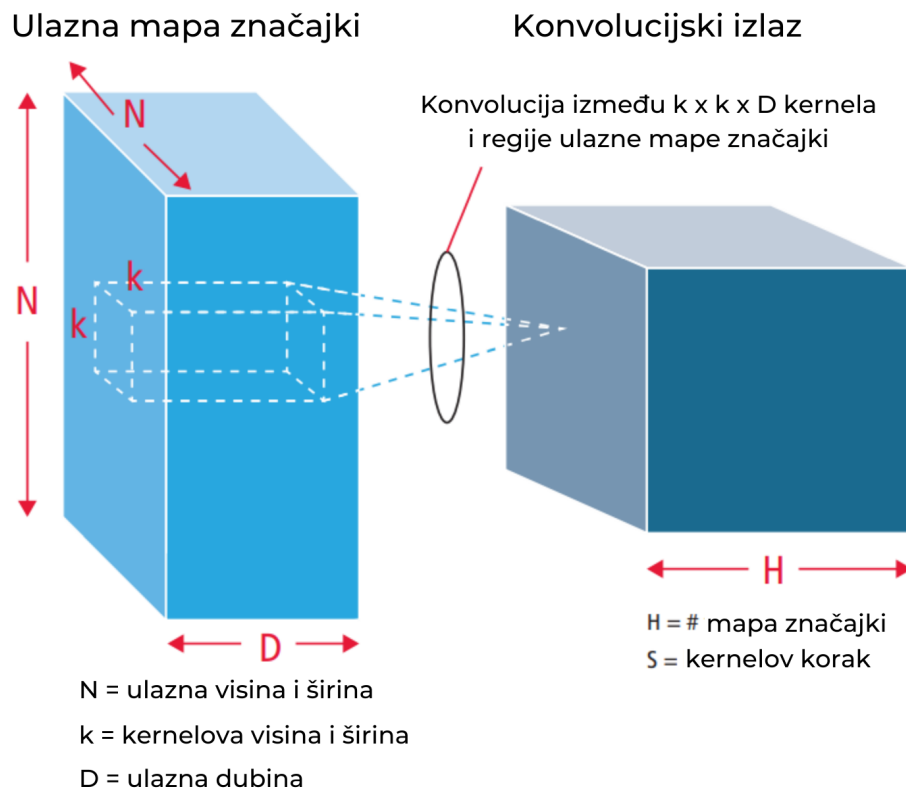
3.2.1. Konvolucijski slojevi

Složeniju arhitekturu CNN mreže dobivamo slaganjem velikog broja različitih slojeva i kao takva najbolje rješava probleme klasifikacije. Četiri najčešća sloja su: konvolucijski sloj, sloj poduzorkovanja/udruživanja, potpuno povezani sloj i nelinearni sloj [9].

- Konvolucijski sloj

Operacija konvolucije izdvaja različite značajke ulaza. Prvi konvolucijski sloj izdvaja niske razine značajke kao što su rubovi, linije i kutovi. Slojevi više razine izdvajaju značajke više razine. Slika 3.4. ilustrira proces 3D konvolucije koji se koristi u CNN-ima. Ulaz je veličine $N \times N \times D$ i savijen je s H *kernelima*, svaka od njih veličine $k \times k \times D$ zasebno. Konvolucija ulaza s jednim *kernelom* proizvodi jednu izlaznu značajku, a s H *kernelima* neovisno proizvodi H značajke. Počevši od gornjeg lijevog kuta unosa, svaki *kernel* se pomiče s lijeva na desno, jedan po jedan element. Kada se dosegne gornji desni kut, *kernel* se pomiče za jedan element prema dolje i opet se *kernel* pomiče s lijeva na desno, jedan po jedan element. Ovaj postupak se ponavlja sve dok *kernel* ne dosegne donji desni kut. Za slučaj kada je $N = 32$ i $k = 5$, postoji 28 jedinstvenih pozicija s lijeva na desno i 28 jedinstvenih pozicija od vrha do dna koje *kernel* može

zauzeti. Svaka značajka u izlazu će sadržavati 28×28 (tj. $(N-k+1) \times (N-k+1)$) elemenata. Za svaku poziciju *kernela* u procesu kliznog prozora, $k \times k \times D$ elementi ulaza i $k \times k \times D$ elementi *kernela* se množe i akumuliraju, element po element [9]. *Kernel* je moguće pomicati i za više piksela (preskakanje) (*eng. stride*).

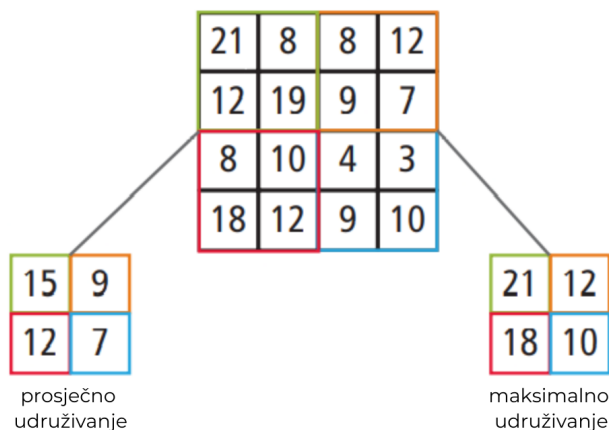


Slika 3.4. Prikaz procesa konvolucije

- Sloj poduzorkovanja/udruživanja

Sloj za poduzorkovanje/udruživanje smanjuje razlučivost značajki. To čini značajke robusnim protiv šuma i izobličenja. Postoje dva načina za udruživanje: maksimalno udruživanje i prosječno udruživanje. U oba slučaja, ulaz je podijeljen na dvodimenzionalne prostore koji se ne preklapaju. Na slici 3.5. "Prikaz maksimalnog i prosječnog udruživanja", sloj 2 je sloj za udruživanje. Svaki ulazna značajka je 28×28 i podijeljena je na 14×14 regija veličine 2×2 . Za prosječno udruživanje, prosjek od četiri izračunavaju se vrijednosti u regiji. Za maksimalno udruživanje odabire se maksimalna vrijednost od četiri vrijednosti. Proces udruživanja se mastavlja tako da je ulaz veličine 4×4 . Za poduzorkovanje 2×2 , slika je 4×4 podijeljen u četiri

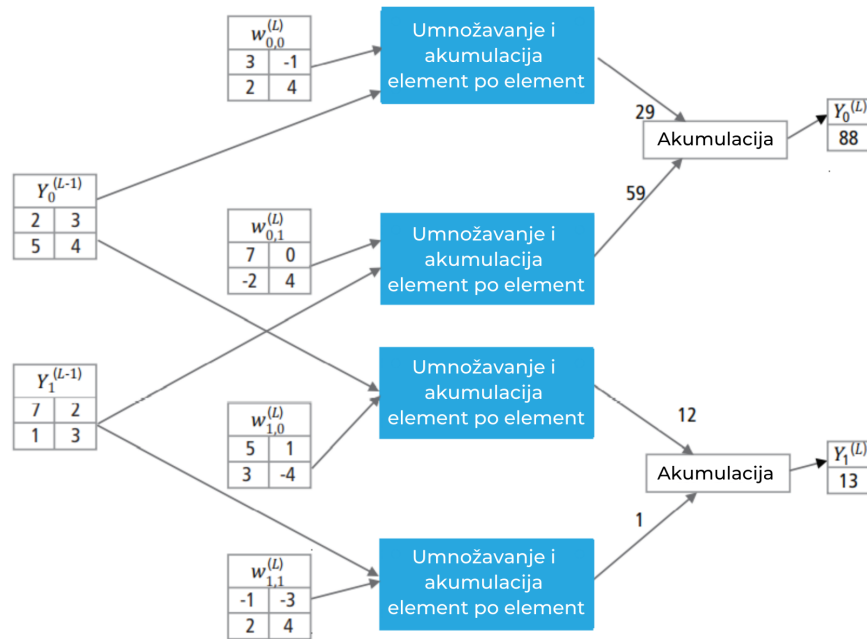
matrice koje se ne preklapaju veličine 2×2 . U slučaju maksimalnog udruživanja, maksimalna vrijednost četiri vrijednosti u matrici 2×2 je izlaz. U slučaju prosječnog združivanja, prosjek od četiri vrijednosti je izlaz. Rezultat s indeksom (2, 2) je rezultat prikazan razlomkom koji je zaokružen na najbliži cijeli broj [9].



Slika 3.5. Prikaz maksimalnog i prosječnog udruživanja

- Potpuno povezani sloj

Potpuno povezani slojevi često se koriste kao završni slojevi CNN-a. Ovi slojevi matematički zbrajaju težinu prethodnog sloja značajki, što ukazuje na preciznu mješavinu "sastojaka" pri određivanju specifičnog ciljanog izlaznog rezultata. U slučaju potpuno povezanog sloja, svi elementi svih značajki prethodnog sloja koriste se u izračunavanju svakog elementa svake izlazne značajke. Slika 3.6. prikazuje potpuno povezani sloj L. Sloj L-1 ima dvije značajke, od kojih je svaka 2×2 , tj. ima četiri elementa. Sloj L ima dvije značajke, od kojih svaka ima jedan element [9].



Slika 3.6. Prikaz potpuno povezanog sloja

- Nelinearni sloj

Neuronske mreže općenito, a posebno CNN-ovi oslanjaju se na nelinearnu funkciju "okidača" za signaliziranje jasne identifikacije vjerojatnih značajki na svakom skrivenom sloju. CNN-i mogu koristiti razne specifične funkcije kao što je ispravljena linearna jedinice (ReLU) i kontinuirane okidačke (nelinearne) funkcije za učinkovitu implementaciju nelinearnog okidanja [9].

3.2.2. Zašto koristiti CNN za prepoznavanje slika

- Otpornost na izobličenje i pomicanje slika

Izobličenja koja nastaju prilikom detekcije objekata događaju se zbog različitih uvjeta osvjetljenja, pojavom djelomičnih okluzija, različitih objektivna kamere, okomitih i vodoravnih pomaka, različitih poza promatranog objekta... CNN je otporan na ova izobličenja jer u cijelom prostoru koristi istu konfiguraciju težine te mu je pomak nepromjenjiv. U teoriji, ako koristimo potpuno povezane slojeve, možemo postići invarijantnost, ali će nam u tom slučaju više jedinica s različitih mjesta ulaza imati istu težinu uzorka. Kako bi mreža naučila što bolje baratati s konfiguracijama težine potreban je veliki broj treninga instanci i što više mogućih varijacija u promatranom prostoru.

- Kapacitet memorije

Zahtjevi za memorijom su drastično smanjeni kod CNN mreža naprema potpuno povezanim neuronskim mrežama. Neuroni nisu potpuno povezani te je zbog toga reduciran broj parametara. Nadalje, prostorna korelacija (veća udaljenost smanjuje korelaciju) i grupiranje neurona (pridruživanje jednog parametra na grupu neurona) omogućavaju smanjenje broja parametara te se time ubrzava treniranje i reducira potreban kapacitet memorije.

- Brže i lakše treniranje

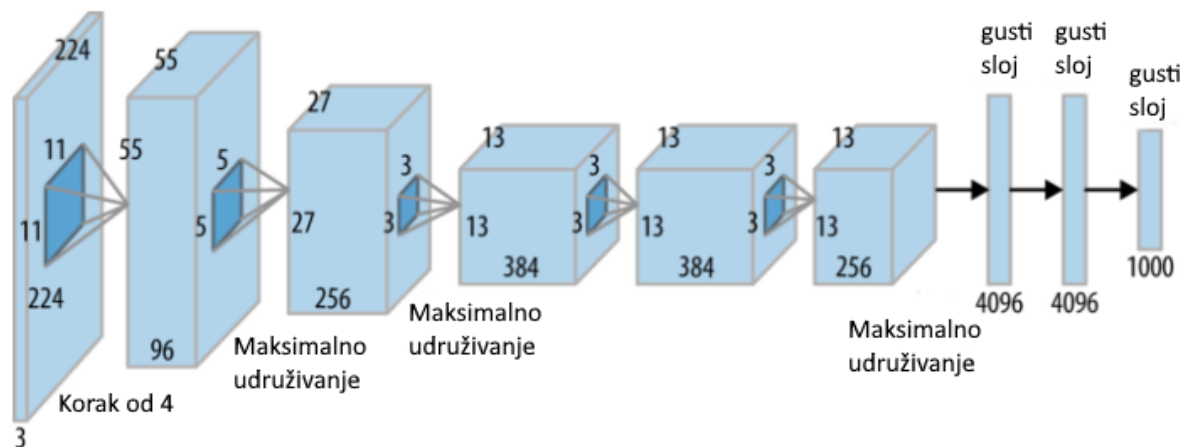
Ukoliko bi se koristila standardna neuronska mreža broj parametara bio bi puno veći za razliku od konvolucijske neuronske mreže pa bi u skladu s tim vrijeme treniranja bilo veće. Kada bi se odlučila provesti savršena obuka i dizajnirala standardna neuronska mreža čija bi izvedba bila identična CNN-u, broj parametara još bi se povećao pa bi tokom procesa treniranja došlo do većeg šuma. Zaključak je da se izvedba standardne neuronske mreže ne može mjeriti s izvedbom konvolucijske neuronske mreže [9].

3.2.3. Klasifikatori

- AlexNet

AlexNet je bila jedna od prvih dubokih konvolucijskih mreža koja je sudjelovala u izazovu ImageNet LSVRC-2012 iz 2012. godine i postigla relevantne rezultate. Ostvarila je najbolju točnost od 84,7% za razliku od drugoplasiranih s točnošću od 73,8%. Mreža se sastoji od 3 potpuno povezana (FC) sloja i 5 konvolucijskih (CONV) slojeva (Slika 3.7. “AlexNet blok dijagram”). Mreža može obučiti ukupno 62 milijuna varijabli. Ulaz u mrežu je serija RGB slika veličine $227 \times 227 \times 3$, a izlaz je vektor vjerojatnosti 1000×1 koji odgovara svakoj klasi. Najčešće korištene aktivacijske funkcije prije AlexNeta su bile sigmoid i tanh funkcija. Negativne strane ovih funkcija su zasićenje, nestajući gradijent (VG) koji otežava obuku mreže stoga AlexNet koristi funkciju ReLU aktivacije. Mreža s ReLU funkcijom je oko 6 puta brža od iste mreže s tanh funkcijom. Iako ReLU pomaže s problemom nestajanja gradijenta, AlexNet je uveo i normalizaciju lokalnog odgovora (LRN) zbog neograničene prirode ReLU-a i nastajanja nepotrebno visokog broja varijabli. LRN provodi normalizaciju u susjedstvu piksela koji pojačavaju aktivaciju promatranog neurona dok istovremeno prigušuju okolne neurone. AlexNet

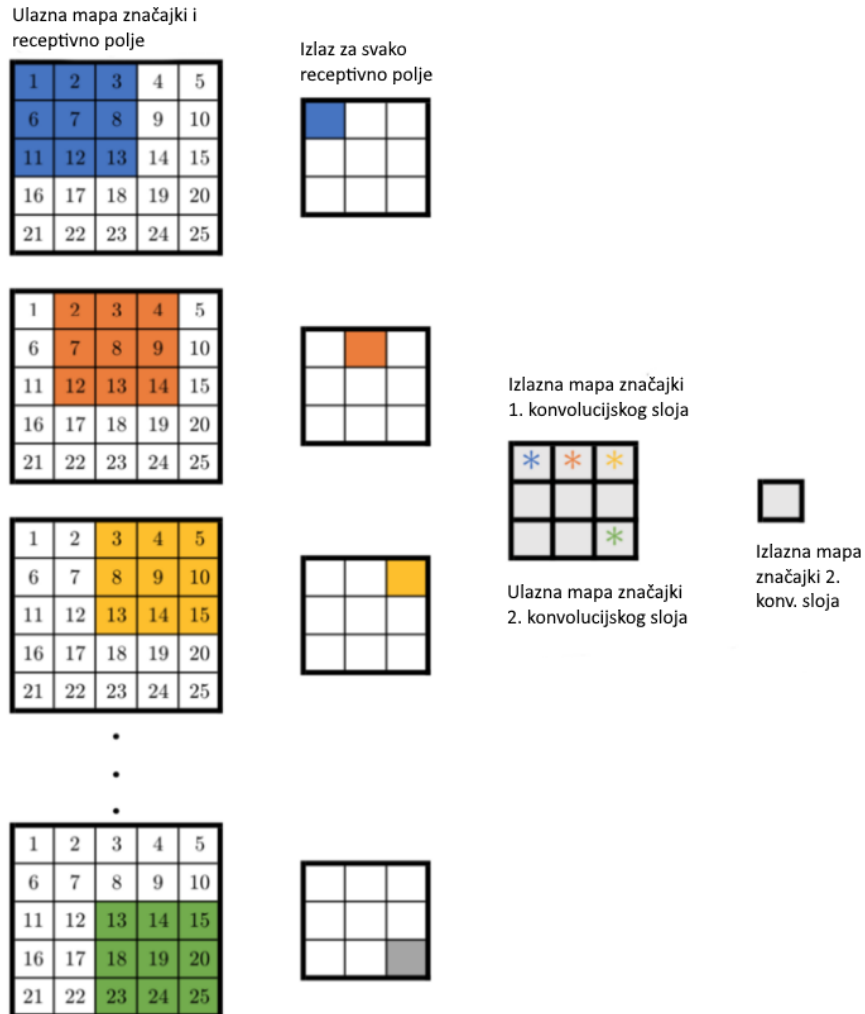
također rješava problem prekomjernog preklapanja, ali se pritom udvostručuje broj iteracija potrebnih za konvergenciju [10].



Slika 3.7. AlexNet blok dijagram

- VGGNet

VGGNet je nastao kako bi smanjio vrijeme treniranja i smanjio broj parametara u konvolucijskim slojevima. Postoji više varijanti VGGNet-a (VGG16, VGG19...) koje se razlikuju samo po ukupnom broju slojeva u mreži. VGG16 ima ukupno 138 milijuna parametara. Svi konvolucijski *kerneli* imaju veličinu 3×3 dok *maxpool kerneli* imaju veličinu 2×2 . VGGNet ima fiksne veličine *kernela* kako bi se sve konvolucijske *kernele* promjenjive veličine, koje se koriste u AlexNetu (11×11 , 5×5 , 3×3), mogle replicirati korištenjem više *kernela* od po 3×3 kao građevnih blokova (Slika 3.8. “Primjer ulaza, obrade i izlaza VGGNet mreže”). Pojam replikacije označava prekrivanje receptivnog polja s *kernelima*. Time se smanjuje broj varijabli koje je moguće brže obučiti i time dolazimo do bržeg učenja i robusnije prilagodbe [11].

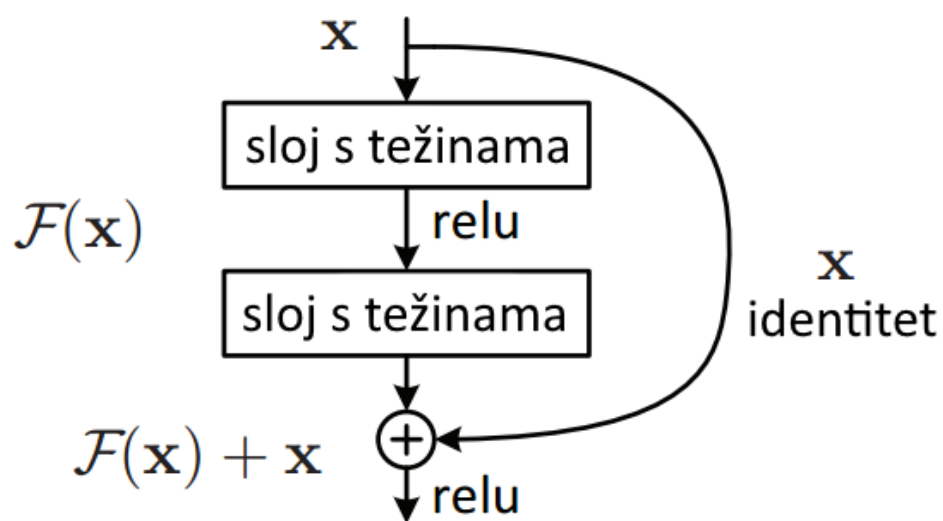


Slika 3.8. Primjer ulaza, obrade i izlaza VGGNet mreže

- ResNet

Neuralne mreže su poznate po tome što ne mogu pronaći jednostavnije mapiranje kada ono postoji. Eksperimenti su pokazali da se točnost smanjuje dodavanjem više slojeva mreži. Spomenuti problemi nastaju zbog problema nestajanja gradijenta. ResNet uvodi dvije vrste ”prečaca veza”: prečac identiteta i prečac za projekciju. Postoji više verzija ResNet XX arhitekture gdje “ XX ” označava broj slojeva. Najčešće korišteni su ResNet50 i ResNet101. Resnet18 ima oko 11 milijuna parametara za obuku. Sastoji se od konvolucijskih slojeva s filterima veličine 3×3 (baš kao VGGNet). U cijeloj mreži koriste se samo dva sloja za spajanje, jedan na početku, a drugi na kraju mreže. Osnovni građevni blok ResNeta je rezidualni blok koji

se ponavlja u cijeloj mreži koji je prikazan na slici 3.9. "Rezidualni blok ResNet mreže". Umjesto da uči preslikavanje iz $x \rightarrow F(x)$, mreža uči preslikavanje iz $x \rightarrow F(x)+G(x)$. Kada je dimenzija ulaza x i izlaza $F(x)$ jednaka, funkcija $G(x) = x$ je funkcija identiteta, a veza prečaca naziva se veza identiteta. Identično mapiranje se uči nuliranjem težina srednjem sloju tijekom treniranja budući da je lakše nulirati težine nego ih gurnuti na jedan. U slučaju kada se dimenzije $F(x)$ razlikuju od x (zbog duljine koraka > 1 između konvolucijskih slojeva), implementira se veza projekcije umjesto veze identiteta. Funkcija $G(x)$ mijenja dimenziju ulaza x u dimenziju izlaza $F(x)$ [12].



Slika 3.9. Rezidualni blok ResNet mreže

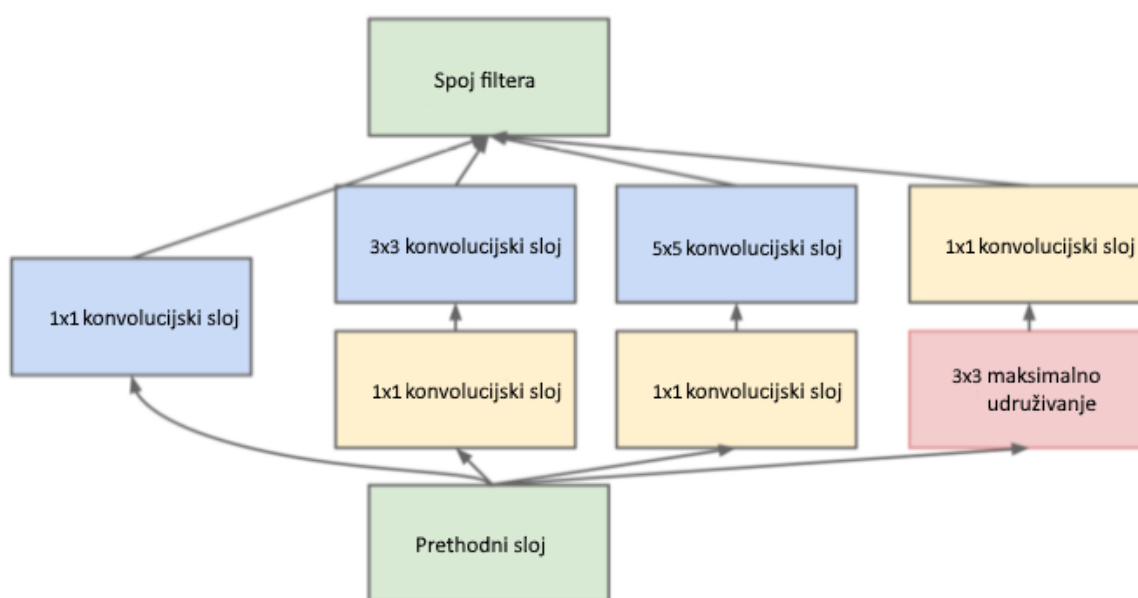
- Inception

Inception, umjesto da jednostavno ide dublje u smislu broja slojeva, ide šire. Više *kernels* različitih veličina implementirano je unutar istog sloja. Arhitektura se sastoji od nekoliko početnih modula sljedeće strukture:

- 1×1 konvolucijskih slojeva,
- 3×3 konvolucijskih slojeva,
- 5×5 konvolucijskih slojeva,
- maksimalno udruživanje.

Za smanjenje dubine koriste se konvolucijski blokovi (1×1) prikazani žutom bojom (Slika 3.10. "Početni modul Inception mreže"). Rezultati iz četiri paralelne operacije zatim se spajaju po

dubini kako bi se formirao “Filter Concatenation” blok (zeleno). Prilikom treniranja Inception povećava mrežni prostor iz kojeg će se odabrati najbolja mreža. Na različitim razinama svaki početni modul može obuhvatiti istaknute značajke. Konvolucijski sloj 5x5 objedinjuje globalne značajke dok je konvolucijski sloj 3x3 sklon hvatanju distribuiranih značajki. Za hvatanje značajki niske razine koje se ističu u susjedstvu zadužena je operacija maksimalnog udruživanja. Sve ove značajke se izdvajaju i spajaju na istoj razini prije nego što se unesu u sljedeći sloj. Prepušta se obuci/mreži da odluči koje značajke imaju najvišu vrijednost i težinu. Najjednostavnija verzija Inceptiona je GoogLeNet [13].



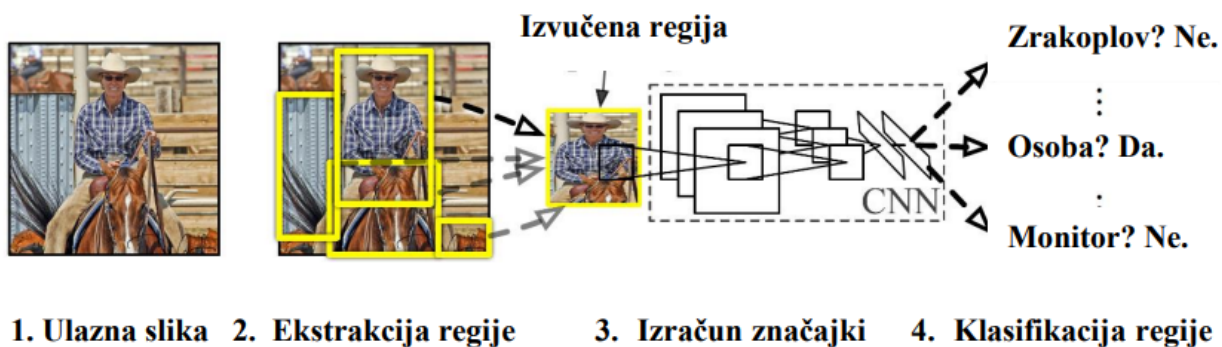
Slika 3.10. Početni modul Inception mreže

3.2.4. Detektori čiji se modeli sastoje od dvije faze

- R-CNN

Kako bi zaobišao problem odabira ogromnog broja regija, R-CNN koristi algoritam selektivnog pretraživanja za izdvajanje 2000 regija iz slike koje se nazivaju predloženim regijama. Izlaz ove neuronske mreže je 4096-dimenzionalni vektor značajki. Konvolucijska neuronska mreža djeluje kao izdavač značajki, a gusti sloj izlaza sastoji se od značajki izdvojenih iz slike (Slika 3.11. ”Arhitektura R-CNN”). Kako bi se klasificirala prisutnost objekta unutar predložene regije sve

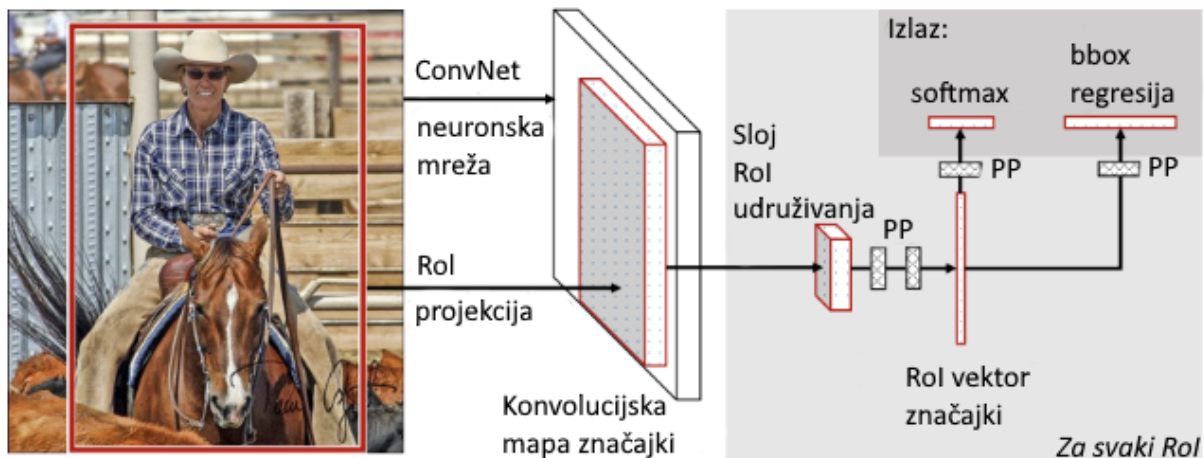
značajke se unose u SVM (*eng. Support Vector Machine*). Kako bi se povećala preciznost graničnog okvira algoritam također predviđa četiri vrijednosti, vrijednosti pomaka. Na primjer, s obzirom na predloženu regiju, algoritam bi predvidio prisutnost osobe, ali bi lice te osobe unutar prijedložene regije moglo biti prepolovljeno. Stoga vrijednosti pomaka pomažu u prilagođavanju graničnog okvira predložene regije. Ono što predstavlja problem pri korištenju R-CNN je to da bismo morali klasificirati 2000 predloženih regija po slici, a to iziskuje puno vremena za treniranje mreže. Ne može se implementirati u stvarnom vremenu i algoritam je fiksni (nema nikakvog učenja) i to bi moglo dovesti do stvaranja loših predloženih regija [14].



Slika 3.11. Arhitektura R-CNN

- Fast R-CNN

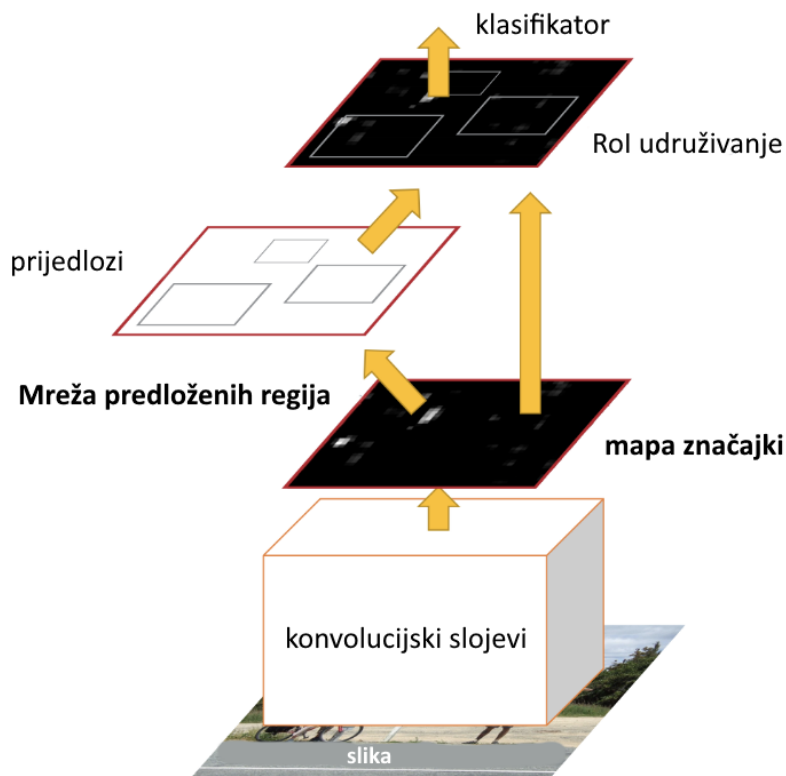
Pristup "Fast R-CNN" mreže sličan je R-CNN algoritmu, no ovdje se generira konvolucijska mapa značajki pomoću ulazne slike umjesto da se CNN-u dostavljaju predložene regije. Identificiraju se predložene regije iz mape značajki i pomoću sloja za prikupljanje RoI-a (*eng. Region of Interest*) preoblikuju se u fiksnu veličinu tako da se mogu uvesti u potpuno povezani sloj. Softmax sloj se koristi za predviđanje klase predložene regije, kao i vrijednosti pomaka za granični okvir (Slika 3.12. "Arhitektura Fast R-CNN"). "Fast R-CNN" brži je od R-CNN-a zato što nije potrebno svaki put nuditi 2000 predloženih regija konvolucijskoj neuronskoj mreži. Umjesto toga, operacija konvolucije se izvodi samo jednom po slici i iz nje se generira mapa značajki [15].



Slika 3.12. Arhitektura Fast R-CNN

- Faster R-CNN

R-CNN i “Fast R-CNN” koriste selektivno pretraživanje kako bi saznali koje su predložene regije. Selektivno pretraživanje je dugotrajan i spor proces koji utječe na performanse mreže. Stoga, osmišljen je algoritam za detekciju objekata koji eliminira algoritam selektivnog pretraživanja i omogućuje mreži da nauči predlagati regije (Slika 3.13.” Algoritam za detekciju objekata Faster R-CNN”). Slika se daje kao ulaz u konvolucijsku mrežu koja pruža konvolucijsku mapu značajki. Za identifikaciju predložene regije i za njeno predviđanje koristi se zasebna mreža umjesto korištenja algoritma selektivnog pretraživanja na mapu značajki. Predviđene predložene regija zatim se preoblikuju s pomoću sloja za udruživanje RoI koji se zatim koristi za klasifikaciju slike unutar predložene regije i predviđa vrijednosti pomaka za granične okvire. Faster R-CNN se može koristiti za detekciju objekata u stvarnom vremenu [16].



Slika 3.13. Algoritam za detekciju objekata Faster R-CNN

3.2.5. Detektori čiji se modeli sastoje od jedne faze

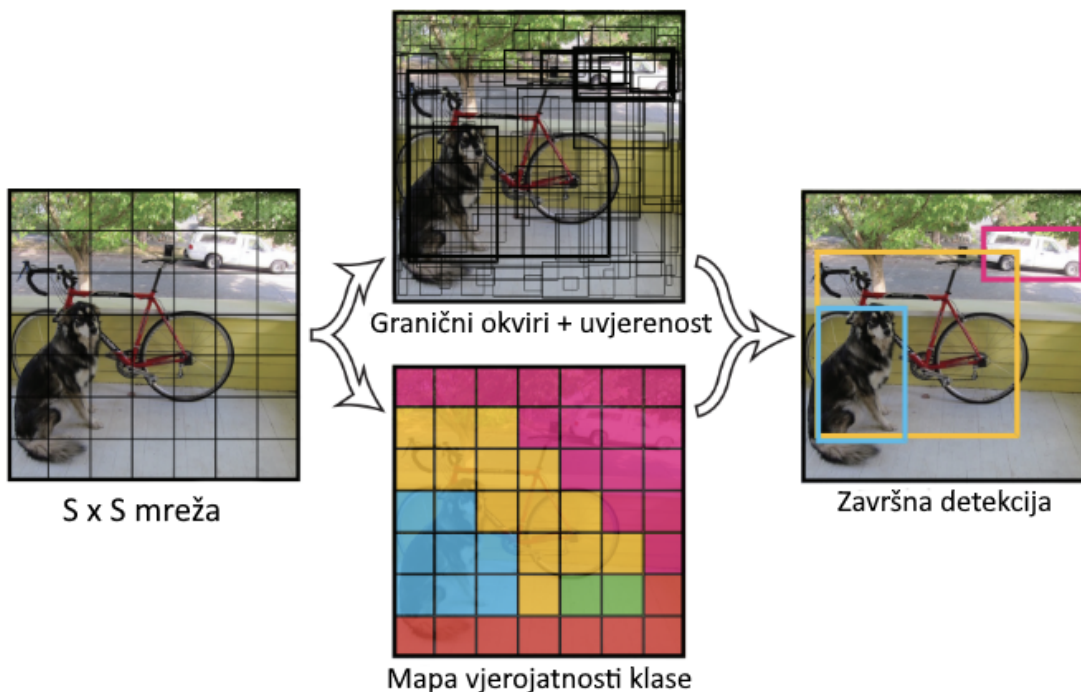
- SSD

SSD (*eng. Single Shot Multibox Detector*) ima dvije komponente, glavu SSD-a (*eng. SSD head*) i okosnicu (*eng. backbone*). Glava SSD-a je odgovorna za izradu mapa značajki koje prikazuju granični okvir i predviđaju klase. Budući da može detektirati objekte različitih veličina algoritam je prostorno nepromjenjiv. Okosnica predstavlja unaprijed treniranu mrežu za klasifikaciju slika kao što je VGG16 ili ResNet koja djeluje kao ekstraktor značajki, bez posljednjeg sloja klasifikacije. Prvi korak SSD-a je dijeljenje slike u mrežu ćelija. Svaka ćelija je odgovorna za predviđanje postoje li koordinate i objekt unutar graničnog okvira ako je objekt prisutan u ćeliji. Svakoj ćeliji je dan broj sidrišta, sidrenih okvira (*eng. anchor box*) različitih omjera i veličina. Ukoliko postoji više objekata unutar jedne ćelije sidrišta omogućuju višestruko otkrivanje. Problem otkrivanja malih objekata SSD rješava tako da SSD glava generira manje mape značajki kako se bliži kraju mreže, a to rezultira većim receptivnim poljima. Na svakom konvolucijskom

sloju sidrišta različitih veličina mogu detektirati objekte različitih veličina počevši od najmanjih. SSD uspijeva otkriti objekte različitih razmjera unutar slike pomoću receptivnih polja. Na svakom konvolucijskom sloju konvolucijske operacije proizvode značajke koje predstavljaju različite regije različitih veličina na slici [17].

- YOLO

YOLO (*eng. You Only Look Once*) je algoritam za otkrivanje objekata koji se razlikuje od algoritama koji se temelje na regiji koji su prethodno prikazani (Slika 3.14. "Algoritam za detekciju objekata YOLO"). Spomenute mreže ne gledaju na cjelovitu sliku već na na dijelove slike koji imaju veliku vjerojatnost da sadrže objekt. U YOLO-u jedna konvolucijska mreža predviđa granične okvire i vjerojatnosti klasa za te okvire. YOLO radi tako da sliku podijeli u $S \times S$ mrežu, unutar svake mreže uzima m graničnih okvira. Za svaki granični okvir, mreža daje vrijednosti pomaka za granični okvir i vjerojatnost klase. Odabiru se granični okviri koji imaju vjerojatnost klase iznad granične vrijednosti i koriste se za lociranje objekta unutar slike. YOLO je brži od drugih algoritama za detekciju objekata. Zbog prostornih ograničenja algoritma YOLO ima probleme s detektiranjem malih objekata unutar slike [18].



Slika 3.14. Algoritam za detekciju objekata YOLO

- YOLOv2

YOLO ima značajne pogreške pri lokalizaciji, stoga se YOLOv2 fokusira na poboljšanje lokalizacije i očuvanje točnosti klasifikacije. Novi dizajn YOLOv2 uklanja potpuno vezani sloj i dodaje sloj grupne normalizacije (eng. *batch normalization*) na sve konvolucijske slojeve te postiže više od 2% poboljšanja kod srednje prosječne preciznosti (mAP). Koristi klasifikator visoke razlučivosti pri treniranju, povećanje razlučivosti od dosadašnje 224 na trenutnu od 448. Mreža mora, pri detekciji novog objekta, simultano učiti i postaviti rezoluciju za novi ulaz. Koristeći Anchor Box, YOLOv2 povećava točnost i izvedbu pri detekciji više objekata odjednom [19]. Osnovna arhitektura koju koristi YOLOv2 je Darknet-19 koja se sastoji od 19 konvolucijskih slojeva. Prednosti YOLOv2 nad njegovim prethodnikom su: poboljšana preciznost detekcije malih objekata, pogreške pri lokalizaciji su smanjene i model razlikuje objekte s različitim omjerima visine i širine [20].

- YOLOv3

YOLOv3 je nastao kao rezultat optimizacije i razvoja YOLOv2. YOLOv3 koristi neovisne logičke klasifikatore dok je YOLOv2 koristio softmax sloj prilikom klasifikacije i time izbjegava složenost izračuna. Objekt se može percipirati kao pješak i dijete u isto vrijeme, ne isključujući jedno drugo. YOLOv3 po svakoj lokaciji predviđa regije na tri različite skale. Svaka predviđena regija se sastoji od graničnih okvira, objekta i 80 rezultata klase; $N \times N \times [3 \times (4 + 1 + 80)]$. Umjesto Darknet-19 YOLOv3 koristi Darknet-53 koja se sastoji od 53 konvolucijska sloja [21]. Detekcija objekata je u realnom vremenu. YOLOv3 postiže bolju detekciju malih objekata, smanjuje pogreške prilikom lokalizacije, ali ima manja postignuća pri detekciji srednjih i velikih objekata za razliku od YOLOv2 [19].

- YOLOv4

YOLOv4 za razliku od YOLOv3 ima povećanje od 12% u broju okvira u sekundi i poboljšanje srednje prosječne preciznosti (mAP) za čak 10%. Arhitektura Yolov4 ima 4 različita bloka: okosnicu (eng. *the backbone*), vrat (eng. *the neck*), sloj gustog predviđanja (eng. *the dense prediction*) i sloj oskudnog predviđanja (eng. *the sparse prediction*). Arhitektura koja je okosnica YOLOv4 je CSPDarknet53 (eng. *Cross-Spatial-Partial Darknet53*) i dijeli trenutni sloj na dva dijela, jedan koji prolazi kroz konvolucijske slojeve i drugi koji ne prolazi kroz konvolucijske

slojeve nakon čega se rezultati agregiraju [22].

- YOLOv5

U roku od samo dva mjeseca nakon izlaska YOLOv4, objavljena je implementacija YOLOv5 na servisu GitHub te je otvorenog izvornog kôda. YOLOv5 nudi arhitekturu za detekciju objekata unaprijed treniranih na skupu podataka MS COCO. Ovo je jedini YOLO detektor objekata koji nema istraživački rad [23]. Razlika YOLOv5 je ta što se radi o PyTorch implementaciji, a ne nadogradnja izvornog Darkneta. Poboljšanje modela se očituje u povećanju mozaičkih podataka i “Anchor Boxa” s automatskim učenjem [24]. Postoje 4 verzije YOLOv5 modela: YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x [25] gdje je prva verzija najmanja i ima najmanju preciznost, a posljednja je najveća i ima najveću preciznost. Usporedba s drugim modelima je loša, dokumentacija je manjkava. Na servisima GitHub [26] i Google Colab [27] moguće je pronaći opis procesa učenja. Iako se YOLOv5 pokazao kao perspektivan model, do velikog ažuriranja YOLOv4 mreže treba proći još vremena [28].

4. DALJINSKO ISTRAŽIVANJE

Daljinsko istraživanje ili prepoznavanje iz daljine (*eng. Remote sensing*) je znanost o proučavanju i otkrivanju fizikalnih svojstava područja dok se sam promatrač ne mora nalaziti na promatranom području (Slika 4.1. “Prikaz dionika procesa daljinskog istraživanja”). Omogućuje vizualizaciju, snimanje i analizu promatranih objekata i značajki površine Zemlje. Ovaj proces je dobio naziv zbog činjenice da koristi senzor (*eng. sensor*) pri prikupljanju informacija iz daljine. Sateliti, zrakoplovi i bespilotne letjelice - dronovi (*eng. Unmanned Aerial Vehicle; UAV*) imaju specijalizirane platforme za nošenje senzora [29]. Zrakoplovi ili sateliti snimaju velika područja na površini Zemlje, percepciju slike morskog dna omogućuju nam sonarni sustavi na brodovima dok promjene temperatura u oceanima možemo mjeriti pomoću satelita. Upotreba procesa prepoznavanja iz daljine može se koristiti i za specifične slučajeve kao što su:

- mapiranje velikih šumskih požara,
- predviđanje vremena promatrajući oblake,
- predviđanje putanje lave i otrovnog dimnog oblaka nakon erupcije vulkana,
- promatranje promjena na poljoprivrednim zemljištima i šumama tokom godina,
- praćenje rasta gradova,
- mapiranje i otkrivanje oceanskog dna. [30]

Pokrivenost, promjene i mapiranje obalnih ekosustava i samo njihovo upravljanje postalo je mnogo lakše i efikasnije primjenom satelitskih snimki i upotrebom procesa prepoznavanja iz daljine. Bespilotne letjelice, dronovi, najčešće se koriste za detekciju makro morskog otpada (*eng. macro litter; ML*). Njihova upotreba potencijalno povećava učestalost istraživanja jer je ljudski napor na terenu puno manji. Takva istraživanja nisu invazivna i smanjuju antropogeni utjecaj na obalu. Moguće je identificirati žarišne točke nakupljanja ML-a s obzirom da UAV slike omogućuju geolokalizaciju detektiranih objekata. Detekcija i kategorizacija ML-ova se u većini slučajeva vrši ručno te je poželjno taj proces automatizirati. Osim što zahtjeva više ljudskog napora, ručna obrada je vrlo subjektivna i kvaliteta procjene ovisi o stručnosti promatrača i njegovom iskustvu. Automatska detekcija i klasifikacija je manje zahtjevna i zamorna, međutim, trenutne metode ne mogu još uvijek kategorizirati ML-ove na detaljan način.

Ono što je “nametnuto” u rješavanju ovog problema, zbog ogromnog rasta i razvoja, je primjena AI algoritma [31].

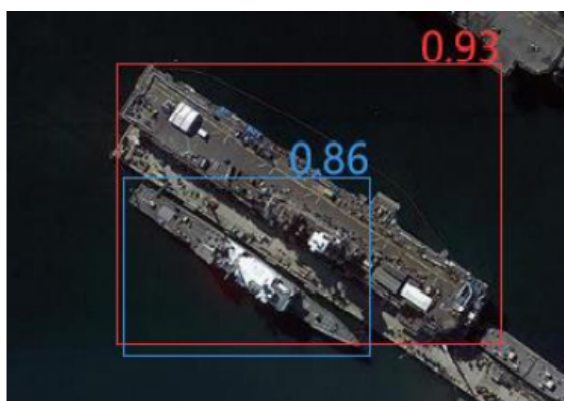


Slika 4.1. Prikaz dionika procesa prepoznavanja iz daljine [29]

5. PREGLED PODRUČJA

5.1. Radovi u kojima se objekt promatranja nalazi na površini mora

Rad [32] predlaže poboljšani algoritam “Faster R-CNN” gdje su objekti grupirani s algoritmom “K-means clustering” i koriste algoritam “Soft NMS” umjesto algoritma “NMS” (*eng. Non - Maximum Suppression*). Obradivali su slike s Google Eartha i natjecateljske platforme Kaggle gdje su promatrani objekti bili brodovi na površini mora. Rezultati su pokazali da poboljšani algoritam “Faster RCNN” ima bolje prosječne točnosti detekcije morskog objekta (Slika 5.1. “Prikaz situacije izostavljanja”) te rješava probleme niske prosječne točnosti i propusta malih objekata u složenom morskome okruženju (Slika 5.2. “Prikaz rezultata s malim objektima”).

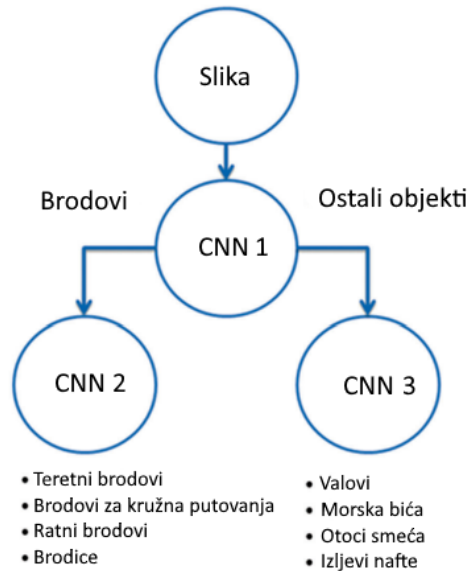


Slika 5.1. Prikaz situacije izostavljanja



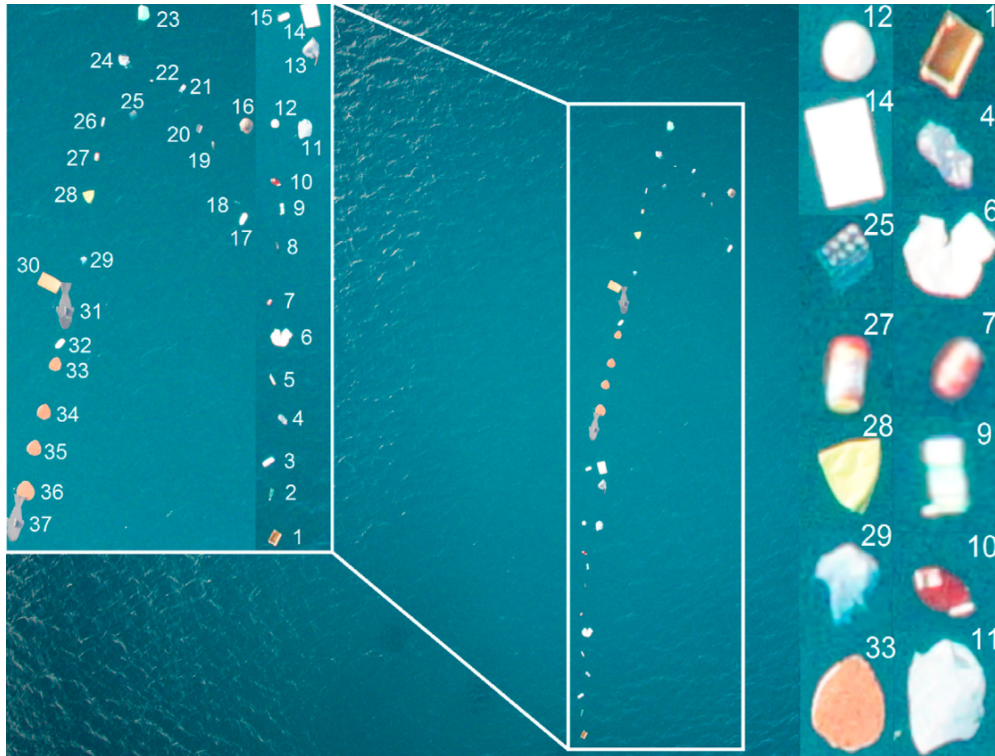
Slika 5.2. Prikaz rezultata s malim objektima

Autori rada [33] prikazuju rezultate na svojoj bazi podataka čije su slike slikane pomoću Google satelita i Google tražilice slika. Promatrani objekti su podijeljeni u dvije kategorije: brodovi i ostali objekti i njihove podkategorije: teretni brodovi, ratni brodovi, brodovi za kružna putovanja, brodice i valovi, otoci smeća, morska bića i izljevi nafte. Predlažu algoritam koji se sastoji od tri konvolucijske neuronske mreže (Slika 5.3. “Dijagram tijeka podataka algoritma za prepoznavanje morskih objekata”). Korištenje CNN mreža se pokazalo kao odgovarajuća metoda umjetne inteligencije pri prepoznavanju morskih objekata iz zraka.



Slika 5.3. Dijagram tijeka podataka algoritma za prepoznavanje morskih objekata

Rad [34] predlaže algoritam koji je sposoban detektirati i kvantificirati plutajući morski makro otpad (eng. *Floating Marine Macro Litter; FMML*) sa slika iz zraka (Slika 5.4. “Slika iz zraka s nizom poznatih predmeta pričvršćenih na užu”). Razvili su web aplikaciju MARLIT [35] koja omogućava korisnicima detektirati FMML s učitanih slika. Predloženi algoritam koristi CNN koji je sposoban učiti iz nestrukturiranih ili neoznačenih podataka. Model je istreniran na 3723 slike, gdje se na 50% njih ne nalazi FMML, a na ostalih 50% se nalazi. Slikane su iz daljine dronom i zrakoplovom iznad Sredozemnog mora. Točnost klasifikacije slika i križne provjere (eng. *cross validation*) bila je 0,85, odnosno 0,81. Otkrivanje i kvantifikaciju FMML-a pojednostavit će implementacija ovog ili sličnih algoritama. Ono što i dalje predstavlja problem je automatizirano praćenje FMML-a na otvorenom moru te su potrebna daljnja istraživanja.



Slika 5.4. Slika iz zraka s nizom poznatih predmeta pričvršćenih na uže (sanduk, boce, limenke, kese, tetrapak, ručnik, balon, karapaks kornjače...)

Autori rada [36] predstavljaju na koji način pratiti prostornu distribuciju plutajuće plastike u svrhu smanjenja negativnih učinaka na okoliš. Bepilotne letjelice (UAV) mogu prikupiti podatke za mapiranje, ali većina metoda zahtijeva ručno označavanje i vizualnu interpretaciju. Cilj ovog rada je testirati prikladne algoritme dubokog učenja za automatsku detekciju plutajuće plastike iz ortofoto snimaka UAV-a, istražiti odnos između veličine plastike i prostorne rezolucije i istrenirati model da razlikuje klase tj. vrstu plastike. Za validaciju i treniranje modela korištena su tri skupa podataka i dva područja istraživanja. Autori predlažu algoritme semantičke segmentacije kao rješenje problema, konkretno algoritme semantičke segmentacije koji rade na principu enkoder-dekoder (*eng. encoder-decoder*). Ovakve mreže rade tako da enkoder generira mapu značajki male rezolucije koristeći standardne CNN slojeve (konvolucijski, udruživanje i aktivacijski). Nakon toga dekoder povećava rezoluciju mapi značajki koristeći podatke iz prolaza kroz enkoder te na kraju prolazi kroz sloj koji klasificira piksele (npr. *softmax*). Rezultat su mape

vjerojatnosti klase za svaki piksel. Tipične arhitekture ovakvog tipa mreže su SegNet i U-Net. U-Net, koji je nastao prije SegNet arhitekture, koristi direktne veze da bi povezao sloj u enkoder dijelu sa slojem u dekodeer dijelu. Mape značajki se spajaju i prenose u cijelosti. U-Net zbog toga troši više memorije, ali mu je potrebno manje trening slika da bi došao do preciznih rezultata.

SegNet ne vezuje cijele mape značajki nego sprema indekse operacije maksimalnog udruživanja (kod korištenja matrice dimenzije $[2 \times 2]$ se spremi indeks u rang $[0, 3]$ na kojem je bio najveći broj). U enkoder dijelu koristi se arhitektura VGG-16 (ali bez potpuno povezanih slojeva na kraju). U dekodeer dijelu se koriste spremljeni indeksi da bi se povećala rezolucija mape te se time iz proriđene dobije gusta mapa značajki. Finalni element dekodeera se šalje u softmax sloj koji klasificira svaki piksel zasebno. Predviđena segmentacija odgovara klasi koja ima najveću vjerojatnost na zadanom pikselu. Autori su odabrali U-Net arhitekturu zbog veće preciznosti. Autori su kreirali vlastiti set slika na kojim su vršili treniranje algoritma. Korišteni su drveni okviri u koje su stavljene boce napravljene od različitih verzija plastike, povezani najlonskim užetom. Odabrani su elementi koji najviše pridonose zagađenju na danim lokacijama. Slike su dobivene koristeći različite visine: 12, 40, 55, 70 i 90 metara. Baze slika ne postoje zbog toga što je ovo prvi rad koji pokušava detektirati tip plastike koristeći UAV slike. Autori su zbog toga morali generirati i vlastiti set podataka. Za svaki prelet UAV-om su generirali ortofotografiju visoke razlučivosti koristeći SFM algoritam (SIFT algoritam detektira ključne značajke, RANSAC algoritam ih poveže kroz više slika). Nakon toga je na svakoj slici izvršeno ručno obilježavanje klasa plastike. Problemi s kojima su se susreli u ovom koraku su što je obilježavanje otežano zbog nemogućnosti operatera da procjeni je li nešto plastika ili nije. Zbog toga su dodali i klasu "možda plastika", koja se koristi kad nisu sigurni u klasu objekta. Drugi problem se događa kod generiranja velike ortofotografije i iskorištavanja iste maske za različitu visinu preleta, naime, kretanje plastičnih objekata između preleta je dovoljno veliko da se jedna maska ne može iskoristiti kod drugog preleta nad istim područjem. Za klasifikaciju plutajuće plastike autori predlažu mrežu koja radi segmentaciju temeljenu na U-Net arhitekturi. U trenutku pisanja rada nije postojala usporedba performansi različitih implementacija, autori su odabrali četiri postojeće implementacije koje su trenirane na ImageNet skupu podataka: ResNet50, ResNeXt50, Inception-ResNet v2 i Xception. ResNet50: koristi konekcije s identitet funkcijom te dodavanjem nekoliko rezidualnih blokova. Inception-ResNet v2 integrira Resnet i Inception v4 koristeći rezidualne konekcije, paralelno radi konvolucije različite veličine i spaja ih u jedan

rezultat. Xception mijenja Inception modul s posebnim konvolucijskim slojevima. ResNeXt50 je sličan Inception modelu gdje je moguće imati više paralelnih primjena *kernels* u konvolucijskom sloju (uz širinu i visinu postoji i dimenzija kardinalnosti). Sve četiri mreže korištene su da bi se usporedili rezultati. Autori su koristili standardne metode da ispituju točnost segmentacije: preciznost (*eng. precision*), opoziv (*eng. recall*) i F1 mjera (*eng. F1 score*). Implementacija je izvršena koristeći navedene mreže te je rezultat šest različitih modela. Autori su testirali različite funkcije gubitka te su zaključili da je najbolja točnost ako se koristi funkcija križne entropije (*eng. cross entropy*). Zbog arhitekture mreže, originalna slika se podijeli u regije od 256×256 piksela te se svaka odvojeno šalje kroz mrežu. Autori spominju problem objekata koji se nalaze na granici regije što je i očekivano kod ovakvog tipa dijeljenja velike slike na regije. Najveća greška se događa u kutovima slika zbog najvećeg nedostatka kontekstualnih informacija. Algoritam ResUNet50, pokazao je najveću točnost pri mapiranju različitih plastičnih materijala (F1-rezultat: najlon: 0,88; orijentirani polistiren (OPS): 0,86; plastika (općenito): 0,78; polietilen tereftalat (PET): 0,92). Kako se smanjivala prostorna razlučivost tako se smanjivala i točnost klasifikacije. Model raspolaže i s podacima procjene površine i volumena plastike što će biti od velike pomoći pri akcijama čišćenja. Slika 5.5. “Rezultati točnosti za različite implementacije mreža” prikazuje usporedbu različitih modela.

	ResUNet50			ResUNext50			XceptionUNet			InceptionResUNet v2		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
OPS	0.86	0.86	0.86	0.99	0.19	0.31	0.81	0.39	0.53	0.01	0.00	0.00
Nylon	0.92	0.85	0.88	0.77	0.96	0.85	0.76	0.87	0.81	0.76	0.74	0.75
PET	0.92	0.92	0.92	0.82	0.96	0.88	0.78	0.75	0.77	0.60	0.72	0.65

Slika 5.5. Rezultati točnosti za različite implementacije mreža

Odabran je ResUNet50 za provedbu daljnjeg istraživanja. Autori su pokazali da visina UAV-a (odnosno, prostorna rezolucija) direktno utječe na rezultate detekcije, kao što slika 5.6. “Rezultati točnosti za različite rezolucije slika” pokazuje.

	13 mm			18 mm			23 mm			30 mm		
	P	R	F1	P	R	F1	P	R	F1	P	R	F1
OPS	0.88	0.77	0.82	0.69	0.71	0.70	0.79	0.31	0.44	0.75	0.45	0.56
Nylon	0.89	0.75	0.82	0.91	0.52	0.66	0.76	0.26	0.39	0.87	0.20	0.33
PET	0.92	0.83	0.87	0.78	0.84	0.81	0.83	0.68	0.75	0.77	0.70	0.73

Slika 5.6. Rezultati točnosti za različite rezolucije slika

Prema CSIRO uputama, minimalna veličina detektirane plastike treba biti 1 cm^2 pa zbog toga autori sugeriraju 4 mm razlučivost (12 m visina UAV-a). Mana je mala površina koja se time obuhvaća s jednom slikom te je potrebno više vremena da se snimi neko područje. Autori ističu da algoritam detektira podvodnu plastiku i u plitkom području rijeke, iako napominju da je izazovno jer prisutnost vodenog sloja iznad plastike povećava reflektivnost te umanjuje točnost detekcije. Algoritam je testiran i na kopnu, no dobiveni rezultati su lošiji u odnosu na površinsku plastiku. Problemi s kojim su se autori susreli je i teško obilježavanje slika na kojima algoritam trenira te predlažu prelet područja s dva drona, jedan na manjoj visini koji hvata detalje, drugi koji generira slike za trening. Uvjeti u kojima se slike trebaju dobiti utječu na rezultate, gdje autori navode da je oblačno bolje od sunčanog (manje refleksije, *sunglint* efekt) i bez vjetra (uzrokuje zamućenost slike, pomak objekata između kadrova).

Autori rada [37] predlažu inventivnu istraživačku primjenu bespilotnih zračnih sustava UAS (*eng. unmanned aerial systems*) i satelitskih slika, otvorenog pristupa, u daljinskom otkrivanju plutajuće plastike u morskom ekosustavu. Izvukli su značajna spektralna mjerenja u gotovo prirodnim uvjetima i usporedili geoprostorne informacije u rasponu od vrlo visoke do umjerene rezolucije. Tri umjetne plutajuće plastične mete od 100 m^2 koje su se sastojale od plastičnih vrećica, najlonskih mreža za ribolov i boca za vodu od 1.5 L postavljene su za daljinsku detekciju u vodama u blizini plaže Tsamakia u Mytilene na otoku Lesvos, Grčka. Spektralna svojstva plutajućih meta i okolnog mora istražena su pomoću satelita Sentinel-2A. Pokazali su kako se UAS može koristiti za procjenu postotka pokrivenosti plastikom sa satelitskih snimka i kako UAS slike vrlo visoke geoprostorne razlučivosti mogu biti korisne u poboljšanju geo-referenciranja satelitskih slika. Uočena ju vrlo jaka do slaba povezanost između spektralne refleksije na p-vrijednosti $< 0,1$ i postotka pokrivenosti pikselima. Pokazali su koliko su važni skupovi podataka UAS-a vrlo visoke geoprostorne rezolucije u potvrđivanju i poboljšanju

geoprostorne točnosti satelitskih podataka za praćenje plastike u vodenom okolišu kao i integrirani paket senzora za praćenje i identifikaciju plastike.

Autori rada [38] uvode novi pristup pri detekciji malih objekata sa satelitskih slika visoke rezolucije. Predlažu korištenje značajnih točaka koje se mogu brzo izdvojiti iz satelitskih slika kao prikladne jedinice promjene za male objekte te kako bi smanjili broj lažnih detekcija. Otkrivaju promjene izdvajanjem značajki iz prethodnih i nedavnih slika i procjenom promjene na temelju veličine promjene značajki. Veličinu procjenjuju izračunavanjem veličine promjene temeljene na pikselima i brojanjem omjera promijenjenih piksela oko izdvojenih značajki. Značajke koje se podudaraju označavaju kao nepromijenjene dok se preostale značajke procjenjuju kao promijenjene ili nepromijenjene na temelju njihove veličine promjene. Novi pristup su testirali na tri kompleta slika s udaljenošću uzorkovanja od 50 cm. Njihov pristup je nadmašio dosadašni pristup, temeljen na pikselima, proizvodeći veću preciznost od 88,7% i točnost od 86,1% pri fiksnoj stopi lažnih detekcija od 10%. Pokazali su kako njihov pristup temeljen na značajkama ima manje pozadinske buke od pristupa temeljenih na pikselima te može kompenzirati nedostatke pristupa temeljenih na nadziranim objektima uspješnim smanjenjem broja kandidata za promjenu. Nedostaci su ti da su objekti bez obilježja ostali neotkriveni.

5.2. Radovi u kojima se objekt promatranja nalazi pod morem

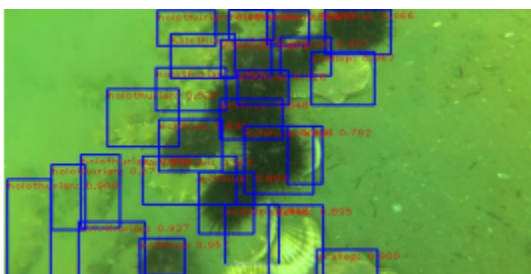
Autori rada [39] koriste dronove s RGB (*eng. red, green, blue*) kamerom kako bi istražili, s pomoću strojnog učenja, utjecaj solarnog kuta i dubine na kvantifikaciju i identifikaciju primarnih bentičkih proizvođača koji se nalaze u rijeci u sjevernoj Kaliforniji. Ciljevi njihovog istraživanja su pomoću algoritma klasifikacije “Random Trees” (“Random Forest”) klasificirati nitaste alge i makrofite, procijeniti točnost algoritma klasifikacije na licu mjesta, istražiti optimalne solarne kutove i maksimalne dubine pri kojima se i dalje može detektirati traženi objekt, kvantificirati postotak pokrivenosti makrofita i nitastih algi na riječnim lokacijama. Koristili su dron “DJI Phantom 4 Pro”, s RGB kamerom od 20 megapiksela. Snimali su na visinama od 19 do 104 metra, letovi su trajali 15-20 minuta iznad svake lokacije, kojih je ukupno bilo 32. Analiza jedne slike je pristup koji je korišten i u istom je analizirana samo jedna slika po mjestu promatranja. Ortomozaici nisu konstruirani zbog poteškoća u povezivanju snimaka iz zraka zbog vode koja je jednolična površina - algoritmi koji bi se koristili za ekstrakciju značajki

i povezivanje slika ne daju rezultate i ne mogu povezati značajke jer se mijenjaju iz slike u sliku (valovi, plutajući objekti, refleksija sunca i slično). Za klasificiranje i treniranje slika korišteno je vizualno promatranje i poznate GPS lokacije te je napravljena kategorizacija od šest klasa: alge, makrofiti, voda, kopno, sjena, odsjaj sunca. Potom je klasifikator "Random Trees" istreniran i primijenjen na sve piksele svih slika u programu "ArcGIS Pro", koji implementira navedeni algoritam. Ukupna točnost svih klasa bila je 82% uzimajući u obzir sve 32 lokacije. Pogreška lažnog prepoznavanja je bila najveća za klasu makrofita (41%), a najmanja za klasu vode (5%) dok je pogreška izostavljanja bila najveća za klasu kopno (32%), a najmanja za klasu sjena (5%). Pokazano je kako na točnost klasifikacije slika uzetih s dronom utječu solarni kut i dubina. Sunčevo podne koje je trajalo 13:13 - 13:20 je utjecalo na kvalitetu slika, stoga slike koje su slikane prije ili poslije Sunčevog podneva pod solarnim kutem 45°-60° su najbolje s najvećom osvijetljenošću podloge i najmanjim Sunčevim blještavilom na površini. Bepilotne letjelice postaju uobičajeni alat koji se koristi u istraživačke svrhe i ima veliki potencijal za povećanje nadzora ekosustava pogotovo na područjima koja prije nisu bila dostupna te je zbog toga od velike važnosti odrediti parametre uzorkovanja kako bi se postigli optimalni rezultati.

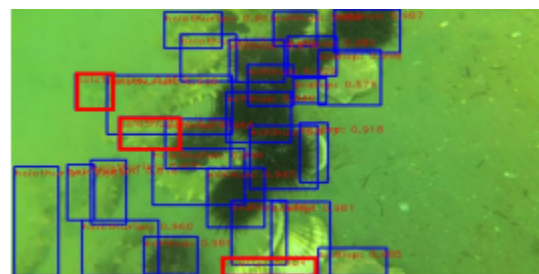
Autori rada [40] imaju za cilj automatski detektirati otpad s morskog dna, koristeći konvolucijske neuralne mreže, u stvarnom okruženju. Neuronska mreža se trenira na setu slika s morskog dna i sadrži jedanaest klasa: plastične boce, plastične vrećice, plastične fragmente, male plastične fragmente, plastične čaše, limenke, gume, plastične čepove, ribolovne mreže, velike objekte i neodređeno. Cilj ovog rada je istražiti kolika je učinkovitost detektiranja objekata na tako složenim slikama dobivenih iz video zapisa pod različitim kutovima, svjetlosnim zasjenjenjem i razinama zumiranja. Podatci korišteni u ovom radu prikupljeni su u okviru programa „Integrirane kampanje informiranja i podizanja svijesti za smanjenje plastičnih vrećica u morskom okolišu” (eng. *LIFE DEBAG - LIFE14 GIE/GR/0011271*). Podvodne slike dobivene su pomoću podvodne kamere (eng. *towed underwater camera, TUC*) koja se nalazi na malom plovilu, koju su vukli po definiranim pravcima zaljeva Ermoupolis u Grčkoj. Skup podataka je uključivao 635 slika izoliranih iz videa razlučivosti 1920×1080 piksela. Ručno je identificirano 1166 objekata, a granični okviri su nacrtani pomoću alata za označavanje slika "LabelImg". Autori su koristili proces augmentacije kako bi povećali skup podataka s obzirom da su imali 635 slika male veličine te koriste Mask R-CNN kao detektor. Mreža je trenirana s pomoću MobileNetV1 okosnice i pokrenuta je korištenjem unaprijed treniranih težina iz COCO (eng.

common objects in context) skupa podataka. Mreža MobileNetV1 je odabrana jer je manje računalo intenzivna u usporedbi s drugim mrežama, a pritom ne gubi svoju točnost te su se tako prevladala ograničenja memorije tijekom obuke i implementacije mreže. Za implementaciju je korišten Keras i TensorFlow. Mreža je postigla ukupnu mAP od 62% po svim klasama, ali je kod određenih klasa ponaosob (ribarske mreže, gume, plastične vrećice i plastični čepovi) bila još učinkovitija i postigla mAP od > 79%. Strojno i duboko učenje relativno je novo istraživačko polje u znanosti o moru koje ima veliki potencijal te je razvoj algoritama za detekciju objekata u realnom vremenu na površini i ispod mora od velike ekološke važnosti. Autori su prikazali i tablicu odabira hiperparametara koje su koristili prilikom treniranja koji su dali najbolje rezultate za podvodne slike. Razloge netočnih predviđanja mreže autori su podijelili u tri kategorije: 1) otpad koji nema jasan geometrijski oblik (poput plastične vrećice) i može se nalaziti u više kategorija istovremeno; 2) polu ili potpuno ukopani otpad u morsko dno koji je toliko vizualno degradiran da i ljudski promatrači imaju problema s detekcijom tog otpada; 3) razni pozadinski elementi i strukture (morska trava, kamenje, razne sjene i izmjena sjena nastala zbog valova). Ovo su glavni razlozi koji utječu na preciznost te još nema jasne metode koja bi se mogla iskoristiti da se ovi problemi riješe.

Rad [41] koristi bazu podataka “URPC 2018” čiji su promatrani objekti na slikama morski krastavci, ježinci, školjke i morske zvijezde. Autori predlažu metodu augmentacije “RoIMix” koja može postići bolju generalizaciju i karakterizira interakcije između slika na kojima se objekti mogu preklapati, biti zamagljeni ili zaklonjeni. Za razliku od dotadašnjih metoda za augmentaciju “RoIMix” obrađuje više slika kako bi imali kvalitetnije i brojnije uzorke za treniranje. Eksperiment pokazuje kako ova predložena metoda poboljšava izvedbu na “Pascal VOC” i “URPC” skupovima podataka kad se radi o detekciji objekata koji su bazirani na regiji (Slika 5.7. “Kontrolna slika” i Slika 5.8. ”Primijenjena metoda RoIMix”).

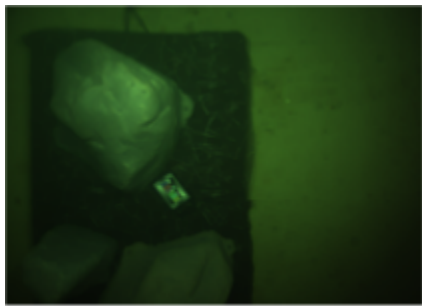


Slika 5.7. Kontrolna slika



Slika 5.8. Primijenjena metoda RoIMix

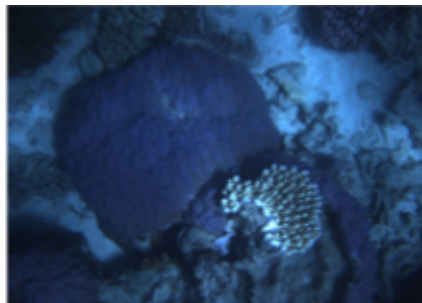
Rad [42] prikazuje rad “WaterGAN” mreže za kreiranje podvodnih, realističnih slika koje su dobivene iz zraka i njihovih parova dobivenih pod vodom povećavajući vidno polje dok je percepcija dubine ograničena. “WaterGAN” mreža korigira boje podvodnih slika koje su izmijenjene zbog fizičkih svojstava mora. Koristeći “WaterGAN” autori su generirali svoj skup podataka, slike iz zraka i njihovi parovi realističnih slika na odgovarajućim dubinama koji su potom trenirali. Dio slika je fotografiran u laboratorijskim uvjetima (Slika 5.9. “Neobrađena podvodna slika”, Slika 5.10. “WaterGAN metoda”), a dio u prirodnom okruženju (Slika 5.11. “Neobrađena podvodna slika”, Slika 5.12. “WaterGAN metoda”).



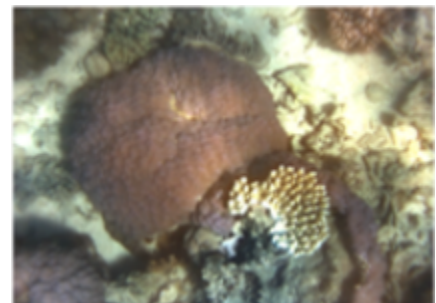
Slika 5.9. Neobrađena podvodna slika



Slika 5.10. WaterGAN metoda



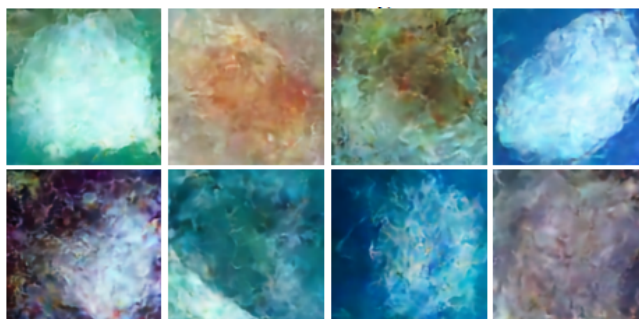
Slika 5.11. Neobrađena podvodna slika



Slika 5.12. WaterGAN metoda

Rad [4] predstavlja probleme loše kvalitete skupova podataka podvodnih slika i njihovu malu brojnost kao i rješenje problema s ciljem što bolje detekcije podvodnih objekata. Koristeći svoju bazu podataka predlažu korištenje dvostupanjskog varijacijskog autoenkodera (eng. *Variational Autoencoder; VAE*) i binarnog klasifikatora. Binarni klasifikator bira slike “dobre kvalitete” i povećava zadani skup podataka (augmentacija) nakon što ih VAE generira (Slika 5.13. “Prva faza”, Slika 5.14. “Druga faza” i Slika 5.15. “Rekonstrukcija”). Završni korak je provjera

klasifikatora i njegova točnost pri detekciji objekata na stvarnim i generiranim slikama. Rezultati pokazuju da tako uvježban klasifikator nadmašuje onaj koji je samo treniran na stvarnim podacima te je ujedno i koristan ukoliko je prikupljanje slika neizvedivo ili izazovno.



Slika 5.13. Prva faza



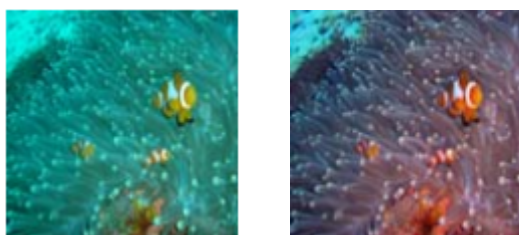
Slika 5.14. Druga faza



Slika 5.15. Rekonstrukcija

Autonomna podvodna vozila (*eng. Autonomous Underwater Vehicle; AUV*) oslanjaju se na razne senzore prilikom donošenja odluka pri detekciji, a onaj koji pruža najviše informacija je "osjet vida" pogotovo pri manjim dubinama. Zbog fizikalnih svojstava mora: lom svjetla, apsorpcija svjetlosti, količina čestica u vodi, dolazi do izobličenja boja promatranih objekata stoga se svaki AUV suočava s teškim izazovom i loše ispunjavaju zadatke ukoliko su vođeni samo senzorom

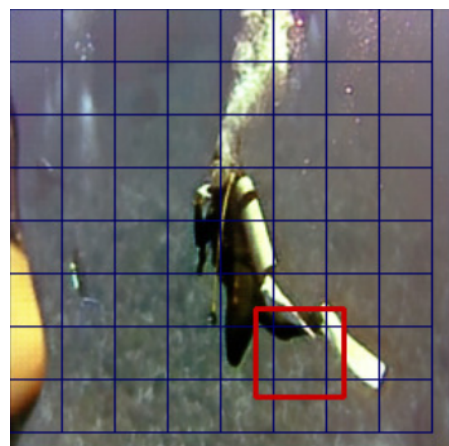
“vida”. Rad [43] predlaže korištenje modela “CycleGAN” s ciljem poboljšanja boja podvodnih slika. Obrađivane su slike iz podskupa baze ImageNet te su izvučeni kadrovi iz videa koji su preuzeti s YouTube kanala. Kvalitativno i kvantitativno rezultati pokazuju učinkovitost ove metode te pokazuju veću točnost pri detekciji morskih živih bića (Slika 5.16. “Original” i Slika 5.17. “CycleGAN”) i ronioaca (Slika 5.18. “Original” i 5.19. “CycleGAN”) za razliku od nekorigitiranih slika.



Slika 5.16. Original Slika 5.17. CycleGAN



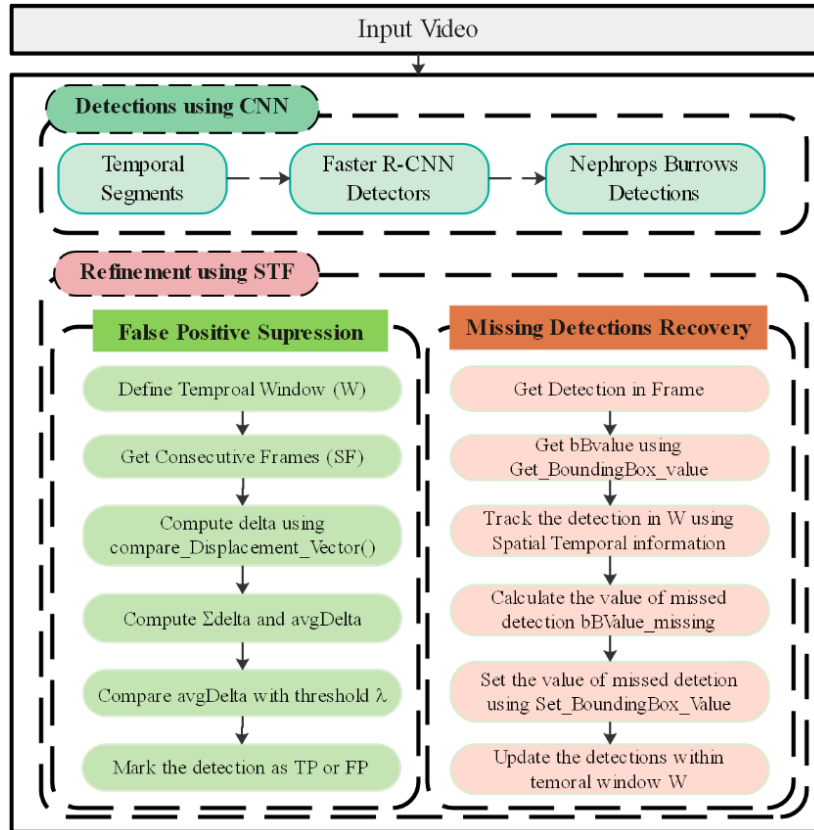
Slika 5.18. Original



Slika 5.19. CycleGAN

Naseer i suradnici [44] predlažu algoritam za precizniju detekciju koji se temelji na prostorno-vremenskoj analizi. Koristeći podvodna videa kao ulaz, algoritam odbacuje lažno pozitivne rezultate te ponovno razmatra i vraća propuštene detekcije podvodnih staništa, jzovina vrste *Nephrops Norvegicus*. Za automatsku klasifikaciju i detekciju koriste najsvremenije duboke neuronske mreže kao što su Inceptionv2 [45], ResNet50 [46], ResNet101 [47] i Faster-RCNN [16]. Predloženi sustav ima dvije faze. U prvoj fazi provodi se detekcija traženog objekta dok se u drugoj fazi provodi rafinirana, detaljnija detekcija. Ulazni video se prvo sekvencionira na vremenske segmente koji se sastoje od N broja okvira koje zatim obrađuju

duboke neuronske mreže. Rezultati se potom prenose u drugu fazu gdje se provodi prostorno-vremensko filtriranje (*eng. spatial-temporal filtering, STF*). Proces poboljšava srednju prosječnu preciznost (*eng. mean average precision, mAP*). Treniranje modela provodi se s pomoću učenja transferiranjem (*eng. transfer learning*) [48] koji se zatim fino padašava u TensorFlowu [49] te se primjenjuju, već navedene, duboke konvolucijske neuronske mreže. Nakon kritičke analize rezultata uočeno je da detektori nailaze na mnogo FP (*eng. false positive*) i propuštaju mnoge TP (*eng. true positive*) čime opada točnost. Modeli ResNet50, ResNet101 i Inception testirani su na videu u trajanju od pet minuta. Predloženi algoritam za precizniju detekciju uzima V , λ i W kao ulaze. V predstavlja video dok λ predstavlja vrijednost praga za vektor pomaka. Vrijednost praga je vrijednost IoU (*eng. intersection over union*) koja se kasnije uspoređuje s IoU otkrivenih staništa. W je veličina vremenskog prozora koji određuje broj okvira u vremenskom prozoru. Ovi modeli pružaju skup TP, FP i propuštenih detekcija. Algoritam svaku detekciju smatra TP ako je njegov prosječni IoU u svim okvirima u vremenskom prozoru veći ili jednak vrijednosti praga λ i ako detektor kontinuirano detektira unutar vremenskog prozora. FP detekcije su one detekcije kojima je kombinirani IoU manji od vrijednosti praga λ i koje se ne detektiraju u uzastopnim okvirima. Propuštene detekcije su one detekcije koje su TP i koje detektor detektira u nekim okvirima, ali su propuštene u nekim srednjim kadrovima zbog vidljivosti staništa ili položaja. Propuštene detekcije vrlo je važno identificirati jer bez njih ne možemo pratiti staništa. Algoritam je podijeljen u dva dijela, suzbijanje lažnih pozitivnih rezultata i prepoznavanje propuštenih detekcija (Slika 5.20. “Algoritam preciznije detekcije”).



Slika 5.20. Algoritam preciznije detekcije

Kada se ugradi predložena metoda s bilo kojim detektorom performanse se uvijek povećaju, a sami učinak se izračuna pomoću mAP-a.

5.3. Ostali relevantni radovi

Rad [50] daje komparativnu analizu šest dobro poznatih dubokih konvolucijskih arhitektura koje se koriste kao ekstraktori značajki za klasifikaciju podvodnih slika. To su VGG19, InceptionV3, ResNet50, Inception-ResNetV2, DenseNet121 i MobileNetV2. U radu se koriste duboki ekstraktor značajki CNN-a na tri različita načina: zamrzavaju se svi njegovi slojevi i treniraju se samo novododani NN (*eng. neural network*) klasifikatori, fino se podešavaju sve težine dubokog CNN-a i treniraju se zajedno s NN klasifikatorom. Prvo se zamrzavaju duboki CNN i obučava se NN klasifikator na vrhu, a nakon toga se odmrzavaju težine dubokog CNN-a kako bi se fino podesili. Glavni ciljevi ovog rada bili su: usporediti performanse istaknutih dubokih konvolucijskih arhitektura, razviti model za autonomnu identifikaciju morskog otpada, usporediti

izvedbu konvencionalnih klasifikatora strojnog učenja obučeni na vektorima značajki ekstrahiranih dubokim konvolucijskim arhitekturama i istražiti različite sheme za korištenje transfera učenja za klasifikaciju morskog otpada; fiksno izdvajanje značajki, fino podešavanje i kombinacija oba. Autori su koristili Adam optimizator sa stopama učenja $\beta_1 = 0.9$, $\beta_2 = 0.999$ i $e = 10^{-7}$. Svi modeli su trenirani u 100 epoha te su koristili mini serije veličine 16 jer zahtijevaju manje memorije za obradu. Baza podataka se sastojala od slika iz baze “Deep-sea Debris Database” japanske agencije JAMSTEC koja je kompletirana slikama iz “Google Images” baze te su koristili proces augmentacije podataka. Skup podataka je sadržavao 2395 slika podijeljenih u šest različitih klasa: metal, plastika, staklo, ostalo smeće, bez smeća i guma. Izgrađena je i jednostavna neuronska mreža kako bi se usporedila izvedba unaprijed obučeni arhitektura. Kada se treniraju s NN klasifikatorom na vrhu najbolje performanse pokazuju FT (*eng. fine tuning*) ekstraktori značajki temeljeni na Inception-u, točnije Inception-ResNetV2 i InceptionV3, postižući ukupnu točnost veću od 90%, 91,40% odnosno 90,57%. Tradicionalni SVM (*eng. support vector machine*) i LR (*eng. logistic regression*) klasifikatori pokazali su se kao vjerodostojne alternative NN klasifikatoru, koji često nadmašuju NN klasifikator. SVM obučen na značajkama Inception-ResNetV2 postiže točnost od 91,61%, dok LR klasifikator obučen na značajkama InceptionV3 postiže točnost od 90,78%. Dobiveni rezultati pokazuju potencijal za daljnje iskorištavanje modela temeljenih na dubokom učenju za identifikaciju i klasifikaciju morskog otpada u stvarnom vremenu u prirodnim vodenim okolišima, uzimajući u obzir izazove koji dolaze s automatskom klasifikacijom morskog otpada u podvodnim slikama.

Autori rada [51] predstavljaju, prvu takvu, bazu podataka “TrashCan” gdje se nalaze slike podvodnog smeća koje su skupljali iz raznih izvora, a ponajviše iz baze podataka “Deep-sea Debris dataset” japanske agencije za znanost i tehnologiju - “Japan agency for marine earth science and technology” (JAMSTEC). Objekti na slikama su označeni i segmentirani kako bi se što bolje istrenirali modeli za detekciju podvodnog smeća (Slika 5.21. “Prikaz rezultata detekcije i segmentacije na skupu podataka TrashCan”). Skup podataka ima dvije verzije “TrashCan - Material” i “TrashCan - Instance”. Predstavljani su i početni rezultati segmentacije koristeći “Mask R-CNN” i rezultati detekcije objekata koristeći “Faster R-CNN” koji ne prikazuju najbolje moguće rezultate, ali pružaju osnovu za rad u budućnosti.

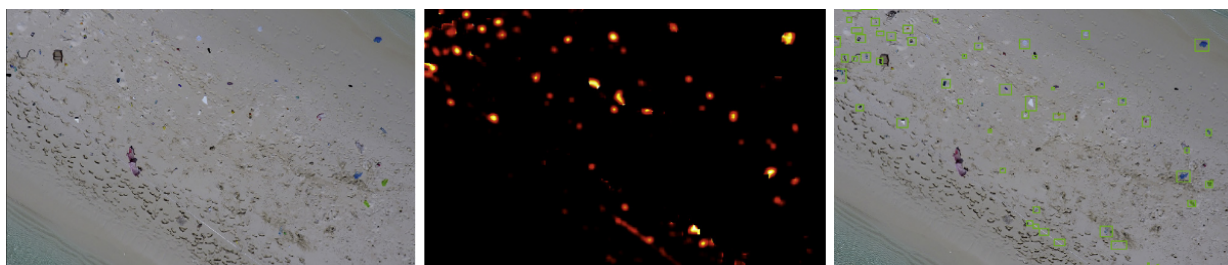


Slika 5.21. Prikaz rezultata detekcije i segmentacije na skupu podataka TrashCan

Autori rada [52] predstavljaju podatke o gustoći i sastavu antropogenog otpada prikupljenih na 44 plaže duž obale Crvenog mora s pomoću prepoznavanja iz daljine. Slike su snimane iz zraka, s visine od 10 metara, komercijalnim dronom “DJI Phantom”. Na snimkama se lako mogu detektirati objekti od nekoliko centimera jer fotografije imaju razlučivost manju od 0,5 cm piksela. Slike su zatim automatski pregledane algoritmom za detekciju objekata “Faster R-CNN”. Algoritam je binarnom klasifikacijom razvrstao objekte na slike sa smećem i slike bez smeća i kategorizirao je objekte u klase. Naišli su na probleme pri kategorizaciji po klasama stoga su izvedene samo na 15 plaža koje su pokazivale najbolje rezultate nakon binarnog razvrstavanja. Prikazani su rezultati gustoće smeća s 44 plaže, sastav smeća s 15 plaža i ukupna masa plastike s 15 plaža. Podaci o vrsti, brzini vjetra i vegetaciji plaže omogućavaju predviđanje putanje i lokacije smeća. Dostupnost takvih podataka (gustoća, sastav i prediktori) mogu pomoći pri izgradnji globalne procjene antropogenog otpada u priobalju.

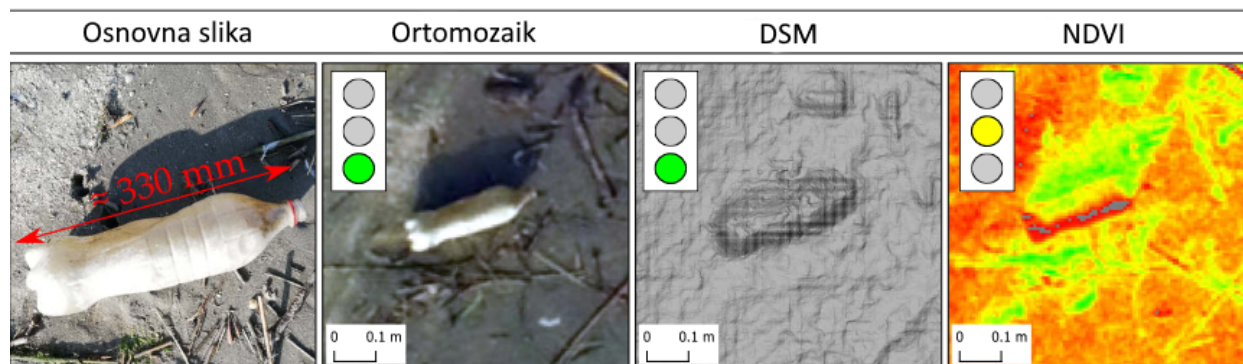
Autori rada [53] predstavljaju *ad hoc* metodologiju za praćenje i automatsko kvantificiranje antropogenog morskog otpada (AMD) čiju glavninu čini plastika, 80%, a jedan je od najvećih ekoloških problema današnjice. Mjesta na kojima se akumulira najviše AMD-a su priobalna područja, plaže. Autori koriste komercijalni bespilotni zrakoplov (UAV) koji je opremljen RGB kamerom visoke razlučivosti i primjenjuju model dubokog učenja kao što je “PlasticFinder” (Slika 5.22. “Prikaz testne slike, PlasticFinder toplinska mapa danih piksela, PlasticFinder mapa s omeđujućim okvirima”). Skupovi podataka za testiranje i treniranje sadržavali su slike visoke razlučivosti koje su omogućile zlatni standard za procjenu izvedbi softvera. S pozitivnom

estimacijom od 94% PlasticFinder dosegao je osjetljivost od 67% prilikom detektiranja AMD-a. Rezultati detekcije pri smanjenju Sunčevog svjetla su lošiji i to je nešto na čemu bi trebalo raditi pri predstavljanju novog modela. Autori su potvrdili učinkovitost korištenja komercijalnih UAV-ova za praćenje AMD-a i po prvi put pokazali potencijal dubokog učenja za automatsku detekciju i kvantificiranje AMD-a.



Slika 5.22. Prikaz testne slike, PlasticFinder toplinska mapa danih piksela, PlasticFinder mapa s omeđujućim okvirima

Rad [54] predstavlja kako locirati i pratiti antropogeni morski otpad (*eng. Anthropogenic Marine Debris; AMD*), poglavito makroplastiku kako bi se spriječila degradacija u mikroplastiku koja je potencijalno štetnija i teža za uklanjanje. Kao brzu metoda za dobivanje slika koriste UAV snimanje i istražuju strategije za mapiranje antropogenih morskih ostataka na plažama koristeći slike različite razlučivosti, informacije o elevaciji i podatke multispektralne kamere. Koristeći niz fotogrametrijskih alata AMD ručno su mapirani duž četiri do pet transekata gdje su prikupljali otpad s plaža i dalje ga obrađivali u laboratoriju gdje im dodjeljuju tri razine povjerenja i njihove klase. Podaci skupljeni s pomoću infracrvenih (*eng. Near Infrared; NIR*) valnih duljina i digitalnim modelima površine (*eng. Digital Surface Model; DSM*) nisu pokazali značajno poboljšanje, čak i kad je ukonjena vegetacija pomoću NDVI karata (*eng. Normalized Difference Vegetation Index*), pri ručnom mapiranju. Pokazali su kako je za mapiranje AMD-a s plaža dobro, brzo, održivo i praktično rješenje korištenje RGB slika s prostornom razlučivosti od oko 200 piksela po metru, pogotovo za makro objekte (Slika 5.23. “Usporedba slikanog objekta dronom su ostalim navedenim metodama”).



Slika 5.23. Usporedba slikanog objekta dronom s ostalim navedenim metodama

5.4. Baze podataka

Duboko učenje pokazalo se kao dobra metoda pri modeliranju nelinearnih složenih sustava. Ovoj metodi je potrebna velika baza podataka kako bi se mogla što bolje istrenirati, međutim, prikupljanje podataka u dubokomorskom okruženju nije lak zadatak [51]. Ono što predstavlja izazov je činjenica da podvodna domena nije toliko istraživana koliko terestrička te su metode vođene računalnim vidom ograničene i nailaze na poteškoće. Neke od njih su: degradacija objekata uslijed dugog izlaganja morskoj vodi, apsorpcija boje se povećava s dubinom, slike postaju izobličene i mutne zbog svjetlosti koja se raspršuje i lomi pod morem, namreškana površina mora uslijed puhanja vjetrova, obraštaj na objektima [4].

Kislik i suradnici su fotografirali područje rijeke s RGB dronovima na visini od 19 do 104 metra te su slikali ukupno 60 slika od kojih su izabrali 32 za obradu. Objekt promatranja bili su nitaste alge i makrofiti [39].

Politikos i suradnici koristili su bazu podataka „Integrirane kampanje informiranja i podizanja svijesti za smanjenje plastičnih vrećica u morskom okolišu” (eng. *LIFE DEBAG - LIFE14 GIE/GR/0011271*). Baza ima ukupno 635 slika, a slike su dobivene s pomoću podvodne kamere (eng. *towed underwater camera, TUC*) koja se nalazi na malom plovilu [40].

Marin i suradnici obradu podataka su provodili na JAMSTEC-ovoj bazi podataka koja je kompletirana slikama iz “Google Images” baze. Baza podataka je sadržavala 2395 slika.

Yabin i suradnici su koristili slike s “Google Eartha” i natjecateljske platforme “Kaggle” [55]. Obrađivali su slike čiji su promatrani objekti bili brodovi na površini mora prepoznati iz daljine s

pomoću satelita. Baza podataka se sastoji od 150 000 slika od kojih se na 42 556 nalazio traženi objekt [32].

Sveučilište Rijeka i suradnici obrađivali su slike brodova s površine mora preuzetih s Google tražilice slika i Google satelita gdje im se baza sastojala od 5608 slika [33].

Lin i suradnici koristili su bazu podataka “URPC 2018.” koja se sastoji od 2901 istreniranih i 800 testnih podvodnih slika prikupljenih autonomnim podvodnim vozilima (AUV). Promatrani objekti su morski krastavci, ježinci, školjke i morske zvijezde [41].

Li i suradnici generirali su svoju bazu podataka s pomoću “WaterGan”-a koja se sastojala od slika fotografiranih iz zraka na odgovarajućoj dubini, u boji, koje su realistične i sadrže floru i faunu [42].

Hong i suradnici predstavili su svoju bazu podataka “TrashCan” [56] koja je nastala prikupljanjem podataka iz različitih izvora poglavito od “Japanske agencije za znanost i tehnologiju (JAMSTEC)” [57] koja posjeduje bazu podataka s velikim brojem podvodnih videa prikupljenih autonomnim podvodnim vozilima (AUV). TrashCan skup podataka trenutno se sastoji od 7212 označenih slika koje sadrže označene objekte podvodnog smeća, faune, flore i vozila na daljinsko upravljanje (ROV). Autori su prikazali dvije verzije navedenog skupa podataka: “TrashCan - Material” i “TrashCan - Instance” [51].

U idućem radu autori koriste bazu podataka koja sadrži slike prikupljene autonomnim podvodnim vozilima (AUV) koje prikazuju podvodno smeće koje je označeno. Slike sadrže objekte raznih materijala, od plastike do metala, a snimane su u raznim podvodnim uvjetima [4].

Fabbi i suradnici su koristili bazu podataka “ImageNet” [58] i videa ronilaca preuzetih s kanala “YouTube” [43]. “ImageNet” je baza slikovnih podataka koji su organizirani prema “WordNet” hijerarhiji. Svaki pojam, skup sinonima tj. sinset (*eng. synset*), kojih ima više od 100000 u ovoj bazi, se nastoji ilustrirati s prosječno 1000 slika. Sve slike su kontrolirane kvalitete i označene. Autori trenutno mogu ponuditi korisnicima 3.2 milijuna slika, a želja im je ponuditi više desetaka milijuna slika koje će biti sortirane i označene za većinu koncepata u “WordNetu” [59].

Baza podataka Martina i suradnika [60] se sastojala od oko tisuću slika uzetih pomoću komercijalnog drona koji je slikavao duž obale [52].

Garcia-Garin i suradnici prikupljali su snimke iz zraka s pomoću bespilotnih letjelica i zrakoplova s posadom, fotografirajući plutajući otpad s površine mora. Baza se sastojala od 3723 slike [34].

Baza podataka “Enhancing underwater visual perception” (EUVP) [61] je skup neuparenih i uparenih slika loše i dobre kvalitete. Autori su stvorili takvu bazu koja olakšava treniranje modela za detekciju objekata s podvodnih slika. Upareni skup podataka pripremljen je postupkom koji su Fabbri i sur. predložili [43] te je poboljšana i ažurirana. Upareni skup sadrži 24840 slika dok neupareni skup podataka sadrži 6665 slika, od toga njih 3195 loše kvalitete, a 3140 dobre kvalitete. Objekti na slikama su flora, fauna i ronjaci.

Naseer i suradnici koristili su bazu podataka “Gulf of Cadiz” vlasništva Španjolskog oceanografskog instituta koja je sadržavala podvodni video te su sa više od 100000 okvira odabrali one koje sadrže podvodna staništa, jazbine vrste *Nephrops Norvegicus*, njih 200.

Primjećujemo da se navedene baze podataka razlikuju po mnogočemu:

- područje uzorkovanja: - površina mora
 - vodeni stupac
 - morsko dno
 - plaža

- način uzorkovanja: - satelitom
 - autonomnim podvodnim vozilom (AUV)
 - vozilom na daljinsko upravljanje (ROV)
 - bespilotnom letjelicom- dronom (UAV)
 - zrakoplovom s posadom

- objekt promatranja: - plovila
 - podmorna flora i fauna
 - morski otpad

JAMSTEC [57] je osnovna baza podataka koja sadrži podvodne slike tj. videa sa morskim otpadom kao objektom promatranja, ali je trenutno nedostupna javnosti zbog incidenta sigurnosti podataka.

TrashCan [56], kao baza koja je potekla od JAMSTEC baze, također možemo navesti kao bazu podvodnih slika gdje je objekt promatranja morski otpad. TrashCan baza je javna i može se koristiti za akademsku i istraživačku upotrebu, ali za komercijalnu upotrebu treba dobiti dopuštenje od JAMSTEC-a. Ono što je još bitno je činjenica da su podaci prikupljeni autonomnim podvodnim vozilom (AUV).

Nadalje, ne postoji ni jedna baza podataka sa slikama podmorja koja je fotografirana bespilotnom letjelicom - dronom, a objekt promatranja je bio morski otpad.

6. ZAKLJUČAK

U svim morima i oceanima pronalazimo morski otpad koji negativno utječe na ekološke, ekonomske, zdravstvene i estetske prilike. Napredak umjetne inteligencije pogotovo konvolucijskih neuronskih mreža uvelike je olakšao istraživanje morskih ekosustava. Ono što i dalje predstavlja problem su fizička i kemijska svojstva mora koja otežavaju način kako prikupiti jasne podatke i slike promatranog područja, uvjetuju na koji način će se slike prikupljati i koje metode će se koristiti u daljnjoj obradi podataka. Postoje različite baze podataka koje se koriste za detekciju objekata pod morem ili na površini mora, ali ne postoji dostupna baza podataka traženih slika gdje su objekti u moru, a fotografirane su bespilotnom letjelicom. Specijalizirani algoritmi i mreže prilagođene rješavanju problema detekcije objekata ispod mora na slikama bespilotne letjelice nisu istražene kao detekcija objekata na površini mora koristeći letjelice ili pod morem koristeći podvodna vozila na daljinsko upravljanje. Nadalje, detekcija objekata na površini mora postiže dobre rezultate zbog manjeg utjecaja vanjskih faktora. Detekcija objekata pod morem nailazi na poteškoće zbog fizikalnih svojstava morske površine, vremenskih uvjeta, nečistoće optičkog medija, bioloških svojstava mora te utjecaja prirode na okoliš pa zbog toga područje nije još dovoljno obrađeno. Pregledom literature može se zaključiti da postoji prostor i potreba za istraživanjem detekcije i klasifikacije objekata pod morem. Iako istraživanja koja su koristila bespilotne letjelice za nadzor i detekciju objekata u plićim dijelovima mora ili rijeka postoje, unapređenje i prijedlog novih rješenja u ovom području čini se obećavajuće.

7. LITERATURA

- [1] L. Jeftic, S. B. Sheavly, E. Adler, and N. Meith, *Marine litter: a global challenge*. Nairobi, Kenya: Regional Seas, United Nations Environment Programme, 2009.
- [2] D. Lithner, Å. Larsson, and G. Dave, “Environmental and health hazard ranking and assessment of plastic polymers based on chemical composition,” *Sci. Total Environ.*, vol. 409, no. 18, pp. 3309–3324, Aug. 2011, doi: 10.1016/j.scitotenv.2011.04.038.
- [3] J. A. Ivar do Sul and M. F. Costa, “The present and future of microplastic pollution in the marine environment,” *Environ. Pollut.*, vol. 185, pp. 352–364, Feb. 2014, doi: 10.1016/j.envpol.2013.10.036.
- [4] J. Hong, M. Fulton, and J. Sattar, “A Generative Approach Towards Improved Robotic Detection of Marine Litter,” in *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France, May 2020, pp. 10525–10531. doi: 10.1109/ICRA40945.2020.9197575.
- [5] European Commission. Joint Research Centre. Institute for Environment and Sustainability. and MSFD Technical Subgroup on Marine Litter., *Guidance on monitoring of marine litter in European seas*. LU: Publications Office, 2013. Accessed: Nov. 02, 2021. [Online]. Available: <https://data.europa.eu/doi/10.2788/99475>
- [6] S. Lippiatt, S. Opfer, and C. Arthur, “Marine Debris Monitoring and Assessment: Recommendations for Monitoring Debris Trends in the Marine Environment.,” NOAA Marine Debris Division, Report, 2013. doi: 10.25607/OBP-727.
- [7] “One Million Less Pieces of Trash in our Ocean | PADI AWARE.” <https://www.diveagainstdebris.org/news/one-million-less-pieces-trash-our-ocean> (accessed Nov. 02, 2021).
- [8] V. Sze, Y.-H. Chen, T.-J. Yang, and J. S. Emer, “Efficient Processing of Deep Neural Networks: A Tutorial and Survey,” *Proc. IEEE*, vol. 105, no. 12, pp. 2295–2329, Dec. 2017, doi: 10.1109/JPROC.2017.2761740.
- [9] S. Hijazi, R. Kumar, and C. Rowen, “Using Convolutional Neural Networks for Image Recognition By,” *undefined*, 2015, Accessed: Apr. 13, 2022. [Online]. Available: <https://www.semanticscholar.org/paper/Using-Convolutional-Neural-Networks-for-Image-By-Hijazi-Kumar/bbf7b5bdc39f9b8849c639c11f4726e36915a0da>
- [10] A. Krizhevsky, I. Sutskever, and G. E. Hinton, “ImageNet Classification with Deep Convolutional Neural Networks,” in *Advances in Neural Information Processing Systems*, 2012, vol. 25. Accessed: Jun. 15, 2022. [Online]. Available: <https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>
- [11] K. Simonyan and A. Zisserman, “Very Deep Convolutional Networks for Large-Scale Image Recognition.” arXiv, Apr. 10, 2015. Accessed: Jun. 15, 2022. [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [12] K. He, X. Zhang, S. Ren, and J. Sun, “Deep Residual Learning for Image Recognition.” arXiv, Dec. 10, 2015. Accessed: Jun. 15, 2022. [Online]. Available: <http://arxiv.org/abs/1512.03385>
- [13] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, “Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning.” arXiv, Aug. 23, 2016. Accessed: Jun. 15, 2022. [Online]. Available: <http://arxiv.org/abs/1602.07261>
- [14] R. Girshick, J. Donahue, T. Darrell, and J. Malik, “Rich feature hierarchies for accurate object detection and semantic segmentation.” arXiv, Oct. 22, 2014. Accessed: Jun. 14,

2022. [Online]. Available: <http://arxiv.org/abs/1311.2524>
- [15] R. Girshick, “Fast R-CNN.” arXiv, Sep. 27, 2015. Accessed: Jun. 14, 2022. [Online]. Available: <http://arxiv.org/abs/1504.08083>
- [16] S. Ren, K. He, R. Girshick, and J. Sun, “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.” arXiv, Jan. 06, 2016. Accessed: Jun. 14, 2022. [Online]. Available: <http://arxiv.org/abs/1506.01497>
- [17] W. Liu *et al.*, “SSD: Single Shot MultiBox Detector,” vol. 9905, 2016, pp. 21–37. doi: 10.1007/978-3-319-46448-0_2.
- [18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection.” arXiv, May 09, 2016. Accessed: Jun. 14, 2022. [Online]. Available: <http://arxiv.org/abs/1506.02640>
- [19] M. E. AtiK, Z. Duran, and R. Özgünlük, “Comparison of YOLO Versions for Object Detection from Aerial Images,” *Int. J. Environ. Geoinformatics*, vol. 9, no. 2, pp. 87–93, Jun. 2022, doi: 10.30897/ijegeo.1010741.
- [20] J. Redmon and A. Farhadi, “YOLO9000: Better, Faster, Stronger,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, Jul. 2017, pp. 6517–6525. doi: 10.1109/CVPR.2017.690.
- [21] J. Hui, “Real-time Object Detection with YOLO, YOLOv2 and now YOLOv3,” *Medium*, Aug. 27, 2019. <https://jonathan-hui.medium.com/real-time-object-detection-with-yolo-yolov2-28b1b93e2088> (accessed Jul. 18, 2022).
- [22] “Yolov3 and Yolov4 in Object Detection.” <https://ai-pool.com/a/s/yolov3-and-yolov4-in-object-detection> (accessed Jul. 18, 2022).
- [23] “Introduction to the YOLO Family,” *PyImageSearch*, Apr. 04, 2022. <https://pyimagesearch.com/2022/04/04/introduction-to-the-yolo-family/> (accessed Jul. 07, 2022).
- [24] C. Supeshala, “YOLO v4 or YOLO v5 or PP-YOLO?,” *Medium*, Aug. 23, 2020. <https://towardsdatascience.com/yolo-v4-or-yolo-v5-or-pp-yolo-dad8e40f7109> (accessed Jul. 12, 2022).
- [25] “ultralytics/yolov5.” Ultralytics, Jul. 12, 2022. Accessed: Jul. 12, 2022. [Online]. Available: <https://github.com/ultralytics/yolov5>
- [26] “Train Custom Data · ultralytics/yolov5 Wiki,” *GitHub*. <https://github.com/ultralytics/yolov5> (accessed Jul. 12, 2022).
- [27] “Google Colaboratory.” <https://colab.research.google.com/drive/1gDZ2xcTOgR39tGGs-EZ6i3RTs16wmzZQ> (accessed Jul. 12, 2022).
- [28] I. I. PhD, “YOLOv4 vs YOLOv5,” *Deelvin Machine Learning*, Jun. 30, 2020. <https://medium.com/deelvin-machine-learning/yolov4-vs-yolov5-db1e0ac7962b> (accessed Jul. 18, 2022).
- [29] GISGeography, “What is Remote Sensing? The Definitive Guide,” *GIS Geography*, Nov. 02, 2016. <https://gisgeography.com/remote-sensing-earth-observation-guide/> (accessed Mar. 15, 2022).
- [30] “What is remote sensing and what is it used for? | U.S. Geological Survey.” <https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used> (accessed Mar. 11, 2022).
- [31] S. Merlino, M. Paterni, M. Locritani, U. Andriolo, G. Gonçalves, and L. Massetti, “Citizen Science for Marine Litter Detection and Classification on Unmanned Aerial Vehicle

- Images,” *Water*, vol. 13, no. 23, p. 3349, Nov. 2021, doi: 10.3390/w13233349.
- [32] L. Yabin, Y. Jun, and H. Zhiyi, “Improved Faster R-CNN Algorithm for Sea Object Detection Under Complex Sea Conditions,” *Int. J. Adv. Netw. Monit. Controls*, vol. 5, no. 2, pp. 76–82, 2020, doi: 10.21307/ijanmc-2020-020.
- [33] I. Lorencin, N. Anđelić, V. Mrzljak, and Z. Car, “Marine Objects Recognition Using Convolutional Neural Networks,” *Naše More*, vol. 66, no. 3, pp. 112–120, Oct. 2019, doi: 10.17818/NM/2019/3.3.
- [34] O. Garcia-Garin *et al.*, “Automatic detection and quantification of floating marine macro-litter in aerial images: Introducing a novel deep learning approach connected to a web application in R,” *Environ. Pollut.*, vol. 273, p. 116490, Mar. 2021, doi: 10.1016/j.envpol.2021.116490.
- [35] “MARLITE.” Feb. 19, 2022. Accessed: Mar. 22, 2022. [Online]. Available: <https://github.com/amonleong/MARLIT>
- [36] G. Jakovljevic, M. Govedarica, and F. Alvarez-Taboada, “A Deep Learning Model for Automatic Plastic Mapping Using Unmanned Aerial Vehicle (UAV) Data,” *Remote Sens.*, vol. 12, no. 9, p. 1515, May 2020, doi: 10.3390/rs12091515.
- [37] K. Topouzelis, A. Papakonstantinou, and S. P. Garaba, “Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018),” *Int. J. Appl. Earth Obs. Geoinformation*, vol. 79, pp. 175–183, Jul. 2019, doi: 10.1016/j.jag.2019.03.011.
- [38] J. Seo, W. Park, and T. Kim, “Feature-Based Approach to Change Detection of Small Objects from High-Resolution Satellite Images,” *Remote Sens.*, vol. 14, no. 3, p. 462, Jan. 2022, doi: 10.3390/rs14030462.
- [39] C. Kislik, L. Genzoli, A. Lyons, and M. Kelly, “Application of UAV Imagery to Detect and Quantify Submerged Filamentous Algae and Rooted Macrophytes in a Non-Wadeable River,” *Remote Sens.*, vol. 12, no. 20, p. 3332, Oct. 2020, doi: 10.3390/rs12203332.
- [40] D. V. Politikos, E. Fakiris, A. Davvetas, I. A. Klampanos, and G. Papatheodorou, “Automatic detection of seafloor marine litter using towed camera images and deep learning,” *Mar. Pollut. Bull.*, vol. 164, p. 111974, Mar. 2021, doi: 10.1016/j.marpolbul.2021.111974.
- [41] W.-H. Lin, J.-X. Zhong, S. Liu, T. Li, and G. Li, “ROIMIX: Proposal-Fusion Among Multiple Images for Underwater Object Detection,” in *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Barcelona, Spain, May 2020, pp. 2588–2592. doi: 10.1109/ICASSP40776.2020.9053829.
- [42] J. Li, K. A. Skinner, R. M. Eustice, and M. Johnson-Roberson, “WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images,” *IEEE Robot. Autom. Lett.*, pp. 1–1, 2017, doi: 10.1109/LRA.2017.2730363.
- [43] C. Fabbri, M. J. Islam, and J. Sattar, “Enhancing Underwater Imagery Using Generative Adversarial Networks,” in *2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, QLD, May 2018, pp. 7159–7165. doi: 10.1109/ICRA.2018.8460552.
- [44] A. Naseer, E. N. Baro, S. D. Khan, and Y. Vila, “A Novel Detection Refinement Technique for Accurate Identification of *Nephtrops norvegicus* Burrows in Underwater Imagery,” *Sensors*, vol. 22, no. 12, Art. no. 12, Jan. 2022, doi: 10.3390/s22124441.
- [45] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the Inception Architecture for Computer Vision,” in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 2818–2826. doi:

- 10.1109/CVPR.2016.308.
- [46] P. Dwivedi, "Understanding and Coding a ResNet in Keras," *Medium*, Mar. 27, 2019. <https://towardsdatascience.com/understanding-and-coding-a-resnet-in-keras-446d7ff84d33> (accessed Jun. 30, 2022).
- [47] "tf.keras.applications.resnet.ResNet101 | TensorFlow Core v2.9.1," *TensorFlow*. https://www.tensorflow.org/api_docs/python/tf/keras/applications/resnet/ResNet101 (accessed Jun. 30, 2022).
- [48] S. J. Pan and Q. Yang, "A Survey on Transfer Learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.
- [49] P. Barham *et al.*, "TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems." arXiv, Mar. 16, 2016. Accessed: Jun. 30, 2022. [Online]. Available: <http://arxiv.org/abs/1603.04467>
- [50] I. Marin, S. Mladenović, S. Gotovac, and G. Zaharija, "Deep-Feature-Based Approach to Marine Debris Classification," *Appl. Sci.*, vol. 11, no. 12, p. 5644, Jun. 2021, doi: 10.3390/app11125644.
- [51] J. Hong, M. Fulton, and J. Sattar, "TrashCan: A Semantically-Segmented Dataset towards Visual Detection of Marine Debris," *ArXiv200708097 Cs*, Jul. 2020, Accessed: Oct. 29, 2021. [Online]. Available: <http://arxiv.org/abs/2007.08097>
- [52] C. Martin, Q. Zhang, D. Zhai, X. Zhang, and C. M. Duarte, "Anthropogenic litter density and composition data acquired flying commercial drones on sandy beaches along the Saudi Arabian Red Sea," *Data Brief*, vol. 36, p. 107056, Jun. 2021, doi: 10.1016/j.dib.2021.107056.
- [53] L. Fallati, A. Polidori, C. Salvatore, L. Saponari, A. Savini, and P. Galli, "Anthropogenic Marine Debris assessment with Unmanned Aerial Vehicle imagery and deep learning: A case study along the beaches of the Republic of Maldives," *Sci. Total Environ.*, vol. 693, p. 133581, Nov. 2019, doi: 10.1016/j.scitotenv.2019.133581.
- [54] Y. Taddia, C. Corbau, J. Buoninsegni, U. Simeoni, and A. Pellegrinelli, "Uav Approach for Detecting Plastic Marine Debris on the Beach: A Case Study in the Po River Delta (Italy)," *Drones*, vol. 5, no. 4, p. 140, Nov. 2021, doi: 10.3390/drones5040140.
- [55] "Ships in Satellite Imagery." <https://www.kaggle.com/rhammell/ships-in-satellite-imagery> (accessed Mar. 31, 2022).
- [56] J. Hong, M. S. Fulton, and J. Sattar, "TrashCan 1.0 An Instance-Segmentation Labeled Dataset of Trash Observations." Jul. 23, 2020. doi: 10.13020/g1gx-y834.
- [57] T. Sasaki *et al.*, "JAMSTEC E-library of Deep-sea Images (J-EDI) Realizes a Virtual Journey to the Earth's Unexplored Deep Ocean," vol. 2016, pp. IN53C-1911, Dec. 2016.
- [58] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei, "ImageNet: A Large-Scale Hierarchical Image Database," p. 8.
- [59] "ImageNet." <https://www.image-net.org/index.php> (accessed Mar. 31, 2022).
- [60] C. Martin, Q. Zhang, D. Zhai, X. Zhang, and C. M. Duarte, "Drone images of sandy beaches and anthropogenic litter along the Saudi Arabian Red Sea," vol. 1, Feb. 2021, doi: 10.17632/gpdsntb3y6.1.
- [61] "The EUVP dataset | Interactive Robotics and Vision Lab." <http://irvlab.cs.umn.edu/resources/euvs-dataset> (accessed Mar. 11, 2022).

PRILOG - TABLICA SLIKA

<i>Slika 2.1. Prikaz deset najčešćih stavki pronađenih u oceanu</i>	3
<i>Slika 3.1. Domena umjetne inteligencije</i>	4
<i>Slika 3.2. Prosljeđujući i povratni sloj</i>	5
<i>Slika 3.3. Tipični dijagram CNN bloka</i>	6
<i>Slika 3.4. Prikaz procesa konvolucije</i>	7
<i>Slika 3.5. Prikaz maksimalnog i prosječnog udruživanja</i>	8
<i>Slika 3.6. Prikaz potpuno povezanog sloja</i>	9
<i>Slika 3.7. AlexNet blok dijagram</i>	11
<i>Slika 3.8. Primjer ulaza, obrade i izlaza VGGNet mreže</i>	12
<i>Slika 3.9. Rezidualni blok ResNet mreže</i>	13
<i>Slika 3.10. Početni modul Inception mreže</i>	14
<i>Slika 3.11. Arhitektura R-CNN</i>	15
<i>Slika 3.12. Arhitektura Fast R-CNN</i>	16
<i>Slika 3.13. Algoritam za detekciju objekata Faster R-CNN</i>	17
<i>Slika 3.14. Algoritam za detekciju objekata YOLO</i>	18
<i>Slika 4.1. Prikaz dionika procesa prepoznavanja iz daljine</i>	22
<i>Slika 5.1. Prikaz situacije izostavljanja</i>	23
<i>Slika 5.2. Prikaz rezultata s malim objektima</i>	23
<i>Slika 5.3. Dijagram tijeka podataka algoritma za prepoznavanje morskih objekata</i>	24
<i>Slika 5.4. Slika iz zraka s nizom poznatih predmeta pričvršćenih na uže (sanduk, boce, limenke, kese, tetrapak, ručnik, balon, karapaks kornjače...)</i>	25
<i>Slika 5.5. Rezultati točnosti za različite implementacije mreža</i>	27
<i>Slika 5.6. Rezultati točnosti za različite rezolucije slika</i>	28
<i>Slika 5.7. Kontrolna slika</i>	31
<i>Slika 5.8. Primijenjena metoda RoIMix</i>	31
<i>Slika 5.9. Neobrađena podvodna slika</i>	32
<i>Slika 5.10. WaterGAN metoda</i>	32
<i>Slika 5.11. Neobrađena podvodna slika</i>	32
<i>Slika 5.12. WaterGAN metoda</i>	32
<i>Slika 5.13. Prva faza</i>	33
<i>Slika 5.14. Druga faza</i>	33
<i>Slika 5.15. Rekonstrukcija</i>	33
<i>Slika 5.16. Original</i>	34
<i>Slika 5.17. CycleGAN</i>	34
<i>Slika 5.18. Original</i>	34
<i>Slika 5.19. CycleGAN</i>	34
<i>Slika 5.20. Algoritam preciznije detekcije</i>	36
<i>Slika 5.21. Prikaz rezultata detekcije i segmentacije na skupu podataka TrashCan</i>	38
<i>Slika 5.22. Prikaz testne slike, PlasticFinder toplinska mapa danih piksela, PlasticFinder mapa s omeđujućim okvirima</i>	39
<i>Slika 5.23. Usporedba slikanog objekta dronom s ostalim navedenim metodama</i>	40