

Klasifikacija riba Jadranskog mora na primjeru vlastitog skupa podataka

Dadić, Petra

Master's thesis / Diplomski rad

2024

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Split, Faculty of Science / Sveučilište u Splitu, Prirodoslovno-matematički fakultet**

Permanent link / Trajna poveznica: <https://um.nsk.hr/um:nbn:hr:166:703686>

Rights / Prava: [In copyright](#)/[Zaštićeno autorskim pravom.](#)

Download date / Datum preuzimanja: **2025-02-20**

Repository / Repozitorij:

[Repository of Faculty of Science](#)



SVEUČILIŠTE U SPLITU
PRIRODOSLOVNO MATEMATIČKI FAKULTET

DIPLOMSKI RAD

**KLASIFIKACIJA RIBE JADRANSKOG MORA
NA PRIMJERU VLASTITOG SKUPA
PODATAKA**

Petra Dadić

Split, rujan 2024.

Temeljna dokumentacijska kartica

Diplomski rad

Sveučilište u Splitu

Prirodoslovno-matematički fakultet

Odjel za informatiku

Ruđera Boškovića 33, 21000 Split, Hrvatska

KLASIFIKACIJA RIBA JADRANSKOG MORA NA PRIMJERU VLASTITOG SKUPA PODATAKA

Petra Dadić

SAŽETAK

Zadatak klasifikacije riba iznimno je važan zbog njegove primjene u industriji i očuvanju biološke raznolikosti. Unatoč dugoj povijesti zadatka klasifikacije ribe, većina istraživanja koristila je različite skupove podataka od kojih svaki ima svoje specifične izazove. U ovom radu predstavljen je novi, vlastiti skup podataka, prikupljen u suradnji s udrugom Oceanus, koji je donio izazov rada s malom količinom podataka i složenim pozadinama na slikama. U radu je razmotreno nekoliko pristupa od tradicionalnih metoda strojnog učenja do naprednih tehnika temeljenih na konvolucijskim neuronskim mrežama (CNN). Vizualizacijom aktivacijskih gradijenata ovih metoda pružen je i uvid u dijelove fotografije koje su imale najviše utjecaja pri detekciji kod svake od metoda.

Ključne riječi: klasifikacija riba, strojno učenje, konvolucijske neuronske mreže

Rad je pohranjen u knjižnici Prirodoslovno-matematičkog fakulteta, Sveučilišta u Splitu

Rad sadrži: 45 stranica, 21 grafički prikaz, 7 tablica i 52 literaturna navoda.
Izvornik je na hrvatskom jeziku.

Mentor: **prof. dr. sc. Saša Mladenović**, redoviti profesor Prirodoslovno-matematičkog fakulteta u Splitu, Sveučilišta u Splitu

Ocjenjivači: **prof. dr. sc. Saša Mladenović**, redoviti profesor Prirodoslovno-matematičkog fakulteta u Splitu, Sveučilišta u Splitu

Antonela Prnjak, asistent Prirodoslovno-matematičkog fakulteta u Splitu, Sveučilišta u Splitu

Nika Jerković, asistent Prirodoslovno-matematičkog fakulteta u Splitu, Sveučilišta u Splitu

Rad prihvaćen: **rujan, 2024**

Basic documentation card

Master Thesis

University of Split

Faculty of Science

Department of Informatics

Ruđera Boškovića 33, 21000 Split, Croatia

CLASSIFICATION OF ADRIATIC SEA FISH USING A CUSTOM DATASET

Petra Dadić

ABSTRACT

Fish classification is a crucial task with numerous applications, ranging from industrial use to biodiversity conservation. Although the topic has been studied for a long time, most research has relied on various datasets, each presenting its own specific challenges. This paper introduces a new, proprietary dataset collected in collaboration with the Oceanus association, which posed the challenge of working with a small dataset and complex backgrounds in the images. Several approaches were considered, from traditional machine learning methods to advanced techniques based on convolutional neural networks (CNNs). Furthermore, by visualizing the activation gradients of these networks, it was possible to gain a better understanding of the parts of the images that had the most significant impact on correct classification for each method.

Key words: Fish Classification, Convolutional Neural Networks (CNNs), Machine Learning
Thesis deposited in library of Faculty of science, University of Split

Thesis consists of: 45 pages, 21 figures, 7 tables and 52 references.
Original language: Croatian.

Mentor: **Saša Mladenović, Ph.D.**, *Full Professor, Faculty of Science and Mathematics, University of Split*

Reviewers: **Saša Mladenović, Ph.D.**, *Full Professor, Faculty of Science and Mathematics, University of Split*

Antonela Prnjak, *Instructor of Faculty of Science, University of Split*

Nika Jerković, *Instructor of Faculty of Science, University of Split*

Thesis accepted: **September, 2024**

IZJAVA

kojom izjavljujem s punom materijalnom i moralnom odgovornošću da sam diplomski rad s naslovom KLASIFIKACIJA RIBA JADRANSKOG MORA NA PRIMJERU VLASTITOG SKUPA PODATAKA izradila samostalno pod voditeljstvom prof. dr. sc. Saša Mladenovića. U radu sam primijenila metodologiju znanstvenoistraživačkog rada i koristila literaturu koja je navedena na kraju diplomskog rada. Tuđe spoznaje, stavove, zaključke, teorije i zakonitosti koje sam izravno ili parafrazirajući navela u diplomskom radu na uobičajen, standardan način citirala sam i povezala s fusnotama s korištenim bibliografskim jedinicama. Rad je pisan u duhu hrvatskog jezika.

Studentica

Petra Dadić



ZAHVALA

Zahvaljujem se svom mentoru prof. dr. sc. Saša Mladenoviću na komentarima, stručnoj pomoći i svim savjetima tijekom izrade ovog diplomskog rada. Također, zahvaljujem svim profesorima i asistentima s PMF-a na prenesenom znanju i podršci tijekom studija.

Hvala mojim prijateljima koji su me kroz sve ove godine motivirali i bili uz mene u svakom trenutku.

Na kraju, zahvaljujem i svojoj obitelji, čija me ljubav i podrška vodila kroz cijelo obrazovanje.

Sadržaj

Uvod	9
1. Pozadina istraživanja	10
1.1. Metodologija.....	10
1.2. PRISMA proces	13
1.3. Zadatak klasifikacije riba.....	13
2. Skupovi podataka	17
2.1. Predobrada podataka.....	18
2.1.1. Uklanjanje pozadine	19
2.1.2. Augmentacije.....	19
2.1.3. Izdvajanje značajki	20
3. Metode klasifikacije riba	22
3.1. Izabrani modeli	23
3.2. Metrike.....	29
4. Eksperiment	30
4.1. Priprema podataka	30
4.2. Treniranje modela.....	33
4.3. Rezultati.....	33
4.3.1. Tradicionalne metode strojnog učenja.....	33
4.3.2. Modeli temeljeni na CNN.....	34
4.4. Uklanjanje pozadine	36
4.5. Tradicionalne metode i segmentacija	39
4.6. Modeli temeljeni na konvolucijskim neuronskim mrežama i segmentacija.....	40
4.7. Tradicionalne metode, segmentacija i ekstrakcija značajki.....	41
4.8. Konačni modeli.....	42
4.9. Predložena primjena modela	43

4.10. Budućnost istraživanja.....	44
Zaključak	45
Literatura	46
Popis slika.....	50
Popis tablica.....	52
Skraćenice.....	53

Uvod

Trenutno u svijetu postoji velika potražnja za ribom i morskim plodovima kao resursom. Velik udio potrošača okreće se ribi zbog njezinih nutritivnih prednosti i važnosti u zdravoj prehrani, a sve veći je i rast skupina koje se odlučuju na pesketarijanstvo, većinski biljnu prehranu u kojoj se, od hrane životinjskog podrijetla, konzumira samo riba.

Velika potražnja, visoka stopa inflacije, visoke cijene i netransparente prakse u ribarskoj industriji ponekad dovode do prevare. Takozvana prevara u ribarstvu (eng. „*Fish fraud*“) je prodaja jeftinije ribe kao da je riječ o skupljij, što može ugroziti zdravlje potrošača kao i ravnotežu pomorskih ekosustava.

U ovom kontekstu, potrebno je osigurati što veću stopu transparentnosti o podrijetlu, način ulova, odnosno uzgoja te utjecaja ribarske industrije na ekosustav. Zbog same količine ribe koja se dnevno lovi i isporučuje, napredni sustavi postaju neophodni. Klasifikacija ribe, ovisno o vrsti, veličini i kvaliteti iznimno je važan korak koji osigurava transparentnost, povjerenje, kvalitetu i odgovornost u ribarskoj industriji.

Cilj ovog rada je trenirati pouzdani model klasifikacije Jadranske ribe i usporediti ga sa ostalim trenutno dostupnim modelima i tehnikama klasifikacije Jadranske ribe na skupu podataka riba sa ribarnice, označenom u suradnji sa udrugom Oceanus.

U radu se međusobno uspoređuju tradicionalni pristupi strojnog učenja sa predtreniranim modelima temeljenim na konvolucijskim neuronskim mrežama, istražuje korištenje segmentacije i različitih augmentacija kako bi se smanjio šum u skupu podataka i poboljšala sama klasifikacija. Rad modela temeljenih na konvoluciji se vizualizira, kako bi se bolje razumjelo koji dijelovi slike su najvažniji prilikom donošenja odluke. Naposljetku se analiziraju podatci te se odabiru najbolji modeli za primjenu u poslovnom okruženju.

Osim navedenog u radu je predstavljen i pregled literature, više informacija o samim modelima te primjeni modela strojnog učenja u ribolovnoj industriji.

1. Pozadina istraživanja

Prije samog istraživanja proveden je pregled postojeće literature na citatnim bazama: Web od Science^[1], Scopus^[2], ArXiv^[3] i IEEE Xplore^[4]. Zbog velike okolišne, ekonomske i industrijske važnosti klasifikacije riba pronađeno je nekoliko preglednih članaka na ovu temu, no po načelu neravnosti i relevantnosti, članci „*Recent advances of machine vision technology in fish classification*“^[5] i „*A survey on fish classification*“^[6] su izabrani su kao polazišna točka ovog istraživanja. Članci pokrivaju dotadašnje standarde i najbolje prakse u klasifikaciji i dotiču se područja segmentacije. U samo područje segmentacije se ne ulazi detaljno ali se pokazalo da je segmentacija slike jedan od koraka koji može doprinijeti poboljšanju same detekcije, koja je temelj ovog istraživanja, pa se segmentacija i u ovom radu obrađuje u tome kapacitetu.

1.1. Metodologija

Prilikom pripreme istraživanja, nekoliko citatnih baza je uzeto u obzir :

Sadržaji sa IEEE XPLORE baze preuzeti su 28. kolovoza 2024. godine i pretraženi po sljedećim kriterijima:

Izabrani članci moraju sadržavati izraze riba „*fish*“ i klasifikacija „*classification*“ u sadržaju ili naslovu rada, moraju biti objavljeni nakon 2022. godine, biti na engleskom jeziku i dostupni za pregled.

Pretraga je dala 934 rezultata, nakon filtera preostalo je 48 članaka čiji su naslovi i sadržaji pročitani, 43 članka su isključena na temelju kriterija isključenja i jednom nije bilo moguće pristupiti. Preostala 4 su pročitana u potpunosti.

Sadržaji sa baze ArXiv preuzeti su 28. kolovoza 2024. godine i pretraženi po sljedećim kriterijima:

Izabrani članci moraju sadržavati izraze riba „*fish*“ i klasifikacija „*classification*“ u sadržaju ili naslovu rada, moraju biti objavljeni nakon 2022. godine, biti na engleskom jeziku i dostupni za pregled.

Pretraga je dala 61 rezultat, nakon filtera preostalo je 9 članaka čiji su naslovi i sadržaji pročitani, 8 članaka je isključeno na temelju kriterija isključenja i jednom nije bilo moguće pristupiti. Preostali članak je pročitani u potpunosti.

Sadržaji sa baze Scopus preuzeti su 28. kolovoza 2024. godine i pretraženi po sljedećim kriterijima:

Izabrani članci moraju sadržavati izraze riba izraze riba „*fish*“ i klasifikacija „*classification*“ u sadržaju ili naslovu rada, moraju biti objavljeni nakon 2022. godine, biti na engleskom jeziku i dostupni za pregled, područje je postavljeno na računalnu znanost „*Computer Science* „.

Pretraga je dala 492 rezultat, nakon čega su dodatno isključeni pregledni članci te svi ostali članci koji su u naslovu sadržavali riječi DNA, RNA, genom „*genome*“, ponašanje „*behaviour*“, na taj način je isključen velik dio medicinskih članaka koji nisu povezani s temom i koriste akronim FISH u medicinskom smislu. Na ovaj način izolirano je 90 članaka na čijim su se naslovima i sadržajima primijenili kriteriji isključenja. Od preostalih 17 članaka , 11 ih nije bilo dostupno te su preostalih 6 pročitani u cijelosti.

Posljednja baza iz koje su preuzeti članici bila je Web of Science. Članci su također preuzeti 28. kolovoza 2024. godine, a pretraživanje je provedeno po sljedećim kriterijima:

Izabrani članci moraju sadržavati izraze riba „*fish*“ i klasifikacija „*classification*“ u sadržaju ili naslovu rada, moraju biti objavljeni nakon 2022. godine, biti na engleskom jeziku i dostupni za pregled, područje je postavljeno na računalnu znanost „*Computer Science* „.

Nakon početnog pronalaska 4671 članka, 8 ih je preostalo nakon primjene filtera i njihovi sadržaji i naslovi su pregledani. Svih 8 je prepoznato kao vrijedno za istraživanje, no kako jednom od njih nije bilo moguće pristupiti preostalih 7 je pročitano u cijelosti.

Kriteriji uključivanja i isključivanja samih publikacija navedeni su u Tablica 1.2 Kriteriji isključivanja publikacija i Tablica 1.1 Kriteriji uključivanja publikacija. Najviše publikacija isključeno je zbog korištenja riječi riba „*fish*“ kao akronima koji se koristi u medicini ili zbog toga što je fokus članka bio samo na podvodnoj detekciji riba. Članci koji su osim podvodne detekcije koristili ili razmatrali detekciju izvan vode su uzeti u obzir.

Kriteriji uključivanja	
	Publikacija je na engleskom jeziku
	Javno dostupan sadržaj
	Rad iz područja tehničkih znanosti
	Rad iz područja računarstva

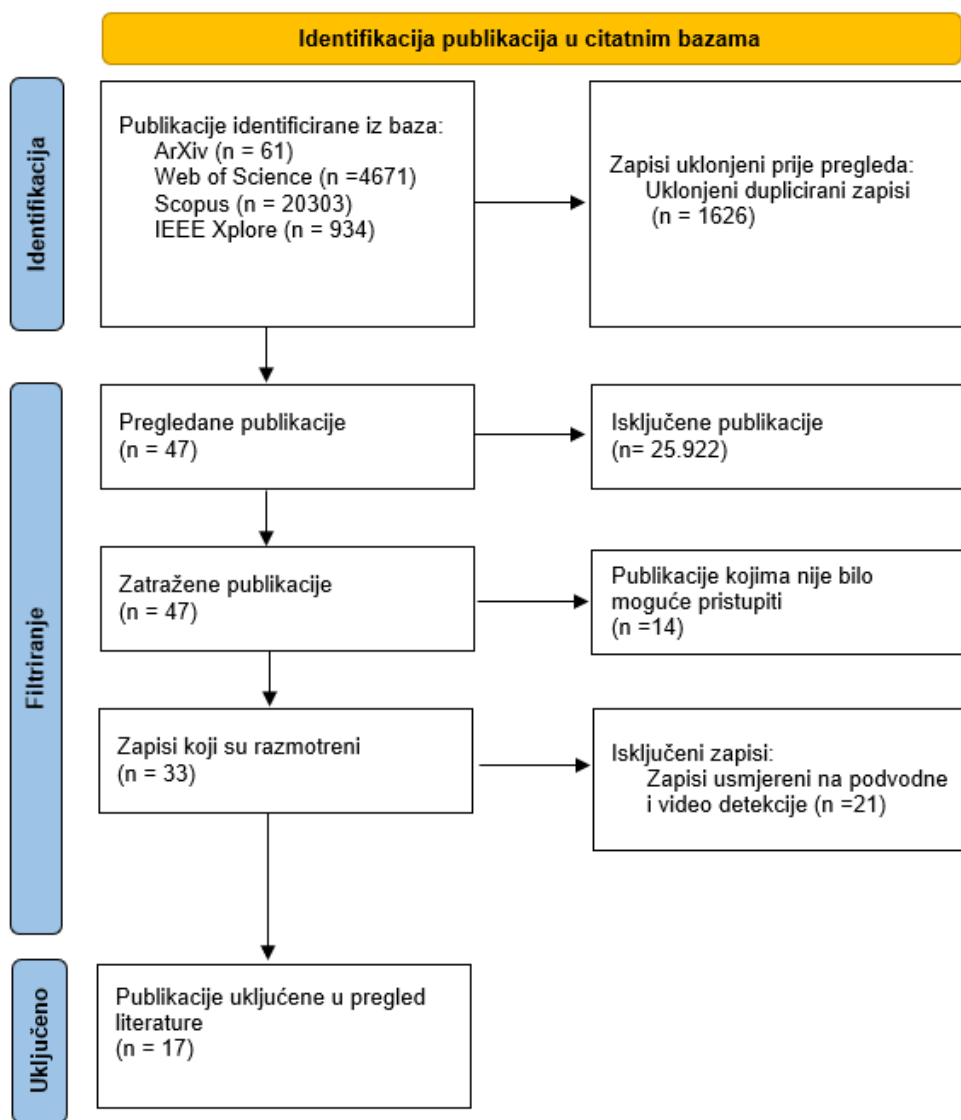
Tablica 1.1 Kriteriji uključivanja publikacija

Kriterij isključivanja	
	Zaključan/nedostupan sadržaj
	Jezik publikacije nije niti engleski niti hrvatski
	Godina objavljivanja je prije 2022
	Medicinski članci koji koriste FISH kao akronim
	Članci koji u naslovu sadrže pojmove DNA, RNA, genom
	Članci koji proučavaju ili identificiraju riblje bolesti
	Članci koji proučavaju ponašanje riba
	Članci koji su usmjereni samo na podvodne fotografije i video zapise

Tablica 1.2 Kriteriji isključivanja publikacija

1.2. PRISMA proces

Na slici 1.1 PRISMA proces prikazan je proces razmatranja, filtriranja i dohvaćanja članaka identificiranih u navedenim citatnim bazama.



Slika 1.1 PRISMA proces

1.3. Zadatak klasifikacije riba

Prepoznavanje riba po vanjskim karakteristikama tradicionalno je vremenski intenzivan posao koji zahtjeva ljudsku intervenciju. Razne grupe ljudi poput znanstvenika koji procjenjuju brojnost i zdravlje ribljih populacija, ribara koji se industrijski bave ulovom i preradom ribe do rekreativaca, imaju potrebu za automatizacijom ovih sustava.

Klasifikacija je jedan od problema nadziranog strojnog učenja koji za cilj ima predvidjeti ispravnu klasu, odnosno oznaku danog ulaza.

Klasifikacija ribe neinvazivnim metodama jedan je od zadataka strojnog učenja koji je dobio na važnosti u novije vrijeme. U ovom poglavlju bit će predstavljene najvažnije metode i dostignuća u rješavanju ovog problema. Ovaj rad se strogo ograničava na klasifikaciju ribe iz fotografija riba izvan mora, isključujući radove koji ribe klasificiraju pomoću video i audio snimaka, kao i one koji se bave specifičnim izazovima povezanim sa podvodnim fotografijama.

Pregledni radovi[5][6] detaljno sumiraju povijest zadatka klasifikacije riba od samih početaka do 2022. godine. U nastavku će biti sumirana njihova najvažnija dostignuća. Nakon sumacije tih pregleda iznijeti će i pregled novih pristupa klasifikaciji razvijenih u razdoblju od 2022. do danas.

Zadatak klasifikacije vrsta ribe od samih početaka suočen je sa mnogim izazovima, najvažniji od kojih su različita orijentacija riba na slikama, razlike između veličine i izgleda pojedinih jedinki unutar vrste, promjene u okolišu, izazovima fotografije u prirodnom okruženju i lošoj kvaliteti slika. Osim različitosti sa jedinkama unutar vrste, jedan od izazova je sličnost sa jedinkama drugih vrsta. Slika 1.2 Jedinka vrste *Diplodus Annularis* (lijevo) i jedinka vrste *Diplodus Vulgaris* (desno) prilično su slične iako prikazuju jedinke različitih vrsta.



Slika 1.2 Jedinka vrste *Diplodus Annularis* (lijevo) i jedinka vrste *Diplodus Vulgaris* (desno)

Već 1994.^[7] napravljeni su prvi pokušaji klasifikacije ali značajna pažnja nije posvećena projektu do 2003.^[8], kada je korištenjem funkcije zakrivljenosti između peraja došlo do značajnog poboljšanja učinkovitosti klasifikacije. Uzevši u obzir značajna poboljšanja klasifikacije postignute korištenjem tehnika ekstrakcije značajki, mnogi istraživači nakon 2003. koriste razne tehnike izdvajanja značajki bazirane na obliku te model izdvajanja

globalnih i lokalnih značajki. Kao što su izdvajanje značajki temeljenih na boji, izdvajanje značajki temeljenih na teksturama i kombinirane metode izdvajanja značajki. Prepoznavanje najboljeg pristupa često ovisi o znanju stručnjaka pa se radi na razvijanju sustava prepoznavanja važnih značajki bez ljudskog utjecaja, najvažnije današnje primjene takvih sustava su korištenjem konvolucijskih neuronskih mreža koje se predtreniraju na velikim skupovima slika i na taj način uče prepoznati relevantne značajke.

Nakon prepoznavanja značajki tradicionalno se koriste metode klasifikacije poput metode naivnog Bayesa, stroja potpornih vektora i jednostavnih neuronskih mreža.

Do značajnog porasta interesa i uspješnosti klasifikacije ribljih vrsta došlo je ranih 2010-ih godina.

Pristupi temeljeni na konvolucijskim neuronskim mrežama, preuzeli su dominaciju u klasifikaciji riba i od tada su najpopularnije tehnike klasifikacije riba. Najutjecajnije pod obitelji konvolucijskih neuronskih mreža korištenih u ovom specifičnom kontekstu su Alex Net, Google LeNet, VGG Net, ResNet, R-CNN, Faster R-CNN i YOLO algoritam.

Alex net je od 2012. godine, jedan od modela koji se najčešće koriste. Koristi Relu aktivacijsku funkciju i sloj sa ispadanjem „*dropout*“ da bi spriječio prenaučenosť. Često uz pomoć augmentacija uvodi nove slike u početni skup podataka. Kombinirani pristupi korištenjem ove obitelji neuronskih mreža u kombinaciji sa SVM modelima za klasifikaciju su imali dobre rezultate u kontroliranim uvjetima^[9] no nisu postigli dobre rezultate na podvodnim slikama. 2014. predložene su izmjenjene modela koje koriste PCA^[10], 2018. godine počinje se koristiti mehanizam pozornosti koji je dodatno pomogao u poboljšanju klasifikacije^[11]. Najnovije metode koriste izmijenjenu mrežu u kombinaciji sa različitim funkcijama aktivacije.

VGGNET^[12] je model koji je poboljšao rezultate klasifikacije produbljivanjem neuronske mreže. Zbog svoje jednostavnosti korištenja i praktičnosti jedan od najpopularnijih i najčešće korištenih pristupa za klasifikaciju ribe. VGG-16 je jedna od implementacija ovog modela pretrenirana na „*Image Net*“^[13] setu podataka. Postiže jako dobre rezultate, no kako je za njegovo korištenje potreban velik set podataka, kod manjih skupova postoji mogućnost pretreniranja modela.

Google LeNet^[14] je osim dubine povećao i širinu neuronske mreže što mu omogućava veći broj parametara. Kako bi se smanjile procesorske potrebe, brža verzija koja optimizira

korištenje resursa - Inception je predložena 2015. godine. Iako je došlo do poboljšanja samih detekcija, prepoznavanje sličnih vrsta je i dalje izazovno za ove modele.

ResNet50^[15] je 2020. godine postigao bolje rezultate klasifikacije riba^[16] u odnosu na VGG-16 i od tada je ResNet obitelj algoritama jedna od najvažnijih u prepoznavanju ribljih vrsta.

YOLO^[17] je obitelj algoritama koja za razliku od ranije spomenutih modela lokalizaciju i detekciju radi u jednom koraku i zbog svoje brzine je popularan za korištenje u modelima koji daju informacije u stvarnom vremenu. Zbog izazova u podvodnoj identifikaciji u novijim radovima istažuju se njegove kombinacije sa drugim neuronskim mrežama^[18]. 2020. godine je korištenjem, CNN-SENet i YOLOv3 postignuta iznimna točnost u klasifikaciji riba.^[19]

Od 2018. godine, zbog količine podataka i resursa potrebnih za treniranje modela, učenje prenošenjem „*transfer learning*“ postaje novi standard klasifikacije ribe. Korištenjem učenja prenošenjem u kombinaciji sa konvolucijskim neuronskim mrežama postignuti su izvrsni rezultati u manje vremena i sa daleko manjim skupovima podataka.

Nedavno su se među State-of-the-art pristupima pojavili i vizualni transformeri^{[20][21]}, koji pokazuju nov i uzbudljiv način razvoja zadatka klasifikacije riba.

2. Skupovi podataka

Definiranje seta podataka prvi je korak klasifikacije. Kako ne postoji standardni podatkovni skup izolirano je nekoliko skupova koji se koriste češće od drugih: Croatian Fish Dataset^[22], Fish4-Knowledge^[23], LifeCLEF14^[24], LifeCLEF15^[25] i FishNet^[26]. U kasnijim radovima često su korišteni i QUT FISH Dataset, The wildfish dataset^[27], A largescale dataset for fish segmentation and classification^[28], Fish Gres^[29], SEAMAPD 21^[30], Wildfish ++, Fish - Pak^[31]. U ovim datasetovima nalaze se različite vrste fotografija i video zapisa. Neki od skupova podataka kao što su A largescale dataset for fish segmentation and classification, QUT FISH Dataset i Fish -Pak sadrže fotografije riba izvan vode, dok su drugi poput Croatian Fish Dataset, Fish4-Knowledge i wildfish skupovi podataka fotografija i video zapisa ispod vode, te nastoje zabilježiti ribe u prirodnom okruženju. U novije vrijeme sve više radova fokusira se na podvodne detekcije zbog jedinstvenih izazova podvodne fotografije i video snimanja, kao što su nepoželjni objekti, loše osvjetljenje i raspoznavanje boje te šumovi. Odabir konačnog skupa podataka ovisi o specifičnoj primjeni rješenja koje se razvija.

Na popisu najutjecajnijih radova na temu klasifikacije riba preglednog članka „*A survey on fish classification techniques*“, od 40 izdvojenih radova čak 19 je koristilo vlastiti skup podataka. U ovom radu će se također koristiti vlastiti skup podataka, koji se sastoji od fotografija riba van vode prikupljenih u suradnji s udrugom Oceanus^[32]. U njemu se nalazi 11 jadranskih vrsta : knez (latinski: *Coris julis*), špar (latinski: *Diplodus annulatus*), fratar (latinski: *Diplodus vulgaris*), batoglavac (latinski: *Pagellus acarne*), arbun (latinski: *Pagellus erythrinus*), pirka (latinski: *Serranus scriba*), komarča (latinski: *Sparus aurata*), gira oštrulja (latinski: *Spicara flexuosa*), lumbrak (latinski: *Symphodus tinca*), pauk bijelac (latinski: *Trachinus draco*) i šaran (latinski: *Trachinus trachinus*). RIBE su fotografirane na splitskoj tržnici. Broj uzoraka između vrsta se razlikuje, a u ovom radu odabran je manji broj svakog podskupa, po 300 fotografija. U nastavku na slici 2.1 Primjeri slika iz skupa podataka prikazan je po jedan primjerak svake riblje vrste uključene u prikupljeni skup podataka.



Slika 2.1 Primjeri slika iz skupa podataka

2.1. Predobrada podataka

Predobrada je skup koraka koji se rade kako bi se fotografije ili videozapisi pripremili za daljnju obradu.

Najčešći primjeri predobrade su izmjena kontrasta, boje, osvjetljenja, skaliranja, izrezivanja dijelova slike, eliminiranja buke i uklanjanja pozadine. Jedna od najčešćih tehnika predobrade je izmjena kanala boje slike iz raznobojne u sive nijanse. Ovi koraci pomažu ujednačiti podatke i ubrzati treniranje modela strojnog učenja.

2.1.1. Uklanjanje pozadine

Uklanjanje pozadine korak je predobrade koji se ovisno o kompleksnosti razdvajanja pozadine i objekta može odraditi na više načina.

Detekcija kontura, jednostavna je metoda za izdvajanje pozadine. Prvo sliku prebacuje u sive tonove a zatim je filtrira Gausovim filtrom kako bi smanjila šum. Gaussov filter zamućuje sliku na kontroliran način kako bi po razlici u pikselima prepoznala rubove objekta. Na taj način uklanja manje šumove a naglašava nagle prijelaze poput rubova. Nakon što prepozna rubove kreira masku.

Malo kompleksnija metoda je korištenje GrabCut^[33] algoritma. GrabCut algoritam prepoznaje objekt unutar pravokutnog okvira. Sve unutar okvira se smatra potencijalnim objektom, dok se sve van okvira tretira kao pozadina. Zatim se kreira maska koja određuje piksele koje pripadaju, odnosno ne pripadaju objektu. Primjenom modela Gaussove mješavine određuje je vjerojatnost pripadanja objekta slici ili pozadini. Proces se ponavlja dok se ne postigne konvergencija, odnosno dok između iteracija nema značajnih razlika.

Segmentacija slike je proces izdvajanja područja slike kako bi se bolje definiralo područje od interesa. Segmentirana slika će razlikovati piksele koji pripadaju objektu od piksela pozadine. Vrlo je važna kod prepoznavanja više objekata na slici ili slika na kojima se nalaze objekti koji nisu važni za detekciju.

Detectron2^[33] je maskirana R konvolucijska neuronska mreža koja sadrži ekstraktore karakteristika za ekstrakciju značajki iz slike. Predlaže područje interesa koje bi moglo sadržavati objekte na temelju značajki. Nakon toga generira masku koja razdvaja objekt od pozadine. Pretreniran je na skupu podataka COCO^[35].

Segmentacija instance^[36] je nadogradnja tehnike segmentacije koja osim razlikovanja piksela objekta i piksela pozadine razlikuje između piksela različitih objekata na istoj slici.

Proces segmentacije posebno se pokazao važan kod klasifikacije slika sa šumom te primjera klasifikacije koji imaju više objekata.

2.1.2. Augmentacije

Prilikom korištenja modela koji su ranije trenirani na velikim skupovima podataka, prilikom dotreniranja često se koriste i augmentacije.

Augmentacije su postupci kojima se povećava broj zapisa u skupu podataka, tako da se na postojeće slike primjene filteri, rotacija ili druge izmjene. Omogućuju modelima strojnog učenja bolju generalizaciju i ponekad se koriste kao tehnika povećanja uzoraka manjinske klase (rad sa nebalansiranim klasama). U najnovijim radovima na temu klasifikacije ribe gotovo su neizostavni a najčešće tehnike augmentacije su zrcaljenje, translacija, izrezivanje dijelova fotografija, dodavanje šuma, izmjene svjetline i boje, rotacija te uvećavanje nasumičnih dijelova slike.

2.1.3. Izdvajanje značajki

Izdvajanje značajki je posljednji korak koji je razmotren prije same primjene klasifikacijskih modela. Dok se u prošlosti ovaj dio većinom obavljao ručno danas je dio procesa samih algoritama strojnog učenja koji se primjenjuju za klasifikaciju.

Metode izdvajanja značajki korištene u ovom radu su, histogram orijentiranih gradijenata, Fourierova analiza i SIFT algoritam, ovi algoritmi su izabrani su zbog toga što su u oba pregledna članka navedeni kao najkorišteniji algoritmi u klasifikaciji riba.

HOG (Histogram orijentiranih gradijenata) je tehnika za izdvajanje značajki, radi tako što prepoznaje rubove i oblike na slici tako što analizira smjerove gradijenata, tj. promjene svjetline između piksela. Odvija se u nekoliko koraka. Prvo transformira sliku u crno bijele tonove, zatim računa promjene susjednih piksela u horizontalnom i vertikalnom smjeru. Na taj način pronalazi rubove. Zatim sliku dijeli na cjeline i za svaku računa u koliko smjerova postoji promjena. Na kraju u obliku vektora prikazuje prikupljene informacije.

SIFT algoritam (Scale-Invariant Feature Transform) metoda je za izdvajanje značajki koja je popularna zbog svoje otpornosti na promjene veličine, rotaciju i osvjetljenje objekta.

Detektira ključne točke „*keypoints*“ koje se najčešće nalaze na rubovima ili kutovima objekata. Nakon prepoznavanja ključnih točaka korištenjem metoda razlike u gausijanima DoG (Difference of Gaussians) pronade točke koje su najznačajnije, neovisno o veličini. Svakoj ključnoj točki se dodjeljuje orijentacija (smjer) temeljena na gradijentima piksela oko te točke. Oko svake ključne točke se analizira mala regija i računa se histogram gradijenata (tj. promjene u smjeru svjetline piksela). Deskriptor je obično vektor koji opisuje lokalne uzorke u toj regiji. Ovi deskriptori su otporni na promjene u skaliranju, rotaciji i osvjetljenju. Kada su deskriptori stvoreni za sve ključne točke na slici, mogu se usporediti s

deskriptorima iz druge slike. Ključne točke koje imaju slične deskriptore smatraju se odgovarajućima, što omogućuje prepoznavanje istih objekata na različitim slikama.

Fourierova transformacija u obradi slika koristi se za analizu i obradu tekstura, uzoraka i frekvencija unutar slike, što je korisno za prepoznavanje ponavljajućih uzoraka, uklanjanje šuma ili poboljšanje kvalitete slike. Fourierova transformacija pretvara sliku iz prostorne domene u frekvencijsku domenu, gdje svaki piksel predstavlja frekvenciju sadržaja slike umjesto intenziteta piksela. Niske frekvencije predstavljaju spore promjene u intenzitetu. Visoke frekvencije predstavljaju brze promjene, tj. rubove. Prepoznavanjem oblika i tekstura pomoću ove transformacije možemo poboljšati detekciju objekata na slikama.

Predtrenirane konvolucijske neuronske mreže se već prilikom inicijalnog treninga uče prepoznavati važne značajke sa slike pa se ove metode najčešće koriste u kombinaciji sa tradicionalnim algoritmima za klasifikaciju.

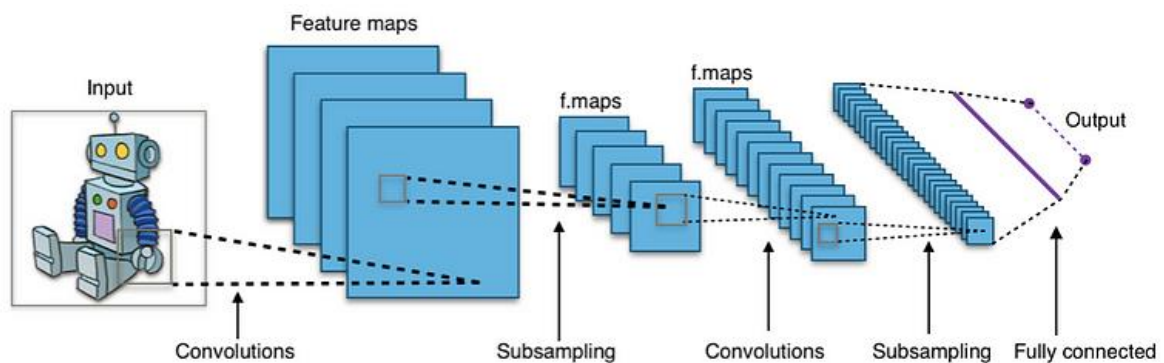
3. Metode klasifikacije riba

Najčešće korištene tradicionalne metode strojnog učenja poput stroja potpornih vektora, metode slučajne šume, k najbližih susjeda i metoda naivnog Bayesa pokazale su dobre rezultate kod detekcije riba u kontroliranim uvjetima. Jednostavne neuronske mreže i Adaboost algoritam zahtijevali su veće skupove podataka i tako postigli još bolji rezultat.

Prilikom zahtjevnije detekcije, pod morem ili sa velikim šumom u podacima, tradicionalne metode ipak nisu bile dovoljne, već se primarno koriste algoritmi temeljeni na konvolucijskim neuronskim mrežama.

Konvolucijske neuronske mreže, CNN su tip duboke neuronske mreže koji primjenom različitih filtera iz slike izvlači značajke poput oblika i rubova, korištenjem aktivacijskih funkcija osiguravaju nelinearnost a korištenjem slojeva za grupiranje smanjuju dimenzionalnost podataka. Pojednostavljeni proces rada konvolucijskih neuronskih mreža prikazan je na slici Slika 3.1 Način rada konvolucijskih neuronskih mreža.

Ulaz u konvolucijsku mrežu je slika koja je u modelu predstavljena mapama značajki koje se iz slike dobivaju korištenjem različitih konvolucija, odnosno filtera. Na ove mape filtera primjenjuju se novi filteri. Potpuno povezani slojevi na kraju pomažu u klasifikaciji i daju konačan izlaz iz modela.



Slika 3.1 Način rada konvolucijskih neuronskih mreža

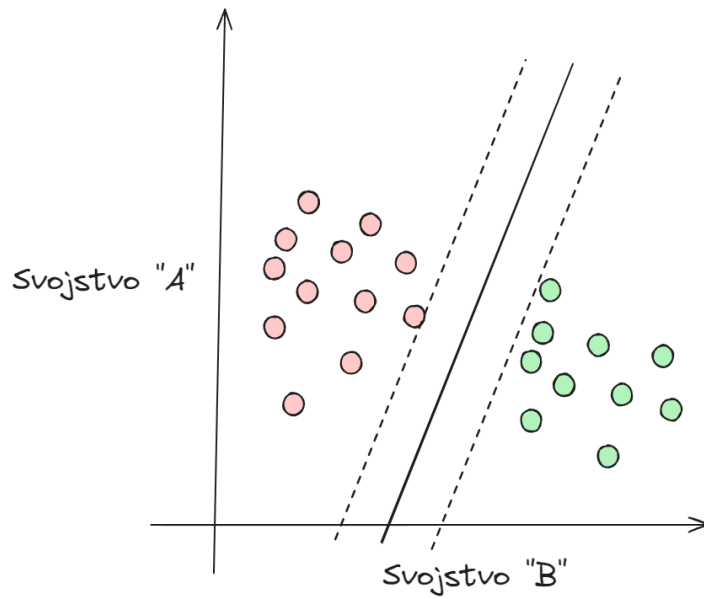
Najutjecajniji poboljšatelji konvolucijskih neuronskih mreža korištenih u ovom specifičnom kontekstu su Alex Net, Google LeNet, VGG Net, ResNet, R-CNN, Faster R-CNN i YOLO algoritam.

3.1. Izabrani modeli

U ovom poglavlju prikazati će se pregled najvažnijih algoritama korištenih u radu. Upravo ovi algoritmi su izabrani zbog njihove korištenosti i uspješnosti u sličnim pod zadacima klasifikacije riba. Na popisu najutjecajnijih članaka iz područja klasifikacije riba preuzetom iz preglednog članka *A survey on fish classification techniques*“, od 39 izdvojenih radova čak ih je deset koristilo stroj potpornih vektora kao klasifikator, četiri su koristila klasifikatore temeljene na stablima, tri su koristila metodu naivnog Bayesa, dva su koristila metodu k-najbližih susjeda te je devet ih je koristilo konvolucijske neuronske mreže. Od preostalih izdvojenih članaka objavljenih nakon 2022. godine, gotovo su svi koristili metode temeljene na konvolucijskim neuronskim mrežama. U ovom radu smo se ograničili na korištenje ResNet, MobileNet i VGG pristupa zbog toga što su jedne od najčešće korištenih metoda ali i zbog ograničenih resursa i nedostatka prikladnih anotacija na podacima.

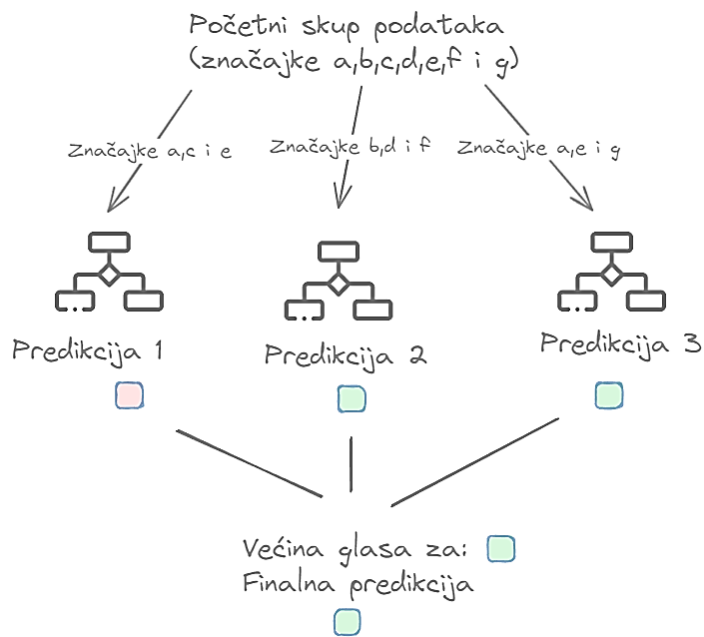
Metoda logističke regresije i metoda višeslojnog perceptrona, izabrane su zbog važnosti u početku razvoja područja klasifikacije riba. U to vrijeme glavni zadatak klasifikacije riba bio je klasifikacija van vode u kontroliranim uvjetima i zbog sličnosti tih uvjeta i slika iz skupa podataka korištenog u ovom radu su i te metode uzete u obzir.

Stroj potpornih vektora (Support Vector Machine, SVM)^[37] nadziran je algoritam strojnog učenja koji klasificira podatke tako da pronalazi višedimenzionalnu ravninu koja maksimizira udaljenost između točaka n-dimenzionalnog prostora. Na slici Slika 3.2 Razdvajanje klasa kod SVM algoritma prikazan je pravac koji odvaja podatke na način da je između njih najveća moguća udaljenost. Kod jednodimenzionalnih podataka, podatci su razdvojeni točkom, kod dvodimenzionalnih podataka pravcem, kod trodimenzionalnih ravninom a kod više dimenzionalnih podataka koriste se hiper-ravnine.



Slika 3.2 Razdvajanje klasa kod SVM algoritma

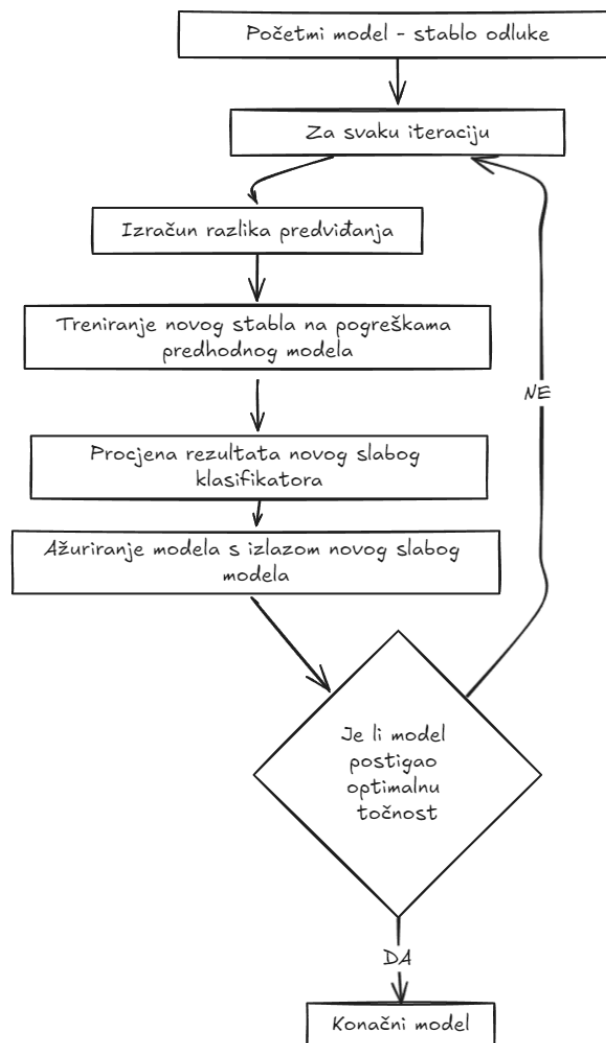
Algoritam slučajne šume^[38] kombinirani je algoritam nadziranog strojnog učenja. Koristi više jednostavnih stabala odluke kojima je na slučajan način određen podskup svojstava na osnovu kojih donose odluku. Svako stablo odluke jednostavno je klasifikator koji pomoću svojstava koja su mu dostupna pokušava naći optimalnu podjelu koja predviđa izlaznu klasu. Ovisno o dubini samog stabla, određen broj podjela izvršava se hijerarhijski prije nego li se uzorku dodijeli klasa. Kada sva stabla dođu do svojih odluka, metodom većinskog glasanja odlučuje se ukupna klasa uzorka.



Slika 3.3 Princip rada algoritma slučajne šume

Shematski prikaz algoritma slučajne šume prikazan je na slici: Slika 3.3 Princip rada algoritma slučajne šume. Početni skup podataka sadrži sedam značajki. Podskup značajki se nasumično dodjeli svakom od n stabala koja se treniraju na tom skupu. Kako ni jedno stablo nema pristup svim značajkama, prilikom predviđanja klase novog primjera, u obzir se uzimaju vrijednosti predikcija svakog od klasifikatora. Na slici su rezultati predviđanja predstavljeni bojom kvadratića. Konačni rezultat jednak je onoj klasi koju je odabrala većina stabala, u ovom slučaju riječ je o zelenom kvadratiću.

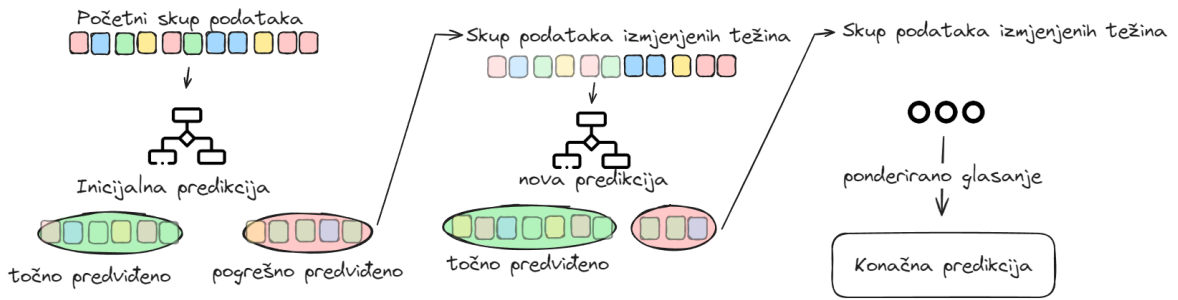
Algoritam Gradient Boosting^[39] također je jedan od kombiniranih metoda nadziranog strojnog učenja. Koristi jednostavna stabla odluke ali na način da pogrešno klasificirane podatke početnih stabala koristi kao ulaz u buduća stabla, pod pretpostavkom da će tako doći do umanjena pogreške klasifikacije. Naziv je dobio po tome što budući klasifikatori svoje predviđanje temelje na promjeni pogreške u prethodnim. Petlja izvršavanja modela prikazan je na slici 3.4 Shematski prikaz rada algoritma Gradient Boosting. Nakon inicijalnog postavljanja slabog klasifikatora, model iterativno na rezidualima uči nove slabe klasifikatore koji svakom iteracijom postaju sve bolji u smanjenju pogreške sve dok ne dođe do konvergencije, točke u kojoj se postigne maksimalno smanjenje pogreške ovog modela.



Slika 3.4 Shematski prikaz rada algoritma Gradient Boosting

Posljednji kombinirani algoritam koji je razmotren u ovom dijelu rada je AdaBoost ^[40], algoritam koji kombinira izlaz više jednostavnih modela i pridodaje im težinsku vrijednost. Teži primjeri dobit će veću važnosti prilikom daljnje klasifikacije i na taj način smanjiti grešku ukupnog sustava. Shematski prikaz rada algoritma prikazan je na slici 3.5 Princip rada algoritma AdaBoost.

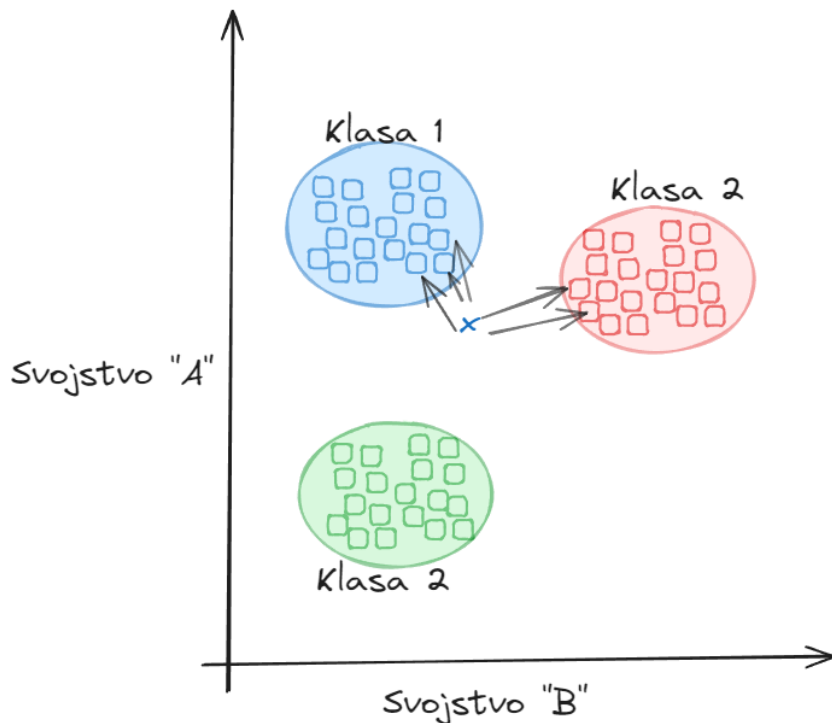
Svaka sljedeća iteracija nastoji ispravno predvidjeti klase početnog skupa dodajući sve veću vrijednost onim primjerima koji se nisu dobro predvidjeli. Nakon predviđenog broja iteracija, postignute željne točnosti ili drugog uvjeta ranog zaustavljanja, predikcije svih klasifikatora uzimaju se u obzir, na način da oni modeli koji su imali veću točnost, imaju veću ulogu prilikom odlučivanja klase uzorka.



Slika 3.5 Princip rada algoritma AdaBoost

Algoritam k najbližih susjeda^[41], algoritam je nadziranog strojnog učenja koji dodjeljuje klasu pojedinom podatku na osnovu njegove blizine drugim označenim primjerima. Parametar k označava koliko se najbližih točaka uzima u obzir prilikom odluke.

Shematski primjer pokazan je na slici: Slika 3.6 Razdvajanje klasa algoritma k - najbližih susjeda. Na slici prikazane su tri klase objekata, označenih različitim bojama. Oznaka „x“ predstavlja novi primjer kojem se određuje klasa. U slučaju na slici, u obzir se uzima 5 najbližih susjeda točke „x“, odnosno primjeri klasa 1 i 2 označeni sivim strelicama. Kako su od pet najbližih susjeda, tri pripadnici klase 1, označene plavom bojom, elementu „x“ predviđa se pripadnost toj klasi.



Slika 3.6 Razdvajanje klasa algoritma k - najbližih susjeda

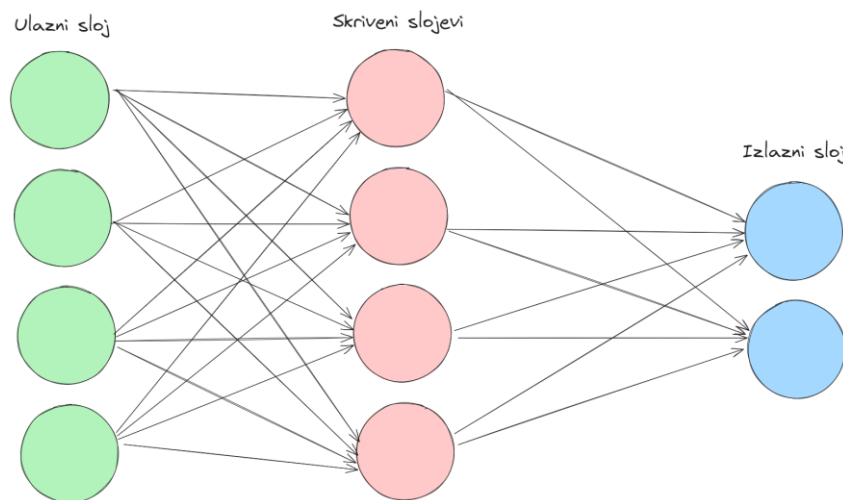
Logistička regresija^[42] algoritam je nadziranog strojnog učenja koji pokušava na osnovu linearnih kombinacija ulaznih značajki predvidjeti kojoj klasi pripada uzorak. Najčešće se koristi kod binarnih klasifikacijskih problema, no moguće ju je koristiti i kod multiklasnog klasifikacijskog problema.

Algoritam naivnog Bayesa^[43] je statistička metoda temeljena na Bayesovoj formuli (1)

$$P(h|D) = \frac{P(D|h)*P(h)}{P(D)} \quad (1)$$

Prvo se izračuna vjerojatnost svake klase, zatim vjerojatnost za svaki atribut da pripada toj klasi. Na temelju formule izračuna vjerojatnost i kao klasu odabire onu koja ima najveću vjerojatnost da joj uzorak pripada.

Višeslojni perceptron^[44] je vrsta unaprijedne neuronske mreže koja se sastoji od slojeva potpuno povezanih neurona. Koristi se za rješavanje problema koji nisu linearno razdvojivi. Sadrži tri osnovne vrste slojeva. Ulazni, koji sadrži čvor za svaku dimenziju ulaznih podataka, skrivene koji nad podacima slijedno vrše razne transformacije i izlazni koji dodjeljuje vjerojatnost pripadanja uzorka svakoj od klasa.



Slika 3.7 Slojevi višeslojnog perceptrona

PCA (Principal Component Analysis)^[45] je tehnika za smanjenje dimenzionalnosti podataka. Pronalazi glavne komponente koje su linearne kombinacije originalnih atributa koje objašnjavaju najveću varijaciju u podacima. Odabirom n takvih komponenti moguće je odraditi klasifikacijski zadatak uz puno manje zauzeće radne memorije.

MobileNet^[46] je obitelj modela strojnog učenja koja je dizajnirana za uređaje s ograničenim resursima. Razvio ih je Google. Temeljene su na konvoluciji koja sadrži različite filtere za svaku komponentu, u slučaju fotografija na svaki kanal zasebno. Pointwise konvolucija

koristi 1*1 konvolucijske filtre za kombiniranje Depthwise Convolution rezultata. Predtreniran je na skupu podataka Image net.

ResNet^[47] je serija modela strojnog učenja koja uz prenošenje jednog sloja u drugi ima i dodatne veze koje mogu preskočiti jedan ili više slojeva u mreži, spojiti sloj sa nekim od slojeva koji su kasnije u mreži. Ovaj sustav omogućava da performanse mreže ne opadaju sa povećanjem broja slojeva. Najčešće korišteni predstavnici ove serije predtrenirani su na skupu podataka Image net.

VGG (Visual Geometry Group)^[48] je arhitektura neuronske mreže koja koristi male konvolucijske filtere kako bi smanjila broj parametara potrebnih za učenje. U zadatku klasifikacije riba često se koristi u kombinaciji sa tradicionalnim algoritmima, preuzimajući dio ekstrakcije značajki.

3.2. Metrike

U klasifikaciji ribe, radovi se najčešće uspoređuju po metrici točnosti.

Točnost je mjera koja odražava koliko je model napravio točnih predikcija. Računa se po formuli (2) i odražava postotkom.

$$\text{Točnost} = \frac{\text{Točna predviđanja}}{\text{Sva predviđanja}} \quad (2)$$

Osim metrike točnosti u novijim radovima su korištene i vizualizacije koje kod konvolucijskih neuronskih mreža pomažu u razumijevanju samog rada mreže. Vizualizacije se kreiraju korištenjem GradCAM^[49] mapiranja (Gradient-weighted Class Activation Mapping).

Grad CAM sliku provodi kroz neuronsku mrežu i izračunava pomoću zadnjeg konvolucijskog sloja koji su dijelovi najviše doprinijeli odluci. Na temelju ovih informacija kreira toplinsku mapu koja vizualno prikazuje važnosti pojedinih dijelova slike.

4. Eksperiment

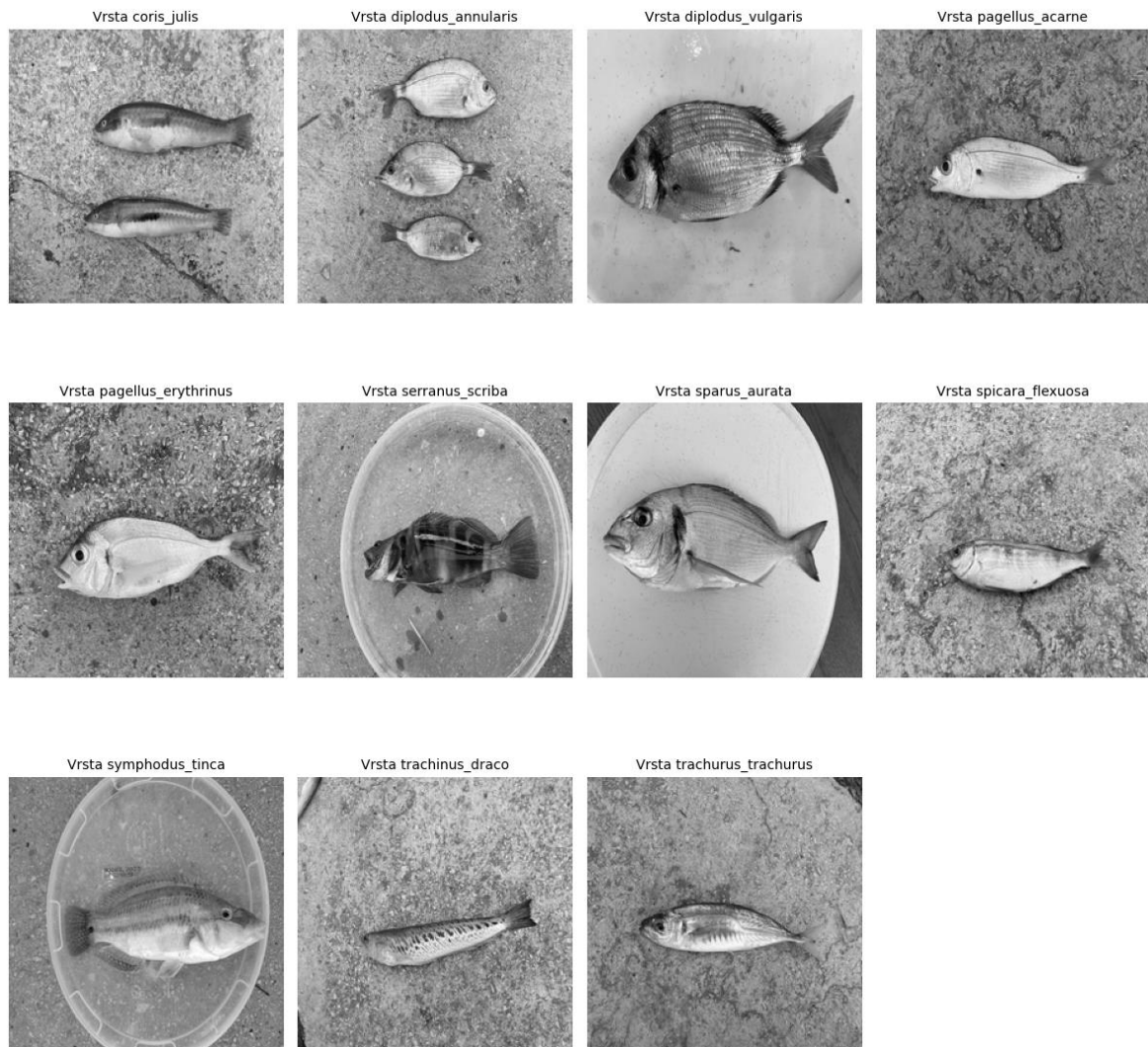
Eksperimentalni dio rada izvršava se treniranjem odabranih modela u Google Colab^[50] okruženju. Besplatna verzija Google Colab okruženja omogućava izvršavanje koda uz korištenje 12,7 GB radnje memorije sustava i 78,2 GB diskovnog prostora. Zbog ovih ograničenja izabrani su modeli koje je moguće pokrenuti na disku, a modeli koji su predtrenirani izvođeni su manji broj epoha.

4.1. Priprema podataka

Jedna od stavki koja se uspoređuje i je način pripreme podataka za sam klasifikacijski algoritam. U radu su iskorištene tri razine pripreme podataka i rezultati svake razine su uspoređeni međusobno kao i sa ukupnim rezultatima eksperimenta.

Inicijalni eksperimenti provode se na skupu na kojem je odrađena transformacija fotografije na dimenzije 224*224 i 3 kanala boje te se pikseli fotografije normaliziraju.

Slike su zatim podijeljene u 2 grupe, na jednu se primjenjuje transformacija u sive tonove, dok druga ostaje u boji. Slika 4.1 Primjeri fotografija nakon prebacivanja u sive tonove prikazuje po jedan primjerak svake vrste nakon ovih transformacija.



Slika 4.1 Primjeri fotografija nakon prebacivanja u sive tonove

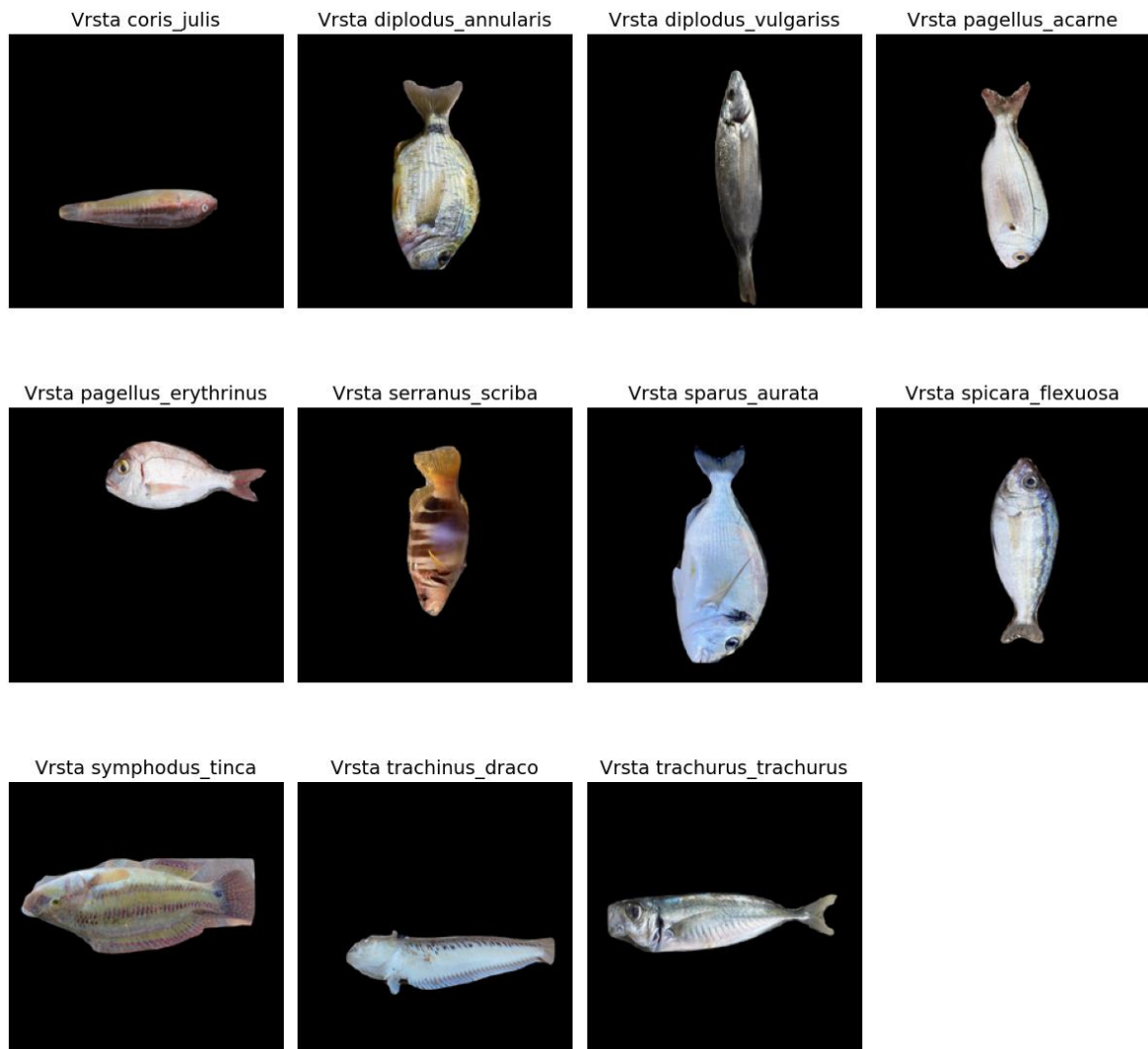
Kako bi se slike mogle koristiti u tradicionalnim metodama strojnog učenja na njih je primijenjena funkcija izravnavanja [51] „*flatten*“ koja je višedimenzionalne podatke pretvorila u niz jednodimenzionalnih podataka.

Na ovim podatcima je zatim primijenjeno smanjenje dimenzionalnosti unosa. Ovaj korak napravljen je zbog ograničenja radne memorije i gdje je bilo moguće provedeni su algoritmi sa i bez korištenja ovog koraka.

Unos u predtrenirane algoritme malo je drukčiji, na njemu se ne primjenjuju transformacije slike u sive tonove kao ni funkcija ravnjanja i smanjenje dimenzionalnosti, kako bi se mreži dalo unos koji je očekivan. Slike dimenzije 224*224, sa 3 kanala boje.

Sljedeća razina obrade podataka bilo je uklanjanje pozadine pristupima detekcije kontura, GrabCat algoritmom te pretreniranim modelom za segmentaciju. Rezultirajuće segmentirane slike prikazane na slici Slike nakon segmentacija zatim su ručno filtrirane kako bi se model

trenirao samo na podacima koji su ispravno segmentirani. Preostali podatci zatim su obrađeni kao i u prethodnoj fazi.



Slika 4.2 Slike nakon segmentacija

Posljednji pristup obrade uz sve navedeno još koristi i algoritme za detekciju značajki. Algoritmi za izdvajanje značajki iskorišteni su sa svim pristupima tradicionalnog strojnog učenja. Nisu korišteni u kombinaciji sa predtreniranim modelima zbog toga što oni u sklopu treninga već uče koje značajke su najbitnije te bi dodatna obrada ovom vrstom algoritama mogla poremetiti njihove rezultate.

4.2. Treniranje modela

Kako bi se modeli istrenirali i usporedili, slike su odijeljene u skupove za trening i testiranje kod klasičnih metoda a kod metoda temeljenih na CNN na trening i validaciju.

U oba slučaja za trening je odabrana većina početnog skupa, a za evaluaciju rezultata se koristi preostalih 20% početnih podataka.

Za modele bazirane na CNN-ovima uz to bilo je potrebno učitati i početne težine, dobijene na treniranju na ImageNet skupu podataka. Modeli su zatim dotrenirani sa po 5 epoha. Ovaj broj epoha pokazao se sasvim dovoljan za dobivanje zadovoljavajućih rezultata detekcije, a zbog manjeg programskog i vremenskog izvršavanja prikladan je i za pokretanje u Colab okruženju.

4.3. Rezultati

4.3.1. Tradicionalne metode strojnog učenja

Treniranjem modela korištenjem tradicionalnih metoda strojnog učenja isprva pokazivalo je dobre rezultate na oba skupa slika, u boji i u nijansama sive, no procjenom rezultata na skupu crno bijelih slika kojima je uklonjena pozadina pokazano je da se detekcija većinski formirala na osnovi pozadine i da su se modeli prenaučili na skup podataka. Skup podataka sa uklonjenom pozadinom u ovom slučaju pomogao je razumjeti kako su ovi podatci formirali odluku, nešto što je ponekad teško razumjeti kod ovakvih modela.

Prenaučenosť je pojava kada se osim na sam objekt klasifikacijski model nauči na podatke koji nisu karakteristični za objekt klasifikacije nego sam skup podataka, u ovom slučaju pozadinu fotografije ribe.

Iz tog razloga niti jedan od modela ne bi bio dobar za daljnju upotrebu. U tablici 4.1 Rezultati modela tradicionalnih metoda strojnog učenja treniranih na slikama s pozadinom prikazani su rezultati klasifikacije na podacima sa uključenom pozadinom te uspješnost istih modela u klasifikaciji slika bez pozadine. Vidljivo je da sama bolja fotografije nije imala velikog utjecaja na klasifikaciju i da su sve predikcije rađene na osnovu pozadine.

Ovi rezultati upućuju na to da su ove uspješnosti varljive i da je potrebno primijeniti daljnje korake predobrade kako bi se dobili bolji rezultati na ovom skupu podataka.

Model	Točnost na skupu sa pozadinama u boji	Točnost na skupu sa pozadinama crno-bijelih slika	Točnost na testnom skupu bez pozadine
Slučajna šuma	96,23%	96,1%	8%
K najbližih susjeda	95%	93,8%	14,4%
Gradient Boosting	93,6%	93,5%	13,1%
Logistička regresija	93,6%	91,3%	6,06%
Stroj potpornih vektora	89,16%	87,88%	10%
Višeslojni perceptron	96,23%	96,01%	10,4%
Metoda naivnog Bayesa	75%	73,6%	10,4%
Ada Boost	21%	21%	11,5%

Tablica 4.1 Rezultati modela tradicionalnih metoda strojnog učenja treniranih na slikama s pozadinom

4.3.2. Modeli temeljeni na CNN

Rezultati modela temeljenih na CNN dotreniranih na slikama sa pozadinom bili su slični rezultatima kod tradicionalnih metoda, unatoč dobrim metrikama na skupu podataka,

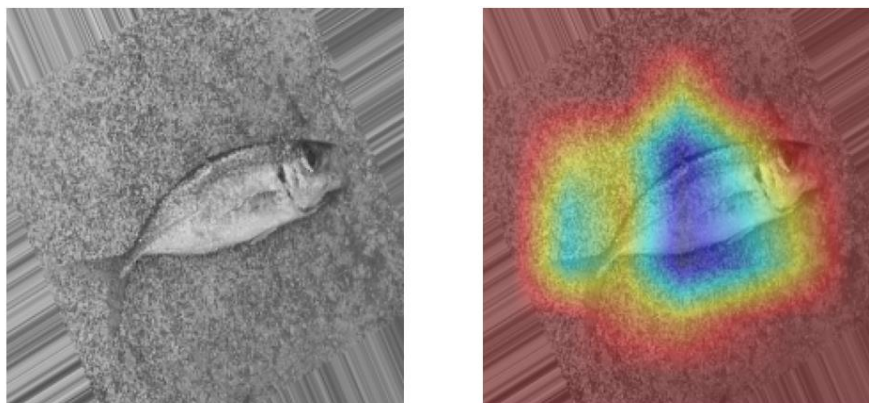
vizualni prikazi pokazali su da se detekcija odvijala samo na osnovu pozadina a ne samog objekta.

Na slikama Slika 4.3 , Slika 4.4 i Slika 4.5 vidimo izlaze različitih modela. Svima je zajedničko da se sam objekt nalazi na području najslabije aktivacije, odnosno da je najmanje bitan za detekciju. Područje najveće aktivacije je područje pozadine.

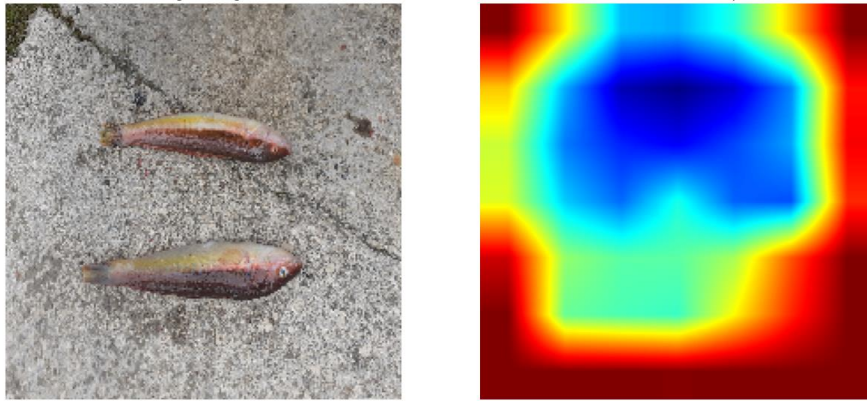
Kod sva tri modela došlo je do pojave ovih rezultata što kao i kod tradicionalnih metoda upućuje da su predikcije napravljene na temelju pozadine a ne samog objekta. Točnosti koje bi u ovom slučaju trebalo uzeti sa zadržkom prikazane su u tablici Tablica 4.2 Rezultati pristupa temeljenih na konvolucijskim neuronskim mrežama na slikama sa pozadinom.

Model	Točnost na skupu slika sa pozadinom
Mobile Net	98,75%
Res Net	75,65%
VGG	99,7%

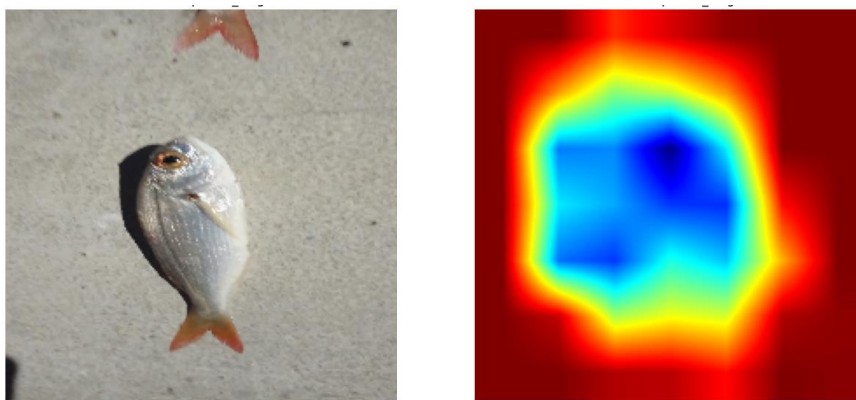
Tablica 4.2 Rezultati pristupa temeljenih na konvolucijskim neuronskim mrežama na slikama sa pozadinom



Slika 4.3 Slika nakon predobrade(lijevo) i aktivacijska toplinska mapa modela ResNet(desno)



Slika 4.4 Ulazna fotografija (lijevo) i rezultirajuća aktivacijska toplinska mapa dobivena korištenjem modela MobileNet(desno)

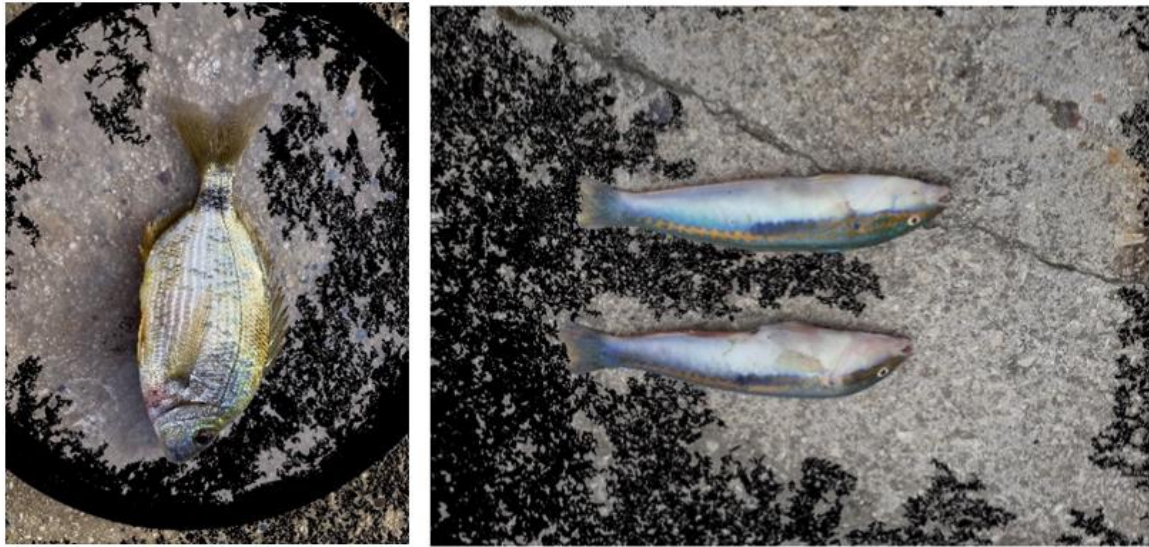


Slika 4.5 Ulazna fotografija (lijevo) i rezultirajuća aktivacijska toplinska mapa dobivena korištenjem modela VGG(desno)

4.4. Uklanjanje pozadine

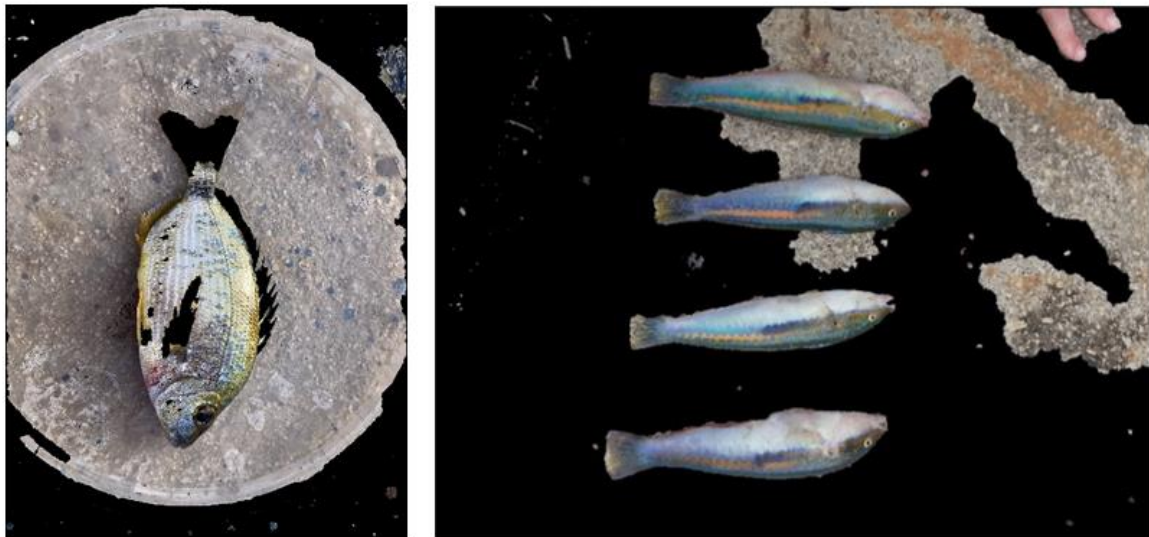
Uvidom u rezultate prethodnog testiranja, potvrđeno je da je sa slika potrebno ukloniti pozadinu kako bi se algoritam usmjerio na prepoznavanje same ribe. Kako bi se pozadina uklonila sa slika primijenjena su tri ranije spomenuta algoritma i uspoređeni su njihovi rezultati.

Najlošije rezultate postigao je pristup detekcije kontura. Zbog kompleksnosti piksela na pozadinama fotografije pozadina na većini slika nije ispravno uklonjena. Primjeri su vidljivi na slici Slika 4.6 Primjer uklanjanja pozadine detekcijom kontura. Zbog velike razlike boje piksela pozadine je došlo do detekcije potpuno pogrešnog područja kao pozadine, i većina pozadine nije uspješno uklonjena.



Slika 4.6 Primjer uklanjanja pozadine detekcijom kontura

Malo bolje rezultate postigao je GrabCut algoritam. Uspješno je detektirao neke dijelove pozadine no na većini fotografija pozadina nije uklonjena u potpunosti. Slika 4.7 GrabCut algoritam za uklanjanje pozadine pokazuje kako su na određenim fotografijama uklonjeni i dijelovi riba koji nisu trebali biti uklonjeni i zato ova metoda nije izabrana za konačni proces obrade slika.



Slika 4.7 GrabCut algoritam za uklanjanje pozadine

Preostala metoda korištenjem predtreniranog segmentacijskog modela, Detectron2 dala je najbolje rezultate, te je pozadina na većini slika ispravno uklonjena. Slika 4.8 Uklanjanje pozadine korištenjem modela Detectron2 prikazuje neke od rezultata iz kojih je vidljivo da

je skoro na svim slikama rub objekta, ribe, dobro prepoznat i da nisu obrisani dijelovi unutar same ribe kao kod korištenja GrabCut algoritma.



Slika 4.8 Uklanjanje pozadine korištenjem modela Detectron2

Iz ranije navedenih razloga, ova metoda uklanjanja pozadine korištena je kao polazna za treniranje modela a slike na kojima segmentacija nije bila uspješna kao na primjerima, ručno su uklonjene iz skupa. Slike na kojima uklanjanje pozadine nije bilo uspješno bilo je lako identificirati jer im za razliku od ostalih algoritama, skoro niti jedan dio pozadine nije bio prepoznat. Slike na kojima je uklonjen dio ribe bile su rjeđe ali se ovaj tip pogreške dogodio u nekoliko slučajeva kao što je vidljivo na slici Slika 4.9 Neuspješno uklanjanje pozadine korištenjem modela Detectron2.



Slika 4.9 Neuspješno uklanjanje pozadine korištenjem modela Detectron2

4.5. Tradicionalne metode i segmentacija

Primjenom segmentacije osigurano je da predikciju nije moguće napraviti po pozadini već samo po samom objektu. Ponovnim treniranjem modela na ovako postavljenom skupu podataka, rezultati su više utemeljeni na samim ribama. Tablica 4.3 Točnost tradicionalnih metoda na skupu bez pozadine pokazuje kako iako su rezultati naizgled lošiji nego u inicijalnom eksperimentu da je već uz pomoć segmentacije moguće postići prilično dobre rezultate tradicionalnim metodama.

Model	Točnost na skupu bez pozadine
Slučajna šuma	88,06%
K najbližih susjeda	83,6%
Gradient Boosting	82,8%.
Logistička regresija	79,7%.
Stroj potpornih vektora	85,3%
Višeslojni perceptron	90,4%
Metoda naivnog Bayesa	68%
AdaBoost	30%

Tablica 4.3 Točnost tradicionalnih metoda na skupu bez pozadine

S rezultatima od 90,4% i 88,04% klasifikacija metodom višeslojnog perceptrona i metoda slučajne šume pokazali su prilično dobre rezultate i dokazali kako su ovi algoritmi dobra alternativa u slučajevima kada zbog veličine skupa podataka i kompleksnosti metoda temeljenih na CNN-ovima njih nije moguće koristiti. Lošije performanse ostalih algoritama ukazuju na to da skup podataka nije dovoljno velik ili da je potrebno implementirati dodatne metode za izvlačenje značajki koje bi pomogle u samoj klasifikaciji. No svakako je pokazano kako i ove metode i dalje imaju prilično važnu ulogu u zadatku klasifikacije ribe.

4.6. Modeli temeljeni na konvolucijskim neuronskim mrežama i segmentacija

Nakon prilično dobrih rezultata korištenjem tradicionalnih metoda, očekivani su još bolji rezultati korištenjem modela temeljenih na konvolucijskim neuronskim mrežama.

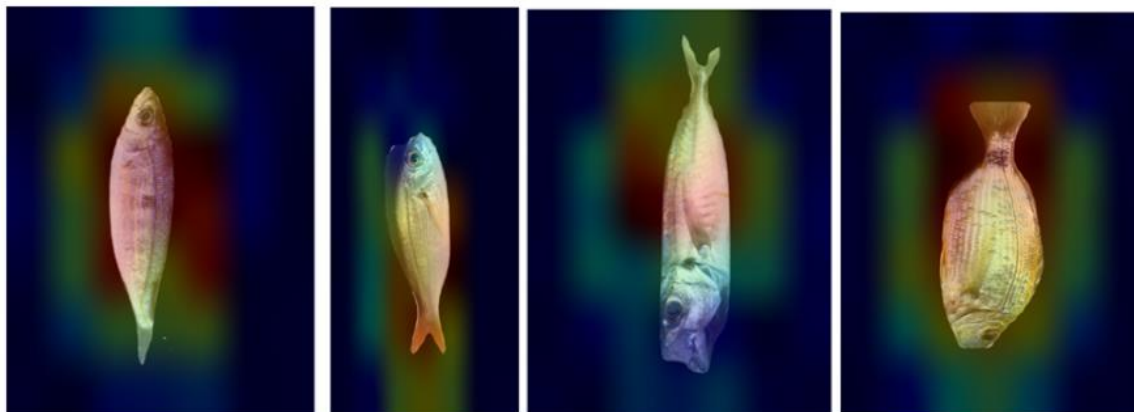
Model	Točnost na skupu segmentiranih slika
Mobile Net	95,3%
Res Net	10,8%
VGG	98,4%

Tablica 4.4 Modeli temeljeni na konvolucijskim neuronskim mrežama trenirani na segmentiranim slikama

Sve metode temeljene na konvolucijskim neuronskim mrežama, osim korištenja ResNet-a postigle su zadovoljavajuće rezultate što je vidljivo i u tablici: Tablica 4.4 Modeli temeljeni na konvolucijskim neuronskim mrežama trenirani na segmentiranim slikama. Res Net unatoč raznim podešavanjima nije uspio generalizirati klase i unatoč dobrim rezultatima na skupu za treniranje nije uspio točno predvidjeti klase skupa za test. Ova metoda bolje bi odgovarala za korištenje na većim skupovima podataka.

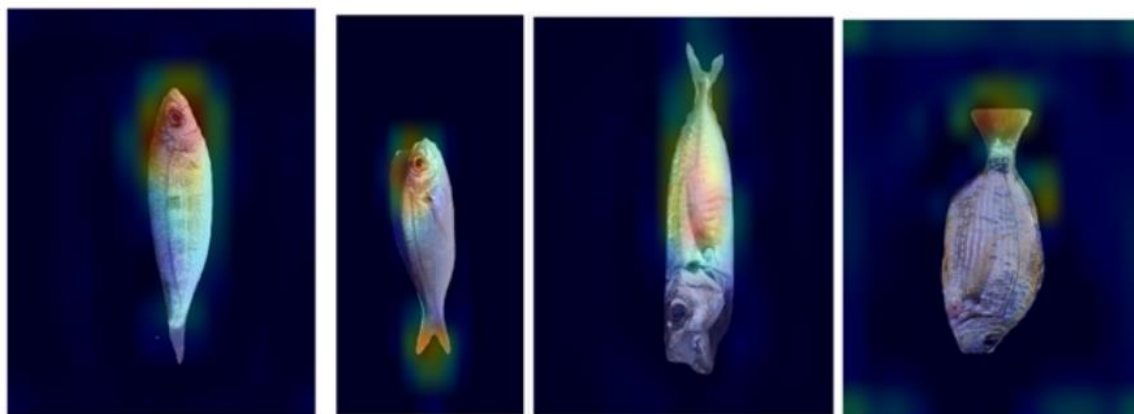
Preostale metode su ostvarile iznimno dobro prepoznavanje već nakon 5 epoha.

Slika 4.10 Prikaz toplinske mape aktivacije prilikom klasifikacije modela MobileNet na nekoliko primjera iz testnog skupa prikazuje najaktivnije regije na osnovu kojih se dodjeljuje klasa. Vidljivo je da je to kod većine riba područje trbuha i u manjoj količini područje repne peraje.



Slika 4.10 Prikaz toplinske mape aktivacije prilikom klasifikacije modela MobileNet

VGG je s druge strane više pozornosti pridijelio glavama i repovima što je vidljivo na Slika 4.11 Prikaz toplinske mape aktivacije prilikom klasifikacije modela VGG, na kojima su prikazane detekcije za nekoliko primjera iz testnog skupa.



Slika 4.11 Prikaz toplinske mape aktivacije prilikom klasifikacije modela VGG

Unatoč fokusu na različite dijelove riba, oba modela su postigla iznimnu točnost.

4.7. Tradicionalne metode, segmentacija i ekstrakcija značajki

Posljednja šansa da se rezultat korištenjem tradicionalnih metoda strojnog učenja približi točnosti postignutoj korištenjem konvolucijskih neuronskih mreža je korištenjem dodatnih algoritama za izdvajanje značajki. Na taj način klasifikacijski algoritam može bolje poopćiti sliku i napraviti predikcije. U tablici Tablica 4.5 Rezultati modela s izdvajanjem značajki prikazani su svi rezultati svih korištenih klasifikatora i metoda izdvajanja značajki. Modeli su postigli različite rezultate korištenjem različitih metoda izdvajanja značajki i svaki model

je imao metodu koja je za njega najbolje funkcionirala. Sveukupno najbolji rezultat postignut je korištenjem Fourierove transformacije i modela k najbližih susjeda. Vrlo blizak rezultat postigao je i stroj potpornih vektora korištenjem histograma orijentiranih gradijenata. Višeslojni perceptron, koji je prije korištenja izdvajanja značajki bio najbolji, imao je nešto lošiju točnost u ovom testiranju. Kontrolna skupina bez ekstrakcije značajki u ovom slučaju bila je u tonovima sive, kako modeli za izdvajanje značajki svakako zahtijevaju prebacivanje slike u sive tonove. Najlošije rezultate su imale metoda naivnog Bayesa i AdaBoost algoritam kojima je ovaj skup podataka najvjerojatnije bio premalen za točnije predikcije.

Model	MLP	SVM	KNN	NB	AdaBoost
HOG	89%	91,1%	81,1%	28,1%	25,6%
SIFT	88%	86,3%	87,7%	64,4%	25,1%
Fourier	86,4%	86,8%	93,1%	47,3%	24,1%
Bez ekstrakcije značajki	88,1%	85,2%	83,2 %	21.5%	25,2%
HOG i PCA	89,7%	91,1%	83,1%	44,5%	28,3%
SIFT i PCA	90,5%	86,4%	86,8%	71,4%	24,1%
Fourierova transformacija i PCA	90,3%	87,3%	93,4%	73,66%	16,2%

Tablica 4.5 Rezultati modela s izdvajanjem značajki

4.8. Konačni modeli

Nakon što su testirani mnogi modeli i pristupi klasifikaciji kao najbolji su odabrana dva.

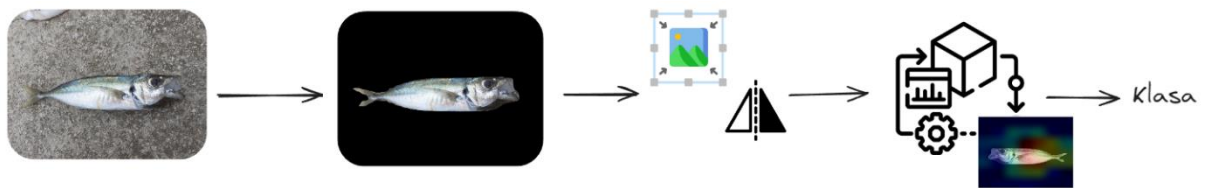
Iako su korištenjem izdvajanja značajki stroj potpornih vektora i metoda k-najbližih susjeda došle jako blizu rezultata MobileNet-a i VGG modela, oni su ipak bili uspješniji i jednostavniji za korištenje. Oba pristupa bila su jako blizu po uspješnosti a kako sama klasifikacija riba ima više primjena, svaki je odabran kao najbolji za jednu od njih. Pristup temeljen na Mobile Net-u pokazao se kao prilično uspješan i uzevši u obzir njegovu kompaktnost i brzinu izvršavanja, odabran je kao najbolji za mobilnu aplikaciju.

Model temeljen na VGG odabran je kao bolji u industrijskoj primjeni klasifikacije zbog veće točnosti.

Kako bi se od testiranih modela dobili konačni sa još boljom mogućnosti detekcije potrebno je dotrenirati na konačnom skupu podataka.

Konačni skup za treniranje najboljih modela formiran je iz početnog cjelovitog skupa. Slike se učitaju i primjenjuje im se segmentacijska maska. Slike se zatim ručno pregledaju i iz skupa podataka se uklanjaju sve one koje nisu ispravno segmentirane. Slike zatim prolaze proces predobrade tako što im se dimenzije podese na zadane i vrijednosti piksela normaliziraju. Zatim slike prolazi proces augmentacija koje ju na slučajan način okreću, prevrću, uvećavaju, zrcale i mijenjaju vrijednost svjetline. Ovako se proširuje skup podataka i osigurava što veća mogućnost generalizacije.

Slika 4.12 Proces treninga konačnih modela razlikuje se od dosadašnjeg treninga jer koristi cijeli skup i samo najbolje metode. Kako bi modeli bili spremni za korištenje potrebno je samo nakon treninga na cijelom predobrađenom skupu spremni njihove težine koje će se u budućnosti koristiti za podstavljanje modela.



Slika 4.12 Proces treninga konačnih modela

4.9. Predložena primjena modela

Nakon pronalaska najboljeg modela važno je razmotriti i njegovu mogućnost primjene. Kako je ranije navedeno postoji više načina korištenja ovakvih modela, kao na primjer za korištenje u ribolovnoj industriji te za prepoznavanje ribe u rekreativne svrhe.

Udruga Oceanus bavi se razvojem aplikacije SeaID u svrhu brze klasifikacije ulovljene ribe za rekreativne ribolovce. Aplikacija će uz hrvatski i latinski naziv prikazivati i minimalnu lovnu duljinu, status ugroženosti i druge informacije za prepoznatu ribu. Ukoliko je vrsta manja od zakonski propisane minimalne duljine ili u lovostaju, aplikacija će povratno upozoriti korisnika i obavijestiti ga da vrstu vrati u more. U sklopu ovog projekta sakupljen je i podatkovni skup korišten u ovom radu. Uloga treniranog modela je sama detekcija ribe,

a predviđena klasa će se zatim koristiti kao ključ pomoću kojeg će se dohvatiti ostali podatci o ulovljenoj ribi. Za ovakav projekt već je pokazan velik interes i odličan je način da se zadatak i primjena klasifikacije ribe pomoću metoda strojnog učenja približi široj javnosti.

Druga mogućnost primjene je u industriji. U pogonima za razvrstavanje ribe postoji velika potreba da se velik ulog brzo i precizno razvrsta. U dijelu pogona to se i dalje radi ručno. Modeli poput opisanih u radu imaju prilike uvelike ubrzati ovaj proces. Pogon bi samo model trebao fino prilagoditi za vrste ribe čijom preradom se bavi kako bi što bolje naučio razlikovati između tih vrsta specifično. Zatim bi prilično brzo i pouzdano dao rezultate. U ovakvoj primjeni model temeljen na VGGu mogao bi predstavljati i privremeno rješenje dok se ne skupi veći set podataka na kojim bi se istrenirao model ResNet, ali za potrebe potvrde ideje VGG model može vrlo dobro pokazati da postoji mjesto za strojno učenje u ribolovnoj industriji.

4.10. Budućnost istraživanja

Područje klasifikacije ribe široko je i omogućava svijetlu budućnost za radove poput ovoga kao i puno načina za nastavak istraživanja.

U ovom radu pokriven je zadatak klasifikacije i u budućnosti daljnjim radom sa ovim skupom podataka moguće je područje istraživanja proširiti i na segmentacijske tehnike koje se koriste za uklanjanje kompleksnih površina kao i usporedbu modela na računalima sa malo boljim memorijskim i procesorskim mogućnostima.

Postoji i prostor za proširenje samog skupa podataka te ručno dodavanje anotacija za modele poput modela obitelji YOLO koji zahtijevaju i lokaciju objekta.

Istraživanje bi se također moglo proširiti i na ostale skupove podataka, istražiti njihove specifičnosti i formirati na jednom mjestu detaljan pregled najboljih pristupa za svaki podtip zadataka klasifikacije riba, bila ona na procesnoj traci u tvornici ili pod morem.

Zaključak

Prepoznavanje i razlikovanje raznih ribljih vrsta zadatak je strojnog učenja koji će sve više i više dobivati na značenju prolaskom vremena. Njegove mnoge primjene, od onih u svrhe identifikacije ribe koju su ulovili ribolovci amateri preko okolišnih i industrijskih zahtijevaju veliku pažnju istraživača.

Kako do sada ne postoji manji broj skupova podataka kao ni jedno jedinstveno područje interesa unutar klasifikacije riba, radovi poput ovog potrebni su da se razne metode detekcije međusobno usporede.

Pregledom literature kao i eksperimentom, pokazano je kako unatoč postojanju novih metoda temeljenih na konvolucijskim neuronskim mrežama i dalje postoji uloga za metode temeljene na tradicionalnim klasifikacijskim metodama strojnog učenja.

Na ovom specifičnom skupu podataka pokazano je da se za njega najbolji rezultati mogu postići dotreniranjem modela temeljenih na neuronskim mrežama uz segmentaciju originalnog skupa slika. Korak uklanjanja pozadine pokazao se ključnim zbog njezinog utjecaja na rezultate detekcije i mogućnosti modela da uoči uzorke koji oku nisu prepoznatljivi.

Vizualizacija samih regija koje su utjecale na predviđanje pomogla je razumijevanju modela i pomaže težnji da se rad umjetne inteligencije u što većoj razini da objasniti.

Literatura

- [1] <https://clarivate.com/>
- [2] <https://www.scopus.com/home.uri>
- [3] <https://arxiv.org/>
- [4] <https://ieeexplore.ieee.org/Xplore/home.jsp>
- [5] Li, D., Wang, Q., Li, X., Niu, M., Wang, H., & Liu, C. (2022). Recent advances of machine vision technology in fish classification. *ICES Journal of Marine Science*, 79(2), 263–284. <https://doi.org/10.1093/icesjms/fsab264>
- [6] Alsmadi, M.K., Almarashdeh, I., A Survey on fish classification Techniques, *Journal of King Saud University - Computer and Information Sciences* (2020), doi: <https://doi.org/10.1016/j.jksuci.2020.07.005>
- [7] Nathalie Castignolles, Michel Cattoen, and M. Larinier "Identification and counting of live fish by image analysis", *Proc. SPIE 2182, Image and Video Processing II*, (23 March 1994); <https://doi.org/10.1117/12.171067>
- [8] D. J. Lee, S. Redd, R. Schoenberger, Xiaoqian Xu and Pengcheng Zhan, "An automated fish species classification and migration monitoring system," *IECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society (IEEE Cat. No.03CH37468)*, Roanoke, VA, USA, 2003, pp. 1080-1085 Vol.2, doi: 10.1109/IECON.2003.1280195
- [9] Yang, X., Zhang, S., Liu, J., Gao, Q., Dong, S., & Zhou, C. (2021). Deep learning for smart fish farming: applications, opportunities and challenges. *Reviews in Aquaculture*, 13(1), 66-90.
- [10] Li, L., & Hong, J. (2014). Identification of fish species based on image processing and statistical analysis research. *2014 IEEE International Conference on Mechatronics and Automation*, 1155-1160.
- [11] Lian, S., Luo, Z., Zhong, Z., Lin, X., Su, S., & Li, S. (2018). Attention guided U-Net for accurate iris segmentation. *Journal of Visual Communication and Image Representation*, 56, 296-304.
- [12] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [13] Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *2009 IEEE conference on computer vision and pattern recognition* (pp. 248–255).
- [14] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Proceedings of the IEEE conference on computer vision and pattern recognition. Going deeper with convolutions, 1-9.
- [15] He K. M., Zhang X. Y., Ren S. Q., Sun J., and IEEE 2016. Deep residual learning for image recognition. In *Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770–778.. IEEE, New York, NY.

- [16] Mathur M., Vasudev D., Sahoo S., Jain D., Goel N. 2020. Crosspooled fishnet: transfer learning based fish species classification model. *Multimedia Tools and Applications*, 79: 31625–31643.
- [17] Redmon J., Divvala S., Girshick R., Farhadi A. 2016. You only look once: unified, real-time object detection. In *Proceedings of the Conference on Computer Vision and Pattern Recognition*, pp. 779–788.
- [18] Jalal A., Salman A., Mian A., Shortis M., Shafait F. 2020. Fish detection and species classification in underwater environments using deep learning with temporal information. *Ecological Informatics*, 57: 101088.
- [19] Knausgård K. M., Wiklund A., Sjørdalen T. K., Halvorsen K. T., Kleiven A. R., Jiao L., Goodwin M. 2021. Temperate fish detection and classification: a deep learning based approach. *Applied Intelligence*. 1–14.
- [20] Veiga, R. J., & Rodrigues, J. M. (2024). Fine-Grained Fish Classification from small to large datasets with Vision Transformers. *IEEE Access*.
- [21] Gong, B., Dai, K., Shao, J., Jing, L., & Chen, Y. (2023). Fish-TViT: A novel fish species classification method in multi water areas based on transfer learning and vision transformer. *Heliyon*, 9(6).
- [22] Jäger, J., Simon, M., Denzler, J., Wolff, V., Fricke-Neuderth, K., & Kruschel, C. (2015). Croatian fish dataset: Fine-grained classification of fish species in their natural habitat. *Swansea: Bmvc*, 2.
- [23] Fisher, R. B., Chen-Burger, Y. H., Giordano, D., Hardman, L., & Lin, F. P. (Eds.). (2016). *Fish4Knowledge: collecting and analyzing massive coral reef fish video data* (Vol. 104, p. 319). Berlin/Heidelberg, Germany: Springer.
- [24] Joly, A. *et al.* (2014). LifeCLEF 2014: Multimedia Life Species Identification Challenges. In: Kanoulas, E., *et al.* Information Access Evaluation. Multilinguality, Multimodality, and Interaction. CLEF 2014. Lecture Notes in Computer Science, vol 8685. Springer, Cham. https://doi.org/10.1007/978-3-319-11382-1_20
- [25] Joly, A. *et al.* (2015). LifeCLEF 2015: Multimedia Life Species Identification Challenges. In: Mothe, J., *et al.* Experimental IR Meets Multilinguality, Multimodality, and Interaction. CLEF 2015. Lecture Notes in Computer Science(), vol 9283. Springer, Cham. https://doi.org/10.1007/978-3-319-24027-5_46
- [26] F. F. Khan, X. Li, A. J. Temple and M. Elhoseiny, "FishNet: A Large-scale Dataset and Benchmark for Fish Recognition, Detection, and Functional Trait Prediction," 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France, 2023, pp. 20439-20449, doi: 10.1109/ICCV51070.2023.01874.
- [27] Zhuang, P., Wang, Y., & Qiao, Y. (2018). WildFish: A Large Benchmark for Fish Recognition in the Wild. *Proceedings of the 26th ACM international conference on Multimedia*.
- [28] O. Ulucan, D. Karakaya and M. Turkan, "A Large-Scale Dataset for Fish Segmentation and Classification," 2020 Innovations in Intelligent Systems and Applications Conference (ASYU), Istanbul, Turkey, 2020, pp. 1-5, doi: 10.1109/ASYU50717.2020.9259867.
- [29] Prasetyo, Eko; Suciati, Nanik; Fatichah, Chastine (2021), "Fish-gres Dataset for Fish Species Classification", Mendeley Data, V1, doi: 10.17632/76cr3wfhff.1

- [30] Boulais, Océane & Alaba, Simegnew & Ball, John & Campbell, Matthew & Iftekhar, Ahmed Tashfin & Moorehead, Robert & Primrose, James & Prior, Jack & Wallace, Farron & Yu, Henry & Zheng, Aotian. (2021). SEAMAPD21: a large-scale reef fish dataset for fine-grained categorization.
- [31] Francescangeli, M., Marini, S., Martínez, E. *et al.* Image dataset for benchmarking automated fish detection and classification algorithms. *Sci Data* **10**, 5 (2023). <https://doi.org/10.1038/s41597-022-01906->
- [32] <https://more.unist.hr/Za-studente/Udruga-Oceanus>
- [33] Rother, C., Kolmogorov, V., & Blake, A. (2004). " GrabCut" interactive foreground extraction using iterated graph cuts. *ACM transactions on graphics (TOG)*, *23*(3), 309-314.
- [34] [Y. Wu](#), [A. Kirillov](#), [F. Massa](#), [W. Lo](#), and [R. Girshick](#). <https://github.com/facebookresearch/detectron2>, (2019)
- [35] Tsung-Yi Lin, Maire, M., Belongie, S. J., Bourdev, L. D., Girshick, R. B., Hays, J., ... Zitnick, C. L. (2014). Microsoft COCO: Common Objects in Context. *CoRR*, *abs/1405.0312*. Retrieved from <http://arxiv.org/abs/1405.0312>
- [36] Climent-Perez P, Galán-Cuenca A, Garcia-d'Urso NE, Saval-Calvo M, Azorin-Lopez J, Fuster-Guillo A. 2024. Simultaneous, vision-based fish instance segmentation, species classification and size regression. *PeerJ Computer Science* *10*:e1770 <https://doi.org/10.7717/peerj-cs.1770>
- [37] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning*, *20*(3), 273–297.
- [38] Ho, T. K. (1995). Random decision forests. In *Proceedings of 3rd international conference on document analysis and recognition* (Vol. 1, pp. 278–282).
- [39] Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of Statistics*, 1189–1232.
- [40] Schapire, R. E. (2013). Explaining adaboost. In *Empirical inference* (pp. 37–52). Springer.
- [41] Fix, Evelyn; Hodges, Joseph L. (1951). Discriminatory Analysis. Nonparametric Discrimination: Consistency Properties (PDF) (Report). USAF School of Aviation Medicine, Randolph Field, Texas
- [42] Cox, D. R. (1958). The regression analysis of binary sequences. *Journal of the Royal Statistical Society: Series B (Methodological)*, *20*(2), 215–232.
- [43] BRADY, N. A statistical Analysis of Use Case. *Proceedings of the 7th International Conference on Telecommunications ConTEL*, Zagreb, (2003), 45-52.
- [44] Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice Hall PTR.
- [45] Karl Pearson F.R.S. . (1901). LIII. On lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), 559–572. <https://doi.org/10.1080/14786440109462720>
- [46] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... Adam, H. (2017). MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. *CoRR*, *abs/1704.04861*. Retrieved from <http://arxiv.org/abs/1704.04861>

- [47] He, K., Zhang, X., Ren, S., & Sun, J. (2015). Deep Residual Learning for Image Recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 770-778.
- [48] S. Liu and W. Deng, "Very deep convolutional neural network based image classification using small training sample size," 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), Kuala Lumpur, Malaysia, 2015, pp. 730-734, doi: 10.1109/ACPR.2015.7486599
- [49] R. R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh and D. Batra, "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," 2017 IEEE International Conference on Computer Vision (ICCV), Venice, Italy, 2017, pp. 618-626, doi: 10.1109/ICCV.2017.74
- [50] <https://colab.research.google.com/>
- [51] <https://numpy.org/doc/stable/reference/generated/numpy.ndarray.flatten.html>
- [52] <https://excalidraw.com/>

Popis slika

Slika 1.1 PRISMA proces izrađeno prema predlošku <https://www.prisma-statement.org/prisma-2020-flow-diagram>

Slika 1.2 Jedinka vrste Diplodus Annularis (lijevo) i jedinka vrste Diplodus Vulgaris (desno) preuzete iz vlastitog skupa podataka

Slika 2.1 Primjeri slika iz skupa podataka preuzete iz vlastitog skupa podataka, slika dobivena korištenjem python skripte

Slika 3.1 Način rada konvolucijskih neuronskih mreža prema <https://medium.com/@himadrisankarchatterjee/a-basic-introduction-to-convolutional-neural-network-8e39019b27c4>

Slika 3.2 Razdvajanje klasa kod SVM algoritma slika izrađena korištenjem online alata Escalidraw [52]

Slika 3.3 Princip rada algoritma slučajne šume izrađena korištenjem online alata Escalidraw [52]

Slika 3.4 Shematski prikaz rada algoritma Gradient Boosting izrađena korištenjem online alata Escalidraw[52]

Slika 3.5 Princip rada algoritma AdaBoost izrađena korištenjem online alata Escalidraw[52]

Slika 3.6 Razdvajanje klasa algoritma k - najbližih susjeda izrađena korištenjem online alata Escalidraw[52]

Slika 3.7 Slojevi višeslojnog perceptrona izrađena korištenjem online alata Escalidraw[52]

Slika 4.1 Primjeri fotografija nakon prebacivanja u sive tonove slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.2 Slike nakon segmentacija slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.3 Slika nakon predobrade(lijevo) i aktivacijska toplinska mapa modela ResNet(desno) slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.4 Ulazna fotografija (lijevo) i rezultirajuća aktivacijska toplinska mapa dobivena korištenjem modela MobileNet(desno) slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.5 Ulazna fotografija (lijevo) i rezultirajuća aktivacijska toplinska mapa dobivena korištenjem modela VGG(desno) slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.6 Primjer uklanjanja pozadine detekcijom kontura slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.7 GrabCut algoritam za uklanjanje pozadine slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.8 Uklanjanje pozadine korištenjem modela Detectron2 slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.10 Prikaz toplinske mape aktivacije prilikom klasifikacije modela MobileNet slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.11 Prikaz toplinske mape aktivacije prilikom klasifikacije modela VGG slika dobivena obradom fotografija vlastitog skupa podataka

Slika 4.12 Proces treniranja konačnih modela izrađena korištenjem online alata Escalidraw[52] na osnovu procesa opisanog u radu

Popis tablica

Tablica 1.1 Kriteriji uključivanja publikacija

Tablica 1.2 Kriteriji isključivanja publikacija

Tablica 4.1 Rezultati modela tradicionalnih metoda strojnog učenja treniranih na slikama s pozadinom

Tablica 4.2 Rezultati pristupa temeljenih na konvolucijskim neuronskim mrežama na slikama sa pozadinom

Tablica 4.3 Točnost tradicionalnih metoda na skupu bez pozadine

Tablica 4.4 Modeli temeljeni na konvolucijskim neuronskim mrežama trenirani na segmentiranim slikama

Tablica 4.5 Rezultati modela s izdvajanjem značajki

Skraćenice

VGG	<i>Visual Geometry Group</i>	obitelj modela strojnog učenja
R-CNN	<i>Region-based Convolutional Neural Network</i>	konvolucijska neuronska mreža bazirana na regijama, arhitektura dubokog učenja koja se koristi u zadacima računalnog vida
YOLO	<i>You Only Look Once</i>	algoritam za detekciju objekata
CNN	<i>Konvolucijska neuronska mreža</i>	arhitektura neuronskih mreža koja koristi konvolucijske slojeve
SVM	<i>Support Vector Machine</i>	stroj potpornih vektora
PCA	<i>Principal Component Analysis</i>	analiza glavnih komponenti
HOG	<i>Histogram of Oriented Gradients</i>	histogram orjentiranih gradijenata, metoda za izdvajanje značajki
SIFT	<i>Scale Invariant Feature Transform</i>	transformacija značajki neovisna o skaliranju, metoda za izdvajanje značajki
DoG	<i>Difference of Gaussians</i>	razlika u gausijanima, metoda naglašavanja značajki
GradCAM	<i>Gradient-weighted Class Activation Mapping</i>	mapiranje aktivacijskih gradijenata, metoda za vizualizaciju rada konvolucijskih neuronskih mreža
MLP	<i>Multi Layer Perceptron</i>	više slojni perceptron
KNN	<i>K-Nearest Neighbours</i>	k-najbližih susjeda, algoritam strojnog učenja
RF	<i>Random Forest</i>	slučajna šuma, algoritam strojnog učenja
NB	<i>Naive Bayes</i>	algoritam naivnog Bayesa