

SVEUČILIŠTE U SPLITU
FAKULTET ELEKTROTEHNIKE, STROJARSTVA I BRODOGRADNJE

Ivan Biliškov

**DETEKCIJA OBJEKATA POD MOREM ANALIZOM
FOTOGRAFIJA S BESPILOTNE LETJELICE**

DOKTORSKA DISERTACIJA

Split, 2025.

SVEUČILIŠTE U SPLITU
FAKULTET ELEKTROTEHNIKE, STROJARSTVA I BRODOGRADNJE

Ivan Biliškov

*Detekcija objekata pod morem analizom
fotografija s bespilotne letjelice*

DOKTORSKA DISERTACIJA

Split 2025.

Doktorska disertacija je izrađena na Zavodu za elektrotehniku i računarstvo, Fakulteta elektrotehnike, strojarstva i brodogradnje.

Mentor: Prof. dr. sc. Vladan Papić

Rad br. _____

PODATCI ZA BIBLIOGRAFSKU KARTICU

Ključne riječi: morski otpad, detekcija podvodnih objekata, pretprocesiranje slika, konvolucijske neuronske mreže, YOLO modifikacija, četverokanalni ulaz, RoWS metoda

Znanstveno područje: Tehničke znanosti

Znanstveno polje: Računarstvo

Znanstvena grana: Umjetna inteligencija

Institucija na kojoj je rad izrađen: Sveučilište u Splitu, Fakultet elektrotehnike, strojarstva i brodogradnje

Mentor rada: Prof. dr. sc. Vladan Papić

Broj stranica: 133

Broj slika: 41

Broj tablica: 6

Broj korištenih bibliografskih jedinica: 202

Povjerenstvo za ocjenu doktorske disertacije:

Osoba 1, Naziv ustanove i grad

Osoba 2, Naziv ustanove i grad

Osoba 3, Naziv ustanove i grad

Povjerenstvo za obranu doktorske disertacije:

Osoba 1, Naziv ustanove i grad

Osoba 2, Naziv ustanove i grad

Osoba 3, Naziv ustanove i grad

Osoba 4, Naziv ustanove i grad

Osoba 5, Naziv ustanove i grad

Disertacije obranjena dana: 2025.

DETEKCIJA OBJEKATA POD MOREM ANALIZOM FOTOGRAFIJA S BESPILOTNE LETJELICE

Sažetak

Morski otpad predstavlja značajnu globalnu prijetnju morskim ekosustavima, prvenstveno uzrokovanu lošim upravljanjem otpadom, neodgovarajućom infrastrukturom i neodgovornim ljudskim aktivnostima. Cilj ovo rada je razvoj sustava za automatsku detekciju objekata na morskom dnu korištenjem fotografija snimljenih bespilotnom letjelicom. U uvodu se ističe problem akumulacije morskog otpada i potreba za neinvazivnim metodama njegove identifikacije. Rad obuhvaća sveobuhvatan pregled literature iz područja računalnog vida – od klasifikatora i detektora temeljenih na konvolucijskim neuronskim mrežama, preko transformera i tehnika daljinskog istraživanja, do specifičnih metoda poboljšanja podvodnih slika (restauracija, prostorna i frekvencijska poboljšanja, modeli zasnovani na boji i dubokom učenju).

U metodologiji je opisan proces prikupljanja vlastitog skupa fotografija visoke rezolucije na Jadranu kroz sve sezone i varijabilne uvjete (visine snimanja 5–15 m, brzina vjetra, položaj sunca) te izrada baze zakrpa od 640×640 piksela s različitim klasama morskog otpada. Za fotografije prikupljene bespilotnom letjelicom su implementirane metode poboljšanja slika koje se koriste na podvodnim slikama. Za pretprocesiranje fotografija su odabrane i primijenjene različite metode, a potom su slike obogaćene dodatnim, četvrtim kanalom (nijansa i procijenjena dubina). YOLOv8 arhitektura je modificirana za četverokanalni ulaz uz prilagodbu težina, nakon čega su modeli trenirani prijenosnim učenjem na prilagođenim skupovima (*OpenImagesV7* i *Selected Patches*). Evaluacija na testnom dijelu skupa koristila je mjere preciznosti, odaziva, F1 i mAP@50. Rezultati pokazuju da kombinacija pretprocesiranja i dodatnih kanala poboljšava detekciju podvodnih objekata u usporedbi s pristupima koji koriste standardne trokanalne arhitekture neuronskih mreža bez pretprocesiranja.

Ključne riječi:

morski otpad, detekcija podvodnih objekata, pretprocesiranje slika, konvolucijske neuronske mreže, YOLO modifikacija, četverokanalni ulaz, RoWS metoda

DETECTION OF OBJECTS UNDER THE SEA BY ANALYSING IMAGES FROM AN UNMANNED AERIAL VEHICLE

Abstract

Marine litter poses a significant global threat to marine ecosystems, primarily caused by poor waste management, inadequate infrastructure, and irresponsible human activities. The aim of this study is development of a system for automatic object detection on the seafloor using images captured by an unmanned aerial vehicle (UAV). The introduction highlights the issue of marine debris accumulation and the need for non-invasive methods of identification. The paper provides a comprehensive literature review in the field of computer vision – covering classifiers and detectors based on convolutional neural networks, transformers, remote sensing techniques, and specific underwater image enhancement methods (including restoration, spatial and frequency-domain enhancements, and approaches based on color and deep learning).

The methodology describes the process of collecting a proprietary dataset of high-resolution images captured over the Adriatic Sea across all seasons and varying conditions (flight altitudes of 5–15 meters, wind speed, sun position), followed by the creation of a patch database with 640×640 pixel images labeled into various marine litter classes. For drone-acquired images, enhancement techniques commonly used in underwater image processing were implemented. Various preprocessing methods were selected and applied to the images, which were subsequently enriched with a fourth channel (hue and estimated depth). The YOLOv8 architecture was modified to accept four-channel input, including weight adaptation, after which the models were trained using transfer learning on custom datasets (*OpenImagesV7* and *Selected Patches*). Evaluation on the test set was performed using precision, recall, F1, and mAP@50 metrics. The results demonstrate that the combination of preprocessing and additional input channels improves underwater object detection compared to standard three-channel neural network architectures without preprocessing.

Keywords:

marine litter, underwater object detection, image preprocessing, convolutional neural networks, YOLO modification, four-channel input, RoWS method

Zahvala

Najveću i iskrenu zahvalu upućujem svom mentoru, prof. dr. sc. Vladanu Papiću, čija je pomoć i podrška tijekom mog doktorskog studija kontinuirana i nemjerljiva, koji je svoje znanje, znanstveno i stručno iskustvo nesebično dijelio.

Dugujem zahvalu i velikom čovjeku i profesoru koji nas je prerano napustio, ali nije zaboravljen, izv. prof. dr. sc. Hrvoju Dujmiću. Počivao u miru.

Veliko hvala i mojim roditeljima, Tončici i Jozi, koji su me uvijek i u svemu podržavali i pružali mi najbolji primjer vrijednosti truda i rada.

Hvala Antoniji na vremenu, savjetima te neprocjenjivoj podršci tijekom pisanja.

Hvala i svima ostalima koji su na bilo koji način bili dio ovog puta.

Posebna zahvala mojoj supruzi Ani na razumijevanju, strpljenju i vjeri u moj uspjeh.

Sadržaj

Sažetak	iv
Abstract	v
Zahvala.....	vii
Sadržaj.....	viii
Popis tablica	xi
Popis slika	xii
Popis oznaka	xiv
1. Uvod.....	1
1.1. Motivacija	3
1.2. Hipoteza	3
1.3. Struktura rada.....	5
2. Opis problema	6
3. Pregled područja	9
3.1. Klasifikacija slika.....	9
3.1.1. Klasifikatori temeljeni na konvolucijskim neuronskim mrežama (CNN)	9
3.1.2. Klasifikatori temeljeni na transformer arhitekturi	10
3.2. Detekcija objekata na slici	10
3.2.1. Detektori koji se sastoje od dvije faze	10
3.2.2. Detektori koji se sastoje od jedne faze.....	11
3.2.3. Detektori temeljeni na transformer arhitekturi	11
3.3. Usporedba CNN i transformer modela	12
3.4. Daljinsko istraživanje.....	12
3.5. Postojeće metode primijenjene na detekciji objekata	13
3.5.1. Radovi u kojima se objekt promatranja nalazi na površini mora.....	13
3.5.2. Radovi u kojima se objekt promatranja nalazi pod morem	19

3.5.3.	Ostali relevantni radovi.....	25
3.6.	Baze podataka	29
3.6.1.	Baze podataka za poboljšanje i obnavljanje podvodnih fotografija	32
3.7.	Ograničenja podvodnog fotografiranja	36
3.7.1.	Uzroci degradacije kvalitete	36
3.7.2.	Posljedice degradacija kvalitete	38
3.8.	Pretprocesiranje fotografija.....	41
3.8.1.	Tehnike poboljšanja kvalitete podvodnih fotografija	41
3.8.2.	Tehnike poboljšanja podvodnih fotografija	49
3.8.3.	Spoj metoda restauriranja i poboljšanja	56
3.8.4.	Relevantne metode pretprocesiranja u kontekstu cilja istraživanja	57
3.9.	Prilagodba arhitekture duboke neuronske mreže	61
3.9.1.	Hue kao četvrti kanal	61
3.9.2.	Dubina kao četvrti kanal	63
4.	Opis i metodologija istraživanja	65
4.1.	Prikupljanje fotografija visoke rezolucije.....	66
4.2.	Kreiranje baze podataka.....	72
4.2.1.	Analiza nove baza podataka.....	75
4.3.	Pretprocesiranje fotografija.....	78
4.3.1.	Primjena pretprocesiranja koristeći duboke neuronske mreža.....	78
4.3.2.	Analiza rezultata pretprocesiranja.....	79
4.4.	Prilagodba arhitekture konvolucijske mreže.....	84
4.4.1.	Uvođenje četvrtog kanala	84
4.4.2.	Prilagodba YOLOv8 arhitekture za četverokanalne fotografije	86
4.4.3.	Integracija metode pretprocesiranja i prilagođene arhitekture.....	87
4.4.4.	Postupak treniranja modela.....	89
4.4.5.	Metodologija evaluacije modela	89

4.5.	Analiza rezultata	89
4.6.	Predloženi postupak za detekciju objekata	92
5.	Zaključak.....	93
5.1.	Doprinos disertacije	94
5.2.	Smjernice za daljnji rad	94
	Literatura.....	96
A	PRILOG A Modifikacije YOLOv8 za podršku četverokanalnih slika.....	118
A.1.	ultralytics.nn.tasks.....	118
A.1.1.	Promjena zadanog broja kanala u DetectionModel	118
A.1.2.	Dodavanje funkcije adapt_weights_to_4ch.....	119
A.1.3.	Prilagodba attempt_load_one_weight.....	120
A.2.	ultralytics.nn.autobackend	121
A.3.	ultralytics.engine.validator.....	122
A.4.	ultralytics.engine.predictor.....	123
A.4.1.	Pretprocesiranje četverokanalnih slika	123
A.4.2.	Zagrijavanje modela u <i>predict</i>	125
A.4.3.	Ažuriranje poruke o brzini	125
A.5.	ultralytics.data.loaders	126
A.6.	ultralytics.data.augment	126
A.6.1.	Mosaic Augmentation.....	127
A.6.2.	RandomPerspective.....	128
A.6.3.	RandomHSV	128
A.6.4.	LetterBox	131
A.7.	ultralytics.utils.plotting	132

Popis tablica

Tablica 3.1 Usporedba postojećih baza podataka i njihovih značajki	31
Tablica 4.1 Lokacija po gradovima i GPS koordinatama	67
Tablica 4.2 Rezultati metoda pretprocesiranja.....	81
Tablica 4.3 Mjerila izvedbe prepoznavanja objekata korištenjem RoWS metoda u različitim klasama objekata	83
Tablica 4.4 Rezultati performansi treniranih YOLOv8n modela na neviđenom testnom skupu podataka	90
Tablica 4.5 Usporedba rezultata performansi treniranih modela na neviđenom testnom skupu podataka	92

Popis slika

Slika 2.1 Izazovi pri detekciji objekata pod morem	6
Slika 3.1 Prikaz situacije djelomične detekcije.....	13
Slika 3.2 Prikaz rezultata detekcije malih objekata na morskoj površini	13
Slika 3.3 Dijagram tijeka podataka algoritma za prepoznavanje morskih objekata	14
Slika 3.4 Fotografija iz zraka s nizom poznatih predmeta pričvršćenih na uža (klase: sanduk, boce, limenke, kese, tetrapak, ručnik, balon, karapaks kornjače...).....	15
Slika 3.5 Rezultati točnosti za različite implementacije mreža [43].....	17
Slika 3.6 Rezultati točnosti za različite prostorne rezolucije fotografija [43]	17
Slika 3.7 Usporedba kontrolne fotografije i primijenjene metode RoIMix	21
Slika 3.8 Usporedba neobrađene podvodne fotografije i primjena WaterGAN metode	21
Slika 3.9 Neobrađena podvodna fotografija i primjena WaterGAN metode.....	22
Slika 3.10 Prva faza generiranja slika koristeći varijacijski autoenkoder	22
Slika 3.11 Druga faza generiranja slika koristeći varijacijski autoenkoder.....	22
Slika 3.12 Uzorak generiranih slika koristeći varijacijski autoenkoder	23
Slika 3.13 Usporedba original i CycleGAN fotografija živih bića u radu [56]	23
Slika 3.14 Usporedba original i CycleGAN fotografija ronjaca u radu [56].....	23
Slika 3.15 Algoritam preciznije detekcije iz rada [57]	25
Slika 3.16 Prikaz rezultata detekcije i segmentacije na skupu podataka TrashCan.....	27
Slika 3.17 Prikaz testne fotografije, PlasticFinder toplinska mapa danih piksela, PlasticFinder mapa s omeđujućim okvirima.....	28
Slika 3.18 Usporedba objekta fotografiranog dronom s ostalim metodama u radu [64]	28
Slika 4.1 Lokacije svih fotografija koje se nalaze u skupu podataka (obalno područje Hrvatske)	67
Slika 4.2 Uvećana lokacija plaže Žnjan, Split, gdje svaka točka predstavlja lokaciju na kojoj je fotografija snimljena	68

Slika 4.3 Distribucija fotografija s objektima po sezoni.....	68
Slika 4.4 Distribucija fotografija po dobu dana	69
Slika 4.5 Distribucija fotografija prema brzini vjetra	70
Slika 4.6 Distribucija vrijednosti solarne elevacije.....	71
Slika 4.7 Distribucija azimuta (kut između sjevera i projicirane lokacije Sunca)	71
Slika 4.8 Dijagram toka procesa kreiranja baze podataka	73
Slika 4.9 Distribucija broja oznaka po zakrpi u bazi podataka Selected Patches	74
Slika 4.10 Označeni uzorci fotografija iz baze podataka Selected Patches	74
Slika 4.11 Distribucija klasa u konačnoj bazi podataka Selected Patches.....	75
Slika 4.12 Metapodaci stanica	76
Slika 4.13 Distribucija kategorija	76
Slika 4.14 Distribucija broja oznaka po jednoj slici	77
Slika 4.15 Uzorci iz baze podataka.....	78
Slika 4.16 „Confusion” matrica za RoWS model pri IoU 0.5	82
Slika 4.17 Krivulja „Preciznost - Odaziv” za model RoWS pri IoU 0.5	83
Slika 4.18 Prikaz originalne fotografije s kanalom Hue	85
Slika 4.19 Prikaz originalne fotografije s kanalom Depth.....	86
Slika 4.20 Proces treniranja baznih modela.....	87
Slika 4.21 Postupak treniranja modela.....	88
Slika 4.22 Krivulja „Preciznost - Odaziv” za model RoWS-4ch-hue pri IoU 0.5.....	91

Popis oznaka

ACCLAHE Adaptively Clipped Contrast Limited Histogram Equalization

AHE Adaptive Histogram Equalization

AI Artificial Intelligence

AMD Anthropogenic Marine Debris

AUC-PR Area Under Curve Precision-Recall

AUV Autonomous Underwater Vehicle

BL Background Light

CHS Contrast histogram stretching

CLAHE Contrast - Limited Adaptive Histogram Equalization

CNN Convolutional Neural Network

COCO Comon Objects in Context

CONV Convolutional

CSIRO Commonwealth Scientific and Industrial Research Organisation

DCP Dark Channel Prior

DPT Dense Prediction Transformer

CDF Cumulative Distribution Function

DECM Discrete Entropy and Contrast Measure

DETR Detection Transformer

DJI Da - Jiang Innovations

DSM Digital Surface Model

DSNMF Deep Sparse Non - negative Matrix Factorization

FMML Floating Marine Macro Litter

FP False Positive

FPS Frames Per Second

FT Fine Tuning

GAN Generative Adversarial Network

GB Green - Blue

GBdehazingRCorrection Image restoration by blue - green channels dehazing and red channel correction

GC Gamma Correction Model

GPU Graphics Processing Unit

GR Green – Red

HF Homomorphism Filtering

HR High Resolution

HSI Hue Saturation Intensity

HSV Hue Saturation Value

ICM Integrated Color Model

IFM Image Formation Models

IoU Intersection over Union

IQA Image Quality Assessment

JAMSTEC Japan Agency for Marine Earth Science and Technology

JTF Joint Trigonometric Filtering

LMG Local Maximum Gradient

LMI Local Maximum Intensity

LR Logistic Regression

MAI Maximum Attenuation Identification

MAP Maximum A Posteriori

mAP mean Average Precision

MARI Marine Autonomous Robotics for Interventions

MCM Multi - Color Model

MILP Minimal Information Loss Principle

MIP Maximum Intensity Prior

ML Macro Litter

MRA Magnetic Resonance Angiogram

MRF Markov random fields

MSE Mean Squared Error

MSRCR Multiscale Retinex Color Restoration

MTF Modulation Transfer Function

MVGB Maximum Value Green Blue

NDVI Normalized Difference Vegetation Index

NIR Near Infrared

NMS Non - Maximum Suppression

NN Neural Network

PSF Point Spreading Function

PSNR Peak Signal to Noise ratio

RAHIM Recursive Adaptive Histogram Modification

RB *Red - Blue*

RCP Red Channel Prior

ReLU Rectified Linear Unit

RGB Red - Green – Blue

RGHS Relative Global Histogram Stretching

RMSE Root Mean Squared Error

ROV Remotely Operated Vehicle

RoWS Removal of Water Scattering

RPN Region Proposal Network

SAUV Soft-Robotic Autonomous Underwater Vehicle

SCM Single - Color Model

SR Super Resolution

SSD Single Shot Detection

STF Spatial - Temporal Filtering

SVM Support Vector Machine

SWT Stationary Wavelet Transform

TM Transmission Map

TP True Positive

TUC Towed Underwater Camera

UAS Unmanned Aerial Systems

UAV Unmanned Aerial Vehicle

UDCP Underwater Dark Channel Prior

ULAP Underwater Light Attenuation Prior

UOI Underwater Optical Imaging

UWCNN UnderWater Convolutional Neural Network

VAE Variational Autoencoder

ViT Vision Transformer

WDF Wave digital filter

YOLO You Only Look Once

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]. Redukcija količine otpada će smanjiti pritisak na morski okoliš. Međutim, moramo misliti i na otpad koji se trenutačno nalazi u njemu te na koji način ga ukloniti kako ne bi došlo do akumulacije i fragmentacije istog.

Detekcija morskog otpada je prvi korak u otklanjanju otpada iz morskog okoliša. U tom kontekstu cilj je automatizirano identificirati instance objekata iz unaprijed definiranih klasa otpada te lokalizirati njihovu prostornu poziciju na slici. Najnovija otkrića u području računalnog vida (*engl. Computer Vision*), pogotovo napredak konvolucijskih neuronskih mreža (*engl. Convolutional Neural Network - CNN*) unaprijedila su područje mobilne robotike i razvoj aplikacija za autonomnu vožnju, nadzor dronom, strojeva za žetvu, strojeva za mužnju, itd. [2]

Cilj ovog rada je istražiti i izraditi platformu koja detektira objekte u moru s boljim rezultatima nego dosadašnje metode koristeći tehnologije računalnog vida i konvolucijskih neuronskih mreža. Da bi se izgradio takav sustav potrebno je osigurati veliki skup podataka na takav način da se što manje narušava morski ekosustav. Kao najmanje invazivan pristup pokazao se proces prikupljanja podataka iz daljine te će se koristiti bespilotna letjelica - dron. Postoje različite baze podataka koje se koriste za detekciju objekata pod morem ili na površini mora, ali ne postoji dostupna baza podataka traženih fotografija gdje su objekti u moru, a fotografirane su bespilotnom letjelicom. Stoga je potrebno izgraditi vlastitu sveobuhvatnu bazu podataka te je objaviti.

Specijalizirani algoritmi i mreže za detekciju objekata ispod površine mora na fotografijama snimljenim bespilotnim letjelicama još su uvijek nedovoljno istraženi. Suprotno tome, postoji više istraživanja vezanih za algoritme za detekciju objekata na površini mora te objekata ispod mora uz pomoć podvodnih 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

optike i raspršenja svjetlosti u vodi pa zbog toga područje istraživanja nije još dovoljno obrađeno.

Fizička, biološka i kemijska svojstva mora [3], [4] otežavaju način prikupljanja jasnih podataka i fotografija promatranog područja, uvjetuju na koji način će se fotografije prikupljati i koje metode će se koristiti u daljnjoj obradi podataka.

S obzirom na to da je sve veća potreba za fotografijama visoke kvalitete radi boljeg razumijevanja i analize podmorskog svijeta, brojna su istraživanja usmjerena na proučavanje tehnika obrade podvodnih fotografija. Obrada podvodne fotografije obično se dijeli na dvije glavne kategorije: restauraciju i unaprjeđenje. Potrebno je istražiti metode pretprocesiranja te ih testirati na vlastitom setu podataka i usporediti dobivene rezultate. Također, potrebno je napraviti istraživanje kako se može modificirati arhitektura postojećih neuronskih mreža te time poboljšati rezultate detekcije objekata.

Ovakva platforma bi mogla služiti 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.

1.1. Motivacija

U posljednjim desetljećima, utjecaj čovjeka na morski okoliš postaje sve vidljiviji, a problem morskog otpada sve izraženiji. Unatoč postojećim mjerama i naporima za smanjenje stvaranja otpada, činjenica je da se značajne količine već nalaze u morima i oceanima, gdje se postupno akumuliraju zbog sporog procesa razgradnje. U tom kontekstu, nužnim se pokazuje ne samo sprječavanje budućeg onečišćenja, već i aktivno djelovanje na identifikaciji i uklanjanju postojećeg otpada.

S obzirom na to da je detekcija prvi korak prema uklanjanju otpada, otvorila se potreba za tehnološkim rješenjima koja omogućuju bolje prepoznavanje morskog otpada, pri čemu se posebna pažnja pridaje neinvazivnosti u odnosu na morski ekosustav. Zbog ograničenja klasičnih metoda, posebno u nedostupnim ili osjetljivim područjima, sve se više istražuju mogućnosti korištenja autonomnih sustava temeljenih na računalnom vidu i strojnome učenju.

Inspiracija za ovo istraživanje nastala je iz uočenog nesrazmjera između postojećih tehnologija i stvarnih potreba na terenu. Nedostatak dostupnih i primjenjivih baza podataka snimljenih iz zraka, specifičnih za detekciju objekata pod morem, postavio je izazov izgradnje vlastitog skupa podataka kao temelja za razvoj specijaliziranih algoritama. Prepoznaje se i potencijal u prilagodbi arhitekture modela i uključivanju dodatnih informacijskih kanala, kao i u primjeni algoritama za pretprocesiranje, što dovodi do poboljšanja performansi detekcije.

Motivacija je također ukorijenjena i u širem društvenom značaju – od konkretne detekcije otpada, preko identifikacije morskih organizama i praćenja ekoloških promjena, do podrške u zaštiti prirodnih resursa. Razvoj pouzdanih i efikasnih alata za ovu svrhu može imati dugoročne koristi za znanstvenu zajednicu, dionike u očuvanju okoliša, ali i lokalne zajednice koje izravno ovise o moru.

1.2. Hipoteza

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 na površini i u plićim dijelovima mora ili rijeka postoje, unapređenje i prijedlog novih rješenja u ovom području čini se obećavajuće.

Fotografije će biti prikupljene dronom, metodom daljinskog istraživanja, pri čemu se detekcija ograničava na dubine do približno 4–5 metara, što odgovara realnom dometu optičke vidljivosti iz zraka u pogodnim uvjetima. Izgradit će se vlastita baza podataka.

Kako bi se unaprijedila detekcija objekata s pomoću neuronskih mreža, bit će istražena implementacija dodatnog, četvrtog kanala u standardni RGB (*engl. Red-Green-Blue*) prostor boja. Očekuje se da će dodatni kanal sadržavati različite vrste informacija koje bi ne samo omogućile bolje razumijevanje vizualnog sadržaja fotografije, već i pružile neuronskim mrežama dodatne podatke koji bi mogli povećati točnost i efikasnost detekcije objekata. Rješenje koje se izabere zahtijevat će pažljivo istraživanje i eksperimentiranje s podacima i arhitekturama neuronskih mreža kako bi se osiguralo da dodatne informacije poboljšavaju performanse, a ne dodaju nepotreban šum ili kompleksnost. Točnost, brzina detekcije i sposobnost generalizacije na nove, neviđene fotografije bit će ključni parametri uspjeha. Osim toga, morat će se razmotriti kompromis između dobivenih poboljšanja i dodatnog računalnog opterećenja koje četvrti kanal donosi, kako u treningu, tako i u primjeni.

Pretpostavka je da će kvaliteta fotografija podvodnih objekata biti narušena upravo zbog fizikalnih svojstava morske površine, vremenskih uvjeta, nečistoće optičkog medija te bioloških i kemijskih svojstava mora. Izgradit će se platforma za pretprocesiranje fotografija koje će poboljšati rezultate detekcije. Koristit će se metode pretprocesiranja fotografija prije samog treniranja modela te će se usporediti rezultati originalnih i pretprocesiranih fotografija. Rezultati detekcije i klasifikacije algoritma na originalnim fotografijama usporediti s fotografijama koje su prošle kroz predloženi proces u ovoj disertaciji bi se trebala razlikovati u koristi potonje.

Osnovne hipoteze ovog istraživanja su:

- a) Formiranjem nove baze fotografija otpada na morskom dnu u priobalnom pojasu, a koje su prikupljene bespilotnom letjelicom, moći će se istrenirati duboka neuronska mreža koja će imati bolje rezultate detekcije i klasifikacije otpada od postojećih.
- b) Korištenjem tehnika pretprocesiranja fotografija gdje su objekti pod morem će se postići bolji rezultati detekcije.
- c) Dodatno poboljšanje će se postići prilagodbom arhitekture duboke neuronske mreže.

1.3. Struktura rada

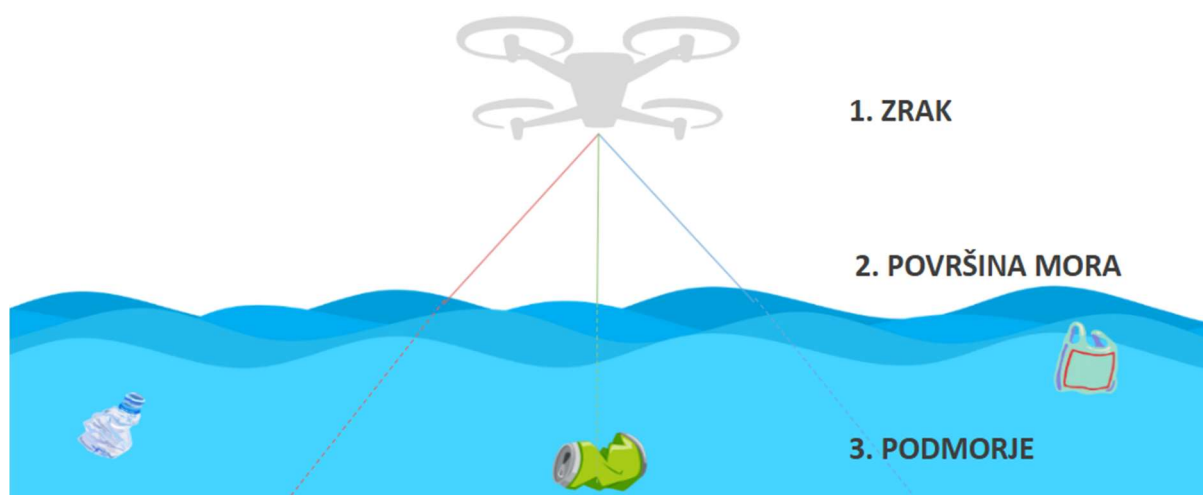
Ova doktorska disertacija strukturirana je u osam poglavlja. U prvom poglavlju uvodi se tema istraživanja, definira motivacija za istraživanje te se postavljaju hipoteze. Poglavlje dva detaljno opisuje problem istraživanja, uključujući izazove i složenost detekcije objekata u podvodnim uvjetima. Treće poglavlje pruža pregled relevantne literature i postojećih metoda koje uključuje pregled konvolucijskih neuronskih mreža, transformera, daljinskog istraživanja, postojećih metoda detekcije objekata (na površini i pod morem), relevantnih baza podataka te ograničenja podvodnog fotografiranja. U četvrtom poglavlju detaljno se opisuje metodologija istraživanja, uključujući prikupljanje fotografija, kreiranje baze podataka, preprocesiranje fotografija različitim metodama, modifikaciju arhitekture modela za detekciju objekata, proces treniranja i evaluacije modela te analizu dobivenih rezultata. Disertacija se zaključuje poglavljem koje sumira njezin doprinos i nudi smjernice za budući rad te pruža pregled korištene literature tijekom istraživanja.

2. Opis problema

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 fotografije promatranog područja, uvjetuju na koji način će se fotografije prikupljati i koje metode će se koristiti u daljnjoj obradi podataka. Specijalizirani algoritmi i mreže prilagođene rješavanju problema detekcije objekata ispod mora na fotografijama 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 dok podvodna detekcija predstavlja izazovniji zadatak.

Može se zaključiti da se problem može razložiti na tri točke, kako prikazuje Slika 2.1:

- 1) Problem prikupljanja fotografija iz zraka u različitim vremenskim uvjetima te različito doba dana i godine s različite visine.
- 2) Problem područja promjene medija iz zraka u more (refleksije, valovi, lom) te problematika načina uklanjanja bljeska, morske pjene te korekcija loma zbog valova.
- 3) Problem raspršivanja svjetlosti, smanjenje vidljivosti te apsorpcija valnih duljina u moru.



Slika 2.1 Izazovi pri detekciji objekata pod morem

Dostupnost, promjene i mapiranje obalnih ekosustava i samo njihovo upravljanje postalo je mnogo lakše i efikasnije primjenom satelitskih snimki i upotrebom procesa prepoznavanja iz

daljine. Bepilotne letjelice, dronovi, najčešće se koriste za detekciju makro morskog otpada (*engl. 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 na to da UAV (*engl. Unmanned Aerial Vehicle*) fotografije 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 zahtjevnja i zamorna, međutim, trenutačne metode ne mogu još uvijek kategorizirati ML-ove na detaljan način. Dominantni pristup u rješavanju ovog problema, zbog ogromnog rasta i razvoja, je primjena AI (*engl. Artificial intelligence*) algoritma [5].

Nadalje, korištenje umjetne inteligencije pri detekciji podvodnih objekata s fotografija nailazi na drugi skup problema. Jedan od ključnih zahtjeva sustava za detekciju objekata na morskom dnu je visoka razina odaziva. Svaki nedetektirani objekt, iako prisutan na fotografiji snimljenog područja, ostaje neevidentiran te kao posljedicu ima izravan negativan utjecaj na okoliš. S druge strane, nužno je izbjeći i velik broj lažno pozitivnih detekcija, budući da takve greške dodatno iscrpljuju dostupne resurse te povećavaju vrijeme potrebno za naknadnu provjeru. Iz tog razloga, cilj je razviti sustav koji će imati optimalan omjer točno pozitivnih i lažno pozitivnih detekcija, kako bi predstavljao pouzdanu i učinkovitu metodu primjenjivu u stvarnim uvjetima detekcije objekata na morskom dnu.

Čest problem kod detekcije objekata, naročito malih dimenzija, predstavlja varijabilnost rezultata ovisno o skali na kojoj se detekcija provodi. Nerijetko se događa da objekt bude uspješno detektiran na jednoj razini detalja, dok na drugoj (većoj ili manjoj) ostane neprepoznat. Iz tog razloga, ključno je prilikom ekstrakcije značajki osigurati njihovu generalizaciju koja omogućava detekciju na različitim skalama. To se u praksi postiže provedbom višeslojne ekstrakcije značajki, odnosno korištenjem piramida značajki ili sličnih metoda, koje sustavu poboljšavaju detekciju bez obzira na veličinu objekta unutar fotografije.

Kvaliteta prikupljenih fotografija značajno utječe na učinkovitost sustava za detekciju objekata na morskom dnu. Fotografije koje su prikupljene korištenjem bepilotnih letjelica predstavljaju osnovni izvor semantičkih informacija o promatranom području. Stoga je za uspješnost detekcije, a time i za povećanje vjerojatnosti pronalaska željenih objekata, ključna upotreba

fotografija visoke prostorne razlučivosti koje omogućuju preciznije prepoznavanje i lokalizaciju objekata od interesa.

U prikupljenim fotografijama posebno se ističe problem nebalansiranosti klasa. Iako nebalansiranost može postojati u različitim oblicima, specifičnost ovog istraživanja očituje se prvenstveno u disproporcionalnosti pozitivnih regija (regija koje sadrže objekte od interesa) i negativnih regija (regija bez objekata od interesa). Naime, većina prikupljenih fotografija ne sadrži nikakve objekte od interesa, dok samo ograničeni broj fotografija prikazuje jedan ili tek nekoliko takvih objekata. U skladu s tim, cijela pozadina, odnosno većina površine fotografije, klasificira se kao negativna regija. Sustav za detekciju stoga mora biti robustan na ovakvu vrstu neravnoteže kako bi se osiguralo uspješno pronalaženje rijetkih objekata u uvjetima gdje dominiraju negativni primjeri.

3. Pregled područja

Pregled područja istraživanja obuhvaća ključne koncepte i postojeće metode relevantne za detekciju objekata u podvodnom okruženju primjenom dubokog učenja, istraživanje dostupnih baza podataka te analiza metoda prilagodbe duboke neuronske mreže.

3.1. Klasifikacija slika

Klasifikatori su modeli dubokog učenja čija je primarna svrha dodjeljivanje slike unaprijed definiranoj klasi. U kontekstu računalnog vida, oni predstavljaju temeljne komponente za razumijevanje vizualnog sadržaja i često se koriste kao polazište za složenije zadatke poput segmentacije ili detekcije objekata. Ovisno o arhitekturi, klasifikatori mogu biti temeljeni na konvolucijskim neuronskim mrežama (CNN) ili transformerskim modelima, pri čemu svaki pristup ima svoje prednosti u pogledu učenja značajki, robusnosti i skalabilnosti.

3.1.1. Klasifikatori temeljeni na konvolucijskim neuronskim mrežama (CNN)

Konvolucijske neuronske mreže su specijalizirane za obradu vizualnih podataka, omogućavaju automatsko učenje značajki iz fotografija. CNN-i smanjuju potrebu za prethodnom obradom i ručno izrađenim filterima jer sami uče filtere tijekom procesa treniranja. Njihova arhitektura sastoji se od konvolucijskih slojeva, slojeva poduzorkovanja, potpuno povezanih slojeva i nelinearnih aktivacijskih funkcija. Konvolucijski slojevi izdvajaju osnovne vizualne značajke poput rubova i kutova, dok poduzorkovanje smanjuje razlučivost značajki radi robusnosti. Potpuno povezani slojevi kombiniraju značajke za konačnu klasifikaciju, a nelinearni slojevi koriste funkcije poput ReLU (*engl. Rectified Linear Unit*) omogućuju složenije reprezentacije. CNN-i su otporni na izobličenja i pomake, zahtijevaju manje memorije i omogućuju brže treniranje u usporedbi s potpuno povezanim neuronskim mrežama [6].

Klasifikatori poput AlexNet [7], VGGNet [8], ResNet [9], Inception [10] značajno su unaprijedili duboko učenje i prepoznavanje uzoraka. AlexNet je pionirska duboka konvolucijska mreža koja je postigla visoku točnost na ImageNet izazovu 2012. godine. Sastoji se od pet konvolucijskih i tri potpuno povezana sloja, koristeći ReLU aktivaciju za brže treniranje i lokalnu normalizaciju radi stabilnosti. VGGNet je pojednostavio arhitekturu korištenjem fiksni 3×3 konvolucijskih kernela, čime se smanjuje broj parametara i vrijeme treniranja. Varijante poput VGG16 i VGG19 omogućuju brže i robusnije učenje. ResNet rješava problem nestajanja gradijenta uvođenjem rezidualnih blokova čime se omogućuje učinkovito treniranje vrlo dubokih mreža. Isto tako koristi prečace identiteta i projekcije kako

bi poboljšao točnost. Inception mreže, poput GoogLeNet, proširuju mrežu horizontalno unutar slojeva korištenjem kernela različitih veličina, što omogućuje paralelno izdvajanje značajki na različitim razinama.

3.1.2. Klasifikatori temeljeni na transformer arhitekturi

Transformeri su se izvorno pojavili u području obrade prirodnog jezika i pokazali su se iznimno uspješnima u zadacima poput strojnog prevođenja i sažimanja teksta. Transformerske arhitekture predstavljaju suvremeni pristup u računalnom vidu koji se, za razliku od klasičnih CNN-a, temelje na mehanizmu vlastite pažnje (*engl. self-attention*). Umjesto konvolucija, Vision Transformer (ViT) dijeli sliku na manje zakrpe, tzv. „*patches*“, (npr. 16×16 piksela) i tretira ih kao sekvencu ulaznih tokena, što omogućuje modelu da uči globalne odnose na slici bez obzira na njihovu udaljenost [11].

3.2. Detekcija objekata na slici

Za razliku od klasifikatora, detektori objekata ne samo da klasificiraju što se nalazi na slici, već i lociraju objekte unutar nje. Ovisno o načinu na koji pristupaju ovom zadatku, detektori se dijele na dvofazne i jednofazne metode. U kontekstu detekcije objekata, faza označava diskretan korak u obradi slike, tijekom kojeg se obavlja određena radnja: prva faza uključuje predlaganje regija interesa dok druga faza obavlja klasifikaciju i preciznu lokalizaciju objekata unutar tih regija. Detektori koji se sastoje od jedne faze obje radnje obavljaju paralelno dok detektori koji se sastoje od dvije faze radnje obavljaju sekvencijalno.

3.2.1. Detektori koji se sastoje od dvije faze

Detektori koji se sastoje od dvije faze, poput R-CNN [12], Fast R-CNN [13] i Faster R-CNN [14] koriste sekvencijalni pristup za identificiranje i klasifikaciju objekata unutar fotografija. U prvoj fazi generiraju se predložene regije interesa (*engl. region proposals*), dok se u drugoj fazi svaka regija klasificira i precizno lokalizira. R-CNN koristi selektivno pretraživanje za izdvajanje predloženih regija što je računalno zahtjevno i nije pogodan za rad u stvarnom vremenu. Fast R-CNN poboljšava brzinu obradom cijele fotografije odjednom i generiranjem mape značajki čime se eliminira potreba za pojedinačnim predlaganjem regija. Faster R-CNN dodatno optimizira proces uklanjanjem selektivnog pretraživanja i integracijom mehanizma za generiranje regija (*engl. Region Proposal Network - RPN*) izravno u mrežu, omogućujući detekciju objekata u stvarnom vremenu.

3.2.2. Detektori koji se sastoje od jedne faze

Detektori u jednoj fazi, poput SSD (*engl. Single Shot Multibox Detector*) [15] i YOLO (*engl. You Only Look Once*) serije [16], [17], [18], [19], [20], [21] i njene inačice [22], [23], [24], [25], [26], [27], [28], [28], [29], [30], [31], [32], [33] omogućuju brzu i učinkovitu detekciju objekata na fotografijama u jednom prolazu mrežom, bez potrebe za eksplicitnim predlaganjem regija. SSD koristi odvojene glave za izradu mapa značajki i okosnicu za ekstrakciju značajki, omogućujući detekciju objekata različitih veličina. YOLO algoritmi fokusiraju se na brzinu i točnost detekcije objekata u stvarnom vremenu, a svaka nova verzija donosi poboljšanja u lokalizaciji, točnosti i brzini detekcije. Ove arhitekture koriste unaprijed trenirane mreže i različite strategije obuke kako bi postigli visoku preciznost i učinkovitost, čineći ih pogodnima za primjene u kojima je ključna obrada u stvarnom vremenu.

3.2.3. Detektori temeljeni na transformer arhitekturi

Detection Transformer (DETR) dodatno prilagođava arhitekturu transformera za detekciju objekata: formulira detekciju kao direktan problem predviđanja skupa objekata i uvodi enkoder - dekodek transformer koji zamjenjuje mnoge ručno dizajnirane komponente klasičnih detektora [34]. DETR može modelirati prostorne odnose (korisno kada je kontekst ključan) te eliminira potrebu za ručnim heuristikama - točke sidrišta i NMS (*engl. Non - Maximum Suppression*), što rezultira konceptualno jednostavnijim postupkom detekcije [34].

Posljednjih godina razvijeni su brojni transformer modeli za računalni vid koji su pronašli široku primjenu. ViT je prvenstveno namijenjen za klasifikaciju slika i kao osnovica za ekstrakciju značajki [11]. Swin Transformer [35] predstavlja hijerarhijski pristup koji uvodi lokaliziranu vlastitu pažnju unutar pomičnih prozora, čime postiže linearnu skalabilnost s veličinom slike. Ovi modeli nalaze primjenu u širokom rasponu zadataka: Google ViT i Swin koriste se za klasifikaciju slika i segmentaciju (od prepoznavanja objekata na fotografijama do medicinske dijagnostike i analize satelitskih snimaka), dok transformer detektori (DETR, Deformable DETR, DINO [36]) postaju sve važniji u domenama poput autonomne vožnje, robotike i video-nadzora – svugdje gdje je mogućnost detekcije uz uvažavanje globalnog konteksta od osobite koristi.

3.3. Usporedba CNN i transformer modela

Transformer modeli mogu nadmašiti CNN modele tek uz dovoljno velike skupove podataka, a na manjim skupovima podataka su skloni prekomjernom prilagođavanju modela (*engl. overfitting*) [11]. Originalni DETR pati od spore konvergencije što su kasniji radovi nastojali poboljšati. CNN pristupi su u tim aspektima robusniji i brže postižu visoku učinkovitost na ograničenim podacima. Poseban izazov je proširenje ovih transformer arhitektura na ulazne slike s više od tri kanala - kod transformera fiksna dimenzionalnost tokena onemogućuje direktno korištenje predtreniranih težina te dramatično povećava duljinu ulazne sekvence. Pristup koji bi tretirao dodatni kanal (npr. dubinu) kao zaseban skup tokena bi povećao duljinu sekvence u transformeru i time kvadratno uvećao trošak mehanizma vlastite pažnje [37]. To značajno povećava memorijske i vremenske zahtjeve modela. Fiksna duljina ulaznih tokena, definirana pozicijskim ugrađivanjem, dodatno ograničava fleksibilnost kod varijabilnih dimenzija slika pa prilagodba modela na podatke različitih dimenzija zahtijeva kompleksne pred i post procesne korake. Konačno, visoka energetska potrošnja i zahtjevi za memorijom imaju ne samo logističke i financijske, nego i ekološke implikacije.

3.4. Daljinsko istraživanje

Daljinsko istraživanje ili prepoznavanje iz daljine (*engl. Remote sensing*) je znanost o proučavanju i otkrivanju fizikalnih svojstava područja dok se sam promatrač ne mora nalaziti na promatranom području. Omogućuje vizualizaciju, snimanje i analizu promatranih objekata i značajki površine Zemlje. Ovaj proces je dobio naziv zbog činjenice da koristi senzor (*engl. sensor*) pri prikupljanju informacija iz daljine. Sateliti, zrakoplovi i bespilotne letjelice imaju specijalizirane platforme za nošenje senzora [29]. Zrakoplovi ili sateliti snimaju velika područja na površini Zemlje, percepciju fotografije morskog dna omogućuju nam sonarni sustavi na brodovima dok promjene temperatura u oceanima možemo mjeriti s 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 tijekom godina
- praćenje rasta gradova

- mapiranje i otkrivanje oceanskog dna [38].

3.5. Postojeće metode primijenjene na detekciji objekata

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

Rad [39] predlaže poboljšani algoritam „Faster R-CNN” gdje su objekti grupirani s algoritmom „K-means clustering” i koriste algoritam „Soft NMS” umjesto algoritma „NMS”. Autori su obrađivali fotografije s Google Earth-a i natjecateljske platforme Kaggle gdje su promatrani objekti bili brodovi na površini mora. Rezultati su pokazali da poboljšani algoritam „Faster RCNN” ima bolje prosječne rezultate detekcije morskog objekta (Slika 3.1) te rješava probleme niske prosječne točnosti i propusta detekcije malih objekata u složenom morskom okruženju (Slika 3.2).



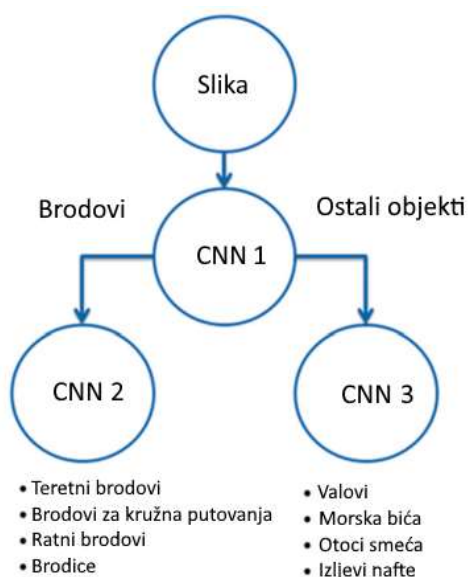
Slika 3.1 Prikaz situacije djelomične detekcije



Slika 3.2 Prikaz rezultata detekcije malih objekata na morskoj površini

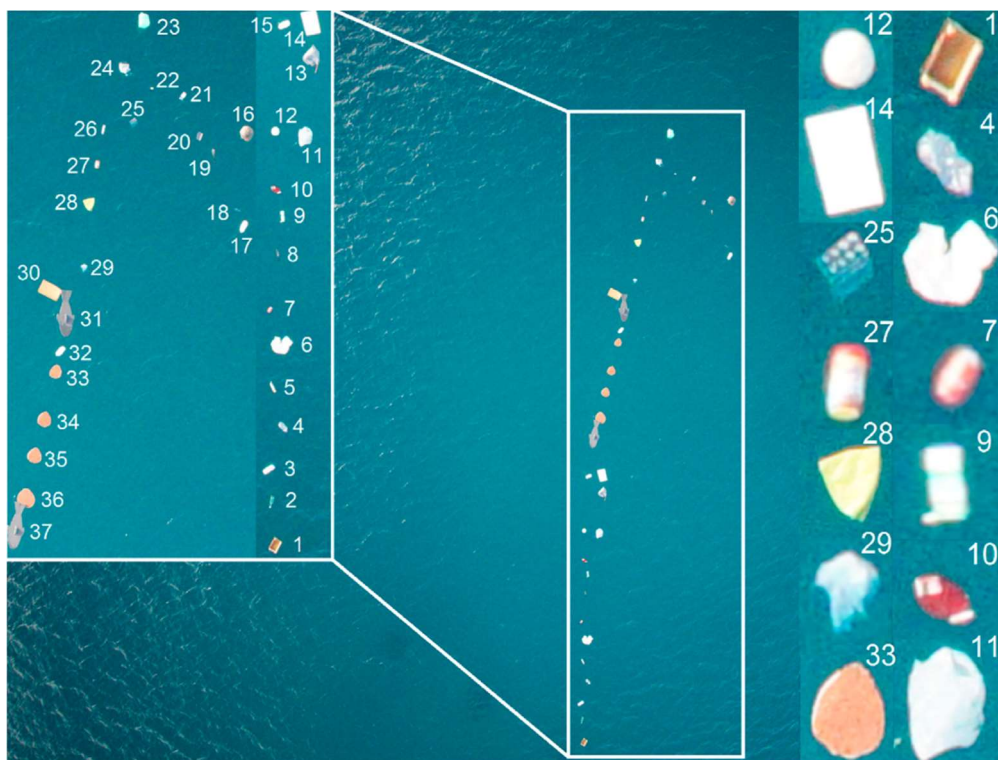
Autori rada [40] prikazuju rezultate na svojoj bazi podataka čije su fotografije dobivene s pomoću Google Karti i Google tražilice fotografija. Promatrani objekti su podijeljeni u dvije kategorije: brodovi i ostali objekti te njihove potkategorije: teretni brodovi, ratni brodovi,

brodovi za kružna putovanja, brodice i valovi, otoci otpada, morska bića i izljevi nafte. Predlažu algoritam koji se sastoji od tri odvojene konvolucijske neuronske mreže – CNN 1, CNN 2 i CNN 3 (Slika 3.3). Korištenje CNN mreža se pokazalo kao odgovarajuća metoda umjetne inteligencije pri prepoznavanju morskih objekata iz zraka.



Slika 3.3 Dijagram tijeka podataka algoritma za prepoznavanje morskih objekata

Rad [41] predlaže algoritam koji je sposoban detektirati i kvantificirati plutajući morski makro otpad (*engl. Floating Marine Macro Litter - FMML*) s fotografija iz zraka (Slika 3.4). Autori prethodnog rada razvili su aplikaciju MARLIT [42] koja omogućava korisnicima detektirati FMML s učitanih fotografija. Predloženi algoritam koristi CNN koji je sposoban učiti iz nestrukturiranih ili neoznačenih podataka. Ukupan skup podataka su 796 fotografije, od kojih 398 nema FMML, a na ostalih 398 se nalazi. Fotografirane su iz daljine, dronom i zrakoplovom iznad Sredozemnog mora. Rezultati klasifikacije fotografija i križne provjere (*engl. cross validation*) bili su: preciznost 0.82 i odaziv 0.84. Zbog nedostatka dovoljnog broja fotografija, autori su koristili proces augmentacije prije treniranja. Navedene mjere dobivene su na skupu u kojem je augmentacija provedena prije podjele na trening i validaciju pa su se u validacijskom dijelu našle fotografije vrlo slične treniranima stoga rezultati prvenstveno odražavaju sposobnost modela da prepozna takve derivacije, a ne nužno njegovu stvarnu generalizacijsku učinkovitost na dosad neviđenom skupu. Ono što i dalje predstavlja problem je automatizirano praćenje FMML-a na otvorenom moru te su potrebna daljnja istraživanja.



Slika 3.4 Fotografija iz zraka s nizom poznatih predmeta pričvršćenih na užice (klase: sanduk, boce, limenke, kese, tetrapak, ručnik, balon, karapaks kornjače...)

Autori rada [43] predstavljaju na koji način pratiti prostornu distribuciju plutajuće plastike u svrhu smanjenja negativnih učinaka na okoliš. UAV-ovi mogu prikupiti podatke za mapiranje, ali većina metoda zahtijeva ručno označavanje i vizualnu interpretaciju. Ciljevi ovog rada su 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 (*engl. 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 mape 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 U-Net i SegNet [44]. U-Net, koji je nastao prije SegNet arhitekture, koristi direktne veze da bi povezao sloj u enkoder dijelu sa slojem u dekoder dijelu. Mape značajki se spajaju i prenose u cijelosti. U-Net zbog toga troši više memorije, ali mu je potrebno manje trening fotografija 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 (bez potpuno povezanih slojeva na kraju). U dekoder dijelu se koriste spremljeni indeksi da bi se povećala rezolucija mape te se time iz prorijeđene dobije gusta mapa značajki. Finalni element dekodera 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 točnosti. Korišteni su drveni okviri u koje su stavljene boce napravljene od različitih verzija plastike, povezani najlonskim užetom. Potom su odabrani elementi koji najviše pridonose zagađenju na danim lokacijama. Fotografije su dobivene koristeći različite visine: 12, 40, 55, 70 i 90 metara. Baze fotografija ne postoje zbog toga što je ovo prvi rad koji pokušava detektirati tip plastike koristeći UAV fotografije. Autori su zbog toga morali kreirati vlastiti skup fotografija. Za svaki prelet UAV-om su generirali ortofotografiju visoke razlučivosti (SIFT [45] algoritam detektira ključne značajke, RANSAC [46] algoritam ih poveže kroz više fotografija). Nakon toga je na svakoj slici izvršeno ručno obilježavanje klasa plastike. Problemi s kojima su se susreli u ovom koraku su otežano obilježavanje 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 različite visine 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 pa su autori odabrali četiri postojeće implementacije koje su trenirane na ImageNet skupu podataka: ResNet50, ResNeXt50, Inception-ResNet v2 i Xception [47]. ResNet50 koristi konekcije s identitet funkcijom te je dodano 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 za ispitivanje performansi segmentacije: preciznost (*engl. precision*), odaziv (*engl. recall*) i F1 mjera (*engl. F1 score*). Implementacija je izvršena koristeći navedene mreže te je rezultat usporedba šest različitih

modela. Autori su testirali različite funkcije gubitka te su zaključili da je najbolja točnost ako se koristi funkcija križne entropije (*engl. cross entropy*). Zbog arhitekture mreže, originalna fotografija 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 fotografije na regije. Najveća greška se događa u kutovima fotografija zbog najvećeg nedostatka kontekstualnih informacija. Algoritam ResUNet50, pokazao je najveću točnost pri mapiranju različitih plastičnih materijala (F1 mjera: najlon: 0.88; orijentirani polistiren (OPS): 0.86; plastika (općenito): 0.78; polietilen tereftalat (PET): 0.92). Kako se smanjivala prostorna razlučivost (autori su koristili razlučivosti 4, 13, 18, 23 i 30 cm/piksel) tako se smanjivala i točnost klasifikacije. Model raspolaže i s podacima procjene površine i volumena plastike što je od velike pomoći pri akcijama čišćenja. Slika 3.5 prikazuje usporedbu različitih modela.

	ResUNet50			ResUNet50			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 3.5 Rezultati točnosti za različite implementacije mreža [43]

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 3.6 prikazuje.

	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 3.6 Rezultati točnosti za različite prostorne rezolucije fotografija [43]

Prema CSIRO (*engl. Commonwealth Scientific and Industrial Research Organisation*) uputama, minimalna veličina detektirane plastike treba biti 1 cm^2 pa zbog toga autori sugeriraju 4 cm/piksel razlučivost (12 m visina letjelice). Mana je mala površina koja se time obuhvaća s jednom fotografijom te je potrebno više vremena da se snimi određeno 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

rezultate 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 fotografija na kojima algoritam trenira te predlažu prelet područja s dva drona, jedan na manjoj visini koji hvata detalje, drugi koji generira fotografije za trening. Uvjeti u kojima se fotografije trebaju dobiti utječu na rezultate gdje autori navode da je oblačno bolje od sunčanog (manje refleksije, *sunglint* efekt) i bez vjetrova (uzrokuje zamućenost fotografije, pomak objekata između kadrova).

Autori rada [48] predlažu inventivnu istraživačku primjenu bespilotnih zračnih sustava UAS (*engl. Unmanned Aerial Systems*) i satelitskih fotografija, otvorenog pristupa, u daljinskom otkrivanju plutajuće plastike u morskom ekosustavu. Ostvarili 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 Lesbos, Grčka. Spektralna svojstva plutajućih meta i okolnog mora istražena su s pomoću satelita Sentinel-2A. Pokazali su kako se UAS može koristiti za procjenu postotka pokrivenosti plastikom sa satelitskih snimka i kako UAS fotografije vrlo visoke geoprostorne razlučivosti mogu biti korisne u poboljšanju geo-referenciranja satelitskih fotografija. Uočena je vrlo slaba do jaka povezanost između spektralne refleksije 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 [49] uvode novi pristup pri detekciji malih objekata sa satelitskih fotografija visoke rezolucije. Predlažu korištenje značajnih točaka koje se mogu brzo izdvojiti iz satelitskih fotografija kao prikladne jedinice promjene za male objekte te kako bi smanjili broj lažnih detekcija. Otkrivaju promjene izdvajanjem značajki iz prethodnih i nedavnih fotografija 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 seta fotografija s rezolucijom uzorkovanja od 50 cm / piksel. Njihov pristup je nadmašio dosadašnji 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 temeljenim na pikselima te može kompenzirati nedostatke pristupa temeljenih na nadziranom objektima uspješnim smanjenjem broja kandidata za detekciju promjene.

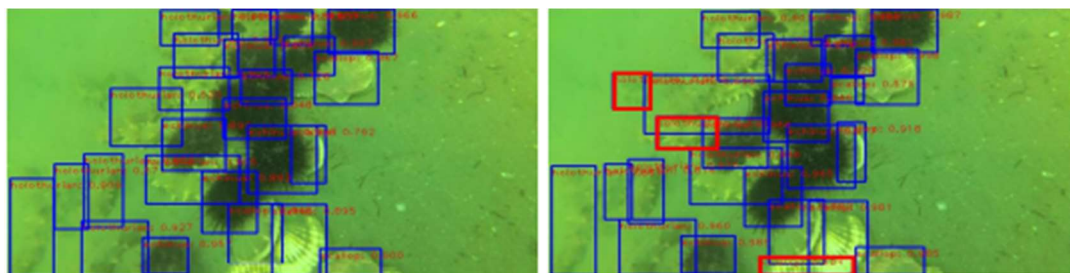
3.5.2. Radovi u kojima se objekt promatranja nalazi pod morem

Autori rada [50] koriste dronove s RGB kamerom kako bi istražili, s pomoću strojnog učenja, utjecaj solarnog kuta i dubine na kvantifikaciju i identifikaciju primarnih bentičkih (bentos je skupni naziv za sve organizme koji žive na dnu vodenih tijela) proizvođača koji se nalaze u rijeci u sjevernoj Kaliforniji. Ciljevi njihovog istraživanja su 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 (*engl. Da-Jiang Innovations*) 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 fotografije je pristup koji je korišten i u istom je analizirana samo jedna fotografija 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 fotografija ne daju rezultate i ne mogu povezati značajke jer se mijenjaju iz fotografije u fotografiju (valovi, plutajući objekti, refleksija sunca i slično). Za klasificiranje i treniranje fotografija 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 fotografija u programu „ArcGIS Pro 2.5” koji implementira navedeni algoritam. Ukupna točnost svih klasa bila je 82 % uzimajući u obzir sve 32 lokacije. Pokazano je kako na točnost klasifikacije fotografija uzetih s dronom utječu solarni kut i dubina. Sunčevo podne koje je trajalo 13:13 - 13:20 je utjecalo na kvalitetu fotografija, stoga fotografije koje su fotografirane prije ili poslije Sunčevog podneva pod solarnim kutem 45°- 60° su najbolje s najvećom osvijetljenošću podloge i najmanjim Sunčevim blještavilom na površini.

Bespilotne 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 [51] imaju za cilj automatski detektirati otpad s morskog dna, koristeći konvolucijske neuronske mreže, u stvarnom okruženju. Neuronska mreža se trenira na setu fotografija 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 fotografijama 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” (*engl. LIFE DEBAG - LIFE14 GIE/GR/0011271*) dostupan na <https://lifedebag.eu> stranici. Podvodne fotografije dobivene su s pomoću podvodne kamere (*engl. towed underwater camera - TUC*) koja se nalazi na malom plovilu koja je vučena po definiranim pravcima zaljeva Ermoupolis u Grčkoj. Skup podataka je uključivao 635 fotografija izoliranih iz videa razlučivosti 1920×1080 piksela. Ručno je identificirano 1166 objekata, a granični okviri su nacrtani s pomoću alata za označavanje fotografija „LabelImg” [52]. Autori su koristili proces augmentacije kako bi povećali skup podataka s obzirom na to da su imali 635 fotografija male veličine te koriste Mask R-CNN kao detektor. Mreža je trenirana s pomoću MobileNetV1 okosnice (*engl. backbone*) i pokrenuta je korištenjem unaprijed treniranih težina iz COCO (*engl. Common Objects in Context*) skupa podataka. Mreža MobileNetV1 je odabrana jer je manje računalno intenzivna u usporedbi s drugim mrežama, a pritom ne gubi svoju točnost za navedenu količinu podataka te su se tako prevladala ograničenja memorije tijekom treniranja i implementacije mreže. Za implementaciju je korišten Keras i TensorFlow [53]. Mreža je postigla ukupnu mAP@50 od 62 % po svim klasama, ali je kod određenih klasa (ribarske mreže, gume, plastične vrećice i plastični čepovi) bila još učinkovitija i postigla mAP@50 (*engl. mean Average Precision*) od 79 %. Autori su prikazali i tablicu odabira hiperparametara koje su koristili prilikom treniranja koji su dali najbolje rezultate za podvodne fotografije. 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). Iako dobrih performansi, mreža MobileNetV1 je zastarjela u usporedbi s modernim arhitekturama (ne koristi novije dizajnerske principe, ima slabu reprezentacijsku snagu za male objekte, slabije povezuje značajke).

Rad [54] koristi bazu podataka „URPC 2018” čiji su promatrani objekti na fotografijama morski krastavci, ježinci, školjke i morske zvijezde. Autori predlažu metodu augmentacije „RoIMix” koja može postići bolju generalizaciju i karakterizira interakcije između fotografija na kojima se objekti mogu preklapati, biti zamagljeni ili zaklonjeni. Za razliku od dotadašnjih metoda za augmentaciju „RoIMix” obrađuje više fotografija 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 3.7).

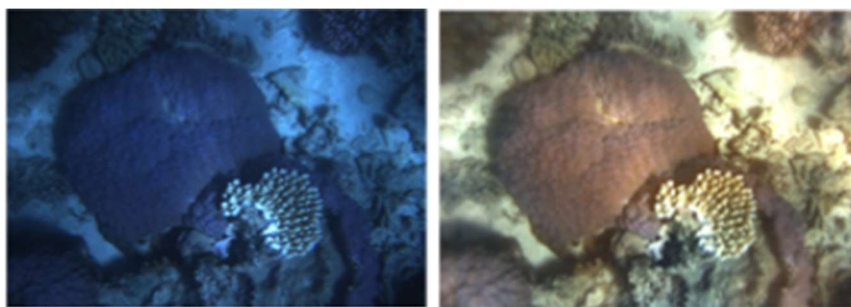


Slika 3.7 Usporedba kontrolne fotografije i primijenjene metode RoIMix

Rad [55] prikazuje rad „WaterGAN” mreže za kreiranje podvodnih, realističnih fotografija koje su dobivene iz zraka i njihovih parova dobivenih pod vodom povećavajući vidno polje dok je percepcija dubine ograničena. „WaterGAN” mreža korigira boje podvodnih fotografija koje su izmijenjene zbog fizičkih svojstava mora. Koristeći „WaterGAN” autori su generirali svoj skup podataka, fotografije iz zraka i njihovi parovi realističnih fotografija na odgovarajućim dubinama koji su potom trenirali. Dio fotografija je fotografiran u laboratorijskim uvjetima (Slika 3.8), a dio u prirodnom okruženju (Slika 3.9). Manjak navedenog skupa podataka jest niska rezolucija fotografija i nedostatak fotografija s otpadom.

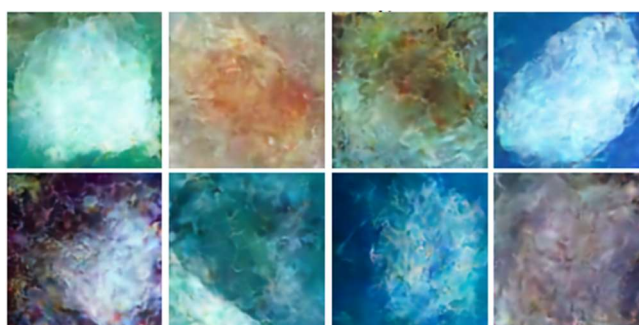


Slika 3.8 Usporedba neobrađene podvodne fotografije i primjena WaterGAN metode



Slika 3.9 Neobrađena podvodna fotografija i primjena WaterGAN metode

Rad [2] naglašava probleme loše kvalitete skupova podataka podvodnih fotografija i njihovu malu brojnost kao i rješenje problema s ciljem što bolje detekcije podvodnih objekata. Koristeći svoju bazu podataka predlažu korištenje dvostupanjskog varijacijskog autoenkodera (*engl. Variational Autoencoder - VAE*) i binarnog klasifikatora. Binarni klasifikator bira fotografije „dobre kvalitete” i povećava zadani skup podataka (augmentacija) nakon što ih VAE generira (Slika 3.10, Slika 3.11 i Slika 3.12). Završni korak je provjera klasifikatora i njegova točnost pri detekciji objekata na stvarnim i generiranim fotografijama. Rezultati pokazuju da tako uvježban klasifikator nadmašuje onaj koji je samo treniran na stvarnim podacima te je ujedno i koristan ako je prikupljanje fotografija neizvedivo ili izazovno.



Slika 3.10 Prva faza generiranja slika koristeći varijacijski autoenkoder

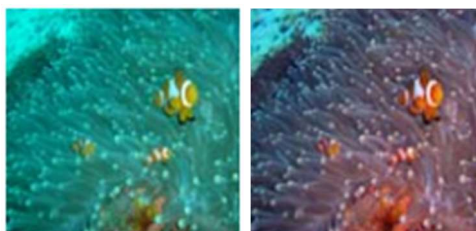


Slika 3.11 Druga faza generiranja slika koristeći varijacijski autoenkoder

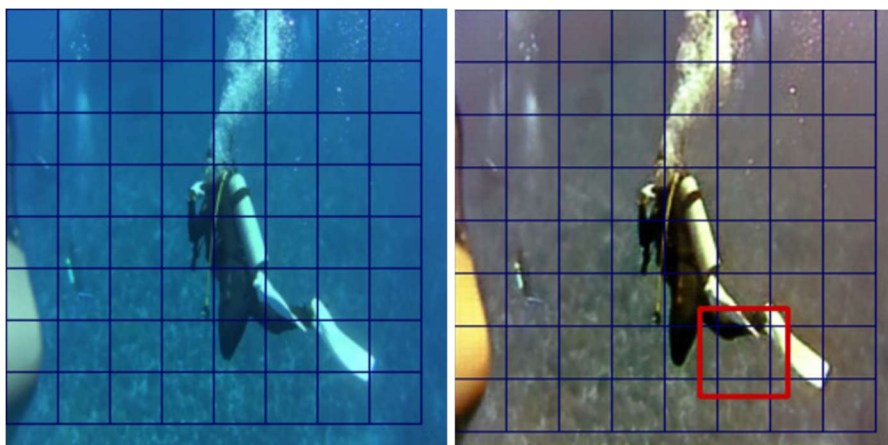


Slika 3.12 Uzorak generiranih slika koristeći varijacijski autoenkoder

Autonomna podvodna vozila (engl. *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 ako su vođeni samo senzorom „vida“. Rad [56] predlaže korištenje modela „CycleGAN“ s ciljem poboljšanja boja podvodnih fotografija. Obrađivane su fotografije 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 3.13) i ronionica (Slika 3.14) za razliku od nekorigiranih fotografija.

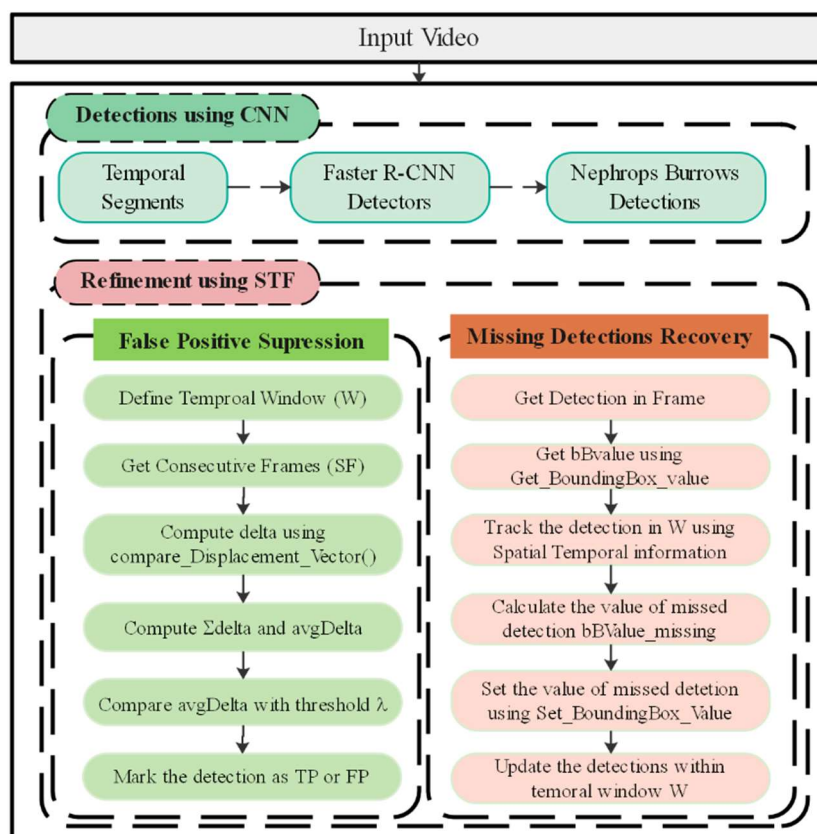


Slika 3.13 Usporedba original i CycleGAN fotografija živih bića u radu [56]



Slika 3.14 Usporedba original i CycleGAN fotografija ronionica u radu [56]

Naseer i suradnici [57] 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, jazbina vrste *Nephrops Norvegicus*. Za automatsku klasifikaciju i detekciju koriste najsvremenije duboke neuronske mreže kao što su Inceptionv2 [58], ResNet50, ResNet101 [47], i Faster-RCNN [14]. 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 (*engl. Spatial-Temporal Filtering - STF*). Proces poboljšava srednju prosječnu preciznost. Treniranje modela provodi se s pomoću prijenosnog učenja (*engl. transfer learning*) [59] koji se zatim fino podešava u TensorFlow-u te se primjenjuju, već navedene, duboke konvolucijske neuronske mreže. Nakon kritičke analize rezultata uočeno je da detektori nailaze na mnogo FP (*engl. False Positive*) detekcija i propuštaju mnoge TP (*engl. True Positive*) objekte čime opadaju performanse. 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 (*engl. Intersection Over Union*) koja se kasnije uspoređuje s IoU otkrivenih staništa. W je veličina vremenskog prozora koji određuje broj okvira u vremenskom intervalu. 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 3.15).



Slika 3.15 Algoritam preciznije detekcije iz rada [57]

Ugrađivanjem predložene metode s navedenim detektorima povećava mAP za 10%.

3.5.3. Ostali relevantni radovi

Rad [60] daje komparativnu analizu šest dobro poznatih dubokih konvolucijskih arhitektura koje se koriste kao ekstraktori značajki za klasifikaciju podvodnih fotografija. 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 (*engl. 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 klasifikaciju morskog otpada u podvodnim fotografijama, usporediti izvedbu konvencionalnih klasifikatora strojnog učenja treniranih na vektorima značajki ekstrahiranih dubokim konvolucijskim arhitekturama i istražiti različite načine za korištenje prijenosnog 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 fotografija iz baze „Deep-sea Debris Database” japanske agencije JAMSTEC (*engl. Japan Agency for Marine Earth Science and Technology*) kojoj su dodane fotografije iz „Google Images” baze te su koristili proces augmentacije podataka. Skup podataka je sadržavao 2395 fotografija podijeljenih u šest različitih klasa: metal, plastika, staklo, ostali otpad, bez otpada i guma. Izgrađena je i jednostavna neuronska mreža kako bi se usporedila izvedba unaprijed treniranih modela. Kada se treniraju s NN klasifikatorom na vrhu najbolje performanse pokazuju FT (*engl. Fine Tuning*) ekstraktori značajki temeljeni na Inceptionu, točnije Inception-ResNetV2 i InceptionV3, postižući ukupnu točnost veću od 90 %, 91,40 % odnosno 90,57 %. Tradicionalni SVM (*engl. Support Vector Machine*) i LR (*engl. Logistic Regression*) klasifikatori pokazali su se kao vjerodostojne alternative NN klasifikatoru koji često nadmašuju NN klasifikator. SVM treniran na značajkama Inception-ResNetV2 postiže točnost od 91.61 %, dok LR klasifikator treniran na značajkama InceptionV3 postiže točnost od 90.78 %. Dobiveni rezultati su dijelom takvi zbog načina snimanja fotografija za bazu podataka (korištena su podvodna vozila s rasvjetom) te pokazuju potencijal za daljnje iskorištavanje modela temeljenih na dubokom učenju za identifikaciju i klasifikaciju morskog otpada na podvodnim fotografijama.

Autori rada [61] predstavljaju, prvu takvu, bazu podataka „TrashCan” gdje se nalaze fotografije otpada na morskom dnu s podvodnih fotografija koje su skupljali iz raznih izvora, a ponajviše iz baze podataka „Deep-sea Debris dataset” japanske agencije za znanost i tehnologiju, JAMSTEC. Objekti na fotografijama su označeni i segmentirani kako bi se što bolje istrenirali modeli za detekciju podvodnog otpada (Slika 3.16). 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 pružaju osnovu za rad u budućnosti.

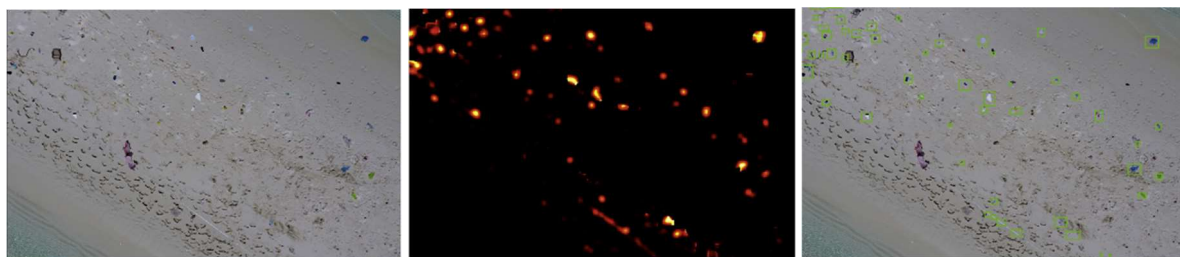


Slika 3.16 Prikaz rezultata detekcije i segmentacije na skupu podataka TrashCan

Autori rada [62] predstavljaju podatke o gustoći i sastavu antropogenog morskog otpada (*engl. Anthropogenic Marine Debris - AMD*) prikupljenih na 44 plaže duž obale Crvenog mora s pomoću prepoznavanja iz daljine. Fotografije su snimane iz zraka, s visine od 10 metara, komercijalnim dronom DJI Phantom. Na snimkama se lako mogu detektirati objekti od nekoliko centimetara jer fotografije imaju razlučivost manju od 0.5 cm / piksela. Fotografije su zatim automatski pregledane algoritmom za detekciju objekata „Faster R-CNN”. Algoritam je binarnom klasifikacijom razvrstao objekte na fotografije s otpadom i fotografije bez otpada 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 otpada s 44 plaže, sastav otpada s 15 plaža i ukupna masa plastike s 15 plaža. Podaci o vrsti, brzini vjetra i vegetaciji plaže omogućavaju predviđanje putanje i lokacije otpada. Dostupnost takvih podataka (gustoća, sastav i prediktori) mogu pomoći pri izgradnji globalne procjene antropogenog otpada u priobalju.

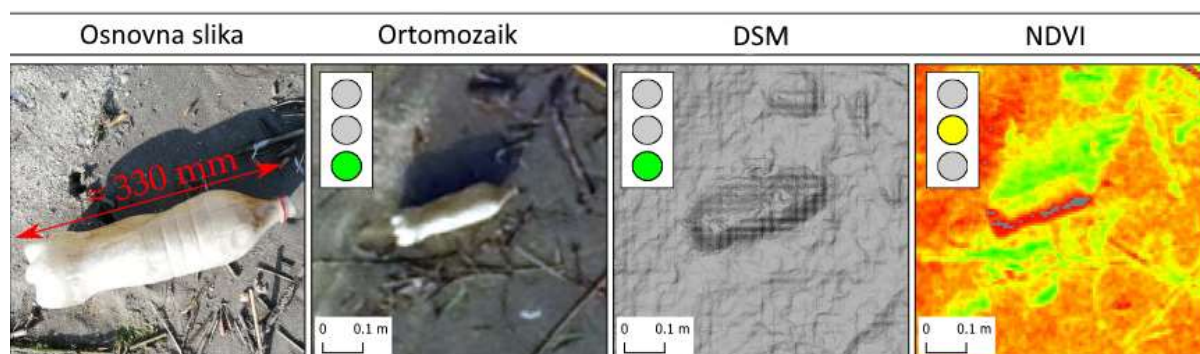
Autori rada [63] predstavljaju *ad hoc* metodologiju za praćenje i automatsko kvantificiranje antropogenog morskog otpada čiju glavninu čini plastika, a jedan je od najvećih ekoloških problema današnjice. Mjesta na kojima se akumulira najviše AMD-a su priobalna područja. Autori koriste komercijalni bespilotni zrakoplov koji je opremljen RGB kamerom visoke razlučivosti i primjenjuju model dubokog učenja kao što je „PlasticFinder” (Slika 3.17). 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, ponajprije zbog svoje male veličine, produljenog trajanja baterije i mogućnosti planiranja autonomnih letova putem korisnički

prilagođenog softvera zemaljske stanice. Autori su isto tako pokazali potencijal dubokog učenja za automatsku detekciju i kvantificiranje AMD-a.



Slika 3.17 Prikaz testne fotografije, PlasticFinder toplinska mapa danih piksela, PlasticFinder mapa s omeđujućim okvirima

Rad [64] predstavlja kako locirati i pratiti antropogeni morski otpad, *AMD*, poglavito makroplastiku kako bi se spriječila degradacija u mikroplastiku koja je potencijalno štetnija i teža za uklanjanje. Kao brzu metoda za dobivanje fotografija koriste UAV snimanje i istražuju strategije za mapiranje antropogenih morskih ostataka na plažama koristeći fotografije različite razlučivosti, informacije o elevaciji i podatke multispektralne kamere. Koristeći niz fotogrametrijskih alata AMD ručno su mapirani duž četiri do pet pojaseva gdje su prikupljali otpad s plaža i dalje ga obrađivali u laboratoriju gdje im dodjeljuju tri razine pouzdanosti i njihove klase. Podaci skupljani s pomoću infracrvenih (*engl. Near Infrared - NIR*) valnih duljina i digitalnim modelima površine (*engl. Digital Surface Model - DSM*) nisu pokazali značajno poboljšanje, čak i kad je uklonjena vegetacija s pomoću NDVI (*engl. Normalized Difference Vegetation Index*) karata, pri ručnom mapiranju. Pokazali su kako je za mapiranje AMD-a s plaža dobro, brzo, održivo i praktično rješenje korištenje RGB fotografija s prostornom razlučivosti od 0.5 cm / piksel, pogotovo za makro objekte (Slika 3.18).



Slika 3.18 Usporedba objekta fotografiranog dronom s ostalim metodama u radu [64]

3.6. 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 [61]. 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, fotografije postaju izobličene i mutne zbog svjetlosti koja se raspršuje i lomi pod morem, namreškana površina mora uslijed puhanja vjetrova, obraštaj na objektima [2].

Trenutno ne postoji ni jedna baza podataka s fotografijama podmorja koja je fotografirana bespilotnom letjelicom, a objekt promatranja je bio morski otpad. U nastavku se opisuju postojeće baze podataka dosadašnjih sličnih istraživanja i njihove značajke.

Kislik i suradnici su fotografirali područje rijeke dronovima s RGB kamerama na visini od 19 do 104 metra te su fotografirali ukupno 60 fotografija od kojih su izabrali 32 za obradu. Objekt promatranja bili su nitaste alge i makrofiti [50].

Politikos i suradnici koristili su bazu podataka „Integrirane kampanje informiranja i podizanja svijesti za smanjenje plastičnih vrećica u morskom okolišu” (*engl. LIFE DEBAG - LIFE14 GIE/GR/0011271*). Baza ima ukupno 635 fotografija, a fotografije su dobivene s pomoću podvodne kamere (TUC) koja se nalazi na malom plovilu [51].

Marin i suradnici [60] obradu podataka su provodili na JAMSTEC-ovoj bazi podataka koja je kompletirana fotografijama iz „Google Images” baze. Baza podataka je sadržavala 2395 fotografija.

Yabin i suradnici su koristili fotografije s „Google Eartha” i natjecateljske platforme „Kaggle” [65]. Obrađivali su fotografije čiji su promatrani objekti bili brodovi na površini mora prepoznati iz daljine s pomoću satelita. Baza podataka se sastoji od 150000 fotografija od kojih se na 42556 nalazio traženi objekt [39].

Lorencin i suradnici obrađivali su fotografije brodova s površine mora preuzetih s Google tražilice fotografija i Google satelita gdje im se baza sastojala od 5608 fotografija [40].

Lin i suradnici koristili su bazu podataka „URPC 2018.” koja se sastoji od 2901 fotografija za trening i 800 testnih podvodnih fotografija prikupljenih autonomnim podvodnim vozilima (AUV). Promatrani objekti su morski krastavci, ježinci, školjke i morske zvijezde [54].

Li i suradnici generirali su svoju bazu podataka s pomoću „WaterGan-a” koja se sastojala od fotografija fotografiranih iz zraka na odgovarajućoj dubini, u boji koje su realistične i sadrže floru i faunu [55].

Hong i suradnici predstavili su svoju bazu podataka „TrashCan” [66] koja je nastala prikupljanjem podataka iz različitih izvora poglavito od JAMSTEC [67] koja posjeduje bazu podataka s velikim brojem podvodnih videa prikupljenih autonomnim podvodnim vozilima (AUV). TrashCan skup podataka trenutačno se sastoji od 7212 označenih fotografija koje sadrže označene objekte podvodnog otpada, faune, flore i vozila na daljinsko upravljanje (*engl. Remotely Operated Vehicle - ROV*). Autori su prikazali dvije verzije navedenog skupa podataka: „TrashCan - Material” i „TrashCan - Instance” [61].

U idućem radu isti autori koriste bazu podataka koja sadrži fotografije prikupljene autonomnim podvodnim vozilima (AUV) koje prikazuju označeni podvodni otpad. Fotografije sadrže objekte raznih materijala, od plastike do metala, a snimane su u raznim podvodnim uvjetima [2].

Fabbri i suradnici [56] su koristili bazu podataka „ImageNet” [7] i videa ronilaca preuzetih s kanala „YouTube”. „ImageNet” je baza slikovnih podataka koji su organizirani prema „WordNet” hijerarhiji. Svaki pojam, skup sinonima tj. sinset (*engl. synset*) kojih ima više od 100000 u ovoj bazi, se nastoji opisati s prosječno 1000 fotografija. Sve fotografije su kontrolirane kvalitete i označene. Autori trenutačno mogu ponuditi korisnicima 3.2 milijuna fotografija, a želja im je ponuditi više desetaka milijuna fotografija koje će biti sortirane i označene za većinu koncepata u „WordNetu” [68].

Baza podataka Martina i suradnika [69] se sastojala od oko tisuću fotografija uzetih s pomoću komercijalnog drona koji je fotografirao plaže duž obale Saudijske Arabije [62].

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 fotografija [41].

Baza podataka „Enhancing underwater visual perception” (EUVP) [70] je skup neuparenih i uparenih fotografija loše i dobre kvalitete. Autori su stvorili takvu bazu koja olakšava treniranje modela za detekciju objekata s podvodnih fotografija. Upareni skup podataka pripremljen je postupkom koji su predložili Fabbri i suradnici [56] te je poboljšan i ažuriran. Upareni skup sadrži 24 840 fotografija dok neupareni skup podataka sadrži 6665 fotografija od toga njih

3195 loše kvalitete, a 3140 dobre kvalitete te 330 testnih fotografija. Objekti na fotografijama su flora, fauna i roniaci.

Naseer i suradnici koristili su bazu podataka „Gulf of Cadiz” vlasništva Španjolskog oceanografskog instituta koja je sadržavala podvodni video. Od 100 000 sličica odabrali su one koje sadrže podvodna staništa, jazbine vrste *Nephrops Norvegicus*, njih 200 [57].

Primjećujemo da se navedene baze podataka razlikuju po mnogočemu:

- a) Području uzorkovanja: površina mora, vodeni stupac, morsko dno, plaža.
- b) Načinu uzorkovanja: satelitom, autonomnim podvodnim vozilom (AUV), podvodnim vozilom na daljinsko upravljanje (ROV), bespilotnom letjelicom - dronom (UAV), zrakoplovom s posadom, manualno.
- c) Objektu promatranja: plovila, podmorna flora i fauna, morski otpad.

Tablica 3.1 sistematizira navedene baze s razlikama u području i načinu uzorkovanja te po objektu promatranja. Za svaku bazu je naveden i broj fotografija koje su autori koristili.

Tablica 3.1 Usporedba postojećih baza podataka i njihovih značajki

Autori	Područje uzorkovanja	Način uzorkovanja	Objekt promatranja	Broj fotografija
Kislik i suradnici[50]	Površina rijeke	Dron s RGB kamerom	Nitaste alge i makrofiti	60
Politikos i suradnici[51]	Vodeni stupac i morsko dno	Vučena podvodna kamera (TUC)	Morski otpad	635
Yabin i suradnici[39]	Površina mora	Satelit	Brodovi	42 556
Lorencin i suradnici[40]	Površina mora	Satelit	Brodovi	5 608
Lin i suradnici[54]	Morsko dno	Podvodno vozilo (AUV)	Podvodna fauna	3 701
Li i suradnici[55]	Morsko dno	Manualno (ronilac)	Podvodna fauna	6 083

Hong i suradnici[66]	Morsko dno	Podvodno vozilo na daljinsko upravljanje (ROV)	Podvodna flora, fauna i otpad	7 212
Hong i suradnici[2]	Morsko dno	Autonomno podvodno vozilo (AUV)	Otpad	1 058
Martin i suradnici[69]	Plaža	Bespilotna letjelica (UAV)	Otpad	~ 1 000
Garcia-Garin i suradnici[41]	Površina mora	Bespilotna letjelica (UAV) i zrakoplov s posadom	Otpad	3 723
Islam i suradnici[70]	Vodeni stupac i morsko dno	Manualno (ronilac)	Podvodna flora, fauna i ljudi	6 665
Naseer i suradnici[57]	Morsko dno	Autonomno podvodno vozilo (AUV)	Podvodna fauna	200

Kompletan i sveobuhvatan skup podataka specifičan za podvodno snimanje ne postoji zbog složenosti procesa prikupljanja podvodnih fotografija. Većina dostupnih podvodnih baza podataka suočava se s ograničenjima poput nedostatka raznolikosti kategorija, usmjerenosti na specifične ciljane objekte te nedostataka popratnih informacija o uvjetima fotografiranja. Nijedna od dostupnih baza podataka, uključujući one koje obuhvaćaju otpad općenito, otpad na morskom dnu te bentičku floru i faunu, ne pruža fotografije koji odgovaraju specifičnim zahtjevima ovog istraživanja – otpad na morskom dnu snimljen bespilotnom letjelicom iz zraka.

3.6.1. Baze podataka za poboljšanje i obnavljanje podvodnih fotografija

Učinkovita detekcija objekata na podvodnim fotografijama zahtijeva pretprocesiranje fotografija. U tu svrhu analizirane su postojeće baze podataka koje se mogu iskoristiti za podvodne metode pretprocesiranja. Međutim, većina dostupnih baza podataka nije primjenjiva

za zadatke detekcije objekata zbog nedostatka pripadajućih oznaka objekata, kao što je prikazano u nastavku:

a) Skup podataka za poboljšanje podvodnih fotografija u stvarnom svijetu (RUIE)

Podaci RUIE [71] za unaprjeđenje stvarnih podvodnih fotografija su obimna zbirka koja broji 4000 fotografija snimljenih iz različitih perspektiva. Ovaj skup podataka je posebno organiziran za korištenje s generativnim suparničkim mrežama za unaprjeđenje podvodnih fotografija (UIE) i podijeljen je u tri manja skupa: „Skup za procjenu kvalitete podvodnih fotografija” (UIQS), „Skup za podvodne boje” (UCCS) i „Podvodni testni set usmjeren na specifične zadatke” (UHTS). Ovi skupovi služe za različite ciljeve poput obnavljanja boja, poboljšanja vizualnog prikaza te podrške u detekciji i klasifikaciji računalnog vida na višem nivo.

b) Skup podataka mjerila poboljšanja podvodne fotografije (UIEB)

Zbirka podataka za procjenu unaprjeđenja podvodnih fotografija (UIEB) [72] uključuje 950 stvarnih podvodnih snimaka od kojih 890 ima pripadajuće referentne fotografije kojima se može uspoređivati kvaliteta. Dodatnih 60 fotografija namijenjeno je za testiranje. Koristi se u algoritmima za kvantitativno i kvalitativno poboljšanje podvodnih fotografija. UIEB kolekcija obuhvaća širok spektar razlučivosti i zastupa različite kategorije poput scenskih kompozicija i primarnih predmeta u kadru.

c) Skup podataka o poboljšanju podvodne vizualne percepcije (EUVP)

Zbirka podataka za unapređenje percepcije podvodne fotografije (EUVP) [70] predstavlja opsežnu kolekciju koja sadrži podvodne snimke niske i visoke razlučivosti, namijenjene za korištenje u okviru suparničkog nadziranog učenja.

d) Skup podataka „Underwater test dataset“ (U-45)

Testna zbirka podataka U-45 [73] obuhvaća 45 pažljivo selektiranih podvodnih snimki, izdvojenih iz veće kolekcije stvarnih podvodnih fotografija. Ovaj javno dostupan skup karakterizira prisutnost slaba kontrasta, izmijenjenih boja i maglovitih efekata koji umanjuju kvalitetu fotografije [74].

e) Skup podataka „Jamaica Port Royal“

Zbirka podataka Jamaica Port Royal [55] sastoji se od 6500 fotografija snimljenih ručnim ronilačkim uređajem u Port Royalu na Jamajci, točnije na lokaciji potonulog grada gdje se

nalaze raznolike prirodne i antropogene strukture. Svi su kadrovi uhvaćeni tijekom jedinstvenog ronilačkog izleta, pri čemu je kamera bila na maksimalnoj udaljenosti od 1.5 metara od morskog dna.

f) Skup podataka morske autonomne robotike za intervencije (MARI)

Zbirka podataka MARI (*engl. Marine Autonomous Robotics for Interventions*) [75] usmjeren je na unaprjeđenje i razvoj suradnje autonomnih podvodnih vozila (AUV) namijenjenih za intervencije na morskim platformama, radove spašavanja, potrage kao i za širok spektar znanstvenih istraživačkih zadataka. Ova zbirka uključuje raznovrsne podvodne videozapise i fotografije zabilježene stereo-vizijskim sustavima postavljenim ispod površine vode.

g) Skup podataka „Modular Optical Underwater Survey System“ (MOUSS)

Zbirka podataka MOUSS [76] nastala je kroz korištenje stacionarne kamere smještene na morskom dnu. Kamera je, u uvjetima ambijentalnog osvjetljenja i na udaljenosti od 1 do 2 metra, zabilježila 159 fotografija riba i drugih značajnih podvodnih objekata. Korištena kombinacija testnih podataka uključivala je kako fotografije prikupljene tijekom faze treniranja tako i one iz novijih prikupljanja.

h) Skup podataka „Monterey Bay Aquarium Research Institute“ (MBARI)

Zbirka podataka MBARI [76], koju je zabilježio „Monterey Bay Aquarium Research Institute“, obuhvaća 666 fotografija riba i drugih bitnih podvodnih entiteta, prikupljenih iz mnoštva različitih područja.

i) Skup podataka „Alaska Fisheries Science Center“ (AFSC)

Zbirka podataka AFSC [76] sastavljena je iz snimaka dobivenih s ROV-a koji su bili smješteni pod vodom te su bili opremljeni RGB kamerama. Zbirka obuhvaća raznovrsne videozapise snimljene tijekom različitih misija ROV-a, uključujući ukupno 571 fotografiju.

j) Skup podataka „Northwest Fisheries Science Center“ (NWFSC)

Zbirka podataka NWFSC [76] nastala je korištenjem daljinski upravljanih vozila usmjerenih prema morskom dnu. U navedenoj zbirci nalazi se 123 fotografije riba i drugih objekata zabilježenih u neposrednoj blizini morskog dna.

k) Skup podataka „Red-Green-Blue depth“ (RGBD)

Zbirka podataka RGBD [76], koja je kreirana s namjenom obnavljanja i optimizacije podvodnih fotografija, uključuje skup od više od 1100 fotografija koje predstavljaju voodootporne karte boja unutar okoliša pod morem.

l) Skup podataka „Fish4knowledge“

Zbirka podataka Fish4knowledge [77] formirana je od snimki riba izvađenih iz video zapisa u stvarnom vremenu i obuhvaća 27 370 fotografija riba. Ova opsežna zbirka segmentirana je u 23 skupine, pri čemu svaki segment odgovara posebnoj vrsti ribe.

m) Skup podataka markera divljih riba

Zbirka podataka o označavanju divljih riba [78] prikupljena je uz pomoć daljinski upravljanih vozila, operirajući u različitim uvjetima oceana. Fotografije riba u ovoj kolekciji su kategorizirane prema kaskadnim klasifikatorima koji koriste Haarove slične karakteristike. Zbirka podataka uključuje označene podatke za trening i validaciju kao i zasebne podatke za testiranje.

n) Skup podataka HabCam

Zbirka podataka HabCam [76] skupljena je kroz analizu podvodnih snimaka morskog dna snimljenih s pomoću vozila HabCam. Ovo vozilo prešlo je preko oceanskog dna snimajući šest fotografija svake sekunde. Fotografije su od izuzetne važnosti za istraživanje ekosustava i napredak oceanografije.

o) Skup podataka o podvodnim fotografijama Port Royala

Zbirka podataka podvodnih fotografija Port Royala [55] oblikovana je upotrebom „Generativnih Suparničkih Mreža” (*engl. Generative Adversarial Network - GAN*) kako bi se stvorile visokokvalitetne, realistične podvodne fotografije. Snimanje ovih fotografija obavljeno je korištenjem kamera na autonomnim i daljinski upravljanim vozilima što je omogućilo dobivanje podvodnih fotografija visoke razlučivosti. Baza se sastoji od umjetno generiranih fotografija te podvodnih fotografija koje su korištene za trening.

p) Skup podataka OUCVISION

Baza podataka OUCVISION [79] je širokoobuhvatna kolekcija korištena za unaprjeđenje i obnovu podvodnih fotografija, s posebnim fokusom na prepoznavanje i detekciju istaknutih objekata. Obuhvaća 4400 fotografija koje prikazuju 220 različitih objekata. Svaki objekt je zabilježen iz četiri različita smjera - desno, lijevo, sa stražnje strane i sprijeda - te iz pet

različitih prostornih pozicija - donji desni kut, donji lijevi kut, centar, gornji desni kut i gornji lijevi kut - rezultirajući ukupno s 20 fotografija po objektu.

q) Baza podataka fotografija podvodnih stijena

Kolekcija podataka podvodnih stijena [76] usmjerena je na poboljšanje i rekonstrukciju podvodnih fotografija. Proces stvaranja uvjerljivih i realističnih podvodnih fotografija temeljio se na upotrebi GAN-a. Baza podataka se sastoji od generiranih i stvarnih fotografija.

r) Baza podataka o podvodnim fotografijama riba

Kolekcija podataka podvodnih fotografija riba [76] sastoji se od snimki koraljnih grebena i morskog života zabilježenih na raznim lokacijama širom svijeta, uključujući Indijski ocean, Crveno more i druge. Ovaj skup podataka uključuje različite vrste grebenskih riba kao što su papagaj-ribe, leptir-ribe, anđeoske ribe i škarpine kao i druge vrste poput koralja i hobotnica.

3.7. Ograničenja podvodnog fotografiranja

Istraživanje podmorskog svijeta putem fotografija je ključno, a provodi se kroz upotrebu iskusnih ronilaca, optičkih kamera, specijalizirane opreme ili ROV-ova. Osim optičkih kamera, spomenuti pristupi imaju niz mana, poput ograničenja u vidnom polju, dubinskim kapacitetima i složenosti operacija. Zbog nepredviđenih nelinearnih hidrodinamičkih utjecaja i nepostojanja preciznog modela upravljanje ROV-om je kompleksno. Podvodna istraživanja prate visoki troškovi zbog potrebe za angažmanom visoko kvalificiranih ronilaca i nadzornika. Također je pod vodom dopušteno provesti samo ograničeno vrijeme što se posebno odnosi na duge inspeksijske zarone. To sve dovodi do produljenja vremena potrebnog za istraživanje. Međutim, korištenje tehnika za unaprjeđenje podvodnih fotografija može značajno smanjiti ove poteškoće. Uzevši u obzir ranije navedene izazove, ograničenja u analizi podvodnih fotografija mogu se podijeliti na one koji proizlaze iz okolišnih faktora i one na samim fotografijama što će biti pobliže razrađeno u sljedećim poglavljima.

3.7.1. Uzroci degradacije kvalitete

Ograničenja povezana s podvodnim uvjetima odnose se na različite aspekte koji utječu na kvalitetu podvodnih fotografija i video materijala. To uključuje faktore poput specifičnosti korištene opreme, pojave refrakcije svjetlosti, problema s neuniformnim osvjetljenjem, učincima kretanja, disperzije svjetlosnih zraka i upijanja svjetlosti u vodi. Svi navedeni elementi mogu rezultirati stvaranjem mutnih i degradiranih vizualnih zapisa pod vodom.

a) Oprema

Snimanje podvodnih fotografija može se postići s pomoću dvije glavne vrste kamera. Prva mogućnost uključuje prilagođavanje tradicionalne kopnene kamere dodavanjem voodootpornog kućišta. Takvo kućište obično se brtvi silikonskim materijalom namijenjenim ronjenju kako bi se osigurala njegova nepropusnost. Ovo je idealna opcija za one koji već posjeduju kvalitetnu kopnenu kameru, ali ne žele ili ne mogu investirati u skupu specijaliziranu podvodnu kameru. Alternativno, fotograf može odabrati specijaliziranu podvodnu kameru koju odlikuje širok raspon kvalitete i cijene, namijenjenu upravo za te svrhe [74].

b) Refrakcija

Refrakcija je pojava savijanja svjetlosti pri njenom prijelazu iz jednog okruženja u drugo [80]. Kada svjetlost koja se odbija od predmeta putuje kroz vodu pa zatim kroz objektiv i zrak unutar podvodnog fotoaparata, dolazi do karakterističnog efekta: predmeti izgledaju oko 25 % veći i bliži, negoli što to stvarnost nalaže. Refrakcija može predstavljati prepreku za precizno fokusiranje na subjekt što može rezultirati zamućenim fotografijama. Međutim, refrakcija se također može koristiti u procesu rekonstrukcije podvodnih scena kako bi se stekao bolji vizualni dojam istih [81], [82].

c) Nejednoliko osvjetljenje

Kako svjetlost prolazi kroz morske dubine, njezin intenzitet slabi s povećanjem dubine. Prirodno svjetlo nije stalno dostupno i varira tijekom dana; primjerice, kada je sunce točno iznad, površina mora reflektira najmanji dio svjetlosti. Vremenski uvjeti također mogu biti ključan faktor u raspoloživosti svjetlosti - olujno vrijeme i nemirno more mogu drastično promijeniti svjetlosne uvjete pod morem. Postoje raznoliki algoritmi dizajnirani za adresiranje problema neujednačenog osvjetljenja [83], [84], [85].

d) Pomicanje

Zamućenje izazvano kretanjem javlja se kada postoji promjena u relativnom položaju između opreme za snimanje i subjekta kojeg se fotografira, uslijed pokreta ili kada podvodne struje utječu na obje komponente. Ovakva dinamika između fotoaparata i subjekta, s pokretnim vodenim masama, rezultira zamućenjem fotografije zbog pokreta [86], [87], [88]. To dovodi do distorzije u podvodnim fotografijama, narušavajući kvalitetu fotografije i raspon tonova. Snimanje statičnih objekata poput koralja ili stijena je stoga obično jednostavnije od

fotografiranja objekata u pokretu. Za ispravljanje zamućenja prouzročenog kretanjem razvijeni su mnogi algoritmi koji pomažu u poboljšanju oštine i kvalitete fotografija [89].

e) Raspršenje

Raspršenje opisuje proces kroz koji lebdeće čestice u vodi odbijaju svjetlost pod različitim kutovima i to na specifičnoj valnoj duljini [90], [91], [92]. Kao rezultat, raspršenje umanjuje kontrast i smanjuje vidljivost, zamagljuje fine detalje i pridonosi generalnoj nejasnoći na podvodnim fotografijama.

f) Apsorpcija

Apsorpcija svjetlosti označava gubitak intenziteta elektromagnetskog zračenja prilikom prolaska kroz medij, pri čemu se dio energije pretvara u toplinu [93], [94]. Zbog svojih fizikalnih svojstava, intenzitet svjetlosti eksponencijalno opada s dubinom - često se navodi da se intenzitet svjetlosti prepolovljuje svakih 10 metara dubine. Dodatno, količina dostupnog svjetla varira ovisno o dobu dana i drugih vanjskih čimbenika poput vremenskih uvjeta i stanja na morskoj površini. Fotografije pod vodom obično se toniraju u zelenkastu ili plavkastu boju zbog upijanja boje, budući da voda apsorbira svjetlost veće valne duljine poput žute, crvene i narančaste, rezultirajući gubitkom tih boja na fotografijama.

3.7.2. Posljedice degradacija kvalitete

Faktori koji ograničavaju kvalitetu podvodnih snimki često rezultiraju degradiranim fotografijama sa smanjenom jasnoćom. Ključno je, stoga upotrijebiti algoritme posebno osmišljene za restauraciju i poboljšanje podvodnih fotografija i videozapisa. Ova ograničenja obuhvaćaju probleme kao što su nizak kontrast, šum, zamućenje, loša vidljivost i prisutnost izmaglice. Pristupanje ovim izazovima s pomoću odgovarajućih tehnika obrade obično je lakše nego mijenjanje samih okolišnih ograničenja koja utječu na snimanje [74].

a) Nizak kontrast

Kontrast predstavlja mjerilo razlike u boji ili osvjetljenju koje omogućava razlučivanje pojedinih objekata unutar istog vidnog polja. Smanjeni kontrast uzrokovan slabim osvjetljenjem, raspršenjem svjetlosti te apsorpcijom negativno utječe na kvalitetu podvodnih snimaka, umanjujući njihovu jasnoću, vidljivost i detalje. Zbog toga je proces poboljšanja kontrasta ključan korak u procesiranju podvodnih fotografija, pri čemu je važno odabrati između poboljšanja globalnog ili lokalnog kontrasta. Lokalno poboljšanje kontrasta uključuje

podjelu fotografije na manje regije i provedbu optimizacije kontrasta unutar svake regije zasebno, dok globalno poboljšanje kontrasta podrazumijeva primjenu tehnika povećanja kontrasta na cijelu fotografiju zajedno [74].

b) Šum

Šum u slici predstavlja neregularne promjene u boji ili osvjetljenju piksela koje nisu dio stvarne scene [95], [96]. On narušava jasnoću podvodnih snimki, a među različitim vrstama šuma izdvajaju se:

- Šum „soli i papra”: Karakterizira ga prisutnost izoliranih svijetlih i tamnih točaka u slici koje odstupaju od okolnih tonova.
- Gaussov šum: To je najčešća vrsta šuma koja ima statističke karakteristike i distribuciju vjerojatnosti koja slijedi normalnu distribuciju.
- Šum u fiksnom načinu rada: Specifične podvodne smetnje poput raspršenja svjetlosti zbog suspendiranih čestica, mutnoće i turbulencije mogu narušiti kvalitetu fotografije.

Zbog negativnog utjecaja šuma na kvalitetu podvodnih fotografija, tehnike za njegovo uklanjanje su ključne za poboljšanje vizualne kvalitete materijala.

c) Zamućenje

U kontekstu obrade fotografija, zamućenje se ponekad namjerno primjenjuje kako bi se omekšali rubovi i postigla glatkoća fotografije. Međutim, u svijetu podvodne fotografije, zamućenje degradira kvalitetu snimka mutnim detaljima i otežava interpretaciju [97], [98]. S obzirom na to da se svjetlost pod vodom raspršuje i apsorbira, podvodne fotografije često pate od ozbiljne deformacije uzrokovane zamućenjem i promjenom boje što drastično utječe na njihovu jasnoću. Ove mane otežavaju bilo kakvu daljnju obradu fotografija kao što je prepoznavanje objekata, klasifikacija i segmentacija. Stoga su tehnike za uklanjanje zamućenja iznimno važne i široko koriste se za čišćenje i obnavljanje podvodnih snimaka.

d) Loša vidljivost

Vidljivost označava stupanj u kojem su objekti uočljivi ljudskim okom ili udaljenost na kojoj se svjetlost i objekti još uvijek mogu razlikovati. Osiguravanje vidljivosti objekata, bilo da se nalaze na malim ili velikim udaljenostima u podvodnim scenama, predstavlja značajan izazov kod analize i obrade fotografija. Svjetlost koja se pod vodom raspršuje te apsorpcija svjetlosti uzrokovana šumom i suspendiranim česticama umanjuju vidljivost, stvarajući velike poteškoće u primjenama računalnog vida. Fenomen prigušenja svjetlosti ograničava efektivni domet

gledanja na približno dvadeset metara u čistim vodama, a u mutnijim uvjetima taj se domet smanjuje na pet metara ili čak manje. Pronalaženje rješenja za probleme u primjeni fotografija nastalih zbog slabe vidljivosti objekata na različitim dometima unutar podvodnih scena je ključno. U tu svrhu provedena su brojna istraživanja usmjerena na poboljšanje vidljivosti na podvodnim fotografijama [95], [99].

e) Zamagljenost

Snimke smanjene vidljivosti, koje se često bilježe u maglovitim ili zamagljenim uvjetima, kompromitirane su zbog apsorpcije svjetlosti i njezina raspršenja. Rezultat su fotografije s niskim kontrastom i ograničenom vidljivošću što otežava identifikaciju objekata ljudskim okom. S obzirom na to da raznoliki efekti koji zamagljuju podvodne snimke, procesi usmjereni na poboljšanje kvalitete navedenih podvodnih fotografija imaju veliku važnost [100], [101].

3.8. Pretprocesiranje fotografija

Sve je veća potreba za fotografijama visoke kvalitete radi boljeg razumijevanja i analize podmorskog svijeta. Brojna istraživanja su usmjerena na proučavanje tehnika obrade podvodnih fotografija. Obrada podvodne fotografije obično se dijeli na dvije glavne kategorije: restauraciju i unaprjeđenje. Ključna razlika među njima leži u tome što se restauracija oslanja na matematički model nastajanja fotografije (*engl. Image Formation Models - IFM*) dok tehnike unaprjeđenja ne uzimaju u obzir izvorni IFM. Jednostavni IFM je poznat kao standardni i efikasan model u restauraciji podvodnih fotografskih snimaka koji se matematički može definirati kao u jednadžbi 3.1:

$$I_c(x) = J_c(x) \cdot t_c(x) + B_c \cdot (1 - t_c(x)), \quad c \in \{r, g, b\} \quad (3.1)$$

gdje je $I_c(x)$ opaženi intenzitet u boji c na poziciji x , $J_c(x)$ stvarna boja (radiancija scene), $t_c(x)$ prijenosna funkcija (koliko svjetla dođe do kamere), B_c pozadinsko (ambijentalno) svjetlo u kanalu c [102].

3.8.1. Tehnike poboljšanja kvalitete podvodnih fotografija

Tehnike poboljšanja kvalitete podvodnih fotografija koristi se oslanjajući na postulate fizičkih modela, gdje se izgrađuje model koji se zasniva na razumijevanju procesa fizičkog degradiranja fotografije i temeljnih zakona o širenju svjetlosti. Ključni korak u ovoj metodi uključuje dedukciju ključnih parametara modela koje podupire postojeće znanje i potom se provodi inverzija procesa degradacije za dobivanje restaurirane fotografije [103]. Tehnike poboljšanja kvalitete se mogu grupirati u dvije osnovne kategorije: hardverske i softverske. Poboljšanje zasnovano na hardveru podijeljeno je na tri grupe koje uključuju polarizacijske metode, stereo fotografiranje i tehnike temeljene na ograničenoj dubini vidnog polja. S druge strane, programski pristupi su podijeljeni na: tehnike bazirane na optičkim karakteristikama fotografija, mjere temeljene na prethodnom znanju i metode restauracije koje koriste algoritme dubokog učenja [74].

3.8.1.1. Hardverske metode poboljšanja kvalitete fotografija

Istraživanje i praćenje podvodnog svijeta često zahtijeva upotrebu niza specijaliziranih hardverskih uređaja koji se također koriste za poboljšanje kvalitete podvodnih fotografija. Metode temeljene na hardveru uključuju upotrebu različitih fizičkih komponenti specifično dizajniranih za restauraciju podvodnih fotografija. Takve metode obuhvaćaju uporabu lasera,

raznih senzora, polarizacijskih filtera, ROV-ova, zrcaljenja, tehnika stereo snimanja i tehnike fotografiranja s vremenski ograničenim rasponom ekspozicija. Polarizacijska obrada, primjerice, koristi se za efikasno smanjenje efekata povratnog raspršenja svjetlosti, a ostvaruje se kroz upotrebu polariziranog izvora svjetla prilikom fotografiranja ili izborom kamere s integriranim polarizacijskim filterima. Laserske metode pomažu u eliminaciji povratnog raspršenja aktiviranjem kamere u sinkronizaciji s laserom koji „prореže“ vodu. Voodootporni senzori se koriste za otkrivanje čestica i organizama koji plutaju kroz vodu kako bi se smanjilo odražavanje svjetlosti s tih objekata. Podvodni stativi se postavljaju na morsko dno radi stabilizacije kamere, omogućavajući time jasnije snimke [74].

3.8.1.2. Softverska restauracija podvodnih fotografija

Softverska restauracija je metoda obrade fotografije koja se ne oslanja direktno na fizičke pristupe, već teži konstruiranju matematičkog modela fotografije i izračunavanju parametara relevantnih za taj model. Algoritmi potrebni za ovu vrstu restauracije implementirani su unutar softvera koji služi za rekonstrukciju i poboljšanje kvalitete podvodnih fotografija. Softverska restauracija je podijeljena u tri ključne kategorije: tehnike temeljene na optičkim karakteristikama, metode koje koriste postojeće znanje o sceni ili objektima i pristupi koji se zasnivaju na dubokom učenju. Nasuprot hardverskim metodama, softverske imaju nekoliko prednosti: tipično zahtijevaju manje vremena za obradu, omogućuju jednostavniju modifikaciju procesa restauracije, nude veću fleksibilnost u dizajnu te obično imaju niže povezane troškove [74].

a) Optički modeli

Model podvodne optičke fotografije (*engl. Underwater Optical Imaging - UOI*) omogućuje dobivanje prirodnih i čistih podvodnih fotografija zahvaljujući kreiranju i primjeni osnovnog modela optičke fotografije, uz korištenje postupka inverznog inženjeringa za poništavanje učinka degradacije fotografije. Ukupni gubitak svjetlosti u vodi pri određenoj valnoj duljini može se opisati koeficijentom oslabljenja $c(\lambda)$, koji se sastoji od koeficijenata apsorpcije $a(\lambda)$ i raspršenja $b(\lambda)$ kao što je prikazano u jednadžbi 3.2:

$$c(\lambda) = a(\lambda) + b(\lambda) \quad (3.2)$$

gdje je λ valna duljina svjetlosti, $a(\lambda)$ koeficijent apsorpcije, $b(\lambda)$ koeficijent raspršenja, $c(\lambda)$ ukupni koeficijent oslabljenja (ukupni gubitak svjetlosti zbog apsorpcije i raspršenja).

Raspršenje se dalje kvantificira funkcijom prostornog raspršenja $\beta(\theta, \lambda)$, koja opisuje kutnu distribuciju raspršene svjetlosti kao u jednadžbi 3.3:

$$\beta(\theta, \lambda) = \lim_{\Delta r \rightarrow 0} \lim_{\Delta \Omega \rightarrow 0} \frac{B(\theta, \lambda)}{\Delta r \Delta \Omega} = \lim_{\Delta r \rightarrow 0} \lim_{\Delta \Omega \rightarrow 0} \frac{\phi_s(\theta, \lambda)}{\phi_i \Delta r \Delta \Omega} \quad (3.3)$$

gdje je $\beta(\theta, \lambda)$ funkcija prostornog raspršenja, θ kut raspršenja, $B(\theta, \lambda)$ količina svjetlosti raspršene u kut θ unutar prostornog kuta $\Delta \Omega$, ϕ_s raspršena energija, ϕ_i incidentna energija, Δr debljina sloja vode.

Integriranjem β po svim smjerovima kao što je prikazano u jednadžbi 3.4:

$$b(\lambda) = \int \beta(\theta, \lambda) d\Omega = 2\pi \int_0^\pi \beta(\theta, \lambda) \sin \theta d\theta \quad (3.4)$$

dobiva se ukupni koeficijent raspršenja $b(\lambda)$ [104].

Postoji mnogo primjena za podvodno optičko fotografiranje kao što su brodske detektorske jedinice, zračne, površinske i podvodne optičke kamere. Trucco i suradnici [105] razvili su tehniku restauracije fotografije koja se automatski prilagođava, bazirajući se na Jaffe-McGlamery UOI modelu [106], [107] gdje se optimalni parametri filtra izračunavaju za pojedinačne fotografije na temelju globalnog poboljšanja kontrasta. Primijenjeni pojednostavljeni fizički model je efikasan za raspršeno svjetlo uz minimalno povratno raspršivanje i razne uvjete snimanja. Tehnika se oslanja na postavku da su podvodne fotografije pod utjecajem i frontalne disperzije te homogenih svjetlosnih uvjeta.

Hou i suradnici [108] predstavili su pristup restauraciji podvodnih fotografija koji se bazira na UOI modelu, polazeći od pretpostavke da zamagljivanje na podvodnim fotografijama proizlazi iz disperzije suspendiranih čestica i organizama. Tehnika uzima u obzir karakteristike podvodnog okruženja prikupljene iz različitih domena kao što su prostorna i frekvencijska te koristi funkciju širenja točke (*engl. Point Spreading Function - PSF*) i funkciju modulacije transfera (*engl. Modulation Transfer Function - MTF*). Dekonvolucijski proces temelji se na procjeni parametara disperzije svjetlosti pri obnovi fotografija. Boffety i suradnici [109] razvili su simulacijski alat za obnovu boje koji koristi podatke dobivene iz podvodnog optičkog fotografiranja, analizirajući utjecaj spektralne diskretizacije na reproduciranje boja i pokazali su da, kad su dostupni samo RGB podaci.

Wen i suradnici [110] predstavili su tehniku podvodnog optičkog fotografiranja koja se temelji na fizikalnim procesima pri formiranju fotografija i nakon toga primijenili su algoritam za

poboljšanje kvalitete fotografija koristeći novi UOI tamni kanal kao metodologiju za izračunavanje brzine raspršivanja i pozadinskog svjetla. Pokazalo se da je ova metoda učinkovita u restauraciji podvodnih snimaka. U kontekstu uzorkovanja, Ahn i suradnici [111] predstavili su SAUV (*engl. Soft-Robotic Autonomous Underwater Vehicle*) sustav za prijenos fotografija te su demonstrirali njegovu efikasnost u stvarnim uvjetima na otvorenom moru koristeći podvodnu optičku tehniku za autonomna vozila čime su poboljšali rezultate detekcije podvodnih objekata.

b) Restauracija fotografija temeljena na prethodnom znanju

Degradacija kvalitete podvodnih fotografija prouzročena je povratnim raspršivanjem svjetlosti, prisutnošću suspendiranih čestica u vodi i apsorpcijom svjetlosti. Brojne tehnike restauracije koriste znanje akumulirano iz prethodnih iskustava kako bi se obnovila kvaliteta fotografija. Takve metode uključuju prilagodbe zasnovane na modelu tamnog kanala (*engl. Dark Channel Prior - DCP*) [112], [113] prilagodbu specifičnu za podvodne uvjete zasnovanu na tamnom kanalu (*engl. Underwater Dark Channel Prior - UDCP*) [114], [115], prioritet najvećeg intenziteta svjetlosti (*engl. Maximum Intensity Prior - MIP*) [116], prilagodbu temeljenu na crvenom kanalu (*engl. Red Channel Prior - RCP*) [117] i prioritet koji uzima u obzir slabljenje podvodne svjetlosti (*engl. Underwater Light Attenuation Prior - ULAP*) [118]. Nadolazeći pododjeljci će detaljnije objasniti različite vrste ovih metodologija baziranih na prethodnom znanju, koje se implementiraju u procesima restauracije podvodnih fotografija [74].

- Metoda tamnog kanala (DCP)

Tehnike odmagljivanja fotografija, kao što su one koje koristi metoda DCP predstavljena od strane [112], alat su za rješavanje problema zamućenosti koji znatno smanjuje vidljivost i mijenja boje na fotografijama. Ovakvo stanje nije samo estetski nezadovoljavajuće već utječe i na pouzdanost primjena računalnog vida, poput detekcije objekata, vanjskog nadzora i zračnih snimki. Zato je uklanjanje zamućenja iz fotografija esencijalan korak u polju računalnog vida i grafike. DCP tehnika se posebno primjenjuje na poboljšanje podvodnih fotografija, koristeći se činjenicom da jasne i čiste podvodne fotografije tipično posjeduju određene piksele s izuzetno niskim intenzitetom u bar jednom kanalu boje. Za svaki piksel x , tamni kanal $J_{dark}(x)$ slike J definiran je kao što je prikazano u jednadžbi 3.5:

$$J_{\text{dark}}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{r, g, b\}} J^c(y) \right) \quad (3.5)$$

gdje je $\Omega(x)$ lokalno područje oko piksela x , c kanal boje, a $J^c(y)$ intezitet boje c na lokaciji y u slici J . Ako se pretpostavi da je $J_{\text{dark}} \approx 0$, može se izvesti procjena prijenosne funkcije $t(x)$ kao što je prikazano u jednadžbi 3.6:

$$t(x) = 1 - \min_{y \in \Omega(x)} \left(\min_{c \in \{r, g, b\}} \frac{I^c(y)}{A^c} \right) \quad (3.6)$$

gdje je $I^c(y)$ intezitet boje u kanalu c , na lokaciji y u ulaznoj slici I , dok je A^c komponenta ambijentalnog svjetla u kanalu c .

Chao i suradnici [119] predstavili su efektivnu metodu temeljenu na DCP-u za smanjenje utjecaja raspršenja i slabljenja svjetlosti na podvodne fotografije, koristeći izračunavanje dubine mutne vode s pomoću pretpostavke o prisutnosti malog broja piksela s niskim intenzitetom boje. Yang i suradnici [120] razvili su pristupačnu i efikasnu DCP metodu za restauraciju podvodnih fotografija, koristeći medijan filtere za izračunavanje dubinskih mapa umjesto kompleksnog postupka mekog matiranja. Također su primijenili korekciju boja za poboljšanje kontrasta čime su rezultati postali znatno bolji u smislu kvalitete i brzine procesiranja. Chiang i suradnici [121] su uveli metodu poboljšanja podvodnih fotografija koja koristi kompenzaciju dužine vala i odmagljivanje, smanjujući razlike u slabljenju svjetla tijekom njegovog širenja i uklanjajući moguće utjecaje na izvor svjetlosti.

Serikawa i suradnici [122] predstavili su novu tehniku koja također kompenzira diskrepanciju u slabljenju svjetlosti i koristi brz algoritam odmagljivanja JTF (*engl. Joint Trigonometric Filtering*). JTF usavršava prijenosnu mapu (*engl. Transmission Map - TM*) temeljenu na DCP procjeni, omogućavajući brojna poboljšanja, kao što su smanjenje raspršivanja, poboljšanje rubnih informacija i kontrasta fotografije. Algoritam se ističe u smanjenju šuma, boljoj izloženosti tamnijih područja i jačanju ukupnog kontrasta.

Peng i suradnici [102] razvili su metodologiju za izračun dubinske mape nužne za restauraciju podvodnih fotografija, bazirajući se na opažanju da što je objekt udaljeniji od kamere, to je mutniji. Integrirali su zamućenost fotografije s IFM modelom, izračunavajući udaljenosti između točaka na sceni i kameri, pokazavši se kao značajno efikasnija od drugih IFM baziranih tehnika poboljšanja. DCP tehnike su prilagođene za specifične izazove podvodnog okruženja

čime su razvijene i implementirane različite podvodne tehnike poboljšanja koje se fokusiraju na DCP [80].

- Podvodna metoda tamnog kanala (UDCP)

Crveni kanal u podvodnim fotografijama često dominira u tamnom kanalu jer se crvena svjetlost brže prigušuje nego plava i zelena dok prolazi kroz vodeni medij. Kako bi se kompenzirao navedeni efekt crvenog kanala, autori rada [114] predstavili su koncept UDCP-a koji za procjenu tamnog kanala koristi samo zeleni i plavi kanal (*engl. Green-Blue - GB*) i time određuje podvodni DCP. U radu [123] predložena je nova metoda koja se suprotstavlja neujednačenosti prigušenja u podvodnim fotografijama u svrhu njihove kompenzacije kroz analizu putanje svjetla. Autori su osmislili tehniku koja koristi procjenu ambijentalnog svjetla i adaptivno filtriranje s ciljem poboljšanja kvalitete fotografija snimljenih u plićacima. Pored toga predstavili su algoritam korekcije boja koji vraća prirodne nijanse fotografije.

Lu i suradnici [90] iznijeli su novu tehniku super-rezolucije (*engl. Super Resoluition - SR*) i raščlanjivanja efekata raspršenja na podvodnim fotografijama. Korištenjem SR algoritma zasnovanog na principu samosličnosti, dobivene su fotografije visoke rezolucije (*engl. High Resolution - HR*) i odvojeni su efekti raspršenja. Nakon toga, primjenom pravila konveksne fuzije, obnovljena je HR fotografija. Ovaj pristup pokazao se vrlo efikasnim u procesu restauriranja podvodnih fotografija. Galdran i suradnici [117] razvili su automatski postupak za obnovu podvodnih fotografija oslonjen na RCP. Ovaj pristup koristi metodu izdvajanja tamnog kanala u kojoj se očekuje da će biti zadržani plavi i crveni kanal. Eksperimenti su pokazali da ova metoda efikasno poboljšava kvalitetu degradiranih podvodnih fotografija.

- Metoda pretpostavke maksimalnog intenziteta (MIP)

Suspendirane čestice u vodi koje uzrokuju efekte zamućenja i zamagljenja mogu znatno umanjiti jasnoću podvodnih fotografija. MIP [124] je koncept koji se koristi za uklanjanje zamućenja fotografije. MIP algoritam izvorno je predložen za prikaz MRA (*engl. Magnetic Resonance Angiogram*) snimanja gdje je interes pokazati krvne žile preko pozadine tkiva na slici. Temelji se na zapažanju da se maksimalna vrijednost lokalnih skupina piksela i gradijenata smanjuje tijekom procesa zamućivanja. MIP je kombinacija lokalnog maksimalnog intenziteta (*engl. Local Maximum Intensity - LMI*) i lokalnog maksimalnog gradijenta (*engl. Local Maximum Gradient - LMG*).

Carlevaris i suradnici [116] razvili su učinkovit algoritam za uklanjanje efekata raspršenja svjetlosti iz podvodnih fotografija, poznat kao tehnika uklanjanja zamagljenosti. Uveli su novi pristup procjeni dubine scene kao usporedbu u razlici intenziteta crvenog kanala u odnosu na vrijednosti između zelenog i plavog kanala. Prva vidljiva točka u prednjem planu predstavlja piksel s najvećom razlikom u intenzitetu između kanala boja.

Zhao i suradnici [125] predstavili su novu metodu koja se koristi za analizu optičkih svojstava vode. Ova tehnika procjenjuje pozadinsko svjetlo (*engl. Background Light - BL*) bazirano na kombinaciji DCP-a i MIP-a. Postupak započinje pronalaskom 0.1 % najsvjetlijih piksela u tamnom kanalu i odabiru piksela s najvećom razlikom u intenzitetima među BG ili GR (*engl. Green - Red*) kanalima. S druge strane, Li i suradnici [126] razvili su novu tehniku za obnovu podvodnih fotografija putem određivanja odabranog pozadinskog svjetla u skladu s pikselima koji pokazuju maksimalnu diskrepanciju u intenzitetima. Metoda se zasniva na odmagljivanju plavo-zelenog kanala i korekciji crvenog kanala. Na početku je implementiran pristup miješanja inspiriran radom Li i suradnika [127], [128], koristeći *quad-tree* podjelu za identifikaciju jednolikih područja pozadine. Sljedeći korak je identifikacija 0.1 % najsvjetlijih kandidata piksela u toj regiji iz tamnog kanala. Naposljetku, izabran je piksel s najvećom diskrepancijom u RB (*engl. Red - Blue*) kanalu kao reprezentativno globalno pozadinsko osvjetljenje.

- Ostale metode temeljene na pretpostavkama

Pored već ranije spomenutih, postoje i druge metode temeljene na pretpostavkama koje nisu u uobičajenoj primjeni, ali pružaju značajne prednosti u procesu restauracije podvodnih fotografija. Peng i suradnici [102] razvili su novu tehniku za izračunavanje dubina podvodnih scena koja se oslanja na apsorpciju svjetla i efekta zamućenja fotografije. Ovaj pristup implementiran je unutar IFM-a i pokazao se vrlo efikasnim prema rezultatima eksperimenata. U drugom radu, Peng i suradnici [129] predstavili su metodologiju za poboljšanje i restauraciju podvodnih snimaka putem smanjenja efekata apsorpcije svjetla, raspršenja, niskog kontrasta i izobličenja boje koja nastaju zbog prolaska svjetlosti kroz mutnu vodu. Ambijentalno svjetlo izračunato je prilagođavanjem boja ovisno o dubini, a prijenos scene procijenjen je temeljem intenziteta svjetla uočenog u odnosu na procijenjeno ambijentalno svjetlo, nakon čega je uslijedila adaptivna korekcija boja.

Li i suradnici [130] razvili su metodologiju koja se zasniva na principu minimalnog gubitka informacija (*engl. Minimal Information Loss Principle - MILP*), primijenivši algoritam za

uklanjanje zamagljenosti kako bi se obnovila boja, prirodni izgled i jasnoća podvodnih fotografija. Kako bi poboljšali kontrast i svjetlinu snimaka te vizualnu kvalitetu, implementirali su i efikasnu strategiju za unaprjeđenje kontrasta.

Wang i suradnici [93] predložili su novu tehniku identifikacije maksimalnog prigušenja (*engl. Maximum Attenuation Identification - MAI*) radi izvođenja dubinskih mapa i utvrđivanja pozadinskog osvjetljenja iz degradiranih podvodnih fotografija. Uključili su i procjenu pozadinske regije kako bi osigurali optimalne rezultate, a metodu su testirali na raznim vrstama fotografija uključujući kalibracijske ploče s bojama i prirodne podvodne scene.

c) Duboko učenje

Obnova kvalitete degradiranih i zamućenih podvodnih fotografija predstavlja veliki izazov. Klasične metode koje se oslanjaju na ručno dizajnirane karakteristike pokazuju slabije izvedbe. Stoga su tehnike dubokog učenja postale ključne u modernim pristupima restauracije. Razvoj dubokog učenja doveo je do značajnih promjena u pristupu restauraciji podvodnih fotografija, od ručnog prilagođavanja parametara preko modela umjetne optimizacije do automatiziranih i efikasnih trening modela. Ovi modeli dubokog učenja koriste se za izvlačenje relevantnih značajki direktno iz podataka.

Ding i suradnici [131] su razvili metodu za rješavanje problema raspršenja svjetla i izmjene boja koje degradiraju podvodne fotografije. Metoda se sastoji od koraka poboljšanja koji uključuju korekciju boja i odmagljivanje zasnovano na modelu atmosferskog raspršenja. Za određene dijelove fotografije izvađene iz fotografija ispravljene boje korištena je konvolucijska neuronska mreža kako bi se predvidjela karta dubine scene. Metoda je pokazala vrlo visoku efikasnost i preciznost u različitim primjenama, uključujući detekciju i prepoznavanje podvodnih objekata.

Cao i suradnici [132] razvili su tehniku za obnovu podvodnih fotografija koristeći dvije različite neuronske mreže za procjenu dubine i pozadinskog osvjetljenja, rješavajući tako problem deformacije boja i niskog kontrasta zbog efekata raspršenja i apsorpcije svjetlosti.

Barbosa i suradnici [133] predstavili su CNN tehnologiju usmjerenu na poboljšanje i restauraciju podvodnih fotografija koja ne zahtijeva standardne podatke stvarnog svijeta već se oslanja na mjere kvalitete fotografija (kontrast, oštrina, cjelovitost rubova i pretpostavka sivog svijeta), postižući time poboljšanje vizualne oštine i očuvanja detalja.

Hou i suradnici [134] razvili su pristup koji koristi rezidualno učenje za obnovu fotografija, temeljen na modelima vođenim podacima. Kroz kvalitativne i kvantitativne analize ovaj pristup se pokazao efikasnim u obradi podvodnih fotografija.

3.8.2. Tehnike poboljšanja podvodnih fotografija

Istraživanja usmjerena na unaprjeđenje kvalitete podvodnih fotografija imaju za cilj poboljšanje boja i kontrasta kroz redistribuciju intenziteta piksela, a ne pribjegavaju analizi principa podvodnog snimanja [135], [136]. Dodatno su razvijene i specijalizirane tehnike poboljšanja koje su usko povezane s karakterističnim izazovima u podvodnoj fotografiji kao što su nizak kontrast i prisustvo izmaglice u slici. Takve tehnike rade promjene na vrijednostima piksela unutar prostorne ili frekvencijske domene kako bi se postigla vizualna poboljšanja.

Konvolucijske neuronske mreže i druge tehnike dubokog učenja primijenjene su za unaprjeđenje podvodnih fotografija koristeći se skrivenim značajkama koje modeli mogu naučiti u cilju poboljšanja fotografije. Tehnike unaprjeđenja podvodnih fotografija kategorizirane su u četiri glavne grupe od kojih svaka koristi različite pristupe: poboljšanje temeljeno na prostornoj domeni, poboljšanje zasnovano na frekvencijskoj domeni, poboljšanje koje se oslanja na postojanost boja i poboljšanje koje implementira metode dubokog učenja [74].

3.8.2.1. Poboljšanje fotografija temeljeno na prostornoj domeni

Tehnika poboljšanja fotografija unutar prostorne domene temelji se na histogramu intenziteta piksela koji modulira nijanse sive ovisno o teoriji mapiranja sivih tonova [137]. Kod podvodnih fotografija, histogrami često prikazuju zbijeniju i jače izraženu distribuciju piksela u usporedbi s onim kod nadzemnih fotografija. Proširivanje dinamičkog opsega ovih histograma dovodi do bolje vidljivosti, bogatijih detalja i jačeg kontrasta unutar fotografija. Tehnike prostorne domene primjenjuju se na distribuciju intenziteta piksela unutar raznih postojećih modela boja, poput RGB, nijansa-zasićenost-intenzitet (*engl. Hue, Saturation, Intensity - HSI*), nijansa-zasićenost-vrijednost (*engl. Hue Saturation Value - HSV*) te CIE-Lab modela. Pristup poboljšanja fotografija putem prostorne domene doživio je značajne napretke u ovom području [138], [139].

Procesi unutar prostorne domene dalje se mogu podijeliti na metode koje se baziraju isključivo na jednoj boji (*engl. Single -Color Model - SCM*) i na one koje koriste kombinacije boja (*engl. Multi-Color Model - MCM*).

a) Metoda jednobojnog modela (SCM)

Ancuti i suradnici [140] uveli su metodu baziranu na fuziji s ciljem unaprjeđenja podvodnih fotografija. Postupak započinje kreiranjem dviju varijacija izvorne fotografije. Prva varijacija prilagođena je balansirajući bijelu boju dok se kontrast druge varijacije povećao korištenjem adaptivne tehnike izjednačavanja histograma. Potom su definirane četiri težinske vrijednosti fuzije koje se oslanjaju na istaknute karakteristike, kontrast i ekspoziciju obiju varijacija fotografija. Na kraju, kombinacijom ta dva prikaza i primijenjenim težinama dobivena je unaprijeđena fotografija.

U drugom radu, Ancuti i suradnici [141] iznijeli su pristup balansiranju boja i poboljšanju kvalitete podvodnih fotografija koji ne zahtijeva primjenu posebnog hardvera niti preliminarnih informacija o sceni ili podvodnim uvjetima. Tehnika se oslanja na fuziju dva izvedena prikaza koji proizlaze iz fotografije prilagođene balansom bijele boje i korigirane originalne degradirane, zamućene fotografije. Ova metoda je doprinijela poboljšanju kontrasta, jasnijih rubova i ukupnoj vidljivosti na podvodnim fotografijama.

Liu i suradnici [142] razvili su tehniku za unaprjeđenje podvodnih fotografija poznatu kao „Deep Sparse Non-negative Matrix Factorization” (DSNMF) koja služi za procjenu osvjetljenja na fotografijama. Fotografije su prvo podijeljene u male segmente gdje je svaki kanal unutar tog segmenta (RGB) obrađen kao matrica koja je dalje raspoređena u slojeve s pomoću DSNMF tehnike. Posljednji sloj ove faktorizirane matrice korišten je za određivanje osvjetljenja fotografije čime je postignuto unaprjeđenje fotografije [74].

b) Metoda modela više boja (MCM)

Torres i suradnici [143] su koristili Markovljeva slučajna polja (*engl. Markov Random Fields - MRF*) za modeliranje odnosa između originalnih i izobličenih podvodnih fotografija i primijenili su maksimalnu posteriornu (*engl. Maximum A Posteriori - MAP*) procjenu za unaprjeđenje boja na podvodnim fotografijama. Pri analizi razlika između fotografija one su konvertirane u CIE-Lab prostor boja kako bi se osigurala percepcijska konzistentnost.

Iqbal i suradnici [144] razvili su efikasnu tehniku za unaprjeđenje podvodnih fotografija koristeći „Integrirani model boja” (*engl. Integrated Color Model - ICM*). Ovaj pristup se

efikasno suprotstavlja problemima koji se pojavljuju kao posljedica raspršenja i apsorpcije svjetlosti. Njihovo rješenje uključivalo je tehniku rastezanja RGB kontrasta za izjednačavanje i pojačanje kontrasta između boja, praćeno primjenom HSI zasićenosti i rastezanja intenziteta za unaprjeđenje boja, svjetline i zasićenosti na degradiranim podvodnim fotografijama.

Ghani i suradnici [145] predstavili su tehniku po imenu „Rekurzivna adaptivna modifikacija histograma” (*engl. Recursive Adaptive Histogram Modification - RAHIM*) s ciljem poboljšanja pozadinskih dijelova podvodnih fotografija. Ovaj pristup uključio je modifikacije svjetlosti i zasićenosti u HSV prostoru boja koristeći se ljudskim vizualnim sistemom i Rayleigh distribucijom nakon čega su izvršene konverzije natrag u RGB model.

Hitam i suradnici [146] razvili su fuzijsku metodu za unaprjeđenje podvodnih snimaka baziranu na CLAHE algoritmu (*engl. Contrast Limited Adaptive Histogram Equalization*) za unaprjeđenje jasnoće fotografije. Tehnika se primijenila na fotografije u RGB i HSV formatu koje su kasnije kombinirane koristeći euklidsku normu za pojačavanje kontrasta i smanjenje šuma.

Huang i suradnici [147] izložili su metodologiju za unaprjeđenje podvodnih fotografija koja se oslanja na RGHS (*engl. Relative Global Histogram Stretching*) u dva prostora boja, RGB i CIE-Lab. Početno pretprocesiranje podvodnih fotografija izveden je s pomoću tehnike prilagodljivog rastezanja histograma u sivoj skali unutar RGB modela boja. Fokus u procesu je bio kontrast i boja izabranih podvodnih fotografija sa selektivno umanjenim širenjem svjetla. U CIE-Lab prostoru boja L i boje a , b koristili su se za optimizaciju krivulje i linearno prilagođavanje rastezanja.

3.8.2.2. Poboljšanje fotografija temeljeno na frekvencijskoj domeni

Tehnike obrade u frekvencijskoj domeni pristupaju poboljšanju fotografija upotrebom prostornih ili konvolucijskih transformacija [148]. U okvirima frekvencijskih metoda razlikujemo dvije glavne komponente: visokofrekventne koje odražavaju dijelove fotografije s velikim promjenama intenziteta piksela, tipično rubove i niskofrekventne koje predstavljaju područja s malim promjenama na fotografiji [149]. Frekvencijske metode se primjenjuju za unaprjeđenje kvalitete podvodnih fotografija putem pojačavanja visokih frekvencija i zatamljavanje niskih frekvencija [150]. Izazov s degradiranim podvodnim fotografijama leži u tome što je razlika između niskofrekventnih i visokofrekventnih komponenti često vrlo mala [151]. Kako bi se riješio ovaj problem, primjenjuju se razne tehnike kao što su homomorfno

filtriranje [152], metode transformacijske domene [153], valične transformacije i tehnike pojačanog filtriranja.

Petit i suradnici [154] predlagali su metodu koji se zasniva na kvaternionima za poboljšanje kontrasta i reprodukcije boje. Ovaj pristup uključuje prethodno sužavanje prostora boja i inverziju prigušenja svjetla. U metodi se koriste niskopropusni filter za smanjenje šuma i visokopropusni filter za čuvanje detalja, reducirajući pri tom niskofrekventne komponente.

Cheng i suradnici [155] razvili su tehniku koja je obuhvaćala Jaffe-McGlamery optički model i predložili precizan algoritam za restauraciju podvodnih fotografija koristeći RCP za izračunavanje transfera svjetla i pozadinskog svjetla. Također su primijenili jednostavan niskopropusni filter za obradu zamagljenih i degradiranih fotografija, analizirajući fizička svojstva funkcije raspršenja.

Feifei i suradnici [156] izložili su tehniku za podvodno unaprjeđenje fotografija baziranu na valičnoj dekompoziciji i visokopropusnom filteru s ciljem smanjenja šuma i obrade fragmenata raspršenja svjetlosti.

Ghani i suradnici [157] predstavili su metodu za povećanje vidljivosti dubljih podvodnih scena koristeći kombinaciju homomorfnog filtriranja, rekurzivnog CLAHE superponiranja i dvostruke valične fuzije fotografija. Homomorfno filtriranje služilo je za osvjetljenje cijele fotografije, a CLAHE je primijenjen na superponirane blokove kako bi se povećale kontrastne razlike.

Priyadhars i suradnici [158] razvili su metodu koja se koristi „Stacionarnom valičnom transformacijom” (*engl. Stationary Wavelet Transform - SWT*) za razlaganje ulazne fotografije na četiri različite komponente za povećanje kontrasta.

3.8.2.3. Poboljšanje fotografija temeljeno na postojanosti boja

Vizualni sustav čovjeka temelji se na principu konstantnosti boje, osiguravajući tako dosljednu percepciju obojenih predmeta pod raznim svjetlosnim uvjetima. Retinex je teorijski model za percepciju boje i osvjetljenja, koji pokušava objasniti kako ljudi održavaju konstantnost boje u promjenjivim uvjetima osvjetljenja. Naziv Retinex je kombinacija riječi *retina* (hrv. mrežnica oka) i *cortex* (hrv. moždana kora) – jer model sugerira da uočavanje boje uključuje i osjetilnu i kognitivnu obradu. Svrha Retinex-a je rekonstrukcija prave reflektirajuće boje objekta uklanjanjem utjecaja promjenjivog osvjetljenja. U obradi slike se koristi za poboljšanje

kontrasta, korekciju boje i normalizaciju osvjetljenja. Pojednostavljena osnovna ideja je da se slika $I(x)$ može promatrati kao što prikazuje jednačina 3.7:

$$I(x) = R(x) \cdot L(x) \quad (3.7)$$

gdje je $I(x)$ opažena vrijednost piksela, $R(x)$ reflektancija, a $L(x)$ osvjetljenje. Retinex algoritmi pokušavaju izračunati $R(x)$ uklanjanjem $L(x)$ iz $I(x)$ [159].

Joshi i suradnici [160] iznijeli su pristup za korigiranje netočnih boja i slabi kontrast na podvodnim fotografijama uzrokovane degradacijom. Korištenje Retinex-a usmjereno je na usklađivanje ljudskog i strojnog vida kroz primjenu modela konstantnosti boje i kompresiju dinamičkog raspona.

Fu i suradnici [161] razvili su metodu za poboljšanje podvodnih fotografija, ciljajući na tipične probleme kao što su zamagljenost, slaba osvjetljenost i izobličenje boja. Tehnika počiva na Retinex modelu koja prvo koristi korekciju boje za ispravljanje distorzija, a zatim Retinex za analizu svjetlosnih efekata, poboljšavajući osvjetljenje i refleksiju kako bi se prevladali defekti ekspozicije.

Zhang i suradnici [162] predstavili su metodu baziranu na Retinex-u koji simulira vizualni sustav čovjeka, kombinirajući različite vrste filtriranja za rješavanje problema ograničenja fotografija pod mutnim uvjetima.

Yong i suradnici [163] uveli su pristup poboljšanja podvodnih prizora konverzijom iz RGB u HSV model boja, a potom prikazujući V kanal kroz Retinex kako bi razdvojili sloj detalja od osvjetljenja, nakon čega poboljšani HSV kanali prelaze natrag u RGB čime se unapređuju fotografije.

Tehnika Zhang i suradnika [164] koja ispravlja degradaciju fotografije oslanja se na *multiscale Retinex* s obnovom boja (*engl. Multiscale Retinex Color Restoration - MSRCR*), sadržeći elemente kao što su procjena osvjetljenja, navođeno filtriranje, rekonstrukcija bez zamućenja i uravnoteženje bijele, efektivno poboljšavajući kontrast i detalje prizora.

Tang i suradnici [165] razradili su metodu za podvodne fotografije zasnivajući se na Retinex-u gdje prilagodba fotografija prethodi korištenju *multiscale Retinex* modela i Gaussovog filtriranja za ubrzavanje obrade.

Konačno, Zhang i suradnici [166] suočili su se s problemom podvodnih fotografija kroz adaptivno poboljšanje kontrasta, primjenjujući specifične algoritme za balansiranje boja i koristeći adaptivno kontrastno poboljšanje sa završnom primjenom maske za oštrinu [74].

3.8.2.4. Poboljšanje fotografija na temelju kontrasta

Kontrast ima ključnu ulogu u subjektivnoj ocjeni kvalitete podvodnih fotografija jer označava varijaciju u svjetlini između tamnih i svijetlih dijelova fotografije. Nerazmjer u svjetlini između susjednih površina stvara kontrast, vizualno obilježje koje doprinosi istaknutosti nekih objekata nad drugima.

Dixit i suradnici [167] predstavili su metodu temeljenu na DCP-u, ACCLAHE (*engl. Adaptively Clipped Contrast Limited Histogram Equalization*) i HF (*engl. Homomorphism Filtering*) algoritmima za poboljšanje fotografije. Prvi korak je bio izračun područja zamućenja kako bi ih DCP uklonio, a zatim je ACCLAHE procijenio maksimalnu količinu piksela u lokalnom histogramu i redistribuirao piksele. HF je dalje korišten za unaprjeđenje fotografija.

Wang i suradnici [168] izložili su svoj pristup koji koristi virtualni Retinex model i procjenu kvalitete fotografije (*engl. Image Quality Assessment - IQA*), optimizirajući kontrast i uklanjajući šum na podvodnim fotografijama. Nakon toga, kvalitetu poboljšanja su mjerili koristeći IQA bez referenci, dokazujući veću efikasnost od drugih metoda.

Bindhu i suradnici [169] predstavili su tehniku unaprjeđenja podvodnih fotografija baziranu na interpolacijskom poboljšanju koja se koncentrirala na povećanje kontrasta. Metoda je rezultirala boljom entropijom, nižim MSE (*engl. Mean Squared Error*) vrijednostima i većim vršnim omjerom signala i šuma.

Guraksin i suradnici [170] koristili su algoritam valićne transformacije i algoritam diferencijalne evolucije za podešavanje kontrasta i normalizaciju svjetlosnih uvjeta na fotografijama. Metoda je ispitivana tako što su usporedili MSE te omjer signala i šuma.

Sankpal i suradnici [171] razvili su tehniku za ispravak efekata oslabljenja svjetlosti u vodi korištenjem Rayleighovog rastezanja za izračunavanje parametra skaliranja metodom najveće vjerojatnosti u svakom kanalu boja.

Azmi i suradnici [172] predložili su četverostupanjsku metodu koja obuhvaća: tehniku neutralizacije boja, primjenu faktora pojačanja temeljenih na razlici između inferiornih i superiornih boja, spajanje dvostrukog intenziteta temeljeno na srednjim vrijednostima i

medijanu, korištenje inteligencije roja za izjednačavanje fotografija i tehnike neoštrog maskiranja za konačno poboljšanje kvaliteta fotografija.

3.8.2.5. Poboljšanje fotografija temeljeno na dubokom učenju

Metode dubokog učenja postižu brze rezultate u ekstrakciji značajki zahvaljujući svojoj složenoj mrežnoj arhitekturi što ih čini široko primjenjivim u različitim područjima kao što su odmagljivanje fotografija [173], detekcija objekata [174] i segmentacija fotografija [175].

Wang, Zhang, Cao i Wang [74] predstavili su efektivan i inovativan CNN pristup pod nazivom „UIE-Net” za poboljšanje kvalitete podvodnih fotografija koje su degradirane zbog disperzije i apsorpcije. Ovaj algoritam se fokusirao na uklanjanje zamagljenja i korekciju boja.

Autori rada [56] su razvili metodu koja ističe probleme u podvodnom okruženju izazvanim suspendiranim česticama, apsorpcijom i refleksijom svjetlosti korištenjem GAN-a za poboljšanje pouzdanosti vizualne percepcije. UWCNN (*engl. UnderWater Convolutional Neural Network*) je visoko učinkovit CNN model koji se koristi za automatsku rekonstrukciju jasnih podvodnih fotografija s visokim kontrastom, treniran s pomoću sintetičkih podataka za podvodne fotografije [176].

Li i suradnici [177] su razvili metodu koja koristi model nadziranog prijenosa boje, uključujući kompleksnu funkciju gubitka čime se omogućuje rekonstrukcija fotografija.

Uplavikar i suradnici [178] razvili su tehniku koja rješava probleme disperzije svjetlosti i slabljenja zahvaljujući učenju i definiranju karakteristika sadržaja podvodnih fotografija.

Li i suradnici [179] kombinirali su fizičke modele podvodnih scena s optičkim svojstvima da bi adresirali probleme kao što su apsorpcija i disperzija svjetlosti i rekonstruirali jasne fotografije visokog kontrasta.

Hu i suradnici [180] razvili su metodu za poboljšanje podvodnih slika degradiranih uslijed raspršenja i apsorpcije svjetlosti. U svrhu poboljšanja slike primijenjen je GAN, koji učinkovito provodi konverziju stila podvodnih slika visoke kvalitete. Iako su GAN-ovi široko korišteni, njihova učinkovitost ovisi o kvaliteti ulaznih podvodnih slika.

Tang i suradnici [181] su predložili generativnu mrežu baziranu na U-Netu s mehanizmom za fokus koji filtrira nebitne karakteristike i čuva informacije o teksturi i stilu, koristeći tri različite funkcije gubitka.

Sve navedene tehnike koriste se za rješavanje različitih problema podvodnih fotografija, uključujući niski kontrast, gubitak boje, zamagljenost i degradaciju fotografije, unapređujući njihovu kvalitetu i poboljšavajući vizualnu percepciju podvodnih fotografija.

3.8.3. Spoj metoda restauriranja i poboljšanja

Posljednjih godina, brojna istraživanja su se okrenula prema kombiniranom pristupu kojim restauriraju i poboljšavaju podvodne fotografije umjesto da se koncentriraju samo na jedno područje. Fuzijske tehnike kombiniraju više modela za povećanje svjetlosti, kontrasta, izoštravanja detalja i smanjenja šuma.

Gao i suradnici [182] razvili su metodu koja koristi DCP za odmagljivanje i procjenu svijetlog kanala fotografije nakon čega su primijenjene tehnike histograma izjednačavanja za poboljšanje fotografije.

Zhou i suradnici [183] predlagali su tehniku koja prvo prilagođava boje, a potom primjenjuje histogram na H kanal (u HSV modelu boja) i završava s tehnikom očuvanja rubova za poboljšanje vizualnog efekta i kvalitete fotografije.

Luo i suradnici [184] izložili su metodu koja kombinira optimizaciju kontrasta, balansiranje boja i istežanje histograma s fokusom na crveni kanal.

Dewangan i suradnici [185] su pokazali primjenu HSV filtera bez potrebe za segmentacijom u cilju poboljšanja jasnoće podvodnih fotografija, dajući bolji uvid u dubinu scene.

Sequeira i suradnici [186] predstavili su efikasnu metodu za restauraciju crvenog kanala s obzirom na to da crvena boja gubi intenzitet pod vodom nakon koje je povećanje kvalitete postignuto upotrebom kombiniranog modela boja.

Daway i suradnici [187] su predstavili metodu koja koristi *multiscale* Retinex s nenadziranom korekcijom boje i Rayleighovim rastežanjem za realističnije rezultate i veći kontrast.

Zhou i suradnici [188] su se bavili niskim kontrastom i izobličenjima boja kroz model Jaffe-McGlamery-a gdje su korekciju boja nadovezali na histogram izjednačavanja, završavajući proces dvorazinskom valićnom transformacijom.

Kombinacija ovih različitih tehnika dokazala je uspjeh u rješavanju specifičnih problema koji podvodne fotografije čine degradiranim kao što su nizak kontrast, gubitak boje i zamagljenost, poboljšavajući tako osvijetljenje, kontrast i izgled fotografija.

3.8.4. Relevantne metode pretprocesiranja u kontekstu cilja istraživanja

3.8.4.1. Removal of water scattering (RoWS) metoda

RoWS [119] je metoda koju su predložili Liu Chao i Meng Wang za vraćanje jasnoće podvodnih fotografija korištenjem modela tamnog kanala. Ova tehnika rješava uobičajeni problem podvodnih fotografija na koje utječe raspršivanje i slabljenje zbog čestica u vodi što je slično učinku guste magle u zraku. Algoritam pretpostavlja da u većini fotografija bar jedan kanal sadrži piksele koji imaju vrlo nizak intenzitet u određenim područjima fotografije. Ova pretpostavka pomaže u procjeni dubine замуćene vode i uklanjanju učinaka замуćenja, otkrivajući izvornu jasnoću fotografija. Osnovna jednačba formiranja podvodne fotografije može se izraziti kao što je prikazano u jednačbi 3.8:

$$I(x) = J(x) \cdot t(x) + A \cdot (1 - t(x)) \quad (3.8)$$

gdje je $I(x)$ opažena (degradirana) fotografija, $J(x)$ stvarna fotografija (bez utjecaja vode), $t(x)$ prijenosna funkcija koja predstavlja količinu svjetlosti koja nije raspršena te A ambijentalno (pozadinsko) svjetlo. Cilj metode je procijeniti $t(x)$ i A kako bi se rekonstruirala stvarna scena $J(x)$ iz opažene slike $I(x)$. Razlika u odnosu na klasični DCP jest da koriste najsvjetliji piksel u tamnom kanalu za detekciju pozadinskog svjetla što je posebno korisno u vodi gdje izmaglica često stvara osjećaj povećanja svjetline.

Rezultati pokazuju poboljšanje u jasnoći fotografije, međutim metoda ne rješava problem upijanja što utječe na restauraciju boje. U radu se zaključuje da iako je metoda učinkovita, možda neće dobro funkcionirati u ekstremnim slučajevima kada je objekt scene sličan pozadinskom svjetlu u velikom području.

3.8.4.2. Integrated Color Model (ICM) metoda

ICM [144] je metodologija koja se koristi za poboljšanje podvodnih fotografija. Pristup uključuje dva glavna koraka: rastezanje kontrasta RGB algoritma i rastezanje zasićenosti i intenziteta HSI. Prvi korak koristi se za izjednačavanje kontrasta boja na fotografijama rastezanjem raspona vrijednosti boja kako bi se iskoristile sve moguće vrijednosti. Svaki piksel se skalira s pomoću funkcije linearnog skaliranja kako bi se održao točan omjer boja. Drugi korak koristi se za povećanje prave boje i rješavanje problema osvjetljenja u podvodnim fotografijama. HSI model pruža širi raspon boja kontroliranjem elemenata boja na slici što pomaže u poboljšanju prave boje i svjetline fotografija. Algoritam rastezanja kontrasta

poboljšava kontrast fotografije širenjem raspona vrijednosti boja kako bi se iskoristio cijeli spektar mogućih vrijednosti. Ovaj proces koristi funkciju linearnog skaliranja za podešavanje vrijednosti piksela. Svaki piksel je skaliran kao što prikazuje jednačba 3.9:

$$P_o = (P_i - c) \cdot \frac{(b - c)}{(d - c)} + a \quad (3.9)$$

gdje je P_o normalizirana vrijednost piksela, P_i je razmatrana vrijednost piksela, a je minimalna vrijednost željenog raspona, b je maksimalna vrijednost željenog raspona, c je najniža vrijednost piksela trenutačno prisutna na slici, d je najveća vrijednost piksela koja je trenutačno prisutna na slici.

3.8.4.3. Contrast Limited Adaptive Histogram Equalization (CLAHE) metoda

CLAHE [189] je tehnika koja se koristi u obradi fotografije za poboljšanje kontrasta fotografija. Za razliku od tradicionalnog izjednačavanja histograma, koje primjenjuje istu transformaciju na sve piksele na slici, CLAHE radi na malim regijama na slici koje se nazivaju pločice. Histogram svake pločice se izjednačava, a susjedne pločice se zatim kombiniraju s pomoću bilinearne interpolacije kako bi se uklonile umjetno izazvane granice. CLAHE je modificirana verzija AHE (*engl. Adaptive Histogram Equalization*) tehnike [190]. Dok AHE može uzrokovati prekomjerno pojačanje šuma na podvodnim fotografijama, CLAHE ublažava ovaj problem dijeljenjem fotografije u nekoliko podblokova i izvođenjem izjednačavanja histograma na svakom dijelu pojedinačno. Međutim, CLAHE ima svoje nedostatke, uključujući stvaranje artefakata i šuma u jednolikim područjima fotografija. Nova vrijednost inteziteta piksela (J) izračunava se pomoću jednačbe 3.10:

$$J = (j_{max} - j_{min}) \cdot P(f) + j_{min} \quad (3.10)$$

gdje je j_{max} maksimalna vrijednost inteziteta, j_{min} minimalna vrijednost inteziteta, dok $P(f)$ predstavlja funkciju kumulativne distribucije (*engl. Cumulative Distribution Function - CDF*) intenziteta piksela na slici koja se koristi za transformaciju vrijednosti piksela za poboljšanje ukupnog kontrasta i vidljivosti detalja na slici [74].

3.8.4.4. Underwater Light Attenuation Prior (ULAP) metoda

Song i suradnici predstavili su ULAP [118], koncept koji se koristi u restauraciji podvodnih fotografija za procjenu udaljenosti objekata od kamere (dubina scene). Načelo iza ULAP-a je

da se dubinska udaljenost povećava s većom vrijednošću razlike između maksimalne vrijednosti zelenog (G) i plavog (B) svjetla (*engl. Maximum Value Green Blue - MVGB*) i vrijednosti crvenog (R) svjetla (pojednostavljeno kao VR) na podvodnoj slici. Ova se razlika koristi za treniranje linearnog modela za procjenu dubine scene što je ključno za obnavljanje podvodnih fotografija procjenom pozadinskog svjetla i prijenosne mape. Na temelju ULAP-a linearni modeli MVGB i VR, za procjenu dubinske mape, razvijeni su na način koji prikazuje jednadžba 3.11:

$$d(x) = \mu_0 + \mu_1 m(x) + \mu_2 v(x) \quad (3.11)$$

gdje x predstavlja piksel, $d(x)$ je dubina podvodne scene u točki x , $m(x)$ je MVGB, $v(x)$ je VR, μ_0 konstanta modela, a μ_1 i μ_2 koeficijenti modela (težine za ulazne značajke).

3.8.4.5. Gamma correction (GC) metoda

GC [191] je tehnika koja se koristi za podešavanje svjetline fotografije kako bi se poboljšala njezina vizualna kvaliteta. Mijenja odnos između ulazne i izlazne svjetline fotografije što rezultira prirodnijim izgledom. To znači da se tamniji dijelovi fotografije mogu posvijetliti, a svjetliji dijelovi potamniti kako bi se postigao uravnotežen prikaz. Primjena gama korekcije može značajno povećati kontrast fotografije, čineći detalje vidljivijima. Ovo je osobito korisno u situacijama kada su fotografije pretamne ili presvijetle. U specifičnim primjenama kao što je podvodna fotografija, gama korekcija može pomoći vratiti prirodne boje koje su izgubljene zbog apsorpcije svjetla u vodi. Relativno je jednostavan za implementaciju i ne zahtijeva složene izračune što ga čini praktičnim za razne aplikacije u stvarnom vremenu.

3.8.4.6. Relative Global Histogram Stretching (RGHS) metoda

RGHS [147] metoda je predložena za poboljšanje fotografija u plitkim vodama, koje često pate od niskog kontrasta, nejasnosti i boje zbog apsorpcije svjetlosti i raspršenja pod vodom. U RGB prostoru boja, RGHS prvo izjednačava zeleni (G) i plavi (B) kanal. Zatim redistribuira svaki histogram RGB kanala koristeći dinamičke parametre koji se odnose na distribuciju intenziteta izvorne fotografije i slabljenje valne duljine različitih boja pod vodom. Bilateralno filtriranje koristi se za uklanjanje šuma uz očuvanje vrijednih detalja i poboljšanje lokalnih informacija fotografije. Korekcija boja izvodi se rastezanjem L komponente i modificiranjem a i b komponenti u CIE-Lab prostoru boja. Ovaj korak ima za cilj poboljšati zasićenost i svjetlinu fotografije kako bi se dobile živopisnije boje. Učinkovito uravnotežuje boju, zasićenost i kontrast poboljšanih podvodnih fotografija.

3.8.4.7. Single underwater image restoration by blue-green channels dehazing and red channel correction (GBdehazingRCorrection) metoda

GBdehazingRCorrection [126] je metoda za poboljšanje kvalitete podvodnih fotografija, koje su često degradirane zbog apsorpcije svjetlosti i raspršenja u vodi. Metoda uključuje dva glavna koraka: odmagljivanje plavo-zelenih kanala i korekciju crvenog kanala. Plavi i zeleni kanali podvodne fotografije obrađuju se algoritmom uklanjanja zamagljivanja. Ovaj se algoritam temelji na proširenju i modifikaciji algoritma DCP. DCP algoritam pomaže u procjeni prijenosne mape i pozadinskog svjetla koji se zatim koriste za oporavak plavih i zelenih kanala fotografije kao što je prikazano u jednadžbi 3.12:

$$t(x) = 1 - \min_{c \in \{g,b\}} \left(\min_{x \in \Omega} \left(\frac{I^c(x)}{B^c} \right) \right) \quad (3.12)$$

gdje je $t(x) \in [0,1]$ prijenosna mapa koja predstavlja postotak zračenja scene koji dopire do kamere, $I^c(x)$ je opaženi intenzitet piksela x u kanalu c , gdje c može biti bilo zeleno g ili plavo b , B^c je intenzitet pozadinskog svjetla za kanal c , $\min_{c \in \{g,b\}}$ odabire minimalnu vrijednost između zelenih i plavih kanala, $\min_{x \in \Omega}$ izračunava minimalni omjer opaženog intenziteta u odnosu na pozadinu svjetlo unutar lokalne zakrpe.

Kanal bez maglice može se obnoviti jednadžbom 3.13:

$$J^c(x) = \frac{I^c(x) - B^c}{t^c(x)} + B^c, \quad c \in g, b \quad (3.13)$$

gdje $J^c(x)$ predstavlja obnovljeni kanal, a $t^c(x) = t(x)$.

Crveni kanal koji više pati od apsorpcije u vodi ispravlja se korištenjem teorije pretpostavke Sivog svijeta. Ova teorija pretpostavlja da bi prosječna boja fotografije trebala biti siva. Prilagođavanjem crvenog kanala na temelju ove pretpostavke, metoda kompenzira izobličenje boje i poboljšava vidljivost crvenog kanala kao što je prikazano u jednadžbi 3.14:

$$\frac{avgRr + avgBr + avgGr}{3} = 0.5 \quad (3.14)$$

gdje su $avgRr$, $avgBr$ i $avgGr$ normalizirane prosječne vrijednosti obnovljenog crvenog, plavog i zelenog kanala.

Prosječna vrijednost obnovljenog crvenog kanala može se procijeniti kao što je prikazano u jednadžbi 3.15:

$$avgRr = 1.5 - avgBr - avgGr \quad (3.15)$$

koeficijent kompenzacije δ može se izračunati kao u jednadžbi 3.16:

$$\delta = \frac{avgRr}{avgR} \quad (3.16)$$

gdje je $avgR$ normalizirana prosječna vrijednost izvornog crvenog kanala.

Oporavljeni crveni kanal $Rrec$ može se dobiti s pomoću jednadžbe 3.17:

$$Rrec = R \cdot \delta \quad (3.17)$$

gdje je R normalizirani izvorni crveni kanal, a δ procijenjeni koeficijent kompenzacije.

Dodatno, metoda uključuje adaptivnu ekspozicijsku mapu za rješavanje problema nekih regija na oporavljenoj slici koja izgledaju pretamno ili presvijetlo. Navedena mapa pomaže u podešavanju ekspozicije različitih područja fotografije kako bi se postigla bolja vizualna kvaliteta. Predložena metoda značajno poboljšava vidljivost, kontrast i ravnotežu boja podvodnih fotografija, nadmašujući postojeće metode u kvalitativnoj i kvantitativnoj procjeni.

3.9. Prilagodba arhitekture duboke neuronske mreže

Uvođenje četvrtog kanala u ulazne podatke konvolucijske mreže predstavlja pristup kojim se proširuje količina dostupne semantičke informacije modelu tijekom procesa učenja. Umjesto ograničavanja na standardne tri kanale, modelu se omogućuje informacije o dodatnoj dimenziji podataka koja može sadržavati spektralne, dubinske ili pak transformirane komponente slike poput nijanse (Hue) iz HSV prostora boja. Na taj način mreža može naučiti kompleksnije obrasce koji nisu vidljivi u osnovnom RGB prikazu. U kontekstu detekcije morskog otpada, dodatni kanal može doprinijeti boljoj separaciji objekata od okoline te povećanju točnosti klasifikacije i lokalizacije.

3.9.1. Hue kao četvrti kanal

Hue je komponenta HSV modela boja koji se koristi za opisivanje boja kroz njihove kromatske karakteristike. U ovom modelu Hue označava vrstu boje koja se opaža poput crvene, plave ili zelene i obično se prikazuje kao kut na krugu boja. Kao jedan od parametara u HSV modelu, Hue je posebno prilagođen ljudskoj percepciji i razlikovanju boja, nijansi. Podešavanje nijanse može značajno unaprijediti fotografije podešavanjem ravnoteže boja i stvaranjem željenog efekta. Ako fotografija ima neželjeni ton boje poput previše plave ili zelene, prilagodba nijanse

može pomoći u neutraliziranju tih tonova i vraćanju boja bliže realnosti. Nijansa se može prilagoditi kako bi se istaknule ili promijenile određene boje [192].

Korekcija Hue kanala na podvodnim fotografijama je bitna zbog specifičnih izazova podvodnog okruženja koji utječu na kvalitetu fotografije. Podvodne fotografije često trpe izobličenja boja i nizak kontrast zbog apsorpcije i raspršenja svjetlosti u vodi što rezultira fotografijama sa zeleno-plavim tonom i manjkom prirodne ravnoteže boja u usporedbi s fotografijama snimljenim iznad vode.

Proces uključuje korištenje naprednih tehnika, poput konvolucijskih neuronskih mreža (CNN), za vraćanje ravnoteže boja i poboljšanje kontrasta fotografija. Korekcijom Hue kanala fotografije mogu postići vjerniju reprodukciju boja što je od velikog značaja za znanstvene i praktične svrhe [193].

Hou i suradnici [194] predstavljaju novu metodu za poboljšanje podvodnih fotografija očuvanjem komponente nijanse (Hue) uz poboljšanje drugih aspekata kao što su kontrast, prikaz boja i osvjetljenje. Autori predlažu pristup s dvostrukim modelom koristeći oba modela boja, Hue-Saturation-Intensity (HSI) i Hue-Saturation-Value (HSV). Metoda uključuje dva glavna koraka: transformacija fotografije iz RGB modela u HSI model, pri čemu se komponenta nijanse (Hue) očuva te primjena algoritma za filtriranje u valnoj domeni (*engl. Wave Digital Filter - WDF*) na komponentama zasićenja i intenziteta kako bi se poboljšao kontrast i smanjio šum. Daljnja transformacija fotografije u HSV model, ponovno čuvajući nijansu te primjena algoritma za ograničeno rastezanje histograma (*engl. Contrast Histogram Stretching- CHS*) na komponentama zasićenja i vrijednosti kako bi se poboljšala vidljivost i vjernost boja. Ključni doprinos rada je mehanizam očuvanja nijanse koji osigurava da se originalne boje zadrže dok se poboljšava kvaliteta fotografije. Eksperimentalni rezultati pokazuju da ovaj pristup nadmašuje nekoliko postojećih metoda u smislu metričkih kvaliteta fotografije kao što su RMSE (*engl. Root Mean Squared Error*), PSNR (*engl. Peak Signal to Noise ratio*) i DECM (*engl. Discrete Entropy and Contrast Measure*) kao i učinkovitost obrade.

Autori rada [195] predlažu rješenje formaliziranjem problema iz geometrijske perspektive i uvođenjem brzog algoritma za njegovo rješavanje. Fokusiraju se na primjenu ovog rješenja za koloriranje fotografija izravno u RGB prostoru za razliku od drugih metoda koje obično koriste prostore svjetline-nijanse poput YUV ili Lab. Predložena metoda izbjegava artefakte koji se mogu pojaviti s drugim prostorima boja i osigurava da konačne RGB vrijednosti ostanu unutar standardnog raspona prikaza. Uvode varijacijski model za bojanje fotografija koji regularizira

fotografije u boji dok rješava problem specifikacije svjetline i nijanse. Ovaj model koristi primal-dual algoritam za optimizaciju što pomaže u postizanju realističnih rezultata bojanja bez promjene izvorne svjetline.

Soni i suradnici [196] predstavljaju novu metodu za poboljšanje podvodnih fotografija koje pate od varijacija boja i loše vidljivosti zbog apsorpcije i raspršenja svjetlosti. Autori uvode brz algoritam u kombinaciji s mehanizmom za očuvanje nijansi kako bi učinkovito riješili ove probleme. Uključuju kompetentnu metodu poboljšanja fotografije koja integrira modele boja HSI i HSV, korištenje nekompleksne logaritamske funkcije kao koraka pretprocesiranja za postizanje izjednačavanja intenziteta i eliminaciju prekomjernih vrijednosti piksela, primjenu tehnika filtriranja u valnoj domeni (WDF) i ograničenog rastezanja histograma (CHS) na HSI i HSV modele boja, pri čemu se očuva komponenta nijanse (H). Predložena metoda pokazuje poboljšanje kvalitete podvodnih fotografija povećanjem vjernosti boja i vidljivosti.

3.9.2. Dubina kao četvrti kanal

Rad [197], „Depth Anything V2“, predstavlja napredni model za procjenu dubine iz fotografija koji značajno nadmašuje svoju prethodnu verziju u smislu robusnosti i detalja predviđanja dubine. Ključne inovacije uključuju zamjenu svih označenih stvarnih fotografija sintetičkim fotografijama, povećanje kapaciteta modela učitelja te korištenje velikog broja pseudo-označenih stvarnih fotografija za obuku modela učenika. U usporedbi s najnovijim modelima temeljenim na „Stable Diffusion“ modelu, „Depth Anything V2“ se ističe svojom učinkovitošću u brzini i točnosti. Koristi DPT (*engl. Dense Prediction Transformer*) kao dekodera dubine, izgrađen na DINOv2 enkoderima. Trening fotografije su rezolucije 518 x 518, za koju autori tvrde da postižu optimalne performanse. Trening proces se odvija u tri koraka: prvo se trenira pouzdan učiteljski model koristeći visoko-kvalitetne sintetičke fotografije; zatim se generiraju precizne pseudo oznake dubine na velikom skupu neoznačenih stvarnih fotografija; konačno, modeli učenika se treniraju na pseudo-označenim stvarnim fotografijama za robusnu generalizaciju.

Optimizacija modela koristi dva termina gubitka: skala i pomak - invarijantni gubitak te gubitak usklađivanja gradijenta koji su prilagođeni za poboljšanje oštrote dubine, osobito kada se koriste umjetne fotografije. Evaluacija modela provedena je na više referentnih setova fotografija, uključujući novonastali DA-2K, koji pokriva raznolike scenarije i pruža definirane oznake dubine. Modeli su dizajnirani da podrže široki spektar primjena s različitim veličinama

modela, od 25 milijuna do 1.3 milijardi parametara te su lako prilagodljivi za daljnje zadatke, pružajući snažnu osnovu za daljnje istraživanje. Ove ključne komponente omogućuju Depth Anything V2 modelu da pruži robusne i detaljno predviđene dubine.

4. Opis i metodologija istraživanja

Istraživanje koje je provedeno sastoji se od sljedećih faza:

1) Prikupljanje fotografija visoke rezolucije

- Potrebno je prikupiti dovoljno veliki i uravnoteženi skup visokokvalitetnih fotografija. Cilj je osigurati minimalno 3000 fotografija snimljenih na bar 15 različitih lokacija, uz varijacije visine od 5 m, 10 m i 15 m te ravnomjerno rasporediti fotografiranje kroz sva godišnja doba, doba dana i različite jačine vjetra.

2) Kreiranje baze podataka

- Označavanje objekata – Ručno označiti objekte interesa tj. vidljiv otpad koji se nalazi na morskom dnu koristeći standardizirani format oznaka.
- Kategorizacija – Objekte je potrebno kategorizirati u skupine: kese, boce, limenke, tkaninu, čaše, gume, građevni materijal, čepove i poklopce.
- Unifikacija – Svakoj fotografiji treba pridružiti meta podatke o uvjetima fotografiranja: lokaciji, dobu dana, godišnjem dobu, kutu upada sunčeve svjetlosti te omogućiti pretragu i filtriranje po navedenim parametrima.
- Stratifikacija – Potrebno je kreirati uravnotežene podskupove s isječcima fotografija koji su ravnomjerno raspoređeni po kategorijama, kako bi se omogućilo konzistentno treniranje i evaluacija modela.

3) Pretprocesiranje fotografija

- Analiziranje postojećih metoda – Potrebno je primijeniti navedene metode za poboljšanje fotografija RoWS, ICM, CLAHE, ULAP, GC, MIP, DCP, RGHS, GBdehazingRCorrection te izgraditi modele za svaku metodu ili kombinaciju metoda.
- Usporedba i analiza rezultata pretprocesiranja – Korištenjem kvantitativnih mjera bi trebalo procijeniti učinkovitost svake od metoda pretprocesiranja.
- Predlaganje metode za daljnje korištenje – Na temelju analize potrebno je odabrati metodu koja će se koristiti u daljnjem istraživanju.

4) Modifikacija dubokih neuronskih mreža

- Eksperimentalna metoda analize postojećih arhitektura – Potrebno je provesti istraživanje na temelju postojeće YOLO arhitekture.

- Prilagodba mreže i testiranje – Potrebno je istražiti mogućnosti poboljšanja poput proširivanja broja ulaznih kanala s dodatnim podacima ili dodavanje slojeva koji poboljšavaju detekciju malih podvodnih objekata.

5) Implementacija predloženih metoda

- Odabir modela – Trebalo bi koristiti najuspješniji model pretprocesiranja i modificirane neuronske mreže u konačnu metodu.
- Treniranje i testiranje – Da bi se dobili ispravni rezultati, potrebno je napraviti treniranje modela na označenim podacima te provesti evaluaciju na neviđenom testnom skupu.

6) Analiza rezultata

- Dobivene rezultate treba analizirati kvantitativno (mAP, preciznost, odaziv) te usporediti s različitim modelima detekcije čime bi se identificirao optimalan pristup za rješavanje zadanog problema.

7) Donošenje zaključka

- Na temelju provedenih eksperimenata i analize rezultata trebalo bi formulirati zaključke istraživanja, s preporukama za budući rad.

4.1. Prikupljanje fotografija visoke rezolucije

Prije prikupljanja fotografija za bazu podataka postavljene su sljedeće smjernice:

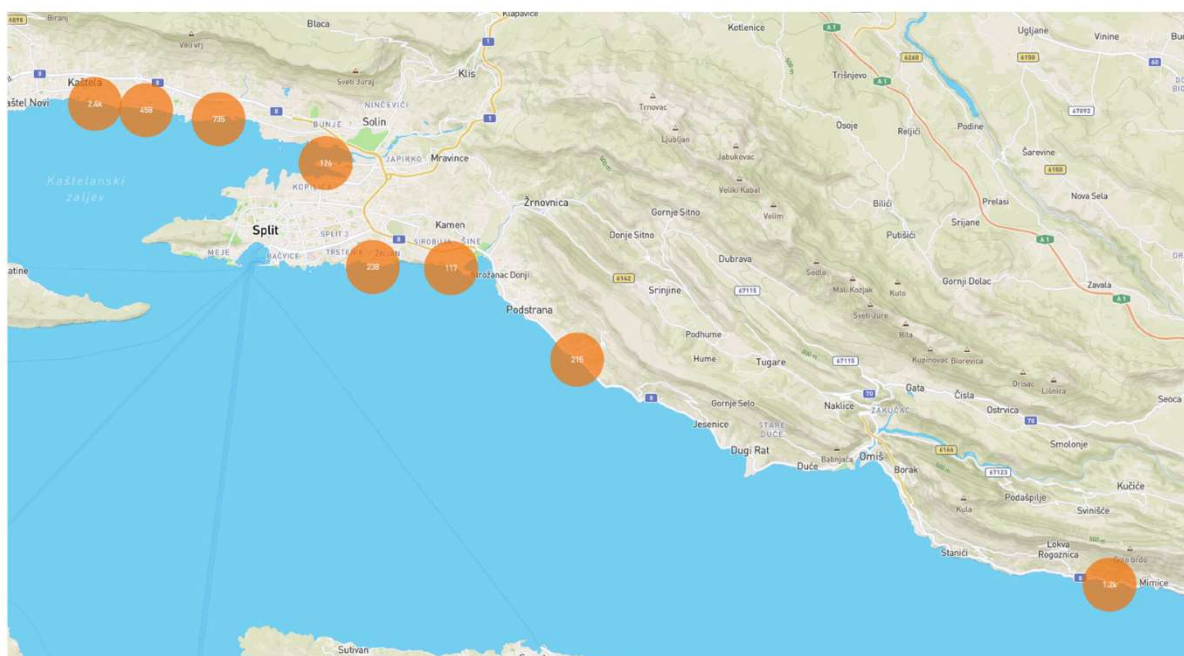
- Metoda prikupljanja podataka mora biti daljinska detekcija s pomoću bespilotne letjelice (UAV).
- Fotografije trebaju biti snimljene s više nadmorskih visina (5 m, 10 m i 15 m).
- Predmet od interesa fotografiranja je morski otpad.
- Morski otpad će se klasificirati na razne nerazgradive predmete.
- Područje snimanja je plitka zona bentosa s vidljivim otpadom, na procijenjenim dubinama do 4–5 metara, gdje je vizualna detekcija otpada iz zraka još moguća.
- Fotografiranje se vrši tokom cijele godine u svim sezonama, od 6:00 do 20:00 sati.
- Parametri koje treba zabilježiti su datum i vrijeme, godišnje doba, GPS koordinate, brzina i smjer vjetra, nadmorska visina i azimut sunca.

Za izradu baze podataka prema gore navedenim pravilima, dronovi DJI Mavic pro 2 kamera Hasselblad 20MP 4K UHD i DJI air 2S kamera 2.4. μ m 5.4 K UHD su korišteni za prikupljanje fotografija s visine od 5 m, 10 m i 15 m. Ove visine odabrane su kao kompromis između

povećanja vidnog polja i očuvanja dovoljne razlučivosti za detekciju malih objekata. Fotografije su snimljene u priobalju Jadranskog mora, u Splitsko-dalmatinskoj županiji. Dvadeset i tri lokacije (stanice) nalaze se u obalnom području sljedećih gradova: Kaštela, Vranjic, Split, Stobreč, Podstrana i Lokva Rogoznica kako prikazuje Tablica 4.1, Slika 4.1 i uvećano Slika 4.2 [198].

Tablica 4.1 Lokacija po gradovima i GPS koordinatama

Lokacija	GPS lokacija
Kaštel Lukšić, Hrvatska	43.550427, 16.366794
Kaštel Kambelovac, Hrvatska	43.548128, 16.385065
Kaštel Gomilica, Hrvatska	43.546852, 16.397127
Kaštel Sućurac, Hrvatska	43.543966, 16.412761
Vranjic, Hrvatska	43.531615, 16.465627
Plaža Žnjan, Split, Hrvatska	43.500656, 16.482890
Stobreč, Hrvatska	43.500646, 16.515163
Podstrana, Hrvatska	43.472911, 16.569938
Lokva Rogoznica, Hrvatska	43.406480, 16.787435

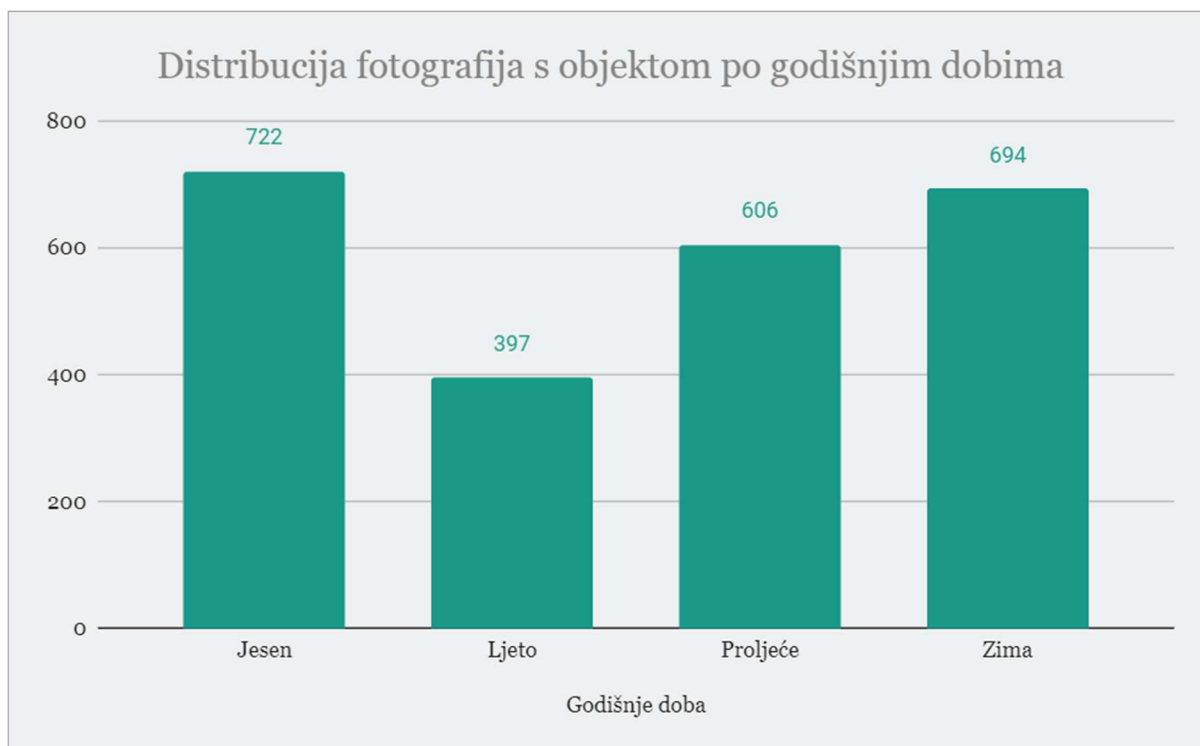


Slika 4.1 Lokacije svih fotografija koje se nalaze u skupu podataka (obalno područje Hrvatske)



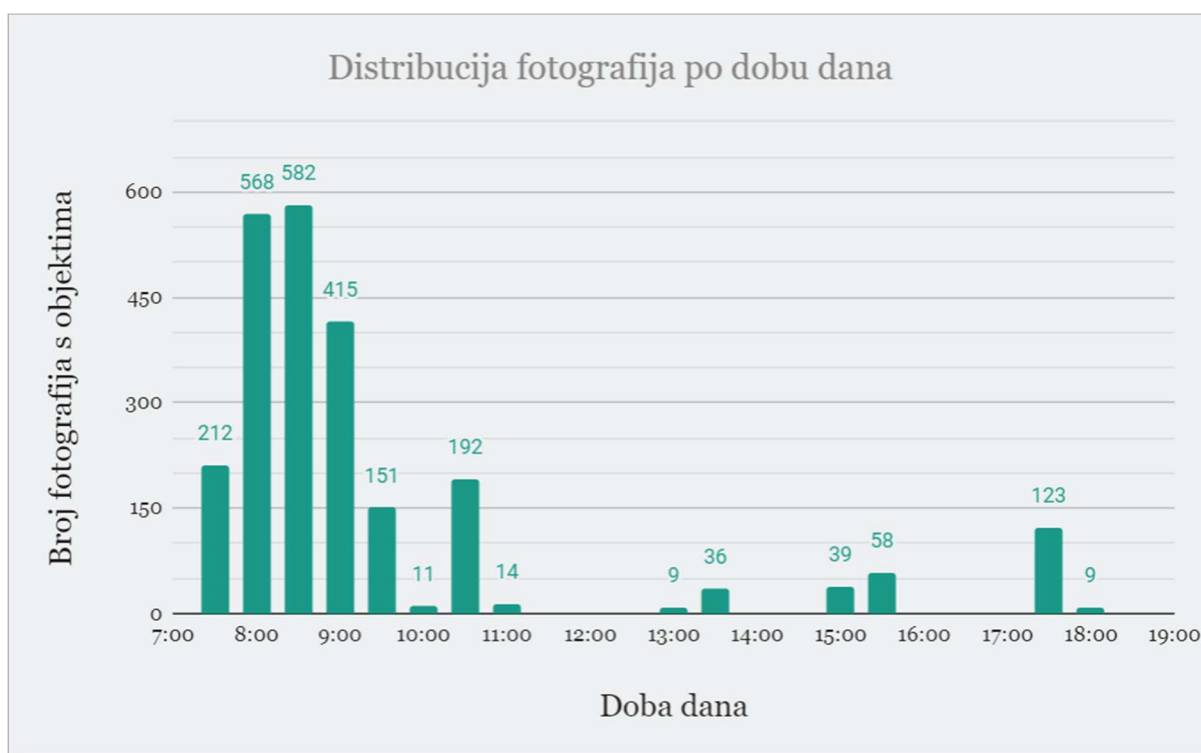
Slika 4.2 Uvećana lokacija plaže Žnjan, Split, gdje svaka točka predstavlja lokaciju na kojoj je fotografija snimljena

Prikupljanje podataka provedeno je kroz sva četiri godišnja doba od prosinca 2021. do ožujka 2023. Tijekom zime obavljena su tri odlaska na teren, pet odlazaka tijekom proljeća, osam tijekom ljeta i četrnaest tijekom jeseni pa je distribucija fotografija zadržana po godišnjim dobima kao što prikazuje Slika 4.3.

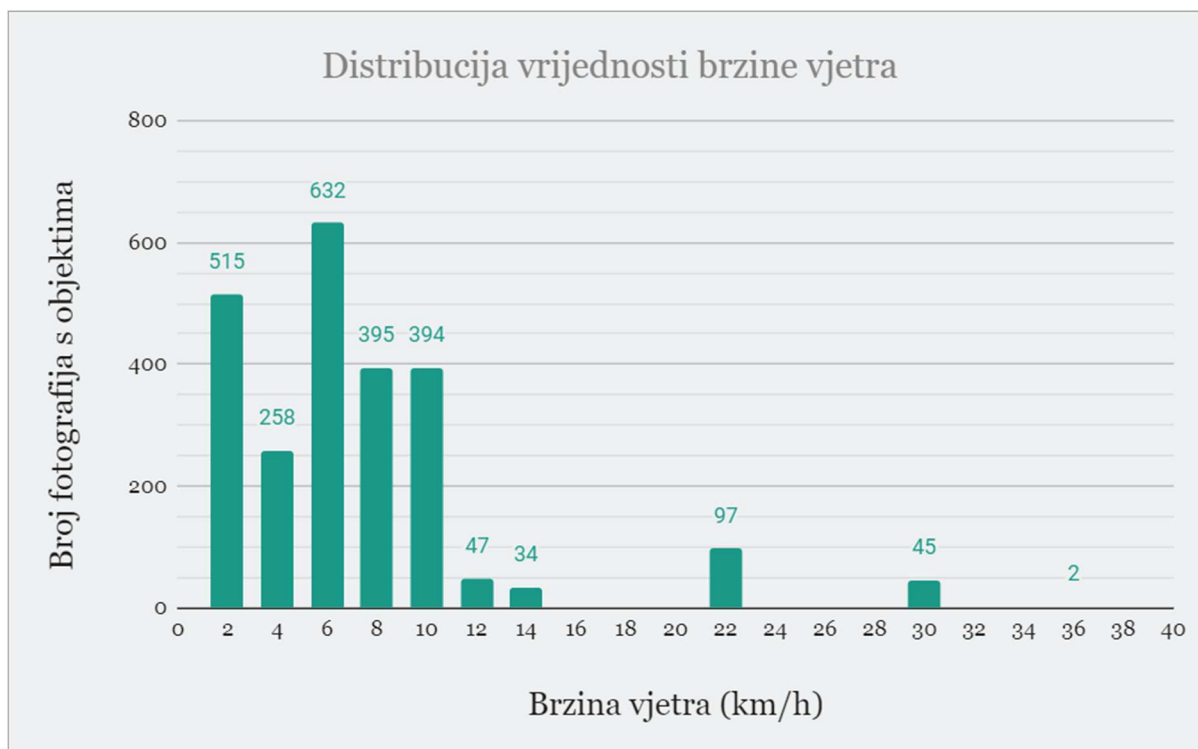


Slika 4.3 Distribucija fotografija s objektima po sezoni

Broj prikupljenih fotografija tijekom ljetnog razdoblja bio je značajno manji u odnosu na ostala godišnja doba zbog otežanih uvjeta rada u priobalnom području. Ljetni mjeseci karakterizirani su povećanom gustoćom turista, pojačanim brodskim prometom te visokim temperaturama, što je ograničavalo mogućnosti pristupa i sigurnog izvođenja UAV snimanja. Kao što prikazuje Slika 4.4, najranije vrijeme za fotografiranje bilo je u 7:10, a najkasnije u 18:30, dok se brzina vjetra kretala od 0 km/h do 35 km/h. Broj fotografija prisutnih u određenom intervalu (prema dobu dana ili brzini vjetra) prikazuju Slika 4.4 i Slika 4.5. Navedeni periodi dana karakterizirani su povoljnijim vremenskim uvjetima za let UAV-a, uključujući slabiji vjetar i manju prisutnost ljudi u priobalnom području.



Slika 4.4 Distribucija fotografija po dobu dana



Slika 4.5 Distribucija fotografija prema brzini vjetra

Još jedan važan faktor u izgradnji ove vrste baze podataka je položaj sunca. Kao što prikazuju Slika 4.6 i Slika 4.7, solarna elevacija i azimut izmjereni su kako bi bili uključeni u bazu podataka (te vrijednosti su prosječne po postaji umjesto po slici).

Solarna elevacija je kut između horizonta i položaja Sunca na nebu te se računa kako je niže navedeno.

Prvo je potrebno izračunati solarnu deklinaciju kao što je prikazano u jednadžbi 4.1:

$$\delta = 23.44^\circ \cdot \sin\left(\frac{360^\circ}{365} \cdot (n - 81) \cdot \frac{\pi}{180}\right) \quad (4.1)$$

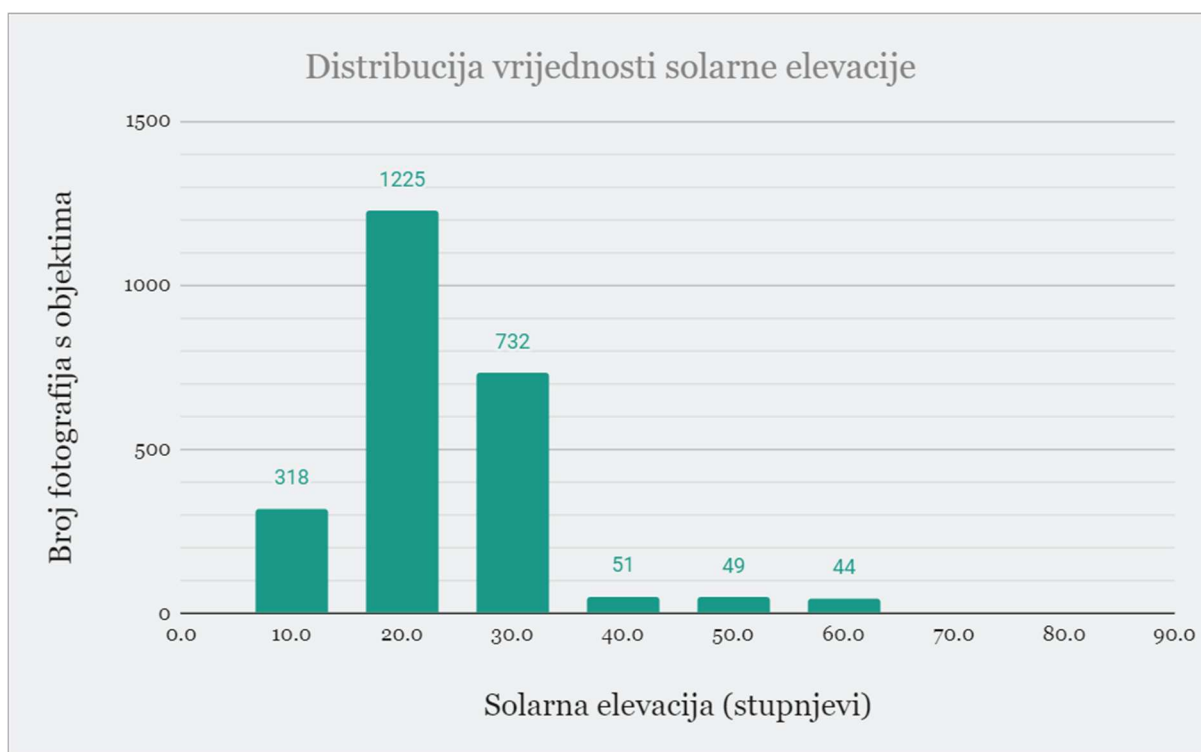
gdje je n redni broj određenog dana u godini, δ deklinacija Sunca (kut između ekvatora i Sunca). Deklinacija je negativna ako se Sunce projicira južno od nebeskog ekvatora.

Zatim je potrebno izračunati solarni satni kut kao što prikazuje jednadžba 4.2:

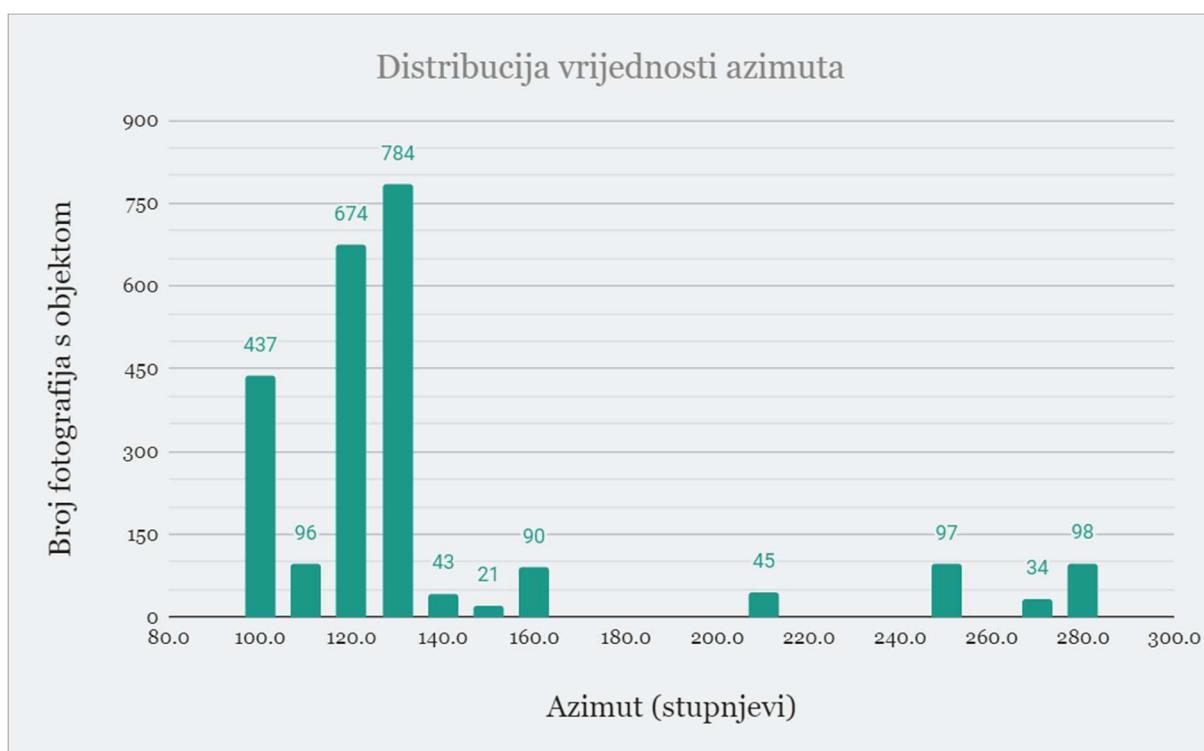
$$h = 15^\circ \cdot (LST - 12) \quad (4.2)$$

gdje je LST lokalno solarno vrijeme (u satima). Nakon toga se solarna elevacija α može izračunati uvrštavanjem zemljopisne širine ϕ u navedenu jednadžbu (4.3):

$$\sin(\alpha) = \sin(\phi) \cdot \sin(\delta) + \cos(\phi) \cdot \cos(\delta) \cdot \cos(h) \quad (4.3)$$



Slika 4.6 Distribucija vrijednosti solarne elevacije

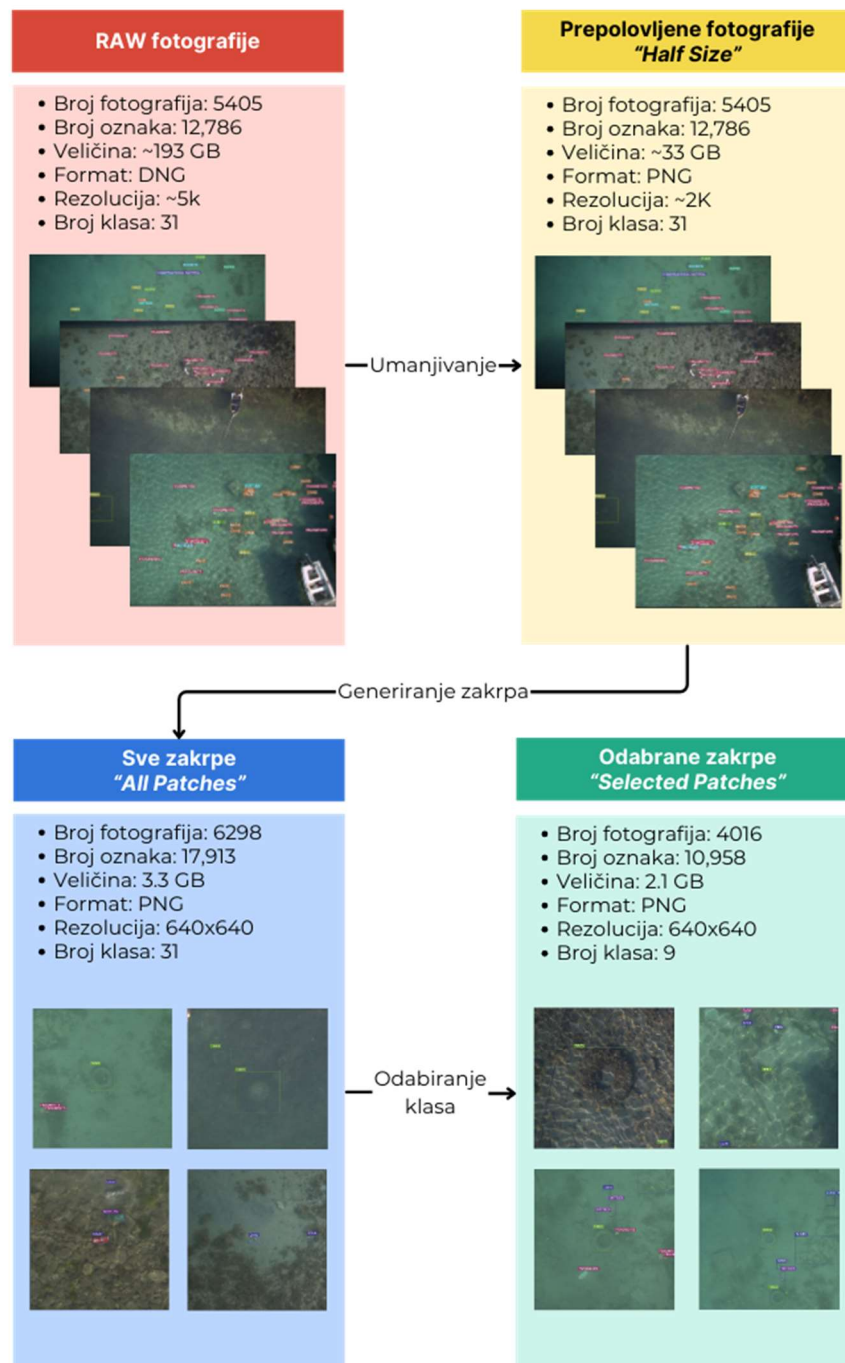


Slika 4.7 Distribucija azimuta (kut između sjevera i projicirane lokacije Sunca)

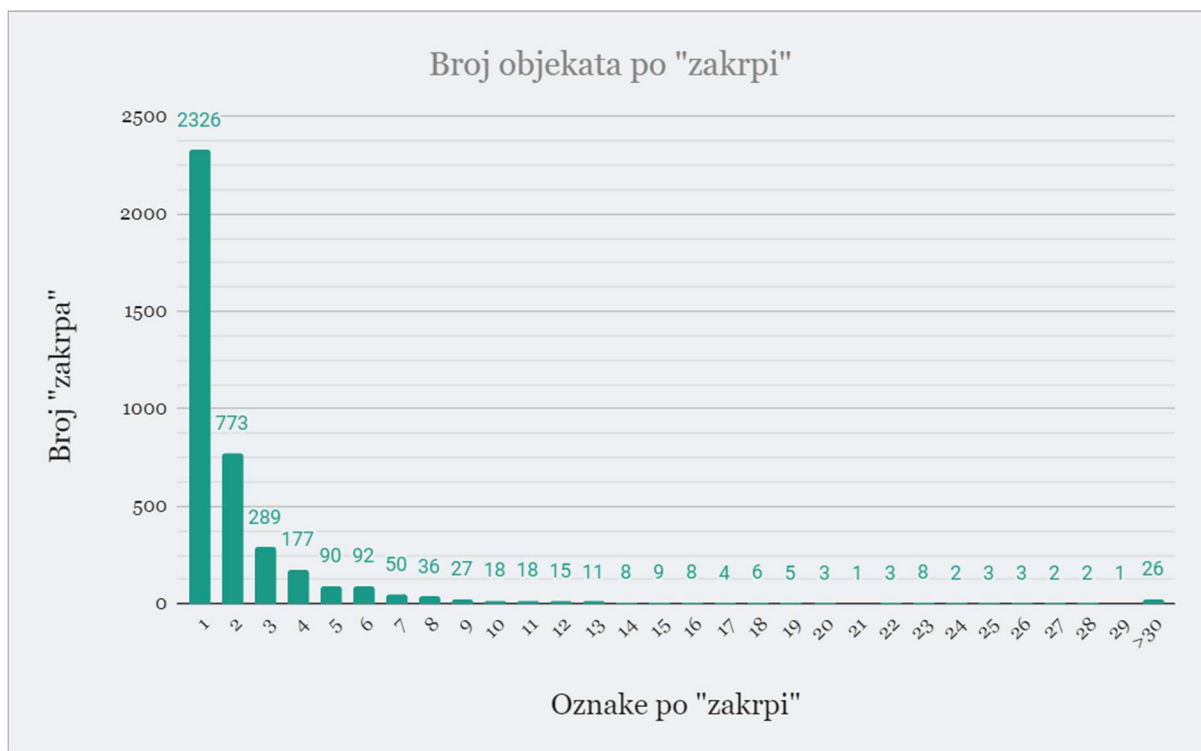
4.2. Kreiranje baze podataka

Bespilotne letjelice korištene za fotografiranje kao format pohrane fotografije koriste *RAW* format, koji sadrži neobrađene senzorske podatke visoke kvalitete. Obrada *RAW* fotografija zahtijeva značajne računalne resurse. Kako bi se omogućila učinkovita priprema podataka za treniranje neuronskih mreža i skratilo ukupno vrijeme obrade tijekom eksperimentiranja, sve su fotografije smanjene za faktor dva u horizontalnoj i vertikalnoj dimenziji, čime se ukupna razlučivost smanjila na jednu četvrtinu izvorne. Smanjenje rezolucije olakšalo je serijsku obradu većeg broja fotografija, pretprocesiranje i treniranje modela na standardnoj računalnoj opremi. Unatoč toj prilagodbi, sve izvorne fotografije visoke razlučivosti pohranjene su i dostupne za potencijalne buduće analize, uključujući obradu manjih objekata koji zahtijevaju veću prostornu razlučivost. Ovaj modificirani skup podataka naziva se *Half Size* baza podataka. Također, većina dubokih neuronskih mreža zahtijeva manje fotografije od onih u *RAW* ili *Half Size* bazi podataka. Promjena veličine ponovno nije bila izvediva opcija budući da fotografije sadrže male objekte, a promjena veličine smanjila bi te objekte na nekoliko piksela. Stoga je baza podataka zakrpa stvorena od *Half Size* baze kao što prikazuje Slika 4.8. Ova baza podataka se naziva *All Patches*. Ako bi se zakrpe izrezivale iz *RAW* fotografija, bilo bi potrebno dodatno skalirati svaku zakrpu na veličinu koja je potrebna neuronskoj mreži. Dijelovi fotografija koji prikazuju objekte izrezani su u zakrpe, odnosno dio fotografije gdje se nalaze označeni objekti. Zakrpe imaju dimenzije 640×640 piksela i sadrže jedan ili više objekata koji su potpuno ili djelomično vidljivi, čime odgovaraju ulaznoj rezoluciji modela neuronske mreže. Navedena baza ima 6298 fotografija smanjene dimenzije čime se ubrzava proces učitavanja i obrade. Zbog preklapanja vidnog polja moguće je da se objekti koji se nalaze blizu granice okvira pojave u dvjema susjednim zakrpama, djelomično ili u cijelosti, čime se objašnjava povećanje broja oznaka u bazi *All Patches* u odnosu na *Half Size* bazu. Budući da su klase objekata nađenih u prirodi neuravnotežene, odabrano je devet klasa za treniranje pokrivajući većinu otpada pronađenog na morskom dnu. Selekcijom odabranih klasa iz baze podataka *All Patches* formirana je konačna baza podataka odabranih zakrpa *Selected Patches*. Podatke o distribuciji broja oznaka po zakrpi prikazuje Slika 4.9. Baza *Selected Patches*, koja se koristi za primjenu pretprocesiranja i treniranje modela, sastoji se od 4016 fotografija dimenzija 640×640 piksela (izvadak baze prikazuje Slika 4.10) koje sadrže klase: fragmenti, građevinski materijal, gume, limenke, boce, tkanina, torbe, čepovi i poklopci te čaše distribuirane kao što prikazuje Slika 4.11. Fragmenti predstavljaju ostatke ili otkinute dijelove većih struktura ili objekata, najčešće nastalih uslijed mehaničkog oštećenja, trošenja ili razgradnje materijala u morskom okolišu,

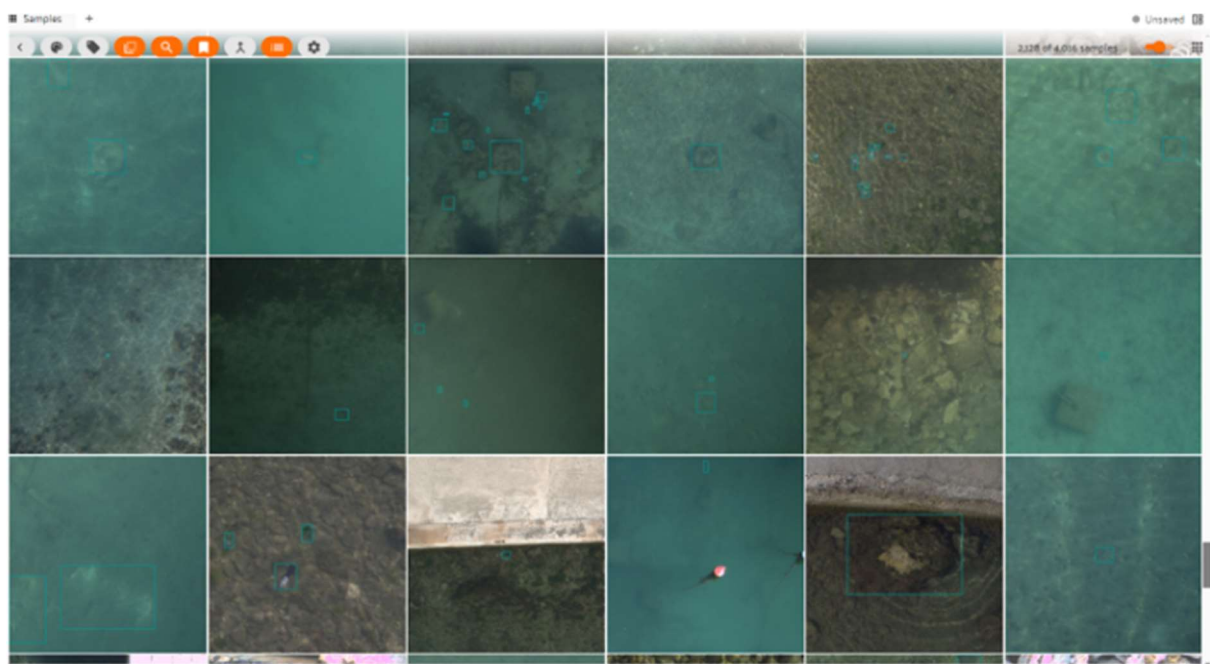
poput dijelova brodskih oplata, otkrnutih komada cigli, betona, plastike ili metala. Iako često neprepoznatljivi kao pripadnici specifičnih funkcionalnih objekata, ovi fragmenti zadržavaju prepoznatljiva fizikalna svojstva i teksture, što ih čini važnom kategorijom prilikom izrade sustava za detekciju otpada.



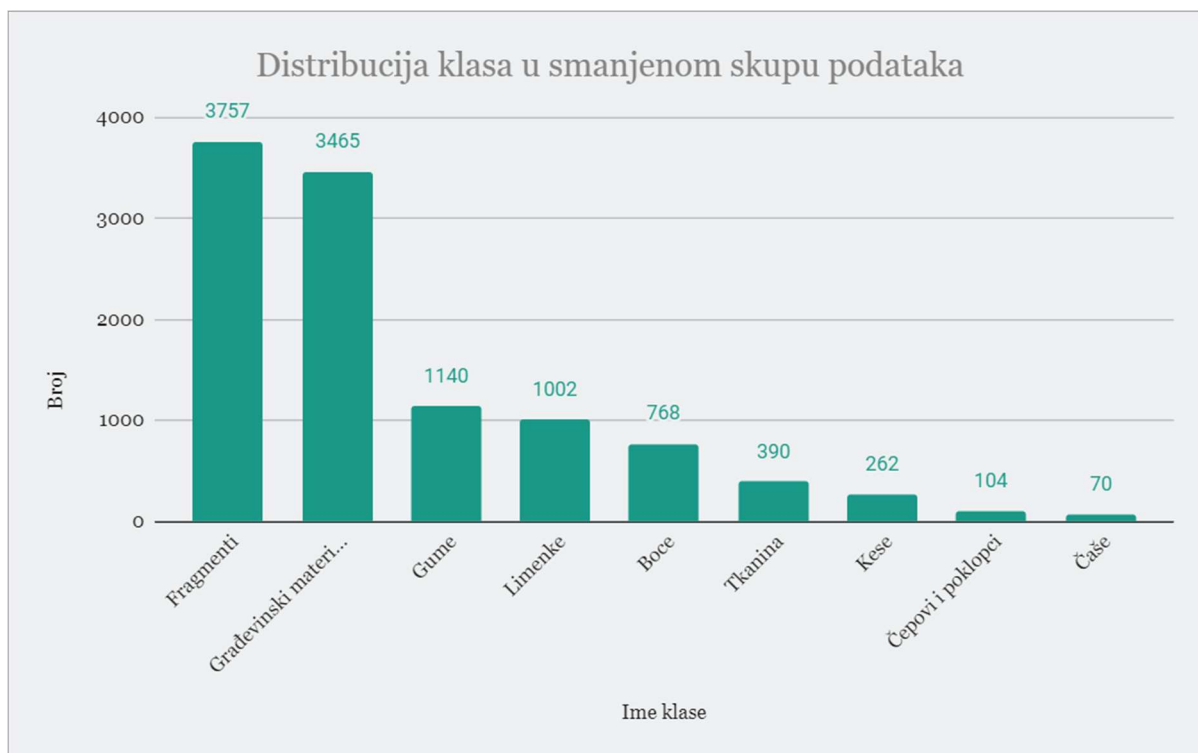
Slika 4.8 Dijagram toka procesa kreiranja baze podataka



Slika 4.9 Distribucija broja oznaka po zakrpi u bazi podataka Selected Patches



Slika 4.10 Označeni uzorci fotografija iz baze podataka Selected Patches



Slika 4.11 Distribucija klasa u konačnoj bazi podataka Selected Patches

4.2.1. Analiza nove baza podataka

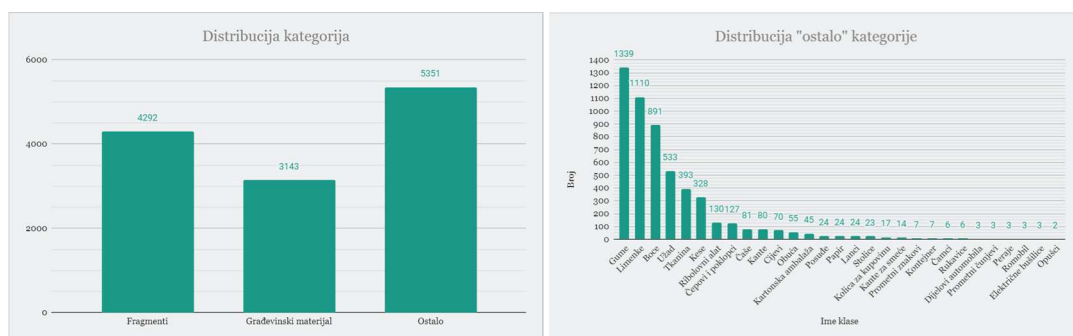
S obzirom na to da baza podvodnih objekata snimljenih bespilotnim letjelicama ne postoji ili nije javna, izgrađena je vlastita baza podataka o kojoj se može saznati više na <https://codeeasy.com/seafloor-litter-database> poveznici.

Baza se sastoji od 5405 originalnih fotografija. Od tih 5405 fotografija 2419 fotografija, 44.8 % ima podvodni otpad, dok ostalih 55.2 % nema. Osim datuma snimanja izračunate su: lokacija, GPS koordinate, brzina vjetra, vrsta vjetra i godišnje doba, kut sunčeve elevacije i azimut. Svakoj postaji dodijeljen je jedinstveni prefiks i serijski broj fotografije kao što prikazuje Slika 4.12.

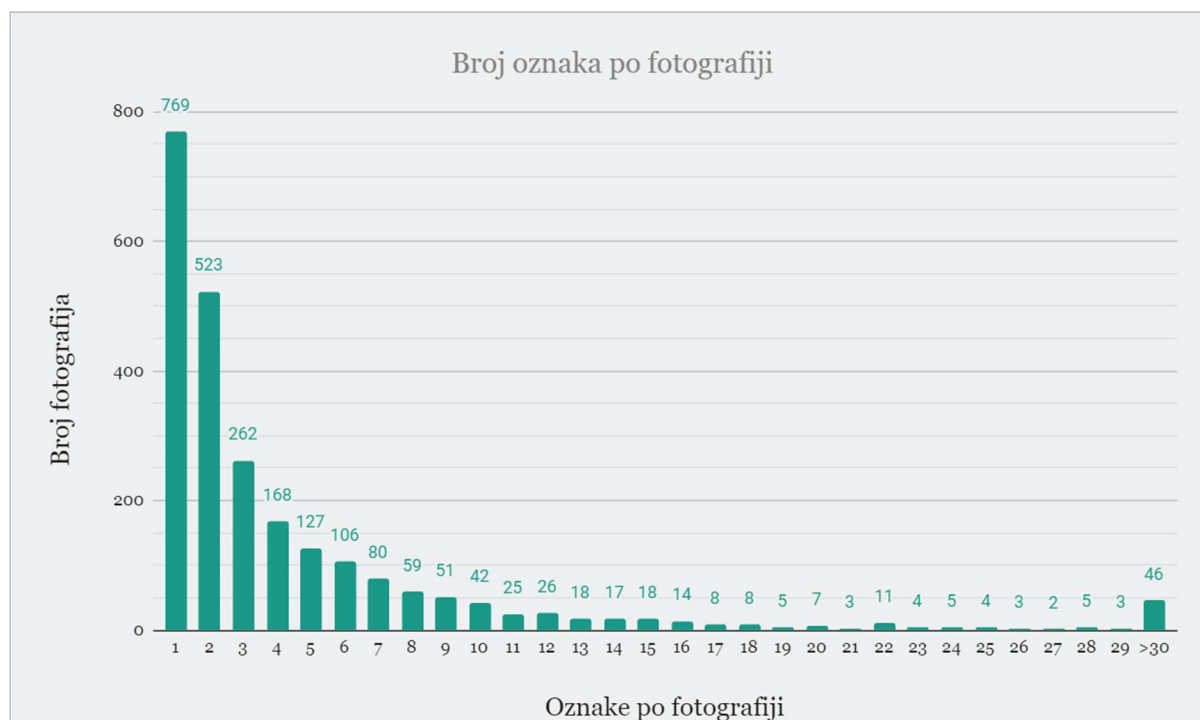
ID lokacije	Datum	Ukupan broj fotografija	Fotografije s objektima	Brzina vjetra (km/h)	Godišnje doba	GPS lokacija	Solarna elevacija (stupnjevi)	Azimut (stupnjevi)
9a9	16.12.2020.	52	39	3.0	Zima	43.40655, 16.78698	21.0	161.3
471	12.5.2021.	35	16	9.5	Proljeće	43.540434, 16.407359	60.6	147.3
831	12.5.2021.	31	5	9.5	Proljeće	43.540434, 16.407359	60.6	147.3
c39	31.5.2021.	23	23	10.5	Proljeće	43.5500965, 16.3642206	59.2	129.5
44d	8.7.2021.	56	56	10.3	Ljeto	43.5481875, 16.387769	18.0	284.3
73c	8.7.2021.	53	42	10.3	Ljeto	43.5481875, 16.387769	18.0	284.3
drb	27.7.2021.	126	49	4.0	Ljeto	43.5316156, 16.4656271	46.2	110.8
jFb	26.8.2021.	35	17	10.8	Ljeto	43.5439668, 16.4127614	26.7	101.8
tOV	26.8.2021.	126	51	10.8	Ljeto	43.5439668, 16.4127614	26.7	101.8
3h7	2.9.2021.	85	85	10.3	Ljeto	43.550427, 16.366794	24.4	103.2
NNm	13.9.2021.	103	52	21.3	Ljeto	43.545650, 16.422538	27.2	246.6
WRu	13.9.2021.	121	45	21.3	Ljeto	43.545650, 16.422538	27.2	246.6
v5s	22.9.2021.	193	17	8.7	Jesen	43.404919, 16.7958076	28.6	121.1
Khy	23.9.2021.	117	27	11.0	Jesen	43.4064802, 16.787435	23.6	115.0
oOR	23.9.2021.	128	20	11.0	Jesen	43.4064802, 16.787435	23.6	115.0
zLm	24.9.2021.	223	48	9.0	Jesen	43.4042214, 16.7957784	28.1	121.8
giS	1.10.2021.	228	51	7.0	Jesen	43.4045246, 16.7907076	40.1	156.2
ohY	3.10.2021.	209	34	14.0	Jesen	43.4039773, 16.7912421	5.7	270.1
E7o	12.10.2021.	145	51	9.3	Jesen	43.5492519, 16.3717885	20.6	124.3
9En	13.10.2021.	218	39	8.8	Jesen	43.5484399, 16.3787327	20.4	124.6
4b8	18.10.2021.	183	95	5.7	Jesen	43.5487482, 16.3757053	20.4	127.2
e4c	18.10.2021.	108	41	5.3	Jesen	43.546852, 16.3971274	26.4	137.4
aAH	19.10.2021.	350	218	6.5	Jesen	43.5455822, 16.417752	22.9	132.9
4Wj	20.10.2021.	238	34	4.0	Jesen	43.500656, 16.4828903	21.0	129.6
d6s	21.10.2021.	117	2	35.0	Jesen	43.50064655, 16.5151637	27.4	141.9
RfY	22.10.2021.	215	45	29.0	Jesen	43.4729118, 16.5699381	31.0	207.9
Aaa	21.2.2023.	310	231	1.0	Zima	43.55140, 16.35986	16.1	123.1
Aab	17.3.2023.	341	288	8.0	Zima	43.545263, 16.360627	21.9	115.6
Aac	18.3.2023.	337	136	3.0	Zima	43.549874, 16.364721	32.1	130.0
Aad	22.3.2023.	573	278	6.1	Proljeće	43.55002, 16.368229	30.6	131.6
Aae	24.3.2023.	326	284	2.0	Proljeće	43.548128, 16.385065	10.8	98.6

Slika 4.12 Metapodaci stanica

Baza se sastoji od 12 786 objekata razvrstanih u 31 klasu: fragmenti, građevinski materijal, gume, limenke, boce, užad, tkanina, torbe, ribolovni alat, čepovi i poklopci, čaše, kante, cijevi, obuća, kartonska ambalaža, posuđe, papir, lanci, stolice, kolica za kupovinu, kante za otpad, prometni znakovi, kontejneri, čamci, rukavice, dijelovi automobila, prometni čunjevi, peraje, romobili, električne bušilice i opušci kao što prikazuje Slika 4.13.

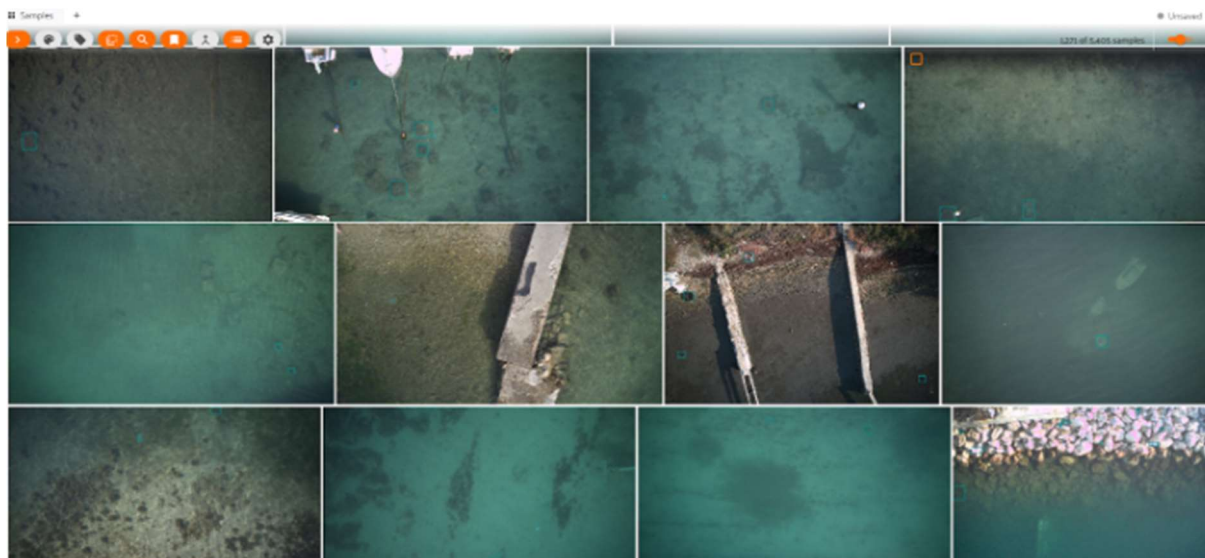


Mnoge fotografije imaju više različitih objekata prisutnih u isto vrijeme te distribuciju objekata po fotografiji prikazuje Slika 4.14. Može se vidjeti da većina fotografija ima jedan ili dva prisutna objekta, ali postoje fotografije s više od 30 prisutnih objekata.



Slika 4.14 Distribucija broja oznaka po jednoj slici

Fotografije su u RAW formatu (DNG) koji je pretvoren u PNG format i promijenjena im je veličina kako bi bile manje (širina i visina su prepolovljene). PNG fotografije trenutačno su u dvije rezolucije (ovisno o tome koja je letjelica korištena): 2732 x 1820 i 2732 x 1535. Sve fotografije su ručno označene i organizirane kao što prikazuje Slika 4.15.



Slika 4.15 Uzorci iz baze podataka

4.3. Pretprocesiranje fotografija

Odabrano je devet metoda pretprocesiranja koje mogu poboljšati kvalitetu fotografije. Metode CLAHE, ICM, ULAP, GC, MIP, DCP, RGHS, RoWS i GBdehazingRCorrection zajednički rješavaju širok raspon problema koji se javljaju u podvodnoj fotografiji, uključujući kontrast, izobličenje boja, izmaglicu i šum. Svaka od ovih metoda pokazala se učinkovitom u poboljšanju kvalitete podvodnih fotografija, nudeći pouzdane rezultate. Ove su metode prilagodljive različitim podvodnim uvjetima, osiguravajući dosljednu izvedbu u različitim scenarijima. Stoga se može osigurati sveobuhvatan i učinkovit pristup pretprocesiranju podvodnih fotografija što dovodi do značajno poboljšane kvalitete fotografija [199], [200]. Iako se u literaturi spomenute metode odnose na poboljšanje fotografija ispod površine mora, pretpostavlja se da bi se njihovom primjenom mogle poboljšati i snimke morskog dna iz zraka snimljene bespilotnom letjelicom. Svaka metoda je implementirana prema opisu rada. Ispitano je i može li kombinacija dvaju metoda pretprocesiranja biti kompatibilna i dati bolje rezultate.

4.3.1. Primjena pretprocesiranja koristeći duboke neuronske mreže

Metode pretprocesiranja fotografija primijenjene na bazu fotografija Selected Patches. Neke od metoda koje se koriste su kombinacija dvije metode, gdje su primijenjene jedna metoda za drugom kako bi se dobila konačna fotografija. Navedene baze podataka i modeli imaju podvlaku u nazivu koje označavaju koje su metode korištene i kojim redoslijedom. Ukupno je

generirano trinaest skupova fotografija koristeći ove metode: CLAHE, CLAHE_GC, CLAHE_ICM, CLAHE_MIP, DCP_MIP, GBdehazingRCorrection_RoWS, GBdehazingRCorrection_UDCP, GC_ICM, GC_RGHS, ICM, RoWS, RoWS_ULAP i ULAP. Zasebni modeli izgrađeni su za svaki od prethodno kreiranih skupova podataka i trenirani korištenjem metode prijenosnog učenja. Primijenjena je arhitektura YOLOv8 s modelom OpenImageV7 koji ima 68.7 milijuna parametara (YOLOv8x-oiv7) koji je služio kao bazni model. YOLOv8 je u trenutku početka istraživanja predstavljao najnoviju i najnapredniju inačicu YOLO serije modela za detekciju objekata. Modularna arhitektura te dostupnost unaprijed treniranih modela na velikim skupovima podataka, poput OpenImageV7, učinila ga je prikladnim odabirom za nadogradnju i prilagodbu specifičnom problemu detekcije morskog otpada. Dodatno su uvježbana još dva modela: model koji sadrži originalne zakrpe (nad kojima nije provedena metoda pretprocesiranja) i model koji ne uključuje metodu prijenosnog učenja. Treniranje modela je provedeno unutar 100 epoha, prag zaustavljanja postavljen na 30 epoha, korištenjem AdamW optimizatora, stopom učenja od 0.000769 i momentom od 0.937 [201]. Navedeni brojevi su predloženi od YOLOv8 znanstvenika bazirani na empirijskom testiranju i dosadašnjim istraživanjima. Svaki skup podataka podijeljen je u tri dijela: treniranje (*engl. train*), testiranje (*engl. test*) i validacija (*engl. validation*), s omjerom 60 / 20 / 20 (2394 / 809 / 813 fotografija) [202]. Ukupno je trenirano petnaest različitih modela – po jedan za svaki pripremljeni skup podataka te dva modela trenirana na neobrađenom skupu. Modeli *None* i *None no transfer* koriste se kao referentna osnova (*engl. baseline*) za usporedbu rezultata ostalih modela. Metode pretprocesiranja RoWS, ICM, CLAHE i ULAP odabrane su jer su prethodna istraživanja ukazala na njihove pozitivne učinke i poboljšane rezultate u kontekstu obrade podvodnih slika. Te su metode primijenjene pojedinačno, bez međusobnog kombiniranja. Kombinacije metoda (npr. CLAHE_GC i druge) odabrane su na temelju zaključaka o učincima pojedinih metoda, pri čemu su one kombinirane na svakoj slici s ciljem postizanja potencijalno boljih rezultata.

4.3.2. Analiza rezultata pretprocesiranja

U ciljanoj detekciji, regije se mogu klasificirati kao točne detekcije ili lažne detekcije što rezultira s četiri moguća ishoda: točno pozitivna detekcija (TP), lažno pozitivna detekcija (FP), točno negativna detekcija (TN) i lažno negativna detekcija (FN). Ako IoU između očekivanog okvira i referentnog okvira premašuje prag (0.5 je korišten u ovom istraživanju), očekivani okvir je označen kao TP. Inače je označen kao FP. Ako nijedan očekivani okvir ne odgovara

referentnom okviru, označava se kao FN. TP označava broj točno identificiranih traženih objekata, FP predstavlja broj netočno identificiranih traženih objekata, a FN označava broj traženih objekata koji nisu otkriveni. Izvedba modela obično se procjenjuje s pomoću mjera preciznost (Pr), odaziv (Re) i F_1 koje se izračunavaju pomoću jednadžbi 4.4, 4.5 i 4.6:

$$Pr = \frac{TP}{FP + TP} \quad (4.4)$$

$$Re = \frac{TP}{FN + TP} \quad (4.5)$$

$$F_1 = 2 \cdot \frac{Pr \cdot Re}{Pr + Re} \quad (4.6)$$

Preciznost i odaziv međusobno su povezane mjere. Kada preciznost ostane visoka dok se odaziv povećava, model radi bolje. Nasuprot tome, model s nižim performansama može doživjeti značajan pad preciznosti kako bi se postigao bolji odaziv. Za opis učinkovitosti modela korištena je krivulja „*Precision-Recall*”. Ocrta odnos između preciznosti (udio pravih pozitivnih predviđanja među svim pozitivnim predviđanjima) i odaziva (udio pravih pozitivnih predviđanja među svim stvarnim pozitivnim predviđanjima). Područje ispod „*Precision-Recall*” krivulje (*engl. Area Under Curve Precision-Recall - AUC-PR*) služi kao pokazatelj sposobnosti modela da održi preciznost preko različitih pragova odaziva. Svi prikazani rezultati odnose se na detekcije koje zadovoljavaju dva kriterija: vrijednost pouzdanosti (*engl. confidence*) veća od 0.45 te IoU veći od 0.5. Navedeni pragovi određuju koje se detekcije smatraju valjanima za evaluaciju.

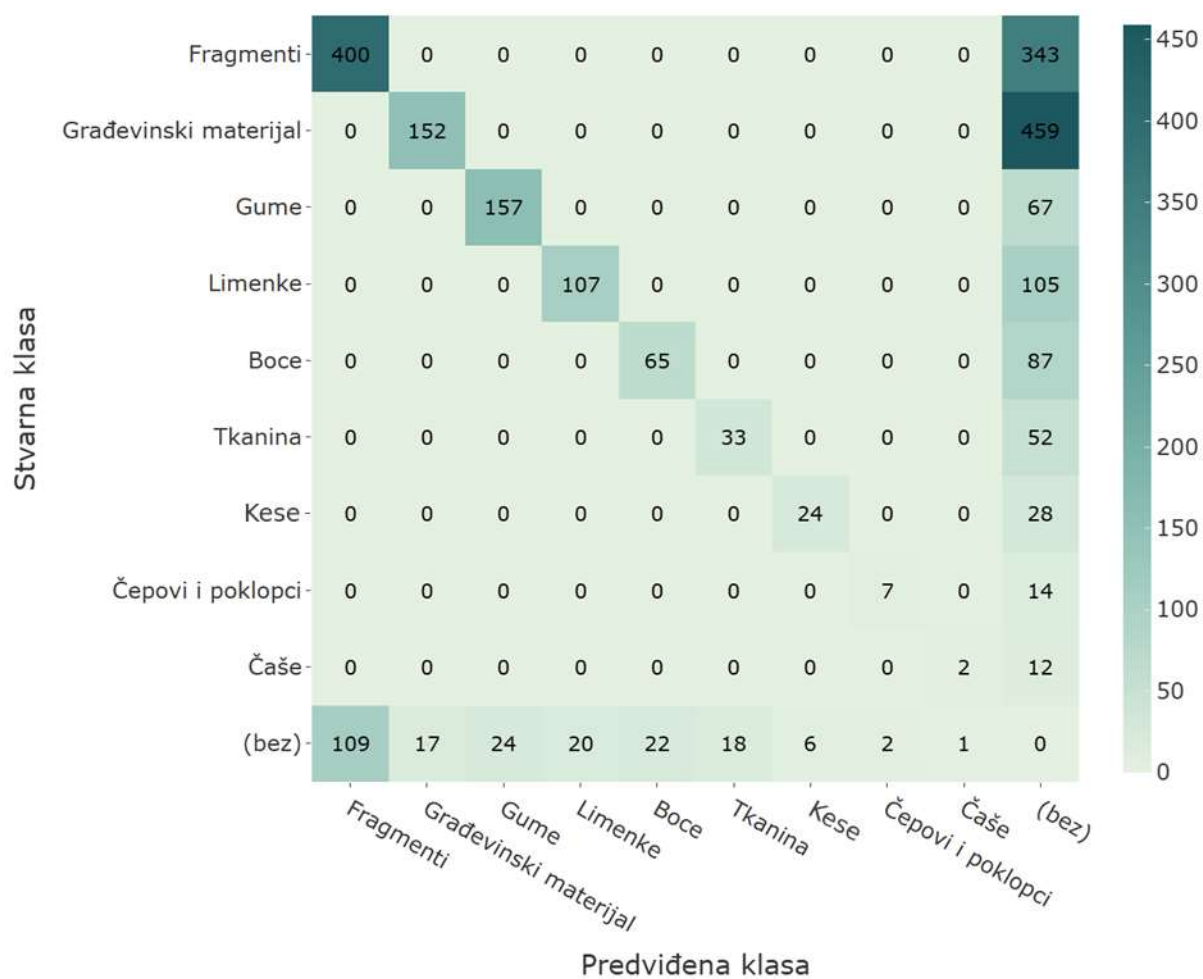
Model „None” označava da nije primijenjeno preprocesiranje već da je korišteno samo prijenosno učenje. Model „None no transfer” odnosi se na model treniran na skupu podataka bez ikakve prethodne obrade i bez korištenja metode prijenosnog učenja. Analizirajući različite modele za obradu fotografije, rezultati za preciznost, odaziv i mjeru F_1 prikazuje Tablica 4.2.

Model RoWS postigao je najvišu F_1 mjeru od 57.74 %, dok je model „None” imao najnižu F_1 mjeru od 42.79 %. Modeli poput CLAHE i ICM pokazali su solidne performanse s F_1 mjerom od 55.68 % odnosno 55.89 %. Modeli kao što su GC__ICM i GC__RGHS imali su slične rezultate s F_1 mjerom od 47.71 %. Ovi podaci ističu varijabilnost u izvedbi različitih modela u obradi fotografije, naglašavajući važnost odabira odgovarajućeg pristupa za specifične potrebe.

Tablica 4.2 Rezultati metoda pretprocesiranja

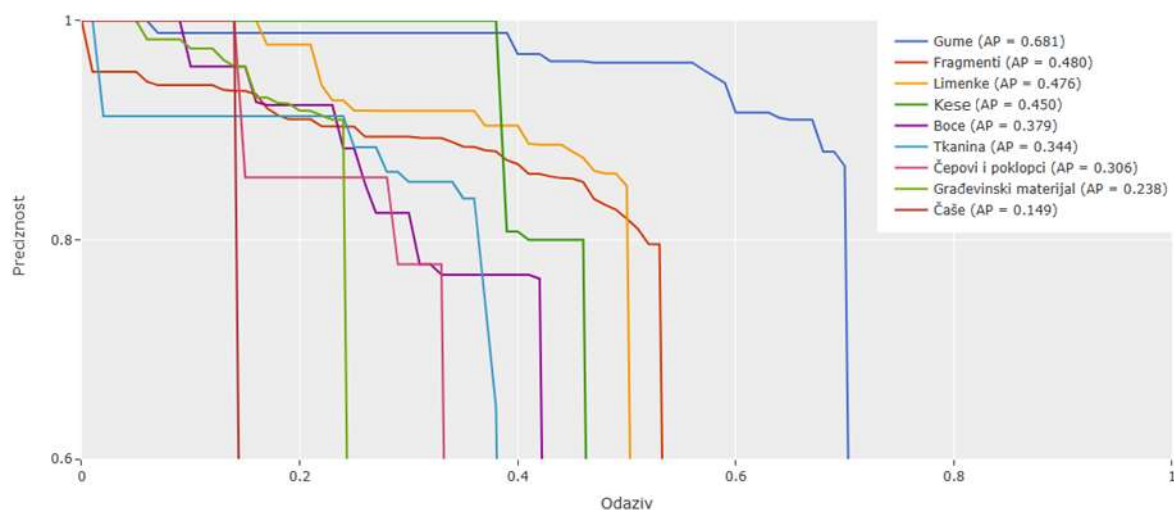
Model	Preciznost (%)	Odaziv (%)	F1 mjera (%)
CLAHE	77.89	43.33	55.68
CLAHE__GC	73.60	42.34	53.75
CLAHE__ICM	66.15	40.21	50.01
CLAHE__MIP	66.15	40.21	50.01
DCP__MIP	75.99	35.34	48.24
GBdehazingRCorrection__RoWS	76.56	29.66	42.75
GBdehazingRCorrection__UDCP	76.56	29.66	42.75
GC__ICM	65.81	37.42	47.71
GC__RGHS	65.81	37.42	47.71
ICM	73.78	44.99	55.89
None	77.42	44.94	56.87
None no transfer	65.06	31.88	42.79
RoWS	81.22	44.80	57.74
RoWS__ULAP	75.04	42.10	53.94
ULAP	72.69	37.65	49.61

Kao što prikazuje Slika 4.16, u *Confusion* matrici, najveće vrijednosti su koncentrirane duž dijagonale matrice što ukazuje da je model ispravno klasificirao većinu slučajeva za svaku kategoriju što sugerira općenito učinkovit model. Klasa „Fragmenti” ima velik broj točno pozitivnih rezultata što ukazuje na to da model ima vrlo dobre rezultate u identificiranju fragmenata. Međutim, postoji velik broj slučajeva u kojima fragmenti ili građevinski materijal uopće nisu identificirani (vidljivo u stupcu „(bez)”). Koncentracija visokih vrijednosti duž dijagonale i nižih vrijednosti izvan dijagonale sugerira da model radi dobro, ali ima problema s propuštanjem detekcija.



Slika 4.16 „Confusion” matrica za RoWS model pri IoU 0.5

Slika 4.17 prikazuje krivulju „Precision - Recall” pri IoU 0.5 i prikazuje sažetak jednog broja „PR-krivulje” (predstavlja područje ispod krivulje, AP). Gume pokazuju najveću prosječnu preciznost što pokazuje da modeli mogu učinkovito razlikovati gume od ostalih kategorija. Fragmenti i limenke pokazuju umjerenu izvedbu. Kесе, boce i druge kategorije pokazuju niži AP. To se može pripisati amorfnoj i poluprozirnoj strukturi koja modelima otežava izdvajanje oblika i značajki objekata.



Slika 4.17 Krivulja „Preciznost - Odaziv” za model RoWS pri IoU 0.5

Kao što prikazuje Tablica 4.3 rezultati otkrivaju značajnu varijabilnost u izvedbi prepoznavanja objekata korištenjem metode RoWS u različitim klasama objekata. Mjerni podaci o preciznosti, odazivu i F1 pokazuju da je metoda posebno učinkovita za prepoznavanje guma čime je postignuta visoka preciznost od 87 %, odaziv od 70 % i F1 mjera od 78 %. Suprotno tome, metoda je imala problema s čašama koje su imale najnižu mjeru s preciznošću od 67 %, odazivom od 14 % i F1 mjerom od 24 % što možemo pripisati bazi podataka koja ne sadrži dovoljno fotografija s različitim varijacijama čaša.

Tablica 4.3 Mjerila izvedbe prepoznavanja objekata korištenjem RoWS metoda u različitim klasama objekata

Klasa	Broj objekata	Preciznost (%)	Odaziv (%)	F1 mjera (%)
Fragmenti	3757	79	54	64
Građevinski materijal	3465	90	25	39
Gume	1140	87	70	78
Limenke	1002	84	50	63
Boce	768	75	43	54
Tkanina	390	65	39	49
Kese	262	80	46	59
Čepovi i poklopci	104	78	33	47
Čaše	74	67	14	24

Ostali predmeti kao što su fragmenti, limenke i boce pokazali su umjerenu izvedbu s preciznošću u rasponu od 75 % do 84 % i odazivom između 43 % i 54 % što je rezultiralo F1 mjerom od oko 54 % do 64 %. Građevinski materijali pokazali su visoku preciznost od 90 %, ali niski odaziv od 25 % što ukazuje da iako je metoda točna kada otkriva te materijale, propušta mnoge slučajeve. Općenito RoWS metoda pokazuje različitu učinkovitost, ističući se u nekim kategorijama kao što su gume dok pokazuje ograničenja u drugima kao što su čaše i tkanina.

4.4. Prilagodba arhitekture konvolucijske mreže

Budući da standardne arhitekture u svojoj osnovnoj formi primarno rade nad RGB slikama, njihova sposobnost generalizacije u izazovnim vizualnim uvjetima može biti ograničena. S ciljem unaprjeđenja performansi detekcije, potrebna je prilagodba strukture konvolucijske neuronske mreže kroz uvođenje dodatnog informacijskog kanala, kao što je komponenta nijanse (Hue) ili dubinske informacije (Depth). Ovakav pristup omogućuje mreži da iskoristi dodatne informacije o sceni koje nisu nužno prisutne u standardnim kanalima fotografije, čime se potencijalno poboljšava sposobnost razlikovanja relevantnih objekata od pozadine.

4.4.1. Uvođenje četvrtog kanala

Za potrebe kreiranja četverokanalnih skupova podataka, kako bi trenirali četverokanalne modele, definiraju se dva načina kreiranja dodatnog četvrtog kanala:

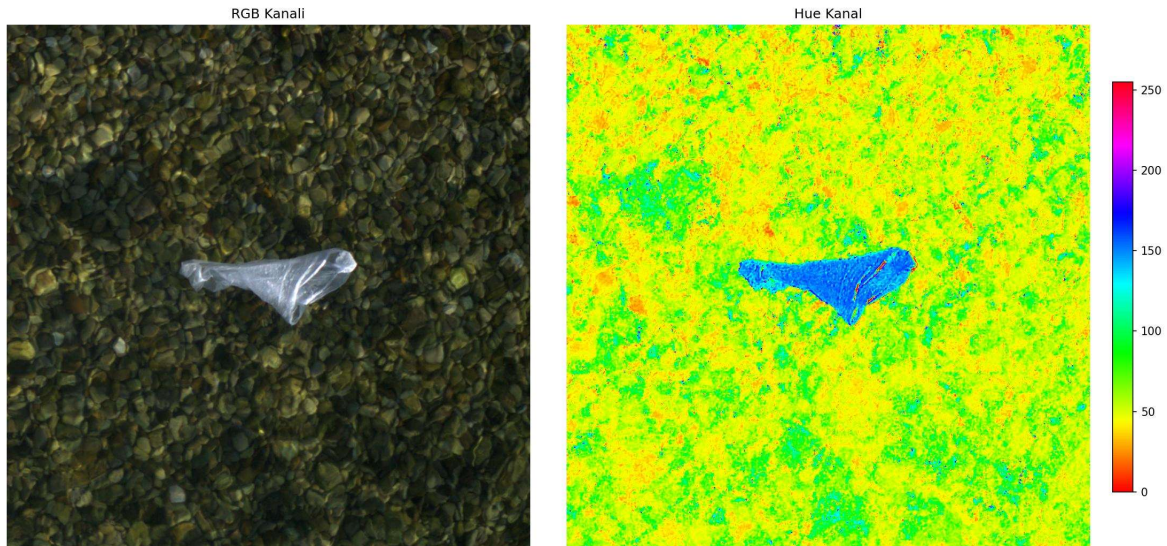
a) Hue kao četvrti kanal

Pretpostavka je da informacije o boji (nijansi) objekata ostaju barem djelomično očuvane unatoč degradaciji fotografije uzrokovanoj prolaskom svjetlosti kroz vodeni stupac. Uvođenje dodatnog kanala koji eksplicitno ističe nijansu (engl. Hue) ima za cilj pružiti modelu dodatnu semantičku informaciju koja može olakšati razlikovanje objekata sličnih oblika, ali različitih boja. Budući da boja objekta može biti značajna za njegovu klasifikaciju (npr. plastična vrećica naspram metalne limenke), izdvajanje nijanse kao zasebne značajke potencijalno doprinosi preciznijoj detekciji i klasifikaciji.

Svaka RGB fotografija konvertirana je u HSL (Hue, Saturation, Lightness) prostor boja korištenjem OpenCV funkcije `cv2.cvtColor(img, cv2.COLOR_BGR2HLS)`. Kanal nijanse (Hue), koji u OpenCV HLS implementaciji ima raspon vrijednosti $[0, 179]$, izoliran je i skaliran na raspon cjelobrojnih vrijednosti $[0, 255]$ prema jednadžbi 4.7:

$$H(x, y) = \text{round}\left(H_{\text{hsl}}(x, y) \cdot \frac{255}{180}\right) \quad (4.7)$$

gdje je $H_{\text{hsl}}(x, y)$ vrijednost piksela na poziciji (x, y) u kanalu nijanse, a *round* funkcija cjelobrojnog zaokruživanja. Ovaj skalirani kanal nijanse (Hue) dodan je kao četvrti kanal originalnoj RGB slici. Slika 4.18 prikazuje originalnu sliku usporedno s vrijednostima Hue kanala.

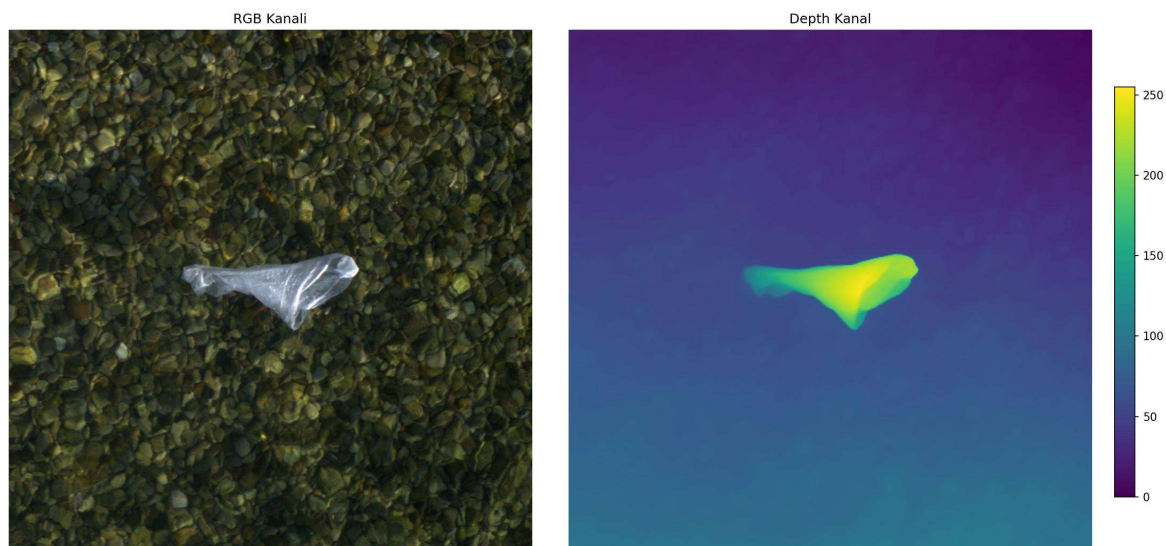


Slika 4.18 Prikaz originalne fotografije s kanalom Hue

b) Dubina kao četvrti kanal

Dodavanje informacije o relativnoj udaljenosti objekata od kamere (prostornoj dubini scene), koja nije direktno prisutna u RGB podacima, može pomoći modelu u razlikovanju objekata od pozadine i razumijevanju trodimenzionalne strukture scene.

Za svaku RGB fotografiju procijenjena je dubinska mapa korištenjem prethodno treniranog modela umjetne inteligencije za monokularnu procjenu dubine, DepthAnything V2 [197]. Izlaz modela je jednokanalna slika (dubinska mapa) gdje vrijednosti piksela predstavljaju relativnu udaljenost od kamere, što je veća vrijednost piksela to se točka nalazi bliže u prostoru. Iako u slučaju objekata na dnu ne dolazi do izražene separacije po dubini, algoritam može identificirati lokalne varijacije u obliku, konturama i sjenama, što se reflektira u vrijednostima dubinske mape. Takve informacije služe za poboljšanje detekcije posebno kod objekata čija se silueta i prostorni kontekst razlikuju od ujednačene teksture morskog dna. Navedena dubinska mapa, normalizirana u rasponu $[0, 255]$, dodana je kao četvrti kanal originalnoj RGB slici. Slika 4.19 prikazuje originalnu sliku usporedno s vrijednostima Depth kanala.



Slika 4.19 Prikaz originalne fotografije s kanalom Depth

4.4.2. Prilagodba YOLOv8 arhitekture za četverokanalne fotografije

YOLOv8 framework modificiran je kako bi omogućio obradu fotografija koje osim standardna tri RGB kanala uključuju i četvrti kanal (Hue ili dubinu). Prilagodba arhitekture sastoji se od modifikacije slojeva neuronske mreže tako da prihvate fotografije s četiri kanala umjesto uobičajena tri. Informacije sadržane u četvrtom kanalu skalirane su u rasponu od 0 do 255, a sama specifična informacija koja se prenosi tim kanalom može varirati. Uz promjenu mreže uvedene su i funkcije kojima se inicijalne težine modela prilagođavaju tako da se ispravno inicijalizira četvrti kanal.

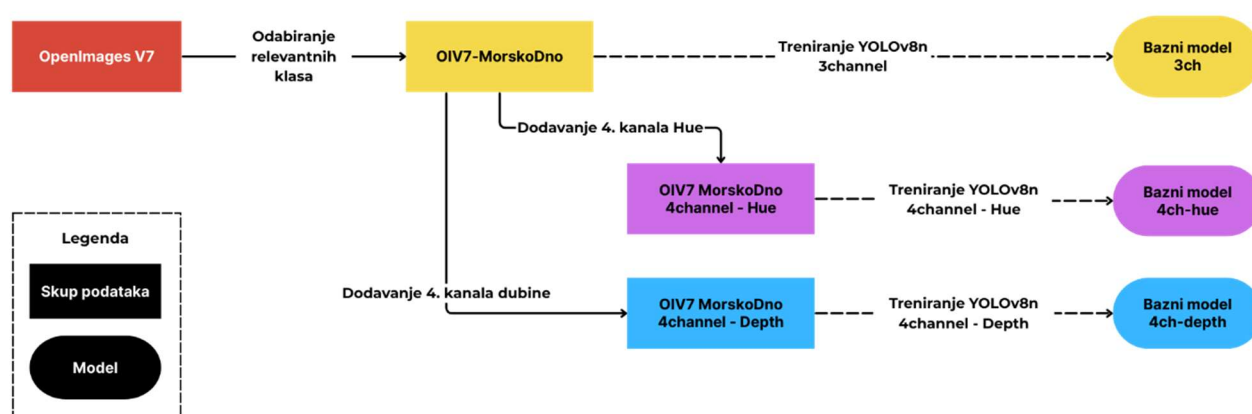
Osim arhitekture modela, prilagođen je i proces učitavanja i augmentacije fotografija. Proces učitavanja modificiran je kako bi se osiguralo očuvanje informacije sadržane u četvrtom kanalu, s obzirom da je originalni YOLOv8 implicitno pretpostavlja da slike imaju isključivo tri kanala. Modificirani sustav ispravno učitava PNG ili TIF fotografije koje imaju četiri kanala. PNG u četvrtom kanalu sprema vrijednost transparentnosti, međutim u ovom primjeru se to koristi kao informacija o Hue ili dubini. Budući da se augmentacija HSV vrijednosti već primjenjuje na RGB sliku, direktna augmentacija na postojećem Hue kanalu dovela bi do nedosljednosti. Stoga se sve promjene u boji primjenjuju isključivo na originalnoj RGB slici, a četvrti kanal (Hue) izračuna iz transformirane RGB slike u svakom koraku augmentacije, čime je očuvan odnos među kanalima. Navedene modifikacije omogućile su uspješno treniranje i evaluaciju YOLOv8 modela korištenjem četverokanalnih fotografija. Detaljan

tehnički opis svih promjena u kodu, uključujući specifične izmjene u pojedinim datotekama i funkcijama dostupne su u PRILOG A Modifikacije YOLOv8 za podršku četverokanalnih slika.

4.4.3. Integracija metode pretprocesiranja i prilagođene arhitekture

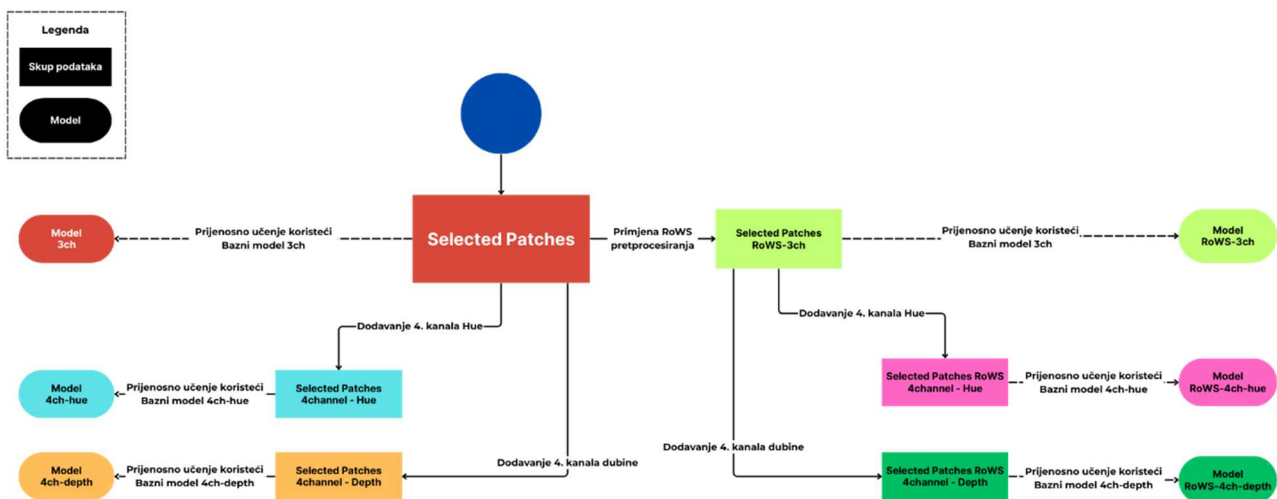
Kako bi se rezultati trokanalnih i četverokanalnih neuronskih mreža mogli ispravno usporediti, nužno je osigurati ujednačene uvjete treniranja. Iako su unaprijed istrenirani (*engl. pre-trained*) trokanalni modeli dostupni, njihova izravna primjena nije moguća u četverokanalnom kontekstu, budući da ne postoje ekvivalentni modeli trenirani na četiri ulazna kanala. Stoga je u ovom istraživanju bilo potrebno konstruirati dvije osnovne baze podataka: prva služi za treniranje baznih modela u trokanalnoj i četverokanalnoj arhitekturi (kao zamjena za standardne predtrenirane modele), a druga je specijalizirana za detekciju otpada na morskom dnu te se koristi za fino podešavanje modela, evaluaciju i usporedbu performansi.

U svrhu kreiranja baznih modela izrađen je skup podataka pod nazivom „OIV7-MorskoDno“, temeljen na odabiru klasa iz skupa OpenImagesV7 koje predstavljaju objekte koji se potencijalno mogu naći na morskom dnu. Skup sadrži 8715 fotografija za trening i 2081 za validaciju, kategoriziranih u 29 klasa poput boca, automobilskih dijelova, plastičnih vrećica i sl. Navedeni skup podataka „OIV7-MorskoDno“ korišten je za treniranje baznog trokanalnog YOLOv8n modela „Bazni model 3ch“. Da bi se ispravno trenirala četverokanalna mreža, iz skupa podataka „OIV7-MorskoDno“ kreirana su dva skupa: „OIV7-MorskoDno 4channel - Hue“ i „OIV7-MorskoDno 4channel - Depth“. Navedena dva skupa su korištena za treniranje dva četverokanalna modela („Bazni model 4ch-hue“ i „Bazni model 4ch-depth“), koji čine bazne modele za prijenosno učenje, kako prikazuje Slika 4.20.



Slika 4.20 Proces treniranja baznih modela

U svrhu kreiranja modela prijenosnog učenja korišten je skup *Selected Patches*. Skup se sastoji od 4016 fotografija veličine 640×640 piksela, koji sadržava 9 specifičnih klasa: fragmenti, građevinski materijal, gume, limenke, boce, tkanine, kese, čepovi i čaše. Kako prikazuje Slika 4.21, navedeni skup podataka *Selected Patches* korišten je za treniranje trokanalnog YOLOv8n modela „Model 3ch“ tehnikom prijenosnog učenja koristeći pred-trenirani „Bazni model 3ch“. Za potrebe treniranja četverokanalnih mreža, iz skupa podataka *Selected Patches* kreirana su dva skupa: „Selected Patches 4channel - Hue“ i „Selected Patches 4channel - Depth“. Navedena dva skupa su korištena za treniranje dva četverokanalna modela tehnikom prijenosnog učenja koristeći pre-trenirane četverokanalne bazne modele, čime su dobiveni modeli „Model 4ch-hue“ i „Model 4ch-depth“.



Slika 4.21 Postupak treniranja modela

Za potrebe ovog istraživanja potrebno je provesti analizu metode pretprocesiranja koristeći prilagođenu četverokanalnu arhitekturu mreže. Metoda s najboljim rezultatima, prikazano u prethodnom poglavlju, je RoWS metoda. Nad skupom podataka *Selected Patches* izvršeno je pretprocesiranje koristeći RoWS metodu te je dobiven skup podataka „Selected Patches RoWS-3ch“, koji je korišten za treniranje trokanalnog YOLOv8n modela „Model RoWS-3ch“ tehnikom prijenosnog učenja koristeći pred-trenirani „Bazni model 3ch“. Za potrebe treniranja četverokanalnih mreža, iz skupa podataka „Selected Patches RoWS-3ch“ kreirana su dva skupa: „Selected Patches RoWS - 4channel - Hue“ i „Selected Patches RoWS - 4channel - Depth“. Navedena dva skupa su korištena za treniranje dva četverokanalna modela tehnikom prijenosnog učenja koristeći pre-trenirane četverokanalne bazne modele, čime su dobiveni modeli „Model RoWS-4ch-hue“ i „Model RoWS-4ch-depth“.

Kompletan proces kreiranja svih skupova podataka i treniranja modela prikazuje Slika 4.21.

4.4.4. Postupak treniranja modela

Za potrebe treniranja modela korišteno je računalo u oblaku Google Cloud Platform Compute Engine s operacijskim sustavom Linux-5.15.0-1078-gcp-x86_64-with-glibc2.29. Računalo ima 2 fizička (4 logička) procesora te četiri Tesla T4 grafičke kartice. Korištena Python verzija je 3.8.10. YOLOv8 framework korišten je u verziji 8.0.235, a PyTorch u verziji 2.4.0.

Praćenje eksperimenata, uključujući mjerenje treniranja i validacije (gubitak, preciznost, odaziv, mAP) provedeno je pomoću platforme Weights & Biases. Za početnu fazu odabrana je YOLOv8n arhitektura zbog kompaktnosti (3.5 milijuna parametara) te manjih računalnih zahtjeva u odnosu na veće varijante YOLOv8 modela, čime je postignuta brza evaluacija učinkovitosti različitih pristupa. Treniranje je ograničeno na 100 epoha s optimizatorom AdamW, početnom stopom učenja od 0.000769 i momentumom od 0.9. Rezultat treniranja jest šest modela koji su nastali tehnikom prijenosnog učenja koji kombiniraju RoWS metodu i dodatni četvrti kanal. Nad tim modelima je provedena evaluacija rezultata. Metoda modela koji pokaže najbolje rezultate će biti korištena za treniranje finalnog modela koristeći YOLOv8x arhitekturu s 68.7 milijuna parametara koristeći već definirani skup podataka *Selected Patches*.

4.4.5. Metodologija evaluacije modela

Performanse konačnih šest modela evaluirane su na izdvojenom testnom dijelu skupa podataka *Selected Patches*. Za svaki model korištena je pripadajuća inačica testnog skupa. Testni skup nije korišten tijekom treniranja ni validacije. Predviđanja na testnim fotografijama generirane su s minimalnim IoU pragom od 0.5 i minimalnim pragom pouzdanosti od 0.3. Za kvantifikaciju performansi korištene su mjere: preciznost (*Precision*), odaziv (*Recall*) i F1 mjera (*F1-Score*). Ukupna mjera modela je izražena korištenjem mAP@50 mjere.

4.5. Analiza rezultata

Ovaj odjeljak potvrđuje da će upotreba odgovarajuće metoda pretprocesiranja u prilagođenoj arhitekturi poboljšati kvalitetu podvodnih fotografija i time poboljšati rezultate algoritma za detekciju podvodnih objekata.

Rezultate evaluacije prikazuje Tablica 4.4. Ono što treba naglasiti prilikom tumačenja rezultata jest da su ovi modeli trenirani na baznom modelu arhitekture YOLOv8n koji ima 3.5 milijuna parametara dok su rezultati evaluacije za metode pretprocesiranja obavljani na modelu YOLOv8x koji ima 68.7 milijuna parametara.

Tablica 4.4 Rezultati performansi treniranih YOLOv8n modela na neviđenom testnom skupu podataka

Model	Preciznost (%)	Odaziv (%)	F1 mjera (%)	mAP@50 (%)
3ch	47.14	23.05	29.21	14.62
4ch-depth	50.29	28.05	34.36	20.45
4ch-hue	57.50	28.43	35.92	21.51
RoWS-3ch	54.92	26.51	33.92	16.46
RoWS-4ch-depth	52.99	28.53	35.64	21.09
RoWS-4ch-hue	59.55	33.05	40.36	25.25

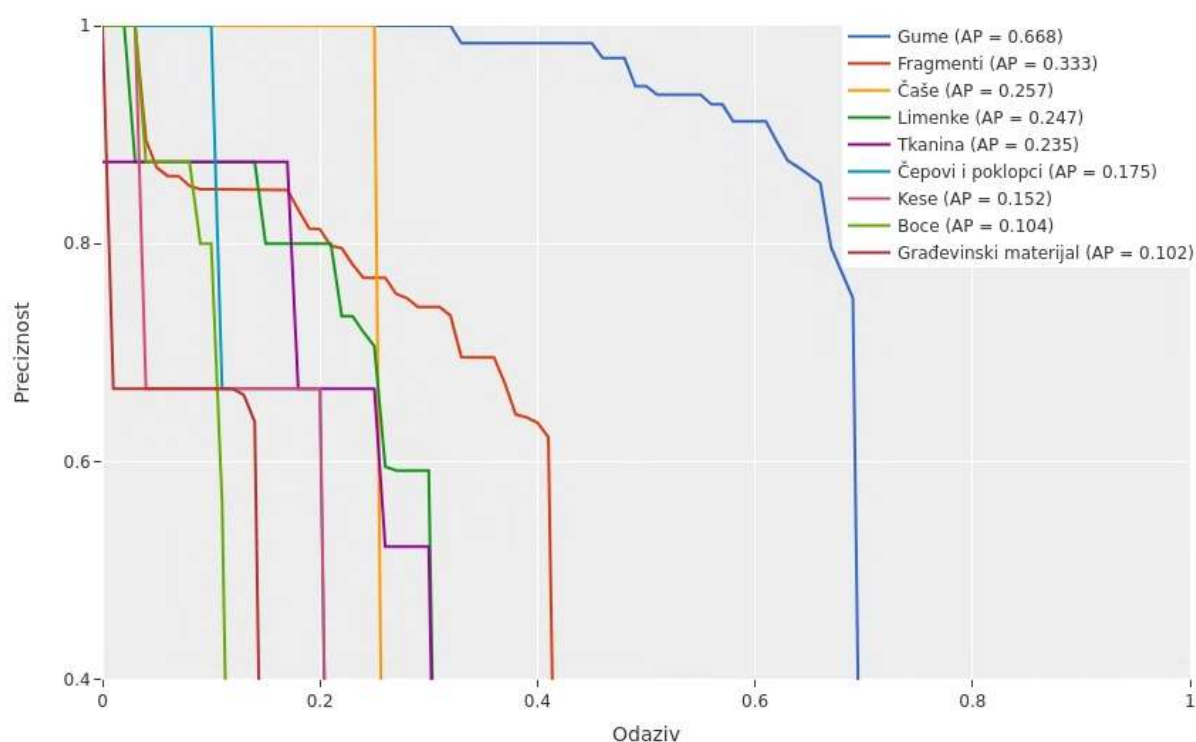
Dobiveni rezultati evaluacije pokazuju jasno povećanje performansi detekcije objekata na morskom dnu primjenom četverokanalnih modela u odnosu na standardni trokanalni model. Trokanalni model (3ch) ostvario je najniže rezultate u svim promatranim mjerama: preciznost od 47.14%, odaziv od 23.05%, F1 mjera od 29.21% te srednju prosječnu preciznost (mAP) od 14.62%.

Primjena četvrtog kanala dubine (Model 4ch-depth) rezultirala je primjetnim poboljšanjem svih mjerenih performansi: preciznost se povećala na 50.29%, odaziv na 28.05%, F1 mjera na 34.36%, dok se mAP povećao na 20.45%. Dodatno, korištenje kanala Hue (Model 4ch-hue) pokazalo se još učinkovitijim, s preciznošću od 57.50%, odazivom od 28.43%, F1 mjerom od 35.92% te mAP vrijednošću od 21.51%. Navedeni rezultati ukazuju na jasnu korist od uvođenja dodatnih kanala, pri čemu kanal Hue daje bolje rezultate u odnosu na kanal dubine.

Pretprocesiranje fotografija metodom RoWS također pozitivno utječe na detekciju, no njegovi su učinci različiti ovisno o korištenom broju kanala. U slučaju trokanalnog RoWS modela (Model RoWS-3ch), ostvarena je preciznost od 54.92%, odaziv 26.51%, F1 mjera 33.92% i mAP od 16.46%, što je poboljšanje u odnosu na trokanalni model, no rezultati su slabiji od četverokanalnih modela bez primjene RoWS metode.

Najbolje ukupne performanse ostvarene su kombinacijom metode RoWS i četvrtog kanala nijanse boje (Model RoWS-4ch-hue). Ovaj model postigao je preciznost od 59.55%, odaziv od 33.05%, F1 mjeru od 40.36% te najvišu srednju prosječnu preciznost (mAP) od 25.25%. Time je potvrđeno kako kombinacija primjene RoWS metode pretprocesiranja i dodatnog Hue kanala, zajedno s prilagođenom modifikacijom ulaznih podataka, pruža optimalnu strategiju za detekciju objekata u podvodnom okruženju.

Slika 4.22 prikazuje krivulju „Precision - Recall” pri IoU 0.5 te legendu koja prikazuje prosječnu preciznost za svaku klasu – AP (*engl. Average Precision*), koja predstavlja površinu područja ispod krivulje. Gume pokazuju najveću prosječnu preciznost što pokazuje da modeli mogu učinkovito razlikovati gume od ostalih kategorija. Fragmenti, čaše, limenke i tkanina pokazuju umjerenu izvedbu. Čepovi, kese, boce i građevinski materijal pokazuju niži AP. To se može pripisati amorfnoj i poluprozirnoj strukturi koja modelima otežava izdvajanje oblika i značajki objekata.



Slika 4.22 Krivulja „Preciznost - Odaziv” za model RoWS-4ch-hue pri IoU 0.5

Iako su razlike umjerene zbog ograničene veličine korištenog YOLOv8n modela, potrebno je napraviti analizu korištenjem veće arhitekture YOLOv8x. Korištenjem iste metodologije kreiranja baznog modela i nakon toga tehnike prijenosnog učenja korištenjem već definiranog skupa podataka „Selected Patches RoWS - 4channel – Hue“ treniran je model YOLOv8x koristeći prethodno navedene parametre. Provedena je evaluacija modela unutar već postavljenih okvira te dobivene rezultate prikazuje Tablica 4.5, zajedno s rezultatima prethodnih analiza.

Tablica 4.5 Usporedba rezultata performansi treniranih modela na neviđenom testnom skupu podataka

Model	Arhitektura	Preciznost (%)	Odaziv (%)	F1 mjera (%)	mAP@50 (%)
RoWS (3ch)	YOLOv8x	81.22	44.80	57.75	38.90
RoWS-4ch-hue	YOLOv8n	59.55	33.05	40.36	25.25
RoWS-4ch-hue	YOLOv8x	84.09	67.54	74.53	71.08

„Model RoWS-4ch-hue“ je postigao preciznost od 84.09%, odaziv od 67.54%, F1 mjeru od 74.53% te najvišu srednju prosječnu preciznost (mAP) od 71.08%. Usporedno s prethodnim najboljim modelom na YOLOv8n, ovaj model postiže 25.54 % veću preciznost, 34.49 % veći odaziv, 34.17 % veću mjeru F1 te mAP@50 veći za 45.83 %. Treba napomenuti da je model „Rows-3ch“ korišten na YOLOv8x verziji modela koja je pred-trenirana na kompletnom OpenImagesV7 setu podataka, dok je za ostala dva modela korišten podskup fotografija iz OpenImagesV7 na kojima se nalaze klase koje predstavljaju objekte koji se potencijalno mogu naći na morskom dnu. Time je potvrđeno kako korištenje YOLOv8x arhitekture s primjenom RoWS metode pretprocesiranja i dodatnog Hue kanala postiže rezultate koji su bolji od primjene samo metode pretprocesiranja na trokanalnim slikama koje su trenirane na YOLOv8x verziji.

4.6. Predloženi postupak za detekciju objekata

Na temelju dobivenih rezultata ovog istraživanja, može se zaključiti kako primjena predložene metode pretprocesiranja fotografija korištenjem RoWS metode, uz istovremenu implementaciju prilagodbe YOLOv8 arhitekture za prihvrat četverokanalnih ulaznih podataka gdje se Hue kanal koristi kao dodatni, četvrti kanal, omogućuje poboljšanje detekcije objekata na fotografijama podvodnog okruženja. S obzirom na utvrđene prednosti ovog pristupa, predlaže se njegova upotreba kao preporučeni standard u odnosu na tradicionalne trokanalne modele bez specijaliziranog pretprocesiranja. Ovaj pristup pruža mogućnost pouzdanije identifikacije objekata te može pozitivno utjecati na operacije poput pretraživanja ili ekološkog monitoringa podvodnih područja, čime doprinosi smanjenju vremena, resursa i ekološki negativnih posljedica povezanih s nezadovoljavajućim metodama detekcije.

5. Zaključak

Ovaj rad predstavlja sveobuhvatan pristup otkrivanju morskog otpada s pomoću bespilotnih letjelica (dronova) putem napredne predobrade fotografije i tehnika dubokog učenja. Značajan doprinos ovog istraživanja je razvoj nove baze podataka s preko 5000 fotografija i više od 12000 objekata kategoriziranih u 31 klasu. Baza podataka obogaćena je metapodacima kao što su GPS lokacija, brzina vjetrova, nadmorska visina sunca i azimut što poboljšava kontekst i upotrebljivost podataka.

Potrebno je naglasiti da se sve navedene metode pretprocesiranja inače koriste na fotografijama koje se prikupljaju snimanjem ispod morske površine. U ovom radu je prikazano kako se mogu koristiti za unaprijeđenje i restauraciju i za fotografije prikupljenih bespilotnom letjelicom iz zraka. U radu se analiziraju različite metode pretprocesiranja fotografija za poboljšanje detekcije podvodnih objekata pri čemu metoda RoWS pokazuje najbolju izvedbu.

Predložena višekanalna arhitektura duboke neuronske mreže značajno poboljšava detekcije u usporedbi s postojećim modelima. Budući rad trebao bi se usredotočiti na dobivanje fotografija s različitih lokacija i stvaranje veće baze podataka. Dio istraživanja trebao bi biti usmjeren na različite tehnike za uklanjanje valova i optičkih izobličenja fotografija po mogućnosti proširenjem postojeće neuronske mreže s posebnim algoritmima koji su specijalizirani za rješavanje navedenih problema, čime bi se uvela poboljšanja u sloju prijelaza zrak – morska površina.

Rezultati potvrđuju da je RoWS pretprocesiranje ključno za poboljšanje performansi detekcije u podvodnim slikama. Hipoteza da prilagodba arhitekture duboke neuronske mreže kao što je dodavanje četvrtog kanala može dodatno poboljšati rezultate je potvrđena, s tim da je kanal Hue pokazao veću korisnost od kanala dubine (Depth) u kombinaciji s metodom RoWS za korištenu YOLOv8n arhitekturu. Četverokanalna metoda s RoWS pretprocesiranjem se preporučuje kao učinkovitiji pristup u usporedbi s tradicionalnim trokanalnim metodama u domeni računalnog vida za podvodne objekte.

Korištenje YOLOv8n modela, kako je prethodno napomenuto u metodologiji istraživanja, predstavlja ograničenje u postizanju optimalnih performansi. Stoga je urađena analiza rezultata, koristeći istu metodologiju, arhitekture YOLOv8x koja posjeduje veći kapacitet te time može efikasnije iskoristiti dodatne informacije iz četvrtog kanala. Sukladno tome, postignuti su bolji rezultati s većom arhitekturom.

Rezultati pokazuju da odgovarajuća upotreba baze podataka i metoda pretprocesiranja može uvelike poboljšati detekciju podvodnih objekata, nudeći obećavajuće rješenje za izazov otkrivanja morskog otpada i praćenja okoliša.

5.1. Doprinos disertacije

Rezultati ovog istraživanja su sukladni očekivanom znanstvenom doprinosu:

- 1) Izrađena je referentna i javno dostupna baza označenih fotografija morskog otpada ispod površine mora prikupljenih bespilotnom letjelicom.
- 2) Dokazano je da kombinacija pretprocesiranja i korištenja modificirane arhitekture poboljšava detekciju morskog otpada u odnosu na do sada najuspješnije pojedinačne metode.

5.2. Smjernice za daljnji rad

Preporučuje se daljnje istraživanje usmjereno na analizu performansi predloženih algoritama na posebno kreiranom skupu podataka. Skup bi trebao biti prikupljen s jedne specifične geografske lokacije, pri različitim visinama leta drona, različitim vremenskim uvjetima (npr. različite brzine i smjerovi vjetrova) te pri različitim oblicima morskih valova. Potrebno je istražiti utječu li vremenski uvjeti (primjerice razdoblje između 8 i 12 sati prijepodne u odnosu na druga razdoblja dana) na rezultate modela te postoji li statistički značajna razlika u detekciji objekata kada se treniranje i evaluacija provode na skupovima fotografija prikupljenih u različitim vremenskim intervalima. Takav skup podataka, iako zahtjevniji za prikupljanje, omogućio bi detaljniju analizu utjecaja specifičnih uvjeta snimanja na performanse modela te istaknuo prednosti i nedostatke primijenjenih metoda pretprocesiranja. Nadalje, omogućio bi i izradu preporuka za pretprocesiranje ovisno o uvjetima u kojima su fotografije snimljene.

Jedan od ključnih problema koji se istaknuo u ovom istraživanju jest izražena neravnoteža klasa u korištenom skupu podataka. Čak i pri ograničenju na samo devet klasa, nebalansiranost je značajna, što može negativno utjecati na rezultate treniranja i evaluacije. Stoga se u daljnjem radu predlaže prikupljanje uravnoteženijih i raznovrsnijih skupova podataka. Također, potrebno je detaljno istražiti različite metode za rješavanje ovog problema, poput tehnika umjetnog povećanja podataka (*engl. oversampling*), tehnika smanjenja broja dominantnih klasa (*engl. undersampling*), primjene sintetskog generiranja podataka (npr. metoda SMOTE) ili korištenja specifičnih funkcija gubitka (poput „focal loss”). Značajnim proširenjem skupa

fotografija ostvarili bi se preduvjeti i za uspješniju primjenu nekih od transformer arhitektura te iskorištavanje njihovih potencijala.

Preporučuje se istražiti alternativne vrste dodatnih kanala ili primijeniti naprednije metode integracije (fuzije) informacija između postojećih i novih kanala, čime bi se dodatno povećala učinkovitost detekcije objekata na morskom dnu.

Iako je fokus ovog istraživanja bio na detekciji objekata na morskom dnu, korištene metode mogu se u budućnosti ispitati i prilagoditi za druge relevantne primjene. Primjerice, sustav bi mogao biti prilagođen za potrebe detekcije osoba u moru (npr. za potrebe spašavanja), detekciju brodova ili plovila na površini mora, identifikaciju zaštićenih vrsta morske flore i faune ili detekciju potencijalnih ekoloških incidenata poput izlivanja opasnih tvari.

Literatura

- [1] L. Jeftic, S. B. Sheavly, E. Adler, i N. Meith, *Marine litter: a global challenge*. Nairobi, Kenya: Regional Seas, United Nations Environment Programme, 2009.
- [2] J. Hong, M. Fulton, i J. Sattar, „A Generative Approach Towards Improved Robotic Detection of Marine Litter“, u *2020 IEEE International Conference on Robotics and Automation (ICRA)*, Paris, France: IEEE, svi. 2020, str. 10525–10531. doi: 10.1109/ICRA40945.2020.9197575.
- [3] J. Darby, M. Clairbaux, A. Bennison, J. L. Quinn, i M. J. Jessopp, „Underwater visibility constrains the foraging behaviour of a diving pelagic seabird“, *Proc. R. Soc. B Biol. Sci.*, sv. 289, izd. 1978, str. 20220862, srp. 2022, doi: 10.1098/rspb.2022.0862.
- [4] Y. Liux, H. Xu, B. Zhang, K. Sun, J. Yang, i B. Li, „Model-Based Underwater Image Simulation and Learning-Based Underwater Image Enhancement Method“, *Information*, sv. 13, izd. 4, str. 187, tra. 2022, doi: 10.3390/info13040187.
- [5] S. Merlino, M. Paterni, M. Locritani, U. Andriolo, G. Gonçalves, i L. Massetti, „Citizen Science for Marine Litter Detection and Classification on Unmanned Aerial Vehicle Images“, *Water*, sv. 13, izd. 23, str. 3349, stu. 2021, doi: 10.3390/w13233349.
- [6] S. Hijazi, R. Kumar, i C. Rowen, „Using Convolutional Neural Networks for Image Recognition“, *Cadence Des. Syst. Inc San Jose*, str. 12, 2015.
- [7] A. Krizhevsky, I. Sutskever, i G. E. Hinton, „ImageNet Classification with Deep Convolutional Neural Networks“, u *Advances in Neural Information Processing Systems*, Curran Associates, Inc., 2012. Pristupljeno: 15. lipanj 2022. [Na internetu]. Dostupno na: <https://proceedings.neurips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>
- [8] K. Simonyan i A. Zisserman, „Very Deep Convolutional Networks for Large-Scale Image Recognition“, 10. travanj 2015., *arXiv*: arXiv:1409.1556. Pristupljeno: 15. lipanj 2022. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1409.1556>
- [9] K. He, X. Zhang, S. Ren, i J. Sun, „Deep Residual Learning for Image Recognition“, 10. prosinac 2015., *arXiv*: arXiv:1512.03385. Pristupljeno: 15. lipanj 2022. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1512.03385>

- [10] C. Szegedy, S. Ioffe, V. Vanhoucke, i A. Alemi, „Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning“, 23. kolovoz 2016., *arXiv*: arXiv:1602.07261. Pristupljeno: 15. lipanj 2022. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1602.07261>
- [11] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, i T. Unterthiner, „An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale“, 03. lipanj 2021., *arXiv*: arXiv:2010.11929. doi: 10.48550/arXiv.2010.11929.
- [12] R. Girshick, J. Donahue, T. Darrell, i J. Malik, „Rich feature hierarchies for accurate object detection and semantic segmentation“, 22. listopad 2014., *arXiv*: arXiv:1311.2524. Pristupljeno: 14. lipanj 2022. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1311.2524>
- [13] R. Girshick, „Fast R-CNN“, 27. rujan 2015., *arXiv*: arXiv:1504.08083. Pristupljeno: 14. lipanj 2022. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1504.08083>
- [14] S. Ren, K. He, R. Girshick, i J. Sun, „Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks“, 06. siječanj 2016., *arXiv*: arXiv:1506.01497. Pristupljeno: 14. lipanj 2022. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1506.01497>
- [15] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, i C.-Y. Fu, „SSD: Single Shot MultiBox Detector“, u *Computer Vision – ECCV 2016*, sv. 9905, Springer International Publishing, 2016, str. 21–37. doi: 10.1007/978-3-319-46448-0_2.
- [16] J. Redmon, S. Divvala, R. Girshick, i A. Farhadi, „You Only Look Once: Unified, Real-Time Object Detection“, 09. svibanj 2016., *arXiv*: arXiv:1506.02640. Pristupljeno: 14. lipanj 2022. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1506.02640>
- [17] M. E. AtiK, Z. Duran, i R. Özgünlük, „Comparison of YOLO Versions for Object Detection from Aerial Images“, *Int. J. Environ. Geoinformatics*, sv. 9, izd. 2, str. 87–93, lip. 2022, doi: 10.30897/ijegeo.1010741.
- [18] J. Redmon i A. Farhadi, „YOLO9000: Better, Faster, Stronger“, u *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI: IEEE, srp. 2017, str. 6517–6525. doi: 10.1109/CVPR.2017.690.

- [19] J. Hui, „Real-time Object Detection with YOLO, YOLOv2 and now YOLOv3“, Medium. Pristupljeno: 08. srpanj 2022. [Na internetu]. Dostupno na: <https://jonathan-hui.medium.com/real-time-object-detection-with-yolo-yolov2-28b1b93e2088>
- [20] „Yolov3 and Yolov4 in Object Detection“. Pristupljeno: 08. srpanj 2022. [Na internetu]. Dostupno na: <https://ai-pool.com/a/s/yolov3-and-yolov4-in-object-detection>
- [21] „Introduction to the YOLO Family“, PyImageSearch. Pristupljeno: 07. srpanj 2022. [Na internetu]. Dostupno na: <https://pyimagesearch.com/2022/04/04/introduction-to-the-yolo-family/>
- [22] C. Supeshala, „YOLO v4 or YOLO v5 or PP-YOLO?“, Medium. Pristupljeno: 12. srpanj 2022. [Na internetu]. Dostupno na: <https://towardsdatascience.com/yolo-v4-or-yolo-v5-or-pp-yolo-dad8e40f7109>
- [23] *YOLOv5*. (12. srpanj 2022.). Python. Ultralytics. Pristupljeno: 12. srpanj 2022. [Na internetu]. Dostupno na: <https://github.com/ultralytics/yolov5>
- [24] „Train Custom Data · ultralytics/yolov5 Wiki“, GitHub. Pristupljeno: 12. srpanj 2022. [Na internetu]. Dostupno na: <https://github.com/ultralytics/yolov5>
- [25] „Google Colaboratory“. Pristupljeno: 12. srpanj 2022. [Na internetu]. Dostupno na: <https://colab.research.google.com/drive/1gDZ2xcTOgR39tGGs-EZ6i3RTs16wmzZQ>
- [26] I. Idrisov, „YOLOv4 vs YOLOv5“, Deelvin Machine Learning. Pristupljeno: 12. srpanj 2022. [Na internetu]. Dostupno na: <https://medium.com/deelvin-machine-learning/yolov4-vs-yolov5-db1e0ac7962b>
- [27] Ultralytics, „YOLOv6“. Pristupljeno: 23. studeni 2023. [Na internetu]. Dostupno na: <https://docs.ultralytics.com/models/yolov6>
- [28] Encord, „YOLOv8 for Object Detection Explained [Practical Example]“, Encord. Pristupljeno: 23. studeni 2023. [Na internetu]. Dostupno na: <https://medium.com/cord-tech/yolov8-for-object-detection-explained-practical-example-23920f77f66a>
- [29] Ultralytics, „YOLOv7“. Pristupljeno: 23. studeni 2023. [Na internetu]. Dostupno na: <https://docs.ultralytics.com/models/yolov7>
- [30] Ultralytics, „Detect“. Pristupljeno: 23. studeni 2023. [Na internetu]. Dostupno na: <https://docs.ultralytics.com/tasks/detect>

- [31] Ultralytics, „YOLOv9“. Pristupljeno: 23. siječanj 2025. [Na internetu]. Dostupno na: <https://docs.ultralytics.com/models/yolov9>
- [32] Ultralytics, „YOLOv10“. Pristupljeno: 23. siječanj 2025. [Na internetu]. Dostupno na: <https://docs.ultralytics.com/models/yolov10>
- [33] Ultralytics, „YOLO11“. Pristupljeno: 23. siječanj 2025. [Na internetu]. Dostupno na: <https://docs.ultralytics.com/models/yolo11>
- [34] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, i S. Zagoruyko, „End-to-End Object Detection with Transformers“, 28. svibanj 2020., *arXiv*: arXiv:2005.12872. doi: 10.48550/arXiv.2005.12872.
- [35] Z. Liu i ostali, „Swin Transformer: Hierarchical Vision Transformer using Shifted Windows“, 17. kolovoz 2021., *arXiv*: arXiv:2103.14030. doi: 10.48550/arXiv.2103.14030.
- [36] H. Zhang i ostali, „DINO: DETR with Improved DeNoising Anchor Boxes for End-to-End Object Detection“, 11. srpanj 2022., *arXiv*: arXiv:2203.03605. doi: 10.48550/arXiv.2203.03605.
- [37] W. Lian, J. Lindblad, P. Micke, i N. Sladoje, „Isolated Channel Vision Transformers: From Single-Channel Pretraining to Multi-Channel Finetuning“, 12. ožujak 2025., *arXiv*: arXiv:2503.09826. doi: 10.48550/arXiv.2503.09826.
- [38] „What is remote sensing and what is it used for? | U.S. Geological Survey“. Pristupljeno: 11. ožujak 2022. [Na internetu]. Dostupno na: <https://www.usgs.gov/faqs/what-remote-sensing-and-what-it-used>
- [39] L. Yabin, Y. Jun, i H. Zhiyi, „Improved Faster R-CNN Algorithm for Sea Object Detection Under Complex Sea Conditions“, *Int. J. Adv. Netw. Monit. Controls*, sv. 5, izd. 2, str. 76–82, 2020, doi: 10.21307/ijanmc-2020-020.
- [40] I. Lorencin, N. Anđelić, V. Mrzljak, i Z. Car, „Marine Objects Recognition Using Convolutional Neural Networks“, *Naše More*, sv. 66, izd. 3, str. 112–120, lis. 2019, doi: 10.17818/NM/2019/3.3.
- [41] O. Garcia-Garin i ostali, „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.*, sv. 273, str. 116490, ožu. 2021, doi: 10.1016/j.envpol.2021.116490.
- [42] *MARLITE*. (19. veljača 2022.). R. Pristupljeno: 22. ožujak 2022. [Na internetu]. Dostupno na: <https://github.com/amonleong/MARLIT>
- [43] G. Jakovljevic, M. Govedarica, i F. Alvarez-Taboada, „A Deep Learning Model for Automatic Plastic Mapping Using Unmanned Aerial Vehicle (UAV) Data“, *Remote Sens.*, sv. 12, izd. 9, str. 1515, svi. 2020, doi: 10.3390/rs12091515.
- [44] V. Badrinarayanan, A. Kendall, i R. Cipolla, „SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation“, 10. listopada 2016., *arXiv:arXiv:1511.00561*. doi: 10.48550/arXiv.1511.00561.
- [45] D. G. Lowe, „Distinctive Image Features from Scale-Invariant Keypoints“, *Int. J. Comput. Vis.*, sv. 60, izd. 2, str. 91–110, stu. 2004, doi: 10.1023/B:VISI.0000029664.99615.94.
- [46] M. A. Fischler i R. C. Bolles, „Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography“, *Commun ACM*, sv. 24, izd. 6, str. 381–395, lip. 1981, doi: 10.1145/358669.358692.
- [47] O. Elharrouss, Y. Akbari, N. Almaadeed, i S. Al-Maadeed, „Backbones-Review: Feature Extraction Networks for Deep Learning and Deep Reinforcement Learning Approaches“, *Comput. Sci. Rev.*, sv. 53, str. 100645, kol. 2024, doi: 10.1016/j.cosrev.2024.100645.
- [48] K. Topouzelis, A. Papakonstantinou, i S. P. Garaba, „Detection of floating plastics from satellite and unmanned aerial systems (Plastic Litter Project 2018)“, *Int. J. Appl. Earth Obs. Geoinformation*, sv. 79, str. 175–183, srp. 2019, doi: 10.1016/j.jag.2019.03.011.
- [49] J. Seo, W. Park, i T. Kim, „Feature-Based Approach to Change Detection of Small Objects from High-Resolution Satellite Images“, *Remote Sens.*, sv. 14, izd. 3, str. 462, sij. 2022, doi: 10.3390/rs14030462.
- [50] C. Kislik, L. Genzoli, A. Lyons, i M. Kelly, „Application of UAV Imagery to Detect and Quantify Submerged Filamentous Algae and Rooted Macrophytes in a Non-Wadeable River“, *Remote Sens.*, sv. 12, izd. 20, str. 3332, lis. 2020, doi: 10.3390/rs12203332.

- [51] D. V. Politikos, E. Fakiris, A. Davvetas, I. A. Klampanos, i G. Papatheodorou, „Automatic detection of seafloor marine litter using towed camera images and deep learning“, *Mar. Pollut. Bull.*, sv. 164, str. 111974, ožu. 2021, doi: 10.1016/j.marpolbul.2021.111974.
- [52] *HumanSignal/labelImg*. (17. svibanj 2025.). Python. HumanSignal. Pristupljeno: 17. svibanj 2025. [Na internetu]. Dostupno na: <https://github.com/HumanSignal/labelImg>
- [53] P. Barham i ostali, „TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems“, 16. ožujak 2016., *arXiv*: arXiv:1603.04467. Pristupljeno: 30. lipanj 2022. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1603.04467>
- [54] W.-H. Lin, J.-X. Zhong, S. Liu, T. Li, i G. Li, „ROIMIX: Proposal-Fusion Among Multiple Images for Underwater Object Detection“, u *ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing*, Barcelona, Spain: IEEE, svi. 2020, str. 2588–2592. doi: 10.1109/ICASSP40776.2020.9053829.
- [55] J. Li, K. A. Skinner, R. M. Eustice, i M. Johnson-Roberson, „WaterGAN: Unsupervised Generative Network to Enable Real-time Color Correction of Monocular Underwater Images“, *IEEE Robot. Autom. Lett.*, str. 1–1, 2017, doi: 10.1109/LRA.2017.2730363.
- [56] C. Fabbri, M. J. Islam, i J. Sattar, „Enhancing Underwater Imagery Using Generative Adversarial Networks“, u *2018 IEEE International Conference on Robotics and Automation (ICRA)*, Brisbane, QLD: IEEE, svi. 2018, str. 7159–7165. doi: 10.1109/ICRA.2018.8460552.
- [57] A. Naseer, E. N. Baro, S. D. Khan, i Y. Vila, „A Novel Detection Refinement Technique for Accurate Identification of Nephrops norvegicus Burrows in Underwater Imagery“, *Sensors*, sv. 22, izd. 12, Art. izd. 12, sij. 2022, doi: 10.3390/s22124441.
- [58] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, i Z. Wojna, „Rethinking the Inception Architecture for Computer Vision“, u *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA: IEEE, lip. 2016, str. 2818–2826. doi: 10.1109/CVPR.2016.308.
- [59] S. J. Pan i Q. Yang, „A Survey on Transfer Learning“, *IEEE Trans. Knowl. Data Eng.*, sv. 22, izd. 10, str. 1345–1359, lis. 2010, doi: 10.1109/TKDE.2009.191.

- [60] I. Marin, S. Mladenović, S. Gotovac, i G. Zaharija, „Deep-Feature-Based Approach to Marine Debris Classification“, *Appl. Sci.*, sv. 11, izd. 12, str. 5644, lip. 2021, doi: 10.3390/app11125644.
- [61] J. Hong, M. Fulton, i J. Sattar, „TrashCan: A Semantically-Segmented Dataset towards Visual Detection of Marine Debris“, *ArXiv200708097 Cs*, srp. 2020, Pristupljeno: 29. listopad 2021. [Na internetu]. Dostupno na: <http://arxiv.org/abs/2007.08097>
- [62] C. Martin, Q. Zhang, D. Zhai, X. Zhang, i C. M. Duarte, „Anthropogenic litter density and composition data acquired flying commercial drones on sandy beaches along the Saudi Arabian Red Sea“, *Data Brief*, sv. 36, str. 107056, lip. 2021, doi: 10.1016/j.dib.2021.107056.
- [63] L. Fallati, A. Polidori, C. Salvatore, L. Saponari, A. Savini, i 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.*, sv. 693, str. 133581, stu. 2019, doi: 10.1016/j.scitotenv.2019.133581.
- [64] Y. Taddia, C. Corbau, J. Buoninsegni, U. Simeoni, i A. Pellegrinelli, „Uav Approach for Detecting Plastic Marine Debris on the Beach: A Case Study in the Po River Delta (Italy)“, *Drones*, sv. 5, izd. 4, str. 140, stu. 2021, doi: 10.3390/drones5040140.
- [65] „Ships in Satellite Imagery“. Pristupljeno: 31. ožujak 2022. [Na internetu]. Dostupno na: <https://www.kaggle.com/rhammell/ships-in-satellite-imagery>
- [66] J. Hong, M. S. Fulton, i J. Sattar, „TrashCan 1.0 An Instance-Segmentation Labeled Dataset of Trash Observations“. 23. srpanj 2020. doi: 10.13020/g1gx-y834.
- [67] T. Sasaki i ostali, „JAMSTEC E-library of Deep-sea Images (J-EDI) Realizes a Virtual Journey to the Earth’s Unexplored Deep Ocean“, sv. 2016, str. IN53C-1911, pros. 2016.
- [68] „ImageNet“. Pristupljeno: 31. ožujak 2022. [Na internetu]. Dostupno na: <https://www.image-net.org/index.php>
- [69] C. Martin, Q. Zhang, D. Zhai, X. Zhang, i C. M. Duarte, „Drone images of sandy beaches and anthropogenic litter along the Saudi Arabian Red Sea“, sv. 1, velj. 2021, doi: 10.17632/gpdsntb3y6.1.

- [70] M. J. Islam, Y. Xia, i J. Sattar, „Fast Underwater Image Enhancement for Improved Visual Perception“, 08. veljača 2020., *arXiv*: arXiv:1903.09766. Pristupljeno: 08. veljača 2024. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1903.09766>
- [71] R. Liu, X. Fan, M. Zhu, M. Hou, i Z. Luo, „Real-world Underwater Enhancement: Challenges, Benchmarks, and Solutions“, 06. ožujak 2019., *arXiv*: arXiv:1901.05320. Pristupljeno: 08. veljača 2024. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1901.05320>
- [72] C. Li i ostali, „An Underwater Image Enhancement Benchmark Dataset and Beyond“, 26. studeni 2019., *arXiv*: arXiv:1901.05495. Pristupljeno: 08. veljača 2024. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1901.05495>
- [73] H. Li, J. Li, i W. Wang, „A Fusion Adversarial Underwater Image Enhancement Network with a Public Test Dataset“, 29. lipanj 2019., *arXiv*: arXiv:1906.06819. Pristupljeno: 08. veljača 2024. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1906.06819>
- [74] Y. M. Alsakar, N. A. Sakr, S. El-Sappagh, T. Abuhmed, i M. Elmogy, „Underwater Image Restoration and Enhancement: A Comprehensive Review of Recent Trends, Challenges, and Applications“, *Vis. Comput.*, sv. 41, izd. 6, str. 3735–3783, tra. 2025.
- [75] F. Oleari, F. Kallasi, D. L. Rizzini, J. Aleotti, i S. Caselli, „An underwater stereo vision system: From design to deployment and dataset acquisition“, u *OCEANS 2015 - Genova*, Genova, Italy: IEEE, svi. 2015, str. 1–6. doi: 10.1109/OCEANS-Genova.2015.7271529.
- [76] M. Yang, J. Hu, C. Li, G. Rohde, Y. Du, i K. Hu, „An In-Depth Survey of Underwater Image Enhancement and Restoration“, *IEEE Access*, sv. 7, str. 123638–123657, 2019, doi: 10.1109/ACCESS.2019.2932611.
- [77] B. J. Boom i ostali, „Long-term underwater camera surveillance for monitoring and analysis of fish populations“.
- [78] R. J. M. Veiga i ostali, „Autonomous Temporal Pseudo-Labeling for Fish Detection“, *Appl. Sci.*, sv. 12, izd. 12, str. 5910, lip. 2022, doi: 10.3390/app12125910.
- [79] L. Chen i ostali, „A Benchmark dataset for both underwater image enhancement and underwater object detection“, 28. lipanj 2020., *arXiv*: arXiv:2006.15789. Pristupljeno: 08. veljača 2024. [Na internetu]. Dostupno na: <http://arxiv.org/abs/2006.15789>

- [80] V. Chari i P. Sturm, „Multi-View Geometry of the Refractive Plane“, u *Proceedings of the British Machine Vision Conference 2009*, London: British Machine Vision Association, 2009, str. 56.1-56.11. doi: 10.5244/C.23.56.
- [81] M.-Y. J. Kuo, S. Murai, R. Kawahara, S. Nobuhara, i K. Nishino, „Surface Normals and Shape From Water“, *IEEE Trans. Pattern Anal. Mach. Intell.*, sv. 44, izd. 12, str. 9150–9162, pros. 2022, doi: 10.1109/TPAMI.2021.3121963.
- [82] F. Chadebecq i ostali, „Refractive Two-View Reconstruction for Underwater 3D Vision“, *Int. J. Comput. Vis.*, sv. 128, izd. 5, str. 1101–1117, svi. 2020, doi: 10.1007/s11263-019-01218-9.
- [83] S. S. Sankpal i S. S. Deshpande, „Nonuniform Illumination Correction Algorithm for Underwater Images Using Maximum Likelihood Estimation Method“, *J. Eng.*, sv. 2016, str. 1–9, 2016, doi: 10.1155/2016/5718297.
- [84] X. Cao, S. Rong, Y. Liu, T. Li, Q. Wang, i B. He, „NUICNet: Non-Uniform Illumination Correction for Underwater Image Using Fully Convolutional Network“, *IEEE Access*, sv. 8, str. 109989–110002, 2020, doi: 10.1109/ACCESS.2020.3002593.
- [85] S. Bazeille, I. Quidu, i L. Jaulin, „Automatic Underwater Image Pre-Processing“, *Proc. CMM06*, lis. 2006.
- [86] Z. Hu, „Fast Non-uniform Deblurring using Constrained Camera Pose Subspace“, *BMVC*, sij. 2012, doi: 10.5244/C.26.136.
- [87] L. Xu i J. Jia, „Two-Phase Kernel Estimation for Robust Motion Deblurring“, u *Computer Vision – ECCV 2010*, K. Daniilidis, P. Maragos, i N. Paragios, Ur., u Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, 2010, str. 157–170. doi: 10.1007/978-3-642-15549-9_12.
- [88] L. Xu, S. Zheng, i J. Jia, „Unnatural L0 Sparse Representation for Natural Image Deblurring“, u *2013 IEEE Conference on Computer Vision and Pattern Recognition*, Portland, OR, USA: IEEE, lip. 2013, str. 1107–1114. doi: 10.1109/CVPR.2013.147.
- [89] M. Vimal Raj i S. Sakthivel Murugan, „Motion Deblurring Analysis for Underwater Image Restoration“, *J. Phys. Conf. Ser.*, sv. 1911, izd. 1, str. 012028, svi. 2021, doi: 10.1088/1742-6596/1911/1/012028.

- [90] H. Lu, Y. Li, Y. Zhang, M. Chen, S. Serikawa, i H. Kim, „Underwater Optical Image Processing: A Comprehensive Review“, 2017, [Na internetu]. Dostupno na: <https://arxiv.org/abs/1702.03600>
- [91] J. Raihan, P. E. Abas, i L. C. De Silva, „Role of Restored Underwater Images in Underwater Imaging Applications“, *Appl. Syst. Innov.*, sv. 4, izd. 96, 2021, doi: 10.3390/asi4040096.
- [92] C. Ancuti, C. Ancuti, T. Haber, i P. Bekaert, „Fusion-based Restoration of the Underwater Images“, *IEEE Int. Conf. Image Process.*, 2011.
- [93] N. Wang, H. Zheng, i B. Zheng, „Underwater Image Restoration via Maximum Attenuation Identification“, sv. 5, 2017.
- [94] H. Lu i ostali, „Underwater image enhancement method using weighted guided trigonometric filtering and artificial light correction“.
- [95] R. Schettini i S. Corchs, „Underwater Image Processing: State of the Art of Restoration and Image Enhancement Methods“, *EURASIP J. Adv. Signal Process.*, sv. 2010, izd. 1, str. 746052, pros. 2010, doi: 10.1155/2010/746052.
- [96] S. Jian i W. Wen, „Study on Underwater Image Denoising Algorithm Based on Wavelet Transform“, *J. Phys. Conf. Ser.*, sv. 806, str. 012006, velj. 2017, doi: 10.1088/1742-6596/806/1/012006.
- [97] Z. Liu, Y. Yu, K. Zhang, i H. Huang, „Underwater image transmission and blurred image restoration“, *Opt. Eng.*, sv. 40, str. 1125–1131, lip. 2001, doi: 10.1117/1.1364500.
- [98] Y. Xu, H. Wang, G. D. Cooper, S. Rong, i W. Sun, „Learning-Based Dark and Blurred Underwater Image Restoration“, *Complexity*, sv. 2020, str. 1–14, kol. 2020, doi: 10.1155/2020/6549410.
- [99] Y. Cho i A. Kim, „Visibility enhancement for underwater visual SLAM based on underwater light scattering model“, u *2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, Singapore: IEEE Press, svi. 2017, str. 710–717. doi: 10.1109/ICRA.2017.7989087.
- [100] S. Emberton, L. Chittka, i A. Cavallaro, „Underwater image and video dehazing with pure haze region segmentation“, *Comput. Vis. Image Underst.*, sv. 168, str. 145–156, ožu. 2018, doi: 10.1016/j.cviu.2017.08.003.

- [101] R. Singh i M. Biswas, „Hazy Underwater Image Enhancement Based on Contrast and Color Improvement Using Fusion Technique“, *Image Process. Commun.*, sv. 22, izd. 3, str. 31–38, ruj. 2017, doi: 10.1515/ipc-2017-0016.
- [102] Y.-T. Peng i P. C. Cosman, „Underwater Image Restoration Based on Image Blurriness and Light Absorption“, *IEEE Trans. Image Process.*, sv. 26, izd. 4, str. 1579–1594, tra. 2017, doi: 10.1109/TIP.2017.2663846.
- [103] H.-H. Chang, „Single Underwater Image Restoration Based on Adaptive Transmission Fusion“, *IEEE Access*, sv. 8, str. 38650–38662, 2020, doi: 10.1109/ACCESS.2020.2971019.
- [104] Y. Shen, C. Zhao, Y. Liu, S. Wang, i F. Huang, „Underwater Optical Imaging: Key Technologies and Applications Review“, *IEEE Access*, sv. 9, str. 85500–85514, 2021, doi: 10.1109/ACCESS.2021.3086820.
- [105] E. Trucco i A. T. Olmos-Antillon, „Self-Tuning Underwater Image Restoration“, *IEEE J. Ocean. Eng.*, sv. 31, izd. 2, str. 511–519, tra. 2006, doi: 10.1109/JOE.2004.836395.
- [106] B. L. McGlamery, „A Computer Model For Underwater Camera Systems“, u *Ocean Optics VI*, S. Q. Duntley, Ur., SPIE, 1980, str. 221–231. doi: 10.1117/12.958279.
- [107] J. S. Jaffe, „Computer modeling and the design of optimal underwater imaging systems“, *IEEE J. Ocean. Eng.*, sv. 15, izd. 2, str. 101–111, tra. 1990, doi: 10.1109/48.50695.
- [108] W. Hou, D. J. Gray, A. D. Weidemann, G. R. Fournier, i J. L. Forand, „Automated underwater image restoration and retrieval of related optical properties“, u *2007 IEEE International Geoscience and Remote Sensing Symposium*, Barcelona, Spain: IEEE, 2007, str. 1889–1892. doi: 10.1109/IGARSS.2007.4423193.
- [109] M. Boffety, F. Galland, i A.-G. Allais, „Color image simulation for underwater optics“, *Appl. Opt.*, sv. 51, izd. 23, str. 5633, kol. 2012, doi: 10.1364/AO.51.005633.
- [110] Haocheng Wen, Yonghong Tian, Tiejun Huang, i Wen Gao, „Single underwater image enhancement with a new optical model“, u *2013 IEEE International Symposium on Circuits and Systems (ISCAS2013)*, Beijing: IEEE, svi. 2013, str. 753–756. doi: 10.1109/ISCAS.2013.6571956.
- [111] J. Ahn, S. Yasukawa, T. Sonoda, Y. Nishida, K. Ishii, i T. Ura, „An Optical Image Transmission System for Deep Sea Creature Sampling Missions Using Autonomous

- Underwater Vehicle“, *IEEE J. Ocean. Eng.*, sv. 45, izd. 2, str. 350–361, tra. 2020, doi: 10.1109/JOE.2018.2872500.
- [112] Kaiming He, Jian Sun, i Xiaoou Tang, „Single Image Haze Removal Using Dark Channel Prior“, *IEEE Trans. Pattern Anal. Mach. Intell.*, sv. 33, izd. 12, str. 2341–2353, pros. 2011, doi: 10.1109/TPAMI.2010.168.
- [113] S. Fayaz, S. Parah, i G. Qureshi, „Efficient underwater image restoration utilizing modified dark channel prior“, *Multimed. Tools Appl.*, sv. 82, lis. 2022, doi: 10.1007/s11042-022-13828-6.
- [114] P. Drews Jr, E. Do Nascimento, F. Moraes, S. Botelho, i M. Campos, „Transmission Estimation in Underwater Single Images“, u *2013 IEEE International Conference on Computer Vision Workshops*, Sydney, Australia: IEEE, pros. 2013, str. 825–830. doi: 10.1109/ICCVW.2013.113.
- [115] P. L. J. Drews, E. R. Nascimento, S. S. C. Botelho, i M. F. Montenegro Campos, „Underwater Depth Estimation and Image Restoration Based on Single Images“, *IEEE Comput. Graph. Appl.*, sv. 36, izd. 2, str. 24–35, ožu. 2016, doi: 10.1109/MCG.2016.26.
- [116] N. Carlevaris-Bianco, A. Mohan, i R. M. Eustice, „Initial results in underwater single image dehazing“, u *OCEANS 2010 MTS/IEEE SEATTLE*, Seattle, WA: IEEE, ruj. 2010, str. 1–8. doi: 10.1109/OCEANS.2010.5664428.
- [117] A. Galdran, D. Pardo, A. Picón, i A. Alvarez-Gila, „Automatic Red-Channel underwater image restoration“, *J. Vis. Commun. Image Represent.*, sv. 26, str. 132–145, sij. 2015, doi: 10.1016/j.jvcir.2014.11.006.
- [118] W. Song, Y. Wang, D. Huang, i D. Tjondronegoro, „A Rapid Scene Depth Estimation Model Based on Underwater Light Attenuation Prior for Underwater Image Restoration“, u *Advances in Multimedia Information Processing – PCM 2018*, sv. 11164, u *Lecture Notes in Computer Science*, vol. 11164. , Cham: Springer International Publishing, 2018, str. 678–688. doi: 10.1007/978-3-030-00776-8_62.
- [119] Liu Chao i Meng Wang, „Removal of water scattering“, u *2010 2nd International Conference on Computer Engineering and Technology*, Chengdu, China: IEEE, 2010, str. V2-35–V2-39. doi: 10.1109/ICCET.2010.5485339.

- [120] H.-Y. Yang, P.-Y. Chen, C.-C. Huang, Y.-Z. Zhuang, i Y.-H. Shiau, „Low Complexity Underwater Image Enhancement Based on Dark Channel Prior“, u *2011 Second International Conference on Innovations in Bio-inspired Computing and Applications*, Shenzhen, Guangdong, China: IEEE, pros. 2011, str. 17–20. doi: 10.1109/IBICA.2011.9.
- [121] J. Y. Chiang i Ying-Ching Chen, „Underwater Image Enhancement by Wavelength Compensation and Dehazing“, *IEEE Trans. Image Process.*, sv. 21, izd. 4, str. 1756–1769, tra. 2012, doi: 10.1109/TIP.2011.2179666.
- [122] S. Serikawa i H. Lu, „Underwater image dehazing using joint trilateral filter“, *Comput. Electr. Eng.*, sv. 40, izd. 1, str. 41–50, sij. 2014, doi: 10.1016/j.compeleceng.2013.10.016.
- [123] H. Lu, Y. Li, L. Zhang, i S. Serikawa, „Contrast enhancement for images in turbid water“, *J. Opt. Soc. Am. A*, sv. 32, izd. 5, str. 886, svi. 2015, doi: 10.1364/JOSAA.32.000886.
- [124] D. Hu, J. Tan, L. Zhang, X. Ge, i J. Liu, „Image deblurring via enhanced local maximum intensity prior“, *Signal Process. Image Commun.*, sv. 96, str. 116311, kol. 2021, doi: 10.1016/j.image.2021.116311.
- [125] X. Zhao, T. Jin, i S. Qu, „Deriving inherent optical properties from background color and underwater image enhancement“, *Ocean Eng.*, sv. 94, str. 163–172, sij. 2015, doi: 10.1016/j.oceaneng.2014.11.036.
- [126] C. Li, J. Quo, Y. Pang, S. Chen, i J. Wang, „Single underwater image restoration by blue-green channels dehazing and red channel correction“, u *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Shanghai: IEEE, ožu. 2016, str. 1731–1735. doi: 10.1109/ICASSP.2016.7471973.
- [127] B. Wang i ostali, „Underwater image enhancement by maximum-likelihood based adaptive color correction and robust scattering removal“, *Front. Comput. Sci.*, sv. 17, izd. 2, str. 172702, srp. 2022, doi: 10.1007/s11704-022-1205-7.
- [128] C. Li, J. Guo, S. Chen, Y. Tang, Y. Pang, i J. Wang, „Underwater image restoration based on minimum information loss principle and optical properties of underwater imaging“, u *2016 IEEE International Conference on Image Processing (ICIP)*, Phoenix, AZ, USA: IEEE, ruj. 2016, str. 1993–1997. doi: 10.1109/ICIP.2016.7532707.

- [129] Y.-T. Peng, K. Cao, i P. C. Cosman, „Generalization of the Dark Channel Prior for Single Image Restoration“, *IEEE Trans. Image Process.*, sv. 27, izd. 6, str. 2856–2868, lip. 2018, doi: 10.1109/TIP.2018.2813092.
- [130] C. Li, J. Guo, R. Cong, Y. Pang, i B. Wang, „Underwater Image Enhancement by Dehazing With Minimum Information Loss and Histogram Distribution Prior“, *IEEE Trans. Image Process.*, sv. PP, str. 1–1, ruj. 2016, doi: 10.1109/TIP.2016.2612882.
- [131] X. Ding, Y. Wang, J. Zhang, i X. Fu, „Underwater image dehaze using scene depth estimation with adaptive color correction“, lip. 2017, str. 1–5. doi: 10.1109/OCEANSE.2017.8084665.
- [132] K. Cao, Y.-T. Peng, i P. C. Cosman, „Underwater Image Restoration using Deep Networks to Estimate Background Light and Scene Depth“, u *2018 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI)*, Las Vegas, NV: IEEE, tra. 2018, str. 1–4. doi: 10.1109/SSIAI.2018.8470347.
- [133] W. V. Barbosa, H. G. B. Amaral, T. L. Rocha, i E. R. Nascimento, „Visual-Quality-Driven Learning for Underwater Vision Enhancement“, u *2018 25th IEEE International Conference on Image Processing (ICIP)*, lis. 2018, str. 3933–3937. doi: 10.1109/ICIP.2018.8451356.
- [134] M. Hou, R. Liu, X. Fan, i Z. Luo, „Joint Residual Learning for Underwater Image Enhancement“, u *2018 25th IEEE International Conference on Image Processing (ICIP)*, lis. 2018, str. 4043–4047. doi: 10.1109/ICIP.2018.8451209.
- [135] Z. Wang, L. Shen, M. Yu, K. Wang, Y. Lin, i M. Xu, „Domain Adaptation for Underwater Image Enhancement“, 22. kolovoz 2021., *arXiv*: arXiv:2108.09650. Pristupljeno: 13. veljača 2024. [Na internetu]. Dostupno na: <http://arxiv.org/abs/2108.09650>
- [136] S. Xu, M. Zhang, W. Song, H. Mei, Q. He, i A. Liotta, „A Systematic Review and Analysis of Deep Learning-based Underwater Object Detection“, *Neurocomputing*, sv. 527, ožu. 2023, doi: 10.1016/j.neucom.2023.01.056.
- [137] Y. Xu, J. Wen, L. Fei, i Z. Zhang, „Review of Video and Image Defogging Algorithms and Related Studies on Image Restoration and Enhancement“, *IEEE Access*, sv. 4, str. 165–188, 2016, doi: 10.1109/ACCESS.2015.2511558.

- [138] M. Abdullah-Al-Wadud, M. Kabir, M. Dewan, i O. Chae, „A Dynamic Histogram Equalization for Image Contrast Enhancement“, *Consum. Electron. IEEE Trans. On*, sv. 53, str. 593–600, lip. 2007, doi: 10.1109/TCE.2007.381734.
- [139] T. Kim i J. Paik, „Adaptive contrast enhancement using gain-controllable clipped histogram equalization“, *IEEE Trans. Consum. Electron.*, sv. 54, izd. 4, str. 1803–1810, stu. 2008, doi: 10.1109/TCE.2008.4711238.
- [140] C. Ancuti, C. O. Ancuti, T. Haber, i P. Bekaert, „Enhancing underwater images and videos by fusion“, u *2012 IEEE Conference on Computer Vision and Pattern Recognition*, Providence, RI: IEEE, lip. 2012, str. 81–88. doi: 10.1109/CVPR.2012.6247661.
- [141] C. O. Ancuti, C. Ancuti, C. De Vleeschouwer, i P. Bekaert, „Color Balance and Fusion for Underwater Image Enhancement“, *IEEE Trans. Image Process.*, sv. 27, izd. 1, str. 379–393, sij. 2018, doi: 10.1109/TIP.2017.2759252.
- [142] X. Liu, G. Zhong, C. Liu, i J. Dong, „Underwater image colour constancy based on DSNMF“, *IET Image Process.*, sv. 11, izd. 1, str. 38–43, 2017, doi: 10.1049/iet-ipr.2016.0543.
- [143] L. A. Torres-Méndez i G. Dudek, „Color Correction of Underwater Images for Aquatic Robot Inspection“, u *Energy Minimization Methods in Computer Vision and Pattern Recognition*, A. Rangarajan, B. Vemuri, i A. L. Yuille, Ur., u *Lecture Notes in Computer Science*. Berlin, Heidelberg: Springer, 2005, str. 60–73. doi: 10.1007/11585978_5.
- [144] K. Iqbal, R. A. Salam, A. Osman, i A. Z. Talib, „Underwater Image Enhancement Using an Integrated Colour Model“, *IAENG Int. J. Comput. Sci.*, sv. 34, sij. 2007.
- [145] A. shahrizan Abdul ghani i N. A. Mat Isa, „Automatic system for improving underwater image contrast and color through recursive adaptive histogram modification“, *Comput. Electron. Agric.*, sv. 141, str. 181–195, ruj. 2017, doi: 10.1016/j.compag.2017.07.021.
- [146] M. Hitam, W. Yussof, E. Awalludin, i Z. Bachok, „Mixture contrast limited adaptive histogram equalization for underwater image enhancement“, predstavljeno na *IEEE International Conference on Computer Applications Technology (ICCAT)*, sij. 2013, str. 1–5. doi: 10.1109/ICCAT.2013.6522017.

- [147] D. Huang, Y. Wang, W. Song, J. Sequeira, i S. Mavromatis, „Shallow-Water Image Enhancement Using Relative Global Histogram Stretching Based on Adaptive Parameter Acquisition“, u *MultiMedia Modeling*, sv. 10704, u *Lecture Notes in Computer Science*, vol. 10704. , Cham: Springer International Publishing, 2018, str. 453–465. doi: 10.1007/978-3-319-73603-7_37.
- [148] S. S. Agaian, K. Panetta, i A. M. Grigoryan, „Transform-based image enhancement algorithms with performance measure“, *IEEE Trans. Image Process.*, sv. 10, izd. 3, str. 367–382, ožu. 2001, doi: 10.1109/83.908502.
- [149] M. H. Asmare, V. S. Asirvadam, i A. F. M. Hani, „Image enhancement based on contourlet transform“, *Signal Image Video Process.*, sv. 7, izd. 9, str. 1679–1690, 2015, doi: 10.1007/s11760-014-0626-7.
- [150] K. Panetta, A. Samani, i S. Agaian, „A Robust No-Reference, No-Parameter, Transform Domain Image Quality Metric for Evaluating the Quality of Color Images“, *IEEE Access*, sv. 6, str. 10979–10985, 2018, doi: 10.1109/ACCESS.2018.2804901.
- [151] Y. Wang, X. Ding, R. Wang, J. Zhang, i X. Fu, „Fusion-based underwater image enhancement by wavelet decomposition“, ožu. 2017, str. 1013–1018. doi: 10.1109/ICIT.2017.7915500.
- [152] A. Grigoryan, „Color Image Enhancement via Combine Homomorphic Ratio and Histogram Equalization Approaches: Using Underwater Images as Illustrative Examples“, 2018. Pristupljeno: 06. ožujak 2024. [Na internetu]. Dostupno na: <https://www.semanticscholar.org/paper/Color-Image-Enhancement-via-Combine-Homomorphic-and-Grigoryan/b3687b1128cbc006db1243201299808d3263b6e8>
- [153] G. Kaur i M. Kaur, „A Study of Transform Domain based Image Enhancement Techniques“, *Int. J. Comput. Appl.*, sv. 152, izd. 9, str. 25–29, lis. 2016, doi: 10.5120/ijca2016911858.
- [154] F. Petit, A.-S. Capelle-Laize, i P. Carre, „Underwater image enhancement by attenuation inversion with quaternions“, u *2009 IEEE International Conference on Acoustics, Speech and Signal Processing*, Taipei, Taiwan: IEEE, tra. 2009, str. 1177–1180. doi: 10.1109/ICASSP.2009.4959799.
- [155] C.-Y. Cheng, C.-C. Sung, i H.-H. Chang, „Underwater image restoration by red-dark channel prior and point spread function deconvolution“, u *2015 IEEE International*

- Conference on Signal and Image Processing Applications (ICSIPA)*, lis. 2015, str. 110–115. doi: 10.1109/ICSIPA.2015.7412173.
- [156] S. Feifei, Z. Xuemeng, i W. Guoyu, „An Approach for Underwater Image Denoising Via Wavelet Decomposition and High-pass Filter“, u *2011 Fourth International Conference on Intelligent Computation Technology and Automation*, ožu. 2011, str. 417–420. doi: 10.1109/ICICTA.2011.388.
- [157] A. S. Abdul Ghani, „Image contrast enhancement using an integration of recursive-overlapped contrast limited adaptive histogram specification and dual-image wavelet fusion for the high visibility of deep underwater image“, *Ocean Eng.*, sv. 162, str. 224–238, kol. 2018, doi: 10.1016/j.oceaneng.2018.05.027.
- [158] P. Ravisankar, T. Sharmila, i V. Rajendran, „A wavelet transform based contrast enhancement method for underwater acoustic images“, *Multidimens. Syst. Signal Process.*, sv. 29, lis. 2018, doi: 10.1007/s11045-017-0533-5.
- [159] E. H. Land i J. J. McCann, „Lightness and Retinex Theory“, *J. Opt. Soc. Am.*, sv. 61, izd. 1, str. 1, sij. 1971, doi: 10.1364/JOSA.61.000001.
- [160] K. R. Joshi i R. S. Kamathe, „Quantification of retinex in enhancement of weather degraded images“, u *2008 International Conference on Audio, Language and Image Processing*, Shanghai, China: IEEE, srp. 2008, str. 1229–1233. doi: 10.1109/ICALIP.2008.4590120.
- [161] X. Fu, P. Zhuang, Y. Huang, Y. Liao, X.-P. Zhang, i X. Ding, „A retinex-based enhancing approach for single underwater image“, u *2014 IEEE International Conference on Image Processing (ICIP)*, Paris, France: IEEE, lis. 2014, str. 4572–4576. doi: 10.1109/ICIP.2014.7025927.
- [162] S. Zhang, T. Wang, J. Dong, i H. Yu, „Underwater image enhancement via extended multi-scale Retinex“, *Neurocomputing*, sv. 245, str. 1–9, srp. 2017, doi: 10.1016/j.neucom.2017.03.029.
- [163] W. Yong-xin, D. Ming, i H. a. N. Chuang, „Underwater Image Enhancement Algorithm Based on Iterative Histogram Equalization with Conventional Light Source“, *Acta Photonica Sin.*, sv. 47, izd. 11, str. 1101002, 2018, doi: 10.3788/gzxb20184711.1101002.

- [164] W. Zhang, L. Dong, X. Pan, J. Zhou, L. Qin, i W. Xu, „Single Image Defogging Based on Multi-Channel Convolutional MSRCR“, *IEEE Access*, sv. 7, str. 72492–72504, 2019, doi: 10.1109/ACCESS.2019.2920403.
- [165] C. Tang, U. Freiherr von Lukas, M. Vahl, S. Wang, Y. Wang, i M. Tan, „Efficient underwater image and video enhancement based on Retinex“, *Signal Image Video Process.*, sv. 13, srp. 2019, doi: 10.1007/s11760-019-01439-y.
- [166] W. Zhang, X. Pan, X. Xie, L. Li, Z. Wang, i C. Han, „Color correction and adaptive contrast enhancement for underwater image enhancement“, *Comput. Electr. Eng.*, sv. 91, str. 106981, svi. 2021, doi: 10.1016/j.compeleceng.2021.106981.
- [167] S. Dixit, S. K. Tiwari, i P. Sharma, „Underwater image enhancement using DCP with ACCLAHE and homomorphism filtering“, u *2016 International Conference on Signal Processing, Communication, Power and Embedded System (SCOPEs)*, lis. 2016, str. 2042–2046. doi: 10.1109/SCOPEs.2016.7955807.
- [168] Y. Wang, R. Chang, RuiNian, B. He, Xunfei Liu, i Jen-Hwa Guo, „Underwater image enhancement strategy with virtual retina model and image quality assessment“, *OCEANS 2016 MTS/IEEE Monterey*, str. 1–5, ruj. 2016, doi: 10.1109/OCEANS.2016.7761381.
- [169] A. Bindhu i O. U. Maheswari, „Under water image enhancement based on linear image interpolation and limited image enhancer techniques“, u *2017 Fourth International Conference on Signal Processing, Communication and Networking (ICSCN)*, Chennai, India: IEEE, ožu. 2017, str. 1–5. doi: 10.1109/ICSCN.2017.8085692.
- [170] G. E. Guraksin, O. Deperlioglu, i U. Kose, „A Novel Underwater Image Enhancement Approach with Wavelet Transform Supported by Differential Evolution Algorithm“, u *Nature Inspired Optimization Techniques for Image Processing Applications*, Cham: Springer International Publishing, 2019, str. 255–278. doi: 10.1007/978-3-319-96002-9_11.
- [171] S. Sankpal i S. Deshpande, „Underwater Image Enhancement by Rayleigh Stretching with Adaptive Scale Parameter and Energy Correction“, u *Computing, Communication and Signal Processing*, sv. 810, u *Advances in Intelligent Systems and Computing*, vol. 810. , Singapore: Springer Singapore, 2019, str. 935–947. doi: 10.1007/978-981-13-1513-8_95.

- [172] K. Azmi, A. shahrizan Abdul ghani, Z. Yusof, i Z. Ibrahim, „Natural-based underwater image color enhancement through fusion of swarm-intelligence algorithm“, *Appl. Soft Comput.*, sv. 85, str. 105810, ruj. 2019, doi: 10.1016/j.asoc.2019.105810.
- [173] W. Ren, S. Liu, H. Zhang, J. Pan, X. Cao, i M.-H. Yang, „Single Image Dehazing via Multi-scale Convolutional Neural Networks“, u *Computer Vision – ECCV 2016*, sv. 9906, B. Leibe, J. Matas, N. Sebe, i M. Welling, Ur., u Lecture Notes in Computer Science, vol. 9906. , Cham: Springer International Publishing, 2016, str. 154–169. doi: 10.1007/978-3-319-46475-6_10.
- [174] X. Li, M. Ye, Y. Liu, i C. Zhu, „Adaptive Deep Convolutional Neural Networks for Scene-Specific Object Detection“, *IEEE Trans. Circuits Syst. Video Technol.*, sv. 29, izd. 9, str. 2538–2551, ruj. 2019, doi: 10.1109/TCSVT.2017.2749620.
- [175] X. Pan i ostali, „Accurate segmentation of nuclei in pathological images via sparse reconstruction and deep convolutional networks“, *Neurocomputing*, sv. 229, stu. 2016, doi: 10.1016/j.neucom.2016.08.103.
- [176] S. Anwar, C. Li, i F. Porikli, „Deep Underwater Image Enhancement“, 10. srpanj 2018., *arXiv*: arXiv:1807.03528. Pristupljeno: 14. veljača 2024. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1807.03528>
- [177] C. Li, J. Guo, i C. Guo, „Emerging from Water: Underwater Image Color Correction Based on Weakly Supervised Color Transfer“, *IEEE Signal Process. Lett.*, sv. 25, izd. 3, str. 323–327, ožu. 2018, doi: 10.1109/LSP.2018.2792050.
- [178] P. Uplavikar, Z. Wu, i Z. Wang, „All-In-One Underwater Image Enhancement using Domain-Adversarial Learning“, 30. svibanj 2019., *arXiv*: arXiv:1905.13342. Pristupljeno: 14. veljača 2024. [Na internetu]. Dostupno na: <http://arxiv.org/abs/1905.13342>
- [179] C. Li, S. Anwar, i F. Porikli, „Underwater scene prior inspired deep underwater image and video enhancement“, *Pattern Recognit.*, sv. 98, str. 107038, velj. 2020, doi: 10.1016/j.patcog.2019.107038.
- [180] K. Hu, Y. Zhang, C. Weng, P. Wang, Z. Deng, i Y. Liu, „An Underwater Image Enhancement Algorithm Based on Generative Adversarial Network and Natural Image Quality Evaluation Index“, *J. Mar. Sci. Eng.*, sv. 9, izd. 7, str. 691, lip. 2021, doi: 10.3390/jmse9070691.

- [181] P. Tang, L. Li, Y. Xue, M. Lv, Z. Jia, i H. Ma, „Real-World Underwater Image Enhancement Based on Attention U-Net“, *J. Mar. Sci. Eng.*, sv. 11, izd. 3, str. 662, ožu. 2023, doi: 10.3390/jmse11030662.
- [182] Y. Gao, H. Li, i S. Wen, „Restoration and Enhancement of Underwater Images Based on Bright Channel Prior“, *Math. Probl. Eng.*, sv. 2016, str. 1–15, 2016, doi: 10.1155/2016/3141478.
- [183] J. Zhou, D. Zhang, i W. Zhang, „Adaptive histogram fusion-based colour restoration and enhancement for underwater images“, *Int. J. Secur. Netw.*, sv. 16, str. 49, sij. 2021, doi: 10.1504/IJSN.2021.112848.
- [184] W. Luo, S. Duan, i J. Zheng, „Underwater Image Restoration and Enhancement Based on a Fusion Algorithm With Color Balance, Contrast Optimization, and Histogram Stretching“, *IEEE Access*, sv. 9, str. 31792–31804, 2021, doi: 10.1109/ACCESS.2021.3060947.
- [185] S. K. Dewangan, „Visual quality restoration & enhancement of underwater images using HSV filter analysis“, u *2017 International Conference on Trends in Electronics and Informatics (ICEI)*, svi. 2017, str. 766–772. doi: 10.1109/ICOEI.2017.8300807.
- [186] G. Sequeira, V. Mekkalki, J. Prabhu, S. Borkar, i M. Desai, „Hybrid Approach for Underwater Image Restoration and Enhancement“, u *2021 International Conference on Emerging Smart Computing and Informatics (ESCI)*, Pune, India: IEEE, ožu. 2021, str. 427–432. doi: 10.1109/ESCI50559.2021.9397058.
- [187] H. H. kareem, H. G. Daway, i E. G. Daway, „Underwater Image Enhancement using Colour Restoration based on YCbCr Colour Model“, *IOP Conf. Ser. Mater. Sci. Eng.*, sv. 571, izd. 1, str. 012125, srp. 2019, doi: 10.1088/1757-899X/571/1/012125.
- [188] J. Zhou, X. Wei, J. Shi, W. Chu, i Y. Lin, „Underwater image enhancement via two-level wavelet decomposition maximum brightness color restoration and edge refinement histogram stretching“, *Opt. Express*, sv. 30, izd. 10, str. 17290, svi. 2022, doi: 10.1364/OE.450858.
- [189] S. M. Pizer, R. E. Johnston, J. P. Ericksen, B. C. Yankaskas, i K. E. Muller, „Contrast-limited adaptive histogram equalization: speed and effectiveness“, u *[1990] Proceedings of the First Conference on Visualization in Biomedical Computing*, Atlanta, GA, USA: IEEE Comput. Soc. Press, 1990, str. 337–345. doi: 10.1109/VBC.1990.109340.

- [190] W. Mustafa i M. Kader, „A Review of Histogram Equalization Techniques in Image Enhancement Application“, *J. Phys. Conf. Ser.*, sv. 1019, str. 012026, lip. 2018, doi: 10.1088/1742-6596/1019/1/012026.
- [191] W. Xiang i ostali, „Underwater image enhancement based on red channel weighted compensation and gamma correction model“, *Opto-Electron. Adv.*, sv. 1, izd. 10, str. 18002401–18002409, 2018, doi: 10.29026/oea.2018.180024.
- [192] M. Loesdau, S. Chabrier, i A. Gabillon, „Hue and Saturation in the RGB Color Space“, u *Image and Signal Processing*, sv. 8509, u *Lecture Notes in Computer Science*, vol. 8509. , Cham: Springer International Publishing, 2014, str. 203–212. doi: 10.1007/978-3-319-07998-1_23.
- [193] J. Narasimharao, „Hue Preservation-Based High-Efficiency Underwater Image Correlation and Enhancement Using Deep Learning Technique“, sv. 12, izd. 03, 2023.
- [194] G. Hou, Z. Pan, B. Huang, G. Wang, i X. Luan, „Hue preserving-based approach for underwater colour image enhancement“, *IET Image Process.*, sv. 12, izd. 2, str. 292–298, velj. 2018, doi: 10.1049/iet-ipr.2017.0359.
- [195] F. Pierre, J.-F. Aujol, A. Bugeau, i V.-T. Ta, „Luminance-Hue Specification in the RGB Space“, u *Scale Space and Variational Methods in Computer Vision*, sv. 9087, u *Lecture Notes in Computer Science*, vol. 9087. , Cham: Springer International Publishing, 2015, str. 413–424. doi: 10.1007/978-3-319-18461-6_33.
- [196] V. Soni, A. Sharma, i J. Rajpurohit, „A Swift Algorithm and Hue-Preserving Based Mechanism for Underwater Image Colour Enhancement“, *Int. J. Intell. Syst. Appl. Eng.*, sv. 12, izd. 1, str. 203–220, pros. 2023.
- [197] L. Yang, B. Kang, Z. Huang, Z. Zhao, X. Xu, i J. Feng, „Depth Anything V2“, 20. listopad 2024., *arXiv*: arXiv:2406.09414. doi: 10.48550/arXiv.2406.09414.
- [198] I. Biliškov i V. Papić, „Development of a Seafloor Litter Database and Application of Image Preprocessing Techniques for UAV-Based Detection of Seafloor Objects“, *Electronics*, sv. 13, izd. 17, 2024, doi: 10.3390/electronics13173524.
- [199] W. Song, Y. Liu, D. Huang, B. Zhang, Z. Shen, i H. Xu, „From shallow sea to deep sea: research progress in underwater image restoration“, *Front. Mar. Sci.*, sv. 10, str. 1163831, svi. 2023, doi: 10.3389/fmars.2023.1163831.

- [200] Y. Wang, W. Song, G. Fortino, L.-Z. Qi, W. Zhang, i A. Liotta, „An Experimental-Based Review of Image Enhancement and Image Restoration Methods for Underwater Imaging“, *IEEE Access*, sv. 7, str. 140233–140251, 2019, doi: 10.1109/ACCESS.2019.2932130.
- [201] „ultralytics/ultralytics/cfg/default.yaml at main · ultralytics/ultralytics“, GitHub. Pristupljeno: 22. kolovoz 2024. [Na internetu]. Dostupno na: <https://github.com/ultralytics/ultralytics/blob/main/ultralytics/cfg/default.yaml>
- [202] „Training, Validation, Test Split for Machine Learning Datasets“. Pristupljeno: 22. kolovoz 2024. [Na internetu]. Dostupno na: <https://encord.com/blog/train-val-test-split/>

PRILOG A Modifikacije YOLOv8 za podršku četverokanalnih slika

Ovaj dodatak detaljno opisuje promjene napravljene na YOLOv8 (specifično na verziji v8.0.235) frameworku kako bi se omogućilo treniranje i inferencija modela korištenjem četverokanalnih slika (RGB + dodatni kanal, u slučaju ove disercije je to kanal Hue). Svako poglavlje opisuje promjene u jednoj specifičnoj datoteci unutar Ultralytics YOLO paketa.

A.1. ultralytics.nn.tasks

Ova datoteka definira osnovne klase modela (poput `DetectionModel`) i pomoćne funkcije za učitavanje modela i težina. Ključne izmjene su napravljene kako bi:

- Model eksplicitno bio definiran za prihvatanje 4 ulazna kanala.
- Omogućilo učitavanje predtreniranih težina (koje su originalno za 3 kanala) u novu četverokanalnu arhitekturu, prilagođavajući težine prvog konvolucijskog sloja.

A.1.1. Promjena zadanog broja kanala u `DetectionModel`

Zadani broj ulaznih kanala (`ch`) u konstruktoru klase `DetectionModel` promijenjen je s 3 na 4. Ovo osigurava da kada se model inicijalizira (npr. prilikom treniranja od početka ili korištenjem YAML konfiguracijske datoteke koja ne specificira eksplicitno parametar `ch`), model će automatski biti konfiguriran za 4 ulazna kanala.

```
# --- a/ultralytics/nn/tasks.py
# +++ b/ultralytics/nn/tasks.py
# @@ -218,7 +218,7 @@ class BaseModel(nn.Module):
#     class DetectionModel(BaseModel):
#         """YOLOv8 detection model."""

#         def __init__(self, cfg='yolov8n.yaml', ch=3, nc=None, verbose=True):
#             # model, input channels, number of classes
#         + def __init__(self, cfg='yolov8n.yaml', ch=4, nc=None, verbose=True):
#             # model, input channels, number of classes
#             """Initialize the YOLOv8 detection model with the given config and
#             parameters."""
#             super().__init__()
#             self.yaml = cfg if isinstance(cfg, dict) else yaml_model_load(cfg)
#             # cfg dict
```

A.1.2. Dodavanje funkcije `adapt_weights_to_4ch`

Dodana je nova funkcija `adapt_weights_to_4ch` koja služi za prilagodbu predtreniranih trokanalnih težina za korištenje u četverokanalnom modelu. Funkcija prima model i njegovo stanje (state dictionary) koje sadrži trokanalne težine.

- Pronalazi prvi konvolucijski sloj modela koji očekuje 4 ulazna kanala.
- Ako težine za taj sloj u stanju imaju dimenziju za 3 kanala:
 - o Stvara novi tenzor za težine s četiri kanala.
 - o Kopira težine za prva tri kanala (RGB) iz predtreniranih težina.
 - o Inicijalizira težine za četvrti kanal dupliciranjem težina zelenog kanala. Zeleni kanal često nosi najviše strukturnih informacija u prirodnim slikama i pretpostavlja se da je dobar početni set težina za novi kanal, minimizirajući potrebu za dodatnim treniranjem ispočetka.
 - o Ažurira stanje modela s novim, prilagođenim četverokanalnim težinama.

```
# --- a/ultralytics/nn/tasks.py
# +++ b/ultralytics/nn/tasks.py
# @@ -637,6 +637,51 @@ def attempt_load_weights(weights, device=None,
inplace=True, fuse=False):
    assert all(ensemble[0].nc == m.nc for m in ensemble), f'Models differ in
class counts {[m.nc for m in ensemble]}'
    return ensemble

+def adapt_weights_to_4ch(model, state_dict):
+    """
+    Adapt 3-channel pretrained weights to 4-channel model.
+
+    Args:
+        model: The 4-channel model
+        state_dict: The state dictionary containing the pretrained weights
+
+    Returns:
+        Updated state dictionary with 4-channel weights
+    """
+    # Find the first conv layer
+    for name, module in model.named_modules():
+        if isinstance(module, nn.Conv2d) and module.in_channels == 4:
+            # Get the corresponding weights from pretrained state dict
```

```

+         if name + '.weight' in state_dict:
+             old_weights = state_dict[name + '.weight']
+             if old_weights.shape[1] == 3:
+                 LOGGER.info(f'Adapting weights of {name} from 3 to 4
channels (green channel duplication strategy)')
+                 # Create new weights with 4 channels
+                 new_weights = torch.zeros((old_weights.shape[0], 4,
*old_weights.shape[2:]),
+                                           device=old_weights.device,
dtype=old_weights.dtype)
+                 # Copy the RGB weights
+                 new_weights[:, :3, :, :] = old_weights
+                 # Initialize the 4th channel using the green channel
(index 1) duplication
+                 new_weights[:, 3, :, :] = old_weights[:, 1, :, :]
+
+                 # Update the weights
+                 state_dict[name + '.weight'] = new_weights
+             break
+
+     return state_dict

```

A.1.3. Prilagodba attempt_load_one_weight

Modificirana je funkcija `attempt_load_one_weight`, koja učitava težine za pojedinačni model. Nakon što se model učitava iz checkpoint (*.pt) datoteke, provjerava se ima li taj model prvi konvolucijski sloj konfiguriran za 4 ulazna kanala. Ako ima (`needs_adaptation` je `True`), tada se poziva novododana funkcija `adapt_weights_to_4ch` kako bi se predtrenirane trokanalne težine prilagodile četverokanalnoj arhitekturi prije nego što se te težine konačno učitaju u model pomoću `model.load_state_dict(state_dict)`.

```

# --- a/ultralytics/nn/tasks.py
# +++ b/ultralytics/nn/tasks.py
# @@ -640,12 +690,23 @@ def attempt_load_one_weight(weight, device=None,
inplace=True, fuse=False):
    ckpt, weight = torch_safe_load(weight) # load ckpt
    args = {**DEFAULT_CFG_DICT, **(ckpt.get('train_args', {}))} # combine
model and default args, preferring model args
    model = (ckpt.get('ema') or ckpt['model']).to(device).float() # FP32
model

```

```

+ # Check if model needs channel adaptation (first conv layer has 4
+ channels)
+ needs_adaptation = False
+ for name, module in model.named_modules():
+     if isinstance(module, nn.Conv2d) and module.in_channels == 4:
+         needs_adaptation = True
+         break
+
+ if needs_adaptation:
+     # Adapt weights for 4-channel input
+     LOGGER.info('Adapting model from 3-channel to 4-channel input')
+     state_dict = model.state_dict()
+     state_dict = adapt_weights_to_4ch(model, state_dict)
+     model.load_state_dict(state_dict)
+
+     # Model compatibility updates
+     model.args = {k: v for k, v in args.items() if k in DEFAULT_CFG_KEYS} #
+     attach args to model
+     model.pt_path = weight # attach *.pt file path to model

```

Ove promjene su fundamentalne za rad s četverokanalnim slikama. Omogućavaju da se:

- Kreiraju modeli koji su strukturno spremni primiti četiri kanala.
- Iskoriste postojeće, prethodno trenirane trokanalne težine kao polazna točka za treniranje četverokanalnog modela, umjesto treniranja potpuno „od nule“. Ovo značajno ubrzava proces treniranja, dok je negativna strana što četvrti kanal ne mora imati informacije koje se mogu naučiti iz zelenog kanala.

A.2. ultralytics.nn.autobackend

Klasa AutoBackend služi kao omotač koji omogućuje rad s modelima u različitim formatima (PyTorch .pt, ONNX, TensorRT engine, itd.) na unificiran način. Metoda warmup se koristi za „zagrijavanje“ modela, tj. izvođenje jednog inferencijskog prolaza s lažnim ulaznim podacima kako bi se alocirale potrebne memorije i optimizirali budući prolazi. Ovo zagrijavanje treba koristiti podatke s ispravnim brojem kanala.

Zadana veličina slike (`imgsz`) koja se koristi kao argument u `warmup` metodi unutar klase `AutoBackend` promijenjena je da koristi 4 kanala umjesto 3. Specifično, dio `(1, 3, 640, 640)` postao je `(1, 4, 640, 640)`. U direktoriju `assets/` se nalaze dvije slike: `bus.jpg` i `zidane.jpg`. Navedene slike je potrebno prebaciti u PNG format i dodati četvrti kanal.

```
# --- a/ultralytics/nn/autobackend.py
# +++ b/ultralytics/nn/autobackend.py
# @@ -465,7 +465,7 @@ class AutoBackend(nn.Module):
#     """
#     return torch.tensor(x).to(self.device) if isinstance(x, np.ndarray) else
x
-     def warmup(self, imgsz=(1, 3, 640, 640)):
+     def warmup(self, imgsz=(1, 4, 640, 640)):
#         """
#         Warm up the model by running one forward pass with a dummy input.
```

Ova promjena osigurava da je model, bez obzira na format u kojem je učitán, ispravno pripremljen (zagrijan) za primanje četverokanalnih slika prije početka stvarne inferencije ili validacije.

A.3. `ultralytics.engine.validator`

Klasa `BaseValidator` upravlja cjelokupnim procesom validacije modela na validacijskom skupu podataka. Slično kao kod `AutoBackend`, ključno je da se model prije početka validacijske petlje pripremi za očekivani četverokanalni ulaz.

Unutar glavne metode (`__call__`) validatora, poziv `model.warmup` je ažuriran tako da specificira četiri kanala. Koristi se veličina skupine (*engl. batch size*) definirana u argumentima (`self.args.batch`), a broj kanala je eksplicitno postavljen na 4.

```
# --- a/ultralytics/engine/validator.py
# +++ b/ultralytics/engine/validator.py
# @@ -143,7 +143,7 @@ class BaseValidator:
#
#     self.dataloader = self.dataloader or
self.get_dataloader(self.data.get(self.args.split), self.args.batch)
```

```

        model.eval()

-         model.warmup(imgsz=(1 if pt else self.args.batch, 3, imgsz,
imgsz)) # warmup
+         model.warmup(imgsz=(1 if pt else self.args.batch, 4, imgsz,
imgsz)) # warmup

        self.run_callbacks('on_val_start')

        dt = Profile(), Profile(), Profile(), Profile()

```

Osigurava da validacijski proces započne s modelom koji je pravilno inicijaliziran i spreman za obradu četverokanalnih slika koje dolaze iz validacijskog skupa podataka (`data_loader`).

A.4. **ultralytics.engine.predictor**

Klasa `BasePredictor` je odgovorna za izvođenje inferencije (predikcije) modela na novim, neviđenim podacima. Ovdje su bile potrebne izmjene na mjestima:

- Predobrada (`preprocess`): Ulazne slike (koje dolaze kao NumPy polja) treba ispravno formatirati za četverokanalni model prije slanja u neuronsku mrežu.
- Zagrijavanje (`warmup`): Model treba zagrijati specifično unutar prediktora koristeći četverokanalni tenzor.

A.4.1. **Preprocesiranje četverokanalnih slika**

Metoda `preprocess` je značajno izmijenjena. Nakon što se ulazne slike (koje mogu biti lista NumPy polja) slože u jedan batch (`np.stack`), provjerava se imaju li slike 4 kanala (`has_4ch = im.shape[3] == 4`).

- Ako slika ima 4 kanala:
 - Prva tri kanala (koji su u BGR poretku iz OpenCV-a) se izdvajaju i konvertiraju u RGB poredak (`rgb = im[..., :3][..., ::-1]`).
 - Četvrti kanal se izdvaja zasebno (`alpha = im[..., 3:4]`). Važno je napomenuti da se ovdje koristi naziv *alpha*, iako u ovom slučaju predstavlja dodatni kanal koji sadržava normalizirane vrijednosti Hue kanala. Naziv

varijable reflektira poziciju unutar PNG datoteke, koje su korištene za treniranje.

- Konvertirani RGB kanali i nepromijenjeni četvrti kanal se spajaju natrag (`im = np.concatenate([rgb, alpha], axis=3)`).
- Ako slika ima 3 kanala:
 - Radi se samo standardna BGR u RGB konverzija (`im = im[..., :-1]`).
- Nakon ove logike specifične za kanale, slika se transponira iz BHWC (Batch, Height, Width, Channel) u BCHW (Batch, Channel, Height, Width) format. Ostatak koda ostaje isti, osigurava se da je slika memorijski kontinuirana (`np.ascontiguousarray`) i konvertira u `torch.Tensor`.

```
# --- a/ultralytics/engine/predictor.py
# +++ b/ultralytics/engine/predictor.py
# @@ -113,11 +113,23 @@ class BasePredictor:
    not_tensor = not isinstance(im, torch.Tensor)
    if not_tensor:
        im = np.stack(self.pre_transform(im))
-        im = im[..., :-1].transpose((0, 3, 1, 2)) # BGR to RGB, BHWC
to BCHW, (n, 3, h, w)
+
+        # Check if image has 4 channels
+        has_4ch = im.shape[3] == 4 if len(im.shape) == 4 else False
+
+        if has_4ch:
+            # Handle 4-channel images properly
+            rgb = im[..., :3][..., :-1] # BGR to RGB for first 3
channels
+            alpha = im[..., 3:4] # Preserve 4th channel
+            im = np.concatenate([rgb, alpha], axis=3)
+        else:
+            # Standard BGR to RGB conversion
+            im = im[..., :-1]
+
+        im = im.transpose((0, 3, 1, 2)) # BHWC to BCHW
+        im = np.ascontiguousarray(im) # contiguous
+        im = torch.from_numpy(im)

    im = im.to(self.device)
```

A.4.2. Zagrijavanje modela u *predict*

Unutar glavne *predict* metode, prije početka petlje po podacima, poziv `model.warmup` je ažuriran da koristi 4 kanala, slično kao u *Validatoru* i *AutoBackendu*.

```
# --- a/ultralytics/engine/predictor.py
# +++ b/ultralytics/engine/predictor.py
# @@ -239,7 +252,7 @@ class BasePredictor:

    # Warmup model
    if not self.done_warmup:
        -         self.model.warmup(imgsz=(1 if self.model.pt or self.model.triton
        else self.dataset.bs, 3, *self.imgsz))
        +         self.model.warmup(imgsz=(1 if self.model.pt or self.model.triton
        else self.dataset.bs, 4, *self.imgsz))
        self.done_warmup = True

        self.seen, self.windows, self.batch, profilers = 0, [], None,
        (ops.Profile(), ops.Profile(), ops.Profile())
```

A.4.3. Ažuriranje poruke o brzini

Na kraju *predict* metode, informativni ispis koji prikazuje prosječne brzine pretprocesiranja, inferencije i postobrade te oblik ulaznog tenzora, ažuriran je da ispravno prikazuje 4 kanala.

```
# --- a/ultralytics/engine/predictor.py
# +++ b/ultralytics/engine/predictor.py
        LOGGER.info(f'Speed: %.1fms preprocess, %.1fms inference, %.1fms
postprocess per image at shape '
-         f'{{(1, 3, *im.shape[2:])}}' % t)
+         f'{{(1, 4, *im.shape[2:])}}' % t)
```

Ove promjene su bitne za ispravno funkcioniranje predviđanja (`model.predict()` ili `model()`). Metoda `preprocess` osigurava da četverokanalna slika (učitana npr. kao BGRa iz OpenCV-a) bude korektno transformirana u format koji neuronska mreža očekuje (RGB+Hue, BCHW poredak, normalizirane vrijednosti). Warmup i logiranje su usklađeni s četverokanalnom obradom.

A.5. ultralytics.data.loaders

Klasa `LoadImages` se koristi kao iterator za učitavanje slika (ili video sličica) jednu po jednu ili u skupini kada se model poziva s izvorom (`source`) koji je direktorij, video datoteka, ili slično. Originalna implementacija koristila je `cv2.imread` bez dodatnih zastavica (*eng. flags*), što bi za PNG slike s alpha kanalom učitalo samo prva tri kanala (BGR), ignorirajući četvrti (*alpha*) kanal.

U `__next__` metodi klase `LoadImages`, koja se poziva za dohvaćanje sljedeće slike, poziv `cv2.imread(path)` zamijenjen je s `cv2.imread(path, cv2.IMREAD_UNCHANGED)`. Zastavica `cv2.IMREAD_UNCHANGED` nalaže OpenCV-u da učitava sliku „takvu kakva jest“, uključujući sve kanale prisutne u datoteci. Ovo je ključno za ispravno učitavanje četverokanalnih PNG slika, gdje se četvrti kanal (koji je tehnički alpha kanal u PNG formatu) koristi za pohranu dodatne informacije.

```
# --- a/ultralytics/data/loaders.py
# +++ b/ultralytics/data/loaders.py
# @@ -336,7 +336,7 @@ class LoadImages:
#     else:
#         # Read image
#         self.count += 1
-         im0 = cv2.imread(path) # BGR
+         im0 = cv2.imread(path, cv2.IMREAD_UNCHANGED) # BGRA
        if im0 is None:
            raise FileNotFoundError(f'Image Not Found {path}')
        s = f'image {self.count}/{self.nf} {path}: '
```

Ova izmjena omogućava da četverokanalne slike (spremljene kao PNG) budu ispravno učitane sa sva četiri kanala.

A.6. ultralytics.data.augment

Ova datoteka sadrži klase za različite tehnike augmentacije podataka koje se primjenjuju tijekom faze treniranja modela (a neke, poput `LetterBox`, i tijekom predikcije). Bilo je nužno osigurati da ove augmentacije ispravno rukuju četverokanalnim slikama, specifično:

- mosaic: Ispravna inicijalizacija pozadinskih slika s 4 kanala.
- randomPerspective: Korištenje ispravne vrijednosti popunjavanja za sva 4 kanala pri geometrijskim transformacijama.
- randomHSV: Primjena promjena na Hue, Saturation i Value kanalima, uz specifično rukovanje četvrtim (Hue) kanalom.
- letterBox: Korištenje ispravne vrijednosti popunjavanja za sva 4 kanala pri prilagođavanju veličine slike.

A.6.1. Mosaic Augmentation

Unutar metoda koje kreiraju mosaic (`_mosaic3`, `_mosaic4`, `_mosaic9`), kod inicijalizacije osnovne (pozadinske) slike (`img3`, `img4`, `img9`) sada se prvo provjerava ima li prva učitana slika 4 kanala (`has_4ch = img.shape[2] == 4`).

- Ako ima 4 kanala, osnovna slika se kreira s 4 kanala (`np.full(..., 4, ...)`).
- Ako ima 3 kanala, kreira se trokanalna slika. U oba slučaja, svi kanali se inicijaliziraju na vrijednost 114, vrijednost koju YOLOv8 koristi za boju pozadine.

```
# --- a/ultralytics/data/augment.py
# +++ b/ultralytics/data/augment.py
# @@ -211,7 +216,12 @@ class Mosaic(BaseMixTransform):

    # Place img in img4
    if i == 0: # top left
-         img4 = np.full((s * 2, s * 2, img.shape[2]), 114,
dtype=np.uint8) # base image with 4 tiles
+         # Check for 4-channel image
+         has_4ch = img.shape[2] == 4
+         if has_4ch:
+             img4 = np.full((s * 2, s * 2, 4), 114, dtype=np.uint8)
# Set all channels to 114
+         else:
+             img4 = np.full((s * 2, s * 2, 3), 114, dtype=np.uint8)
+             x1a, y1a, x2a, y2a = max(xc - w, 0), max(yc - h, 0), xc, yc
# xmin, ymin, xmax, ymax (large image)
+             x1b, y1b, x2b, y2b = w - (x2a - x1a), h - (y2a - y1a), w, h
# xmin, ymin, xmax, ymax (small image)
+         elif i == 1: # top right

# (Slične promjene su napravljene i u _mosaic3 i _mosaic9)
```

A.6.2. RandomPerspective

Kod primjene affine ili perspektivne transformacije pomoću `cv2.warpAffine` ili `cv2.warpPerspective`, parametar `borderValue` definira boju kojom se popunjavaju novi pikseli koji nastanu transformacijom izvan granica originalne slike. Ako slika ima 4 kanala (`has_4ch`), koristi se (114, 114, 114, 114); inače se koristi standardno (114, 114, 114).

```
# --- a/ultralitics/data/augment.py
# +++ b/ultralitics/data/augment.py
# @@ -423,11 +438,14 @@ class RandomPerspective:

    # Combined rotation matrix
    M = T @ S @ R @ P @ C # order of operations (right to left) is
IMPORTANT
    # Affine image
+    has_4ch = img.shape[2] == 4
+    borderValue = (114, 114, 114, 114) if has_4ch else (114, 114, 114)
# Use 114 for all channels
+
    if (border[0] != 0) or (border[1] != 0) or (M != np.eye(3)).any():
# image changed
        if self.perspective:
-            img = cv2.warpPerspective(img, M, dsize=self.size,
borderValue=(114, 114, 114))
+            img = cv2.warpPerspective(img, M, dsize=self.size,
borderValue=borderValue)
        else: # affine
-            img = cv2.warpAffine(img, M[:2], dsize=self.size,
borderValue=(114, 114, 114))
+            img = cv2.warpAffine(img, M[:2], dsize=self.size,
borderValue=borderValue)
        return img, M, s
```

A.6.3. RandomHSV

Augmentacija `RandomHSV` mijenja nasumično nijansu (Hue), zasićenost (Saturation) i vrijednost (Value) slike. Izmjene su napravljene da se ispravno obrade četverokanalne slike:

- Provjerava se ima li slika 4 kanala (`has_4ch`).
- Ako ima 4 kanala:

- Slika se razdvaja na BGR dio (`bgr = img[..., :3]`) i četvrti kanal (`h_channel = img[..., 3:4]`).
- Standardna HSV augmentacija (izračunavanje pomaka za H, S, V pomoću lookup tablica `lut_hue`, `lut_sat`, `lut_val`) primjenjuje se samo na BGR dio.
- Potrebno je primijeniti isti Hue pomak i na četvrti kanal.
- Augmentirani BGR dio se spaja s nepromijenjenim četvrtim kanalom (`img = np.concatenate(...)`).

Ako slika ima 3 kanala, primjenjuje se standardna HSV augmentacija na cijelu sliku.

```
# --- a/ultralalytics/data/augment.py
# +++ b/ultralalytics/data/augment.py
# @@ -597,40 +615,77 @@ class RandomHSV:
    self.sgain = sgain
    self.vgain = vgain
-
    def __call__(self, labels):
        """
        Applies random HSV augmentation to an image within the predefined
        limits.
        + For 4-channel images, the 4th channel (Hue) must be shifted
        consistently
        + with the Hue shift applied to the RGB channels. # <- Komentar je
        ostao, ali kod ne radi to.

        The modified image replaces the original image in the input 'labels'
        dict.
        """
        img = labels['img']
        if self.hgain or self.sgain or self.vgain:
            r = np.random.uniform(-1, 1, 3) * [self.hgain, self.sgain,
self.vgain] + 1 # random gains
-            hue, sat, val = cv2.split(cv2.cvtColor(img, cv2.COLOR_BGR2HSV))
+
+            # Check if this is a 4-channel image
+            has_4ch = img.shape[2] == 4
+            if has_4ch:
+                # Split into BGR and Hue channels
+                bgr = img[..., :3]
+                h_channel = img[..., 3:4]
```

```

+
+         # Process BGR channels as usual
+         hue, sat, val = cv2.split(cv2.cvtColor(bgr,
cv2.COLOR_BGR2HSV))
+     else:
+         # Standard 3-channel processing
+         hue, sat, val = cv2.split(cv2.cvtColor(img,
cv2.COLOR_BGR2HSV))
+
+         dtype = img.dtype # uint8
-
+
+         # Calculate the hue shift
+         h_shift = int(180 * r[0] - 180) if r[0] > 1 else int(180 * r[0])
+
+         x = np.arange(0, 256, dtype=r.dtype)
+         lut_hue = ((x * r[0]) % 180).astype(dtype)
+         lut_sat = np.clip(x * r[1], 0, 255).astype(dtype)
+         lut_val = np.clip(x * r[2], 0, 255).astype(dtype)
-
+
+         # Process BGR channels
+         im_hsv = cv2.merge((cv2.LUT(hue, lut_hue), cv2.LUT(sat,
lut_sat), cv2.LUT(val, lut_val)))
-         cv2.cvtColor(im_hsv, cv2.COLOR_HSV2BGR, dst=img) # no return
needed
+
+         if has_4ch:
+             adjusted_h_channel = cv2.LUT(h_channel.reshape(-1),
+                                     ((np.arange(0, 256) * 180 /
255 + h_shift) % 180) * 255 / 180).astype(dtype)
+             adjusted_h_channel =
adjusted_h_channel.reshape(h_channel.shape)
+
+             # Update the BGR channels
+             bgr_augmented = cv2.cvtColor(im_hsv, cv2.COLOR_HSV2BGR)
+
+             # Merge the augmented BGR with the adjusted Hue channel
+             img = np.concatenate([bgr_augmented, adjusted_h_channel],
axis=2)
+         else:

```

```

+             # Standard 3-channel update
+             cv2.cvtColor(im_hsv, cv2.COLOR_HSV2BGR, dst=img) # no
return needed
+
+     labels['img'] = img
    return labels

```

A.6.4. LetterBox

Klasa LetterBox mijenja veličinu slike na ciljanu veličinu uz očuvanje omjera stranica, dodajući sive rubove (padding) po potrebi. Vrijednost piksela za te rubove (value parametar u cv2.copyMakeBorder) sada se dinamički određuje: ako slika ima 4 kanala (has_4ch), koristi se (114, 114, 114, 114); inače se koristi standardno (114, 114, 114).

```

# --- a/ultralytics/data/augment.py
# +++ b/ultralytics/data/augment.py
# @@ -723,11 +778,16 @@ class LetterBox:
    img = cv2.resize(img, new_unpad, interpolation=cv2.INTER_LINEAR)
    top, bottom = int(round(dh - 0.1)) if self.center else 0,
int(round(dh + 0.1))
    left, right = int(round(dw - 0.1)) if self.center else 0,
int(round(dw + 0.1))
+
+     # Use appropriate padding value based on channel count
+     has_4ch = img.shape[2] == 4
+     padding_value = (114, 114, 114, 114) if has_4ch else (114, 114, 114)
# Use 114 for all channels
+
    img = cv2.copyMakeBorder(img, top, bottom, left, right,
cv2.BORDER_CONSTANT,
-                               value=(114, 114, 114)) # add border
+                               value=padding_value) # add border
    if labels.get('ratio_pad'):
        labels['ratio_pad'] = (labels['ratio_pad'], (left, top)) # for
evaluation

```

Ove izmjene su vitalne za ispravno provođenje augmentacije podataka tijekom treniranja modela. Osiguravaju da se četverokanalne slike konzistentno obrađuju u svim koracima augmentacije, od kreiranja mozaika, preko geometrijskih i HSV transformacija, do finalnog prilagođavanja veličine slike za ulazak u mrežu.

A.7. ultralytics.utils.plotting

Funkcija `plot_images` se koristi interno unutar YOLOv8 frameworka za vizualizaciju slika iz batcha tijekom treniranja ili validacije (npr. sprema datoteke `batch_0_labels.jpg`, `batch_0_pred.jpg`). Budući da standardni slikovni formati poput JPG ne podržavaju 4 kanala, i alati za prikazivanje slika očekuju 3 kanala (RGB), funkcija je modificirana tako da za prikaz koristi samo prva tri kanala, čak i ako ulazne slike imaju 4 kanala.

Funkcija sada eksplicitno određuje broj kanala ulaznih slika (`num_channels`). Zatim definira broj kanala koji će se koristiti za prikaz (`display_channels = min(3, num_channels)`), osiguravajući da se ne koristi više od 3 kanala. Prilikom kreiranja praznog mozaika (`mosaic = np.full(...)`) i prilikom kopiranja svake pojedinačne slike iz batcha u mozaik (`mosaic[y:y + h, x:x + w, :] = im[:, :, :display_channels]`), koristi se samo `display_channels` (tj. prva 3 kanala) iz ulazne slike `im`, efektivno odbacujući četvrti kanal za potrebe vizualizacije.

```
# --- a/ultralytics/utils/plotting.py
# +++ b/ultralytics/utils/plotting.py
# @@ -596,14 +596,23 @@ def plot_images(images,
    bs, _, h, w = images.shape # batch size, _, height, width
    bs = min(bs, max_subplots) # limit plot images
    ns = np.ceil(bs ** 0.5) # number of subplots (square)
-   if np.max(images[0]) <= 1:
-       images *= 255 # de-normalise (optional)
+   # ... (denormalization code)

    # Build Image
-   mosaic = np.full((int(ns * h), int(ns * w), 3), 255, dtype=np.uint8) #
init
+   # Determine number of channels in the images
+   num_channels = images.shape[1] # Could be 3 or 4 for RGB/RGBA
+
+   # Use only first 3 channels for visualization (since we're displaying as
RGB)
+   display_channels = min(3, num_channels)
+
```

```

+     # Initialize mosaic with appropriate channels
+     mosaic = np.full((int(ns * h), int(ns * w), display_channels), 255,
dtype=np.uint8) # init
+
+     for i, im in enumerate(images):
+         if i == max_subplots: # if last batch has fewer images than we
expect
+             break
+             x, y = int(w * (i // ns)), int(h * (i % ns)) # block origin
+             im = im.transpose(1, 2, 0)
-         mosaic[y:y + h, x:x + w, :] = im
+         # Only use the first 3 channels for display (RGB)
+         mosaic[y:y + h, x:x + w, :] = im[:, :, :display_channels]

+         # Resize (optional)
+         scale = max_size / ns / max(h, w)

```

Ova promjena osigurava da vizualizacije koje generira YOLOv8 (poput slika batch-eva spremljenih tijekom treniranja) budu uvijek ispravne trokanalne RGB slike, bez obzira na to što model interno radi s 4 kanala. Ovo sprječava greške prilikom spremanja ili prikazivanja tih vizualizacija u standardnim alatima.

Životopis

Ivan Biliškov rođen je 07. svibnja 1987. godine u Splitu, Republika Hrvatska. Stručnjak je u području računarstva s bogatim radnim iskustvom. Trenutačno je na poslijediplomskom studiju računarstva na Fakultetu elektrotehnike, strojarstva i brodogradnje u Splitu gdje se bavi analizom i obradom fotografija. Diplomirao je računarstvo na istom fakultetu 2010. godine. Također je certificiran za SAP Boot Kamp (SAP SAP01, SAP BC400, SAP PLM200) koji je završio u Zagrebu 2010. godine.

Ivanova profesionalna karijera započela je kao programer namjenskog softvera u tvrtki Venio Indicium d.o.o. od 2006. do 2010. godine. Nakon toga radio je kao viši programer u tvrtkama Invarion Ltd i Neogov. Od 2014. do 2018. godine bio je tehnički direktor u CultureConnect LLC, a potom tehnički voditelj u PatientPal LLC. Također je radio kao full stack programer u 6Crickets i tehnički voditelj u OfflinePal. Trenutačno je glavni izvršni direktor u LOOP d.o.o. (Codeasy).

Ivan je sudjelovao u brojnim projektima u sferi umjetne inteligencije, uključujući BrAIIn i Pomozi.hr. U sferi razvoja softvera radio je na projektima kao što su Voice Of Norway, Swiftly, Enjoy.ly, PatientPal, Invarion, CultureConnect, Hajduk, Kimbell Art Museum i mnogim drugim.

Član je stručnih udruženja Split Tech City i ML/AI Split. Predavao je na međunarodnim konferencijama kao što su Data Science Europe 2023 i Innovate Now (Smart Innovation Norway). Ivan je također organizirao skupove i radionice, uključujući The Geek Gathering u Splitu. Njegove publikacije u posljednjih pet godina su „Players Detection using U-Net based Fully Convolutional Network“, „Development of a Seafloor Litter Database and Application of Image Preprocessing Techniques for UAV-Based Detection of Seafloor Objects” i „Patients prefer human psychiatrists over chatbots: a cross-sectional study“.

Curriculum Vitae

Ivan Biliškov was born on May 7, 1987, in Split, Republic of Croatia. He is an expert in the field of computer science with extensive professional experience. He is currently pursuing a postgraduate degree in Computer Science at the Faculty of Electrical Engineering, Mechanical Engineering and Naval Architecture in Split, where he focuses on image analysis and processing. He earned his undergraduate degree in Computer Science from the same faculty in 2010. He is also certified in SAP Boot Camp (SAP SAP01, SAP BC400, SAP PLM200), which he completed in Zagreb in 2010.

Ivan's professional career began as a custom software developer at Venio Indicium d.o.o. from 2006 to 2010. He then worked as a senior developer at Invarion Ltd and Neogov. From 2014 to 2018, he served as Technical Director at CultureConnect LLC, and subsequently as Technical Lead at PatientPal LLC. He also worked as a full-stack developer at 6Crickets and as Technical Lead at OfflinePal. He is currently the Chief Executive Officer at LOOP d.o.o. (Codeasy).

Ivan has participated in numerous artificial intelligence projects, including BrAIIn and Pomozi.hr. In the field of software development, he has worked on projects such as Voice Of Norway, Swiftly, Enjoy.ly, PatientPal, Invarion, CultureConnect, Hajduk, Kimbell Art Museum, and many others.

He is a member of professional organizations Split Tech City and ML/AI Split. He has lectured at international conferences such as Data Science Europe 2023 and Innovate Now (Smart Innovation Norway). Ivan has also organized meetups and workshops, including The Geek Gathering in Split.

His publications in the past five years include „Players Detection using U-Net based Fully Convolutional Network“, „Development of a Seafloor Litter Database and Application of Image Preprocessing Techniques for UAV-Based Detection of Seafloor Objects“ i „Patients prefer human psychiatrists over chatbots: a cross-sectional study“.