Automatska detekcija plivačkih tehnika i analiza plivačkih stilova

Šimac, Ivan

Master's thesis / Diplomski rad

2018

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Rijeka / Sveučilište u Rijeci**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:195:270285

Rights / Prava: In copyright / Zaštićeno autorskim pravom.

Download date / Datum preuzimanja: **2024-11-19**

Sveučilište u Rijeci
Fakultet informatike
i digitalnih tehnologija

Repository / Repozitorij:

Repository of the University of Rijeka, Faculty of Informatics and Digital Technologies - INFORI Repository

zir.nsk.hr

DIGITALNA KNJIŽNICA

DIGITALNI AKADEMSKI ARHIVI I REPOZITORIJI

Sveučilište u Rijeci – Odjel za informatiku

Diplomski studij informatike – nastavnički smjer

Ivan Šimac

Automatska detekcija plivačkih tehnika i analiza plivačkih stilova

Diplomski rad

Mentor: doc. dr. sc. Marina Ivašić - Kos

Rijeka, rujan 2018. godine

Sadržaj

1. Sažetak ... 3
2. Uvod ... 4
3. (Umjetna) inteligencija ... 5
3.1. Ekspertni sustavi ... 6
3.2. Rudarenje podataka ... 6
3.3. Umjetna neuronska mreža ... 7
3.4. Strojno učenje i raspoznavanje uzoraka ... 7
3.5. Konvolucijske neuronske mreže ... 8
4. Domena eksperimenta i definicija problemskog zadatka ... 11
4.1. Plivanje ... 11
4.1.1. Tehnika delfin (po novoj nomenklaturi dupin) ... 12
4.1.2. Tehnika leđno ... 13
4.1.3. Tehnika prsno ... 14
4.1.4. Tehnika kraul ... 15
5. Snimanje i prikupljanje uzoraka ... 17
5.1. Tijek snimanja ... 19
5.2. Pozicije kamere ... 20
5.3. Problemi kod snimanja na bazenu ... 22
5.4. Predobrada podataka ... 23
6. YOLO i Mask R-CNN ... 25
7. Analiza rezultata ... 27
8. Daljnje učenje mreže ... 36
9. Zaključak ... 38
10. Literatura ... 39

1. Sažetak

Cilj rada jest predstaviti mogućnosti detekcije akcija odnosno tehnike plivača koji plivaju u vodenom mediju.

U radu je opisan i predstavljen postupak snimanja plivača u bazenu, obrade snimljenog materijala te postupak detekcije i klasifikacije slika uz pomoć konvolucijskih neuronskih mreža.

Istaknuti su i opisani problemi kod detekcije i klasifikacije uslijed neodgovarajućeg položaja kamere i smetnji uzrokovanih vodom.

Predstavljeni su dobiveni rezultati te prijedlog metoda za poboljšanje detekcije i klasifikacije na temelju istog snimljenog materijala.

Ključne riječi: neuronske mreže, konvolucijske neuronske mreže, detekcija, klasifikacija, Yolo, Mask R-CNN, plivanje, tehnika, analiza, stil, akcija, GoPro kamera

2. Uvod

U informatičkoj znanosti, unazad nekoliko godina veliki se značaj pridaje umjetnoj inteligenciji, odnosno želimo postići da računala „razmišljaju“, da uspijevaju riješiti pojedine, da ne kažemo i sve zadatke. Iste one koji su stavljeni pred ljude.

U ljudskom životu, unazad velikog broja godina, veliki se značaj pridaje sportu, rekreaciji, kretanju.

Ovaj će rad spojiti oboje. I informatiku (računalstvo), odnosno polje umjetne inteligencije kao i sport, i to plivanje.

Intencija je autora u radu sastaviti ekspertni sustav koji bi mogao detektirati plivačke tehnike i analizirati plivačke stilove. Pojednostavljeno, želimo napraviti sustav koji je u mogućnosti „vidjeti“ plivača te ga upozoriti na nedostatke u njegovom stilu plivanja. Proces izrade takvog ekspertnog, inteligentnog sustava svakako je složen te će u ovom radu biti prikazan samo dio tog procesa uz naznake za daljnje korake pri izradi potpunog sustava za analizu plivanja.

Iako je riječ o složenom dijelu informatike i računalstva, u radu je cilj sve postupke i elemente predstaviti na jednostavan način kako bi svi zainteresirani čitatelji mogli dobiti uvid u specifičnost dobivenih rezultata te možda i sami nastaviti u istom segmentu razvoja računalnih sustava.

3. (Umjetna) inteligencija

Psiholozi inteligenciju, kao ljudsku karakteristiku, opisuju kao sposobnost snalaženja u novim i nepoznatim situacijama. Nadalje, napominju kako je i ranije stečeno znanje važno za dobivanje rezultata u novoj okolini u kojoj ne pomaže nagonsko ponašanje i rješavanje problema.

Pojam umjetne inteligencije (UI) u računalnu tehnologiju uveo je Marvin Minsky[1] [1] koji se bavio istraživanjem neuronskih mreža te strojnog učenja. Pojam označava sve one radnje koje računalo može izvesti a da su one ranije bile osigurane samo za čovjeka. Pod time ne smatramo matematičko računanje ili izvođenje velikog broja operacija nekog tipa, već sposobnost računala da predvidi neke situacije (recimo naše osjećaje s obzirom na način pisanja, glazbu koju bismo slušali s obzirom na dosad odslušanu), da samostalno vozi automobile, da kreira najbrži put do neke destinacije odnosno da se snađe i djeluje u novim i nepoznatim situacijama.

Veliko je polje primjene računala u svakodnevnom životu i dajemo im sve veću autonomiju. Mnogi se zapravo tog razvoja računala i boje. Neki da će ostati bez posla (što se masovno i događa u jednostavnijim zanimanjima) do toga da bi umjetna inteligencija mogla ovladati i nama ljudima koji je i stvaramo. Ovdje svakako moraju vrijedi i tri zakona robotike (robote smatramo inteligentnima), a koji nam kazuju kako robot nikada ne smije naštetiti čovjeku, te ga slušati, osim u naredbi da čovjeku naškodi [2]. Vojna se industrija orijentirala na umjetnu inteligenciju te ju koristi na razne načine. Ulazak u tu temu previše je filozofske naravi, tako da nećemo filozofski pristupiti upotrebi umjetne inteligencije, već smo samo naveli određene paradigme.

[1] U nekoj literaturi, zasluge za kreiranje pojma UI uz Minskog idu i Johnu McCarhtyu

Nameće se pitanje kako učiniti da računalo, koji mnogi smatraju samo skupom elektroničkih dijelova putem kojih mogu pretraživati internet, zaista nešto i nauči, odnosno da nam podastre svoje „viđenje“ stvari.

Nekoliko je načina, ukratko ćemo ih opisati.

3.1. Ekspertni sustavi

Ekspert je osoba koja suvereno vlada područjem i koja zna „gotovo“ sve iz domene za koju je smatramo ekspertom. Ovdje želimo da računalo radi i „razmišlja“ upravo poput eksperta. Računalu treba baza znanja te domene, pravila koja slijedi kada se pojavi nova informacija te način komunikacije s korisnikom sustava. Većina se ljudskog znanja može predstaviti pravilima (if-then-else). Primjeri takvih ekspertnih sustava su medicinske naravi. Od korisnika se traži da unese simptome bolesti, a računalo na temelju simptoma pretražuje bazu bolesti i javlja korisniku o kojoj bi bolesti moglo biti riječ. Uspješnost samog sustava (računala) omogućava ako su već ti navedeni simptomi upućivali na bolest koja je potvrđena [3].

3.2. Rudarenje podataka

U današnjem svijetu, podataka imamo napretek. Kreiraju se neprestano, svaka naša prijava u neki sustav stvara zapis, svaki naš klik na internet tražilici za nama ostavlja određene podatke. Ukoliko te podatke krenemo obrađivati, možemo dobiti vrlo dobre i valjanje zaključke o interesima pojedinog korisnika. Tipa, redovito kupujete u nekoj trgovini, trgovina vam izdaje njihovu karticu kao vjernom kupcu kako biste ostvarili brojne pogodnosti. Na temelju kupljenih artikala, kroz nekoliko dana na vašu adresu stižu kuponi koji nude povoljniju

cijenu baš onih artikala za koje je došlo vrijeme da ih kupite. I time trgovine osiguravaju da ćete opet upravo kod njih izvršiti kupovinu. Tako radi i popularni YouTube. Pregledom raznih video uradaka, YouTube slaže popis video zapisa koji bi vas mogli zanimati kako biste čim duže ostali upravo na tom servisu odnosno mrežnoj stranici.

Računala imaju sposobnost pohrane velike količine podataka te iz te velike količine podataka mogu pronaći svakakve veze iz čega ova vrsta učenja postaje vrlo popularna.

3.3. Umjetna neuronska mreža

Ljudski živčani sustav građen je od neurona, najmanje građevne jedinice. Ljudski mozak može odraditi više raznovrsnih procesa odjednom uz veliku brzinu obrade. Tu leži ideja u primjeni neuronskih mreža u umjetnoj inteligenciji. Neuronske mreže mogu obrađivati nekoliko procesa podataka, mogu raditi s nepotpunim podacima, stvaraju vlastite odnose među podacima iako oni nisu eksplicitno zadani. Primjenjuju se u domeni raspoznavanja uzoraka, obradi slike, govora te raznih simulacija.

3.4. Strojno učenje i raspoznavanje uzoraka

Raspoznavanje uzoraka jedna je od grana strojnog učenja. Postoje dvije vrste učenja: nadzirano i nenadzirano. Kod nadziranog učenja postoji skup podataka iz kojih stroj (računalo) uči i traži vezu među podacima radi lakšeg zaključivanja na podacima na kojim ćemo testirati. Nenadzirano učenje nema „učitelja“, odnosno nema točnog rješenja i prethodnog svrstavanja u klasu. Ovdje stroj (računalo)

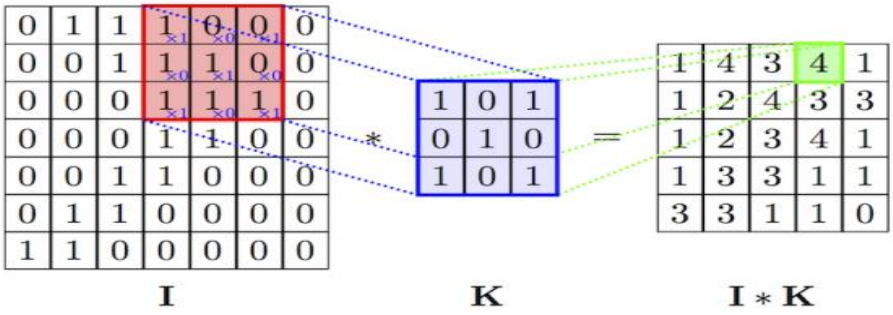
„samo“ organizira podatke na temelju strukture i značajki koje pronalazi u podacima.

3.5. Konvolucijske neuronske mreže

Konvolucijske mreže [4] su tip neuronske mreže se sastoje od jednog ulaznog, jednog izlaznog i jednog ili više skrivenih slojeva od kojih je barem jedan konvolucijski sloj. Osim konvolucijskog sloja za konvolucijske neuronske mreže specifičan je i sloj sažimanja.

Koriste se najčešće za probleme detekcije, raspoznavanja i klasifikacije slikovnih podataka. Zadaci detekcije i raspoznavanja objekata na slikama su posebice složeni zadaci jer čak i isti objekt može biti smješten na različitim pozicijama na sceni, može se preklapati s bilo kojim drugim objektom, može biti snimljen s različitim povećanjem, iz različite perspektive, pod različitim uvjetima osvjetljenja i slično. Iz tog razloga niti pojavljivanje istog objekta na sceni u različitim uvjetima nije jednostavan zadatak raspoznavanja jer se pojavljivanja objekta mogu međusobno razlikovati npr. u poziciji, veličini, boji, obliku i perspektivi.

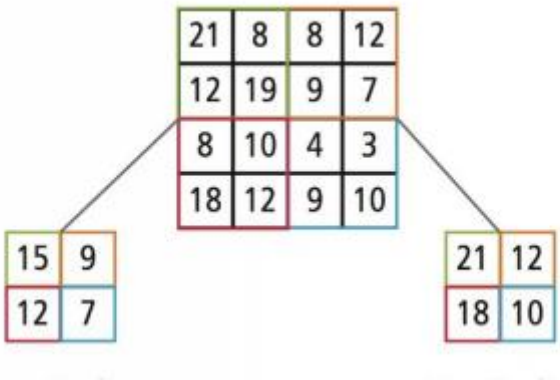
Konvolucijski sloj stvaraju mape značajki tako da uzimaju mapu na ulazu sloja i rade dvodimenzionalnu konvoluciju. Sastoji se od filtara koje sadrže težine koje je potrebno naučiti kako bi mreža davala dobre rezultate. Filtri su najčešće manjih prostornih dimenzija od ulaza, no uvijek su jednake dubine kao i ulaz. Tako ako se na primjer na ulazu konvolucijskog sloja nalazi slika s tri komponente boja (RGB) i veličine je 25x25 točaka, filtar može biti matrica težina veličine 5x5 točaka, ali potrebno je imati tri takve matrice, odnosno jednu za svaku komponentu boje. Svaki neuron konvolucijske mreže djeluje lokalno na samo jedan dio slike.



Slika 1. Operacije konvolucije između originalne slike (I) i jezgre (K)

Na slici 1. vidljiva je konvolucija originalne slike I s jezgrom konvolucije K. I*K je transformirana slika. U matrici na slici I odabire se centralni piksel. Okolni se pikseli tada množe s težinama u matrici jezgre konvolucije te se dobivena težina smješta u novu matricu odnosno u transformiranu sliku.

Sloj sažimanja smanjuje veličinu mapa značajki te povećava neosjetljivost na manje pomake značajki. Sažimanje usrednjavanjem traži aritmetičku sredinu mapa značajki. Kasnijim testiranjem utvrdilo se da bolje vrijednosti dobivamo uzimanjem maksimalne vrijednosti - sažimanje maksimalnom vrijednošću.



Slika 2. Primjer sažimanja aritmetičkom sredinom (lijevo) te sažimanje maksimalnom vrijednošću (desno)

Najpoznatije arhitekture su LeNet-5, AlexNet, VGG 16, Inception, ResNet, ResNeXt, DenseNet, Mask R-CNN. Ovdje je Mask R-CNN ujedno i detektor. Ovdje istaknimo razliku između detekcije i klasifikacije. Detekcija je prolaz kroz sliku i traženje značajki objekata na slici. Klasifikacija je stavljanje slika u određenu grupu na temelju nađenih značajki objekata. Klasifikacija sadrži slike sličnih (istih) objekata.

U praktičnom djelu rada bavit ćemo se nadziranim učenjem i to konvolucijskom neuronskom mrežom koje su već ranije naučene na ogromnom skupu fotografija.

4. Domena eksperimenta i definicija problemskog zadatka

Cilj rada je detekcija plivačkih tehnika i analiza plivačkih stilova korištenjem metoda strojnog i dubokog učenja, tj. korištenjem konvolucijske neuronske mreže.

4.1. Plivanje

Ovdje ćemo plivanje promatrati kao rekreativnu aktivnost, kao dio svakodnevnog života, ne samo kao sport. Plivanje kao (vrhunski) sport zahvaća jako mali dio populacije.

Plivanje kao rekreacija zahvaća jako veliki broj ljudi, praktički sve one koji gravitiraju bazenima ili moru i vole prakticirati plivanje. Na pitanje je li to ustvari plivanje, odgovorit ćemo da nije. Ljudi najčešće ne znaju plivati. Oni plutaju, održavaju se na površini i tehnike plivanje tumače na način koji nije točan, već su tehnike savladali opservacijskom metodom.

Osobe koje se žele rekreativno baviti plivanjem iz nekih razloga ne žele privatnog trenera (moguće da je taj razlog financijski). Na taj način ne ovladavaju plivanjem, vrlo brzo postaju demotivirani radi izostanka rezultata te odustaju.

Redovito odlaze na kopnene aktivnosti tipa trčanje, badminton, tenis, nogomet. Kopnene aktivnosti također imaju zamku. Zamka je tvrda podloga na kojoj se mnogi ozljeđuju. Nakon ozljede, propisuje se plivanje. I opet smo na početku.

Iz tog razloga navedeno je kako je intencija baviti se upravo tim ljudima – rekreativcima i napraviti sustav koji njima može pomoći u učenju plivanja na način da sustav daje povratne informacije o samom stilu plivanja. Vrhunski plivači imaju slične sustave, ali nijedan nema mogućnost davanje samostalne podrške odnosno savjeta. Najčešće su to sustavi koji mogu snimku usporiti i razlomiti na dijelove nakon što trener da određene parametre (kao npr. gdje je 15

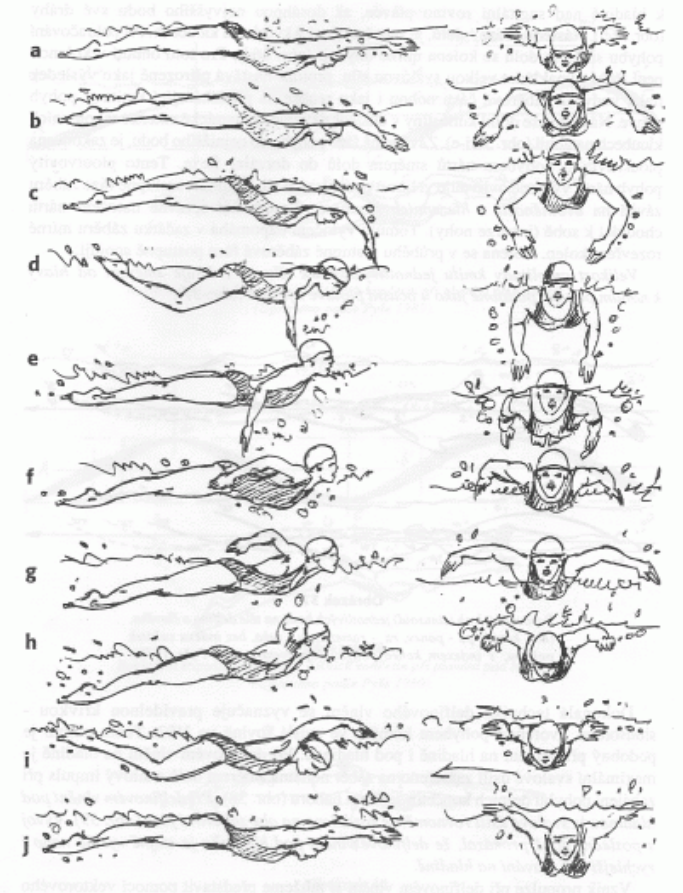
metara, gdje je plivač napustio ronjenje i izašao na površinu vode). Ti su sustavi redovito i skupocjeni, razvijeni su u Sjedinjenim Američkim Državama i najčešće nedostupni i presloženi za krajnjeg korisnika koji nema ambicije postati vrhunski plivač ili učiti kako sustavom operirati.

Kod ovog ekspertnog sustava, jedino što vam treba je kamera (čak i obična funkcionira – iako što kvalitetnija, kvalitetniji je i rezultat), ova aplikacija koja se razvija te vaša dobra volja za plivanjem i napredovanjem.

Za daljnje razumijevanje teksta, napravit će razliku između tehnike i stila plivanja. Naime, tehnika plivanja je biomehaničko savršenstvo u načinu plivanja. Tehnika plivanja nam govori kako se najlakše kretati kroz vodu (minimizirajući otpor vode) te kako iskoristiti najveći opseg pokreta svakog zgloba bez mogućnosti da će to utjecati na ozljedu [5]. Pravilna je tehnika ono što svaki plivač (sportaš) želi dostići i postići. Recimo kako postoje četiri tehnike plivanja (delfin, leđno, prsno i kraul).

4.1.1. Tehnika delfin (po novoj nomenklaturi dupin)

Tehnika delfin najteža je tehnika za savladavanje. Ne zbog njezine kompleksnosti, već zbog snage ramenog zgloba koju plivač mora imati kako bi tehniku izveo pravilno. Tehnika započinje paralelnim radom nogu u smjeru gore-dolje. Ruke plivača također se kreću istovremeno prema bokovima. Uz jak istovremeni udarac obje noge, plivač prebacuje ruke u početni položaj gdje su ruke ispružene ispred glave. Kada ruke dođu do ravnine bokova, plivač uzima zrak. Zbog pomicanja težišta, tijelo tone i tu nastaje najveći problem povrata ruku u početni položaj zbog nedostatka snage. Rekreativni plivači u gotovo svim slučajevima izbjegavaju ovu tehniku.

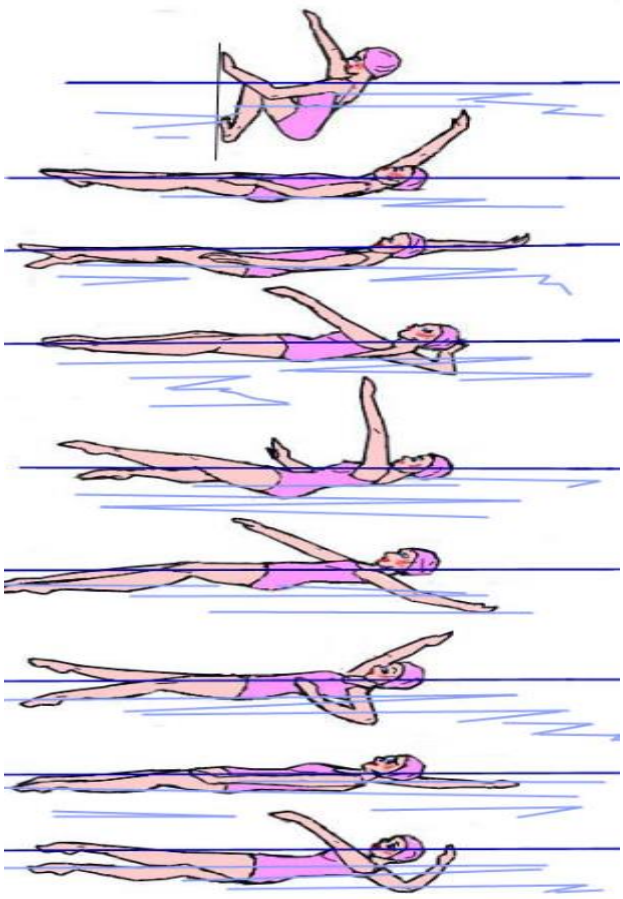


Slika 3. Tehnika plivanja delfin načinom *[6]*

4.1.2. Tehnika leđno

Tehnika leđno najjednostavnija je za početnika iz razloga što plivač početnik ne treba misliti na jednu vrlo važnu stvar – disanje. Također, plivanje ovom tehnikom se preporučuje onim osobama koje imaju problem s leđima zbog pravilnog položaja glave. Tehnika se izvodi na leđima. Plivač izvodi udarce nogama jednom pa drugom nogom. Tako rade i ruke koje iz početnog položaja (ruke u položaju iznad glave) kreću (prvo jedna pa druga) do kukova kroz vodu, a potom kroz zrak do početnog položaja. Kaže se da je entropijski[2] najzahtjevnija tehnika (zbog suprotnog položaja tijela u odnosu na druge tehnike).

[2] Veličina kojom se izražava neuređenost sustava

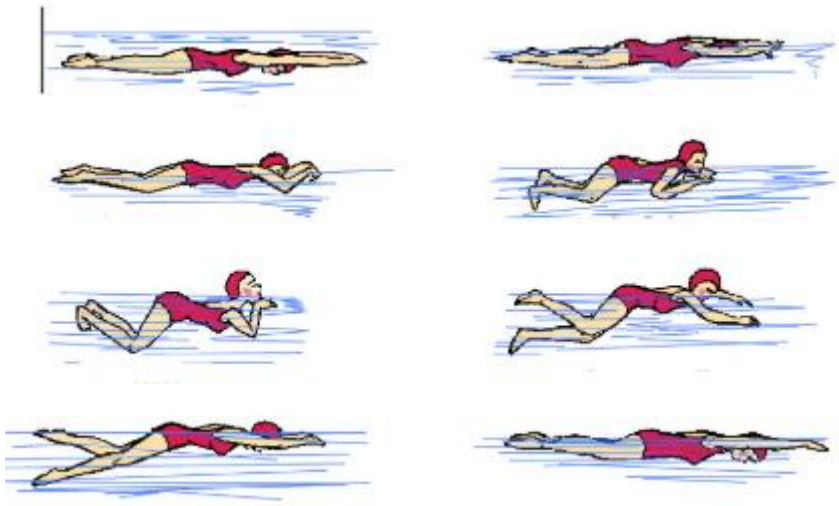


Slika 4 Tehnika leđno *[7]*

4.1.3. Tehnika prsno

Ovaj je način plivanja tehnički najzahtjevniji. Često se kod plivača može uočiti sljedeća zakonitost. Ili znaš plivati prsno ili ne znaš (plivač koji ima prirodne predispozicije za dobro plivanje prsnim načinom, najčešće ne može pravilno izvesti ostale tehnike, ma koliko se trudio; i obrnuto, plivač koji dobro „barata“ preostalim tehnikama u prsnom biva prosječan). Tehnika prsno proizašla je iz tehnike delfin. Plivač je u potrbušnom položaju ispruženih ruku i nogu. Plivanje započinje razdvajanjem ruku lijevo i desno te podvlačenjem ruku pod sebe. U tom trenutku plivač privlači noge prema stražnjici te generira snažnu silu kojom će noge ispucati nazad te time ostvariti propulziju. Prije potpunog pružanja nogu, plivač vraća ruke u početni položaj. Ova je tehnika zbog privlačenja nogu

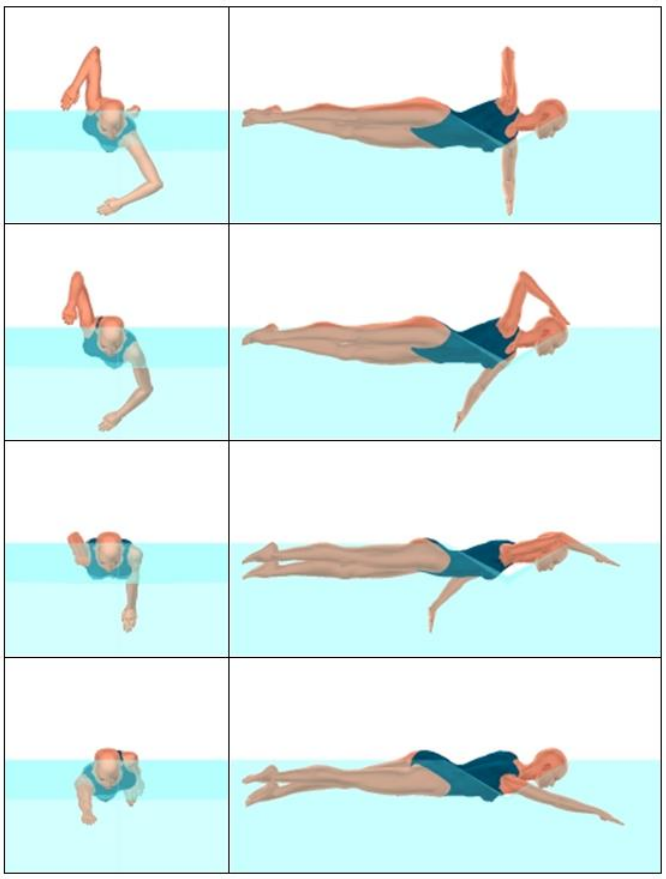
u suprotnom smjeru od kretanja te vraćanju ruku kroz vodu (a ne kroz zrak kako je slučaj kod ostalih tehnika) najsporija tehnika.



Slika 5. Tehnika prsno *[8]*

4.1.4. Tehnika kraul

Tehnika kraul najbrža je tehnika te biomehanički najjednostavija (problem u savladavanju stvara disanje). Većinu vremena plivač provodi plivajući kraul. Upravo zbog svojih biomehaničkih svojstava, ovdje dolazi do najmanje ozljeda. Laički, možemo reći kako je ta tehnika obrnuta od leđne tehnike. Plivač je u potrbušnom položaju gdje ruke i noge rade naizmjence. Glava se okreće u stranu kako bi plivač uzeo zrak.



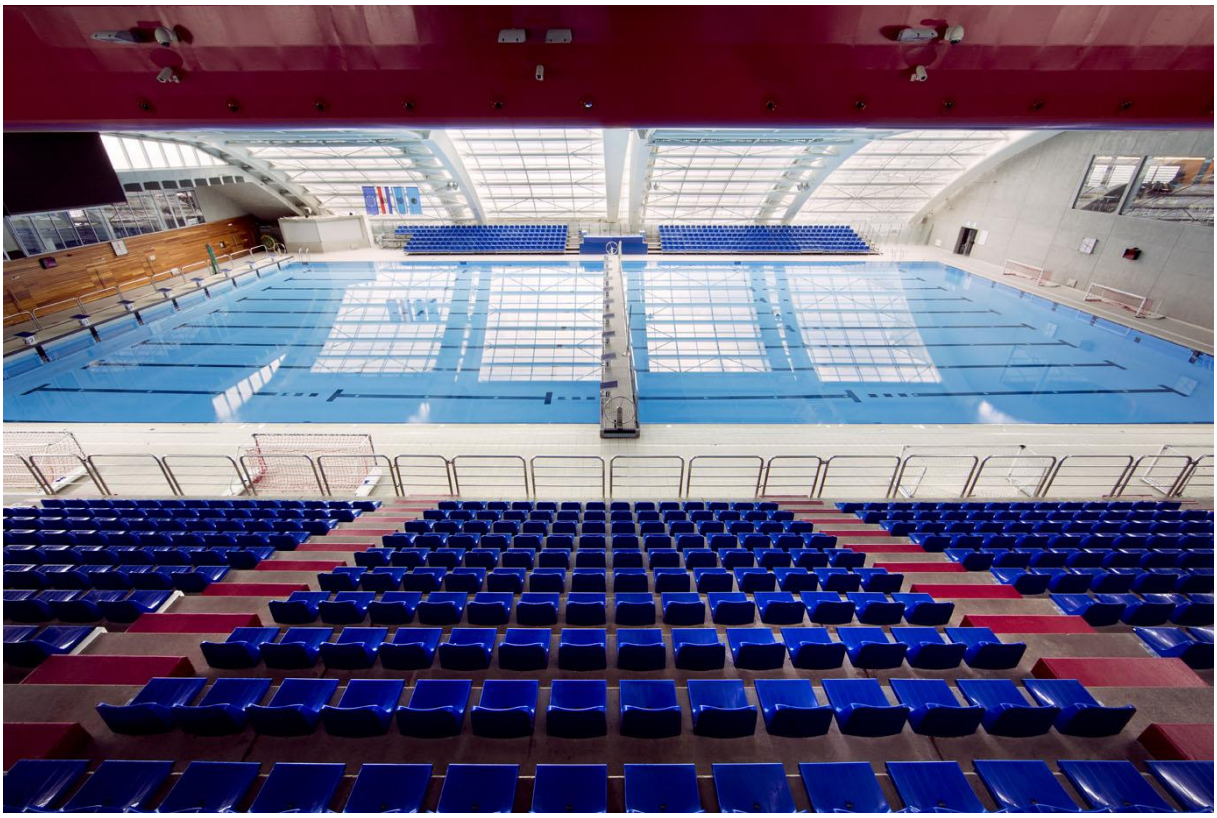
Slika 6. Tehnika kraul *[9]*

Stil plivanja je tumačenje tehnike od strane plivača koji je izvodi [10]. Dakle, stil je vezan uz svakog plivača i na način na koji on zapravo pliva. Samim time, stil je trenutno stanje i cilj je stil izjednačiti s tehnikom plivanja. U daljnjem ćemo dijelu analizirati stil plivanja (što ne valja, gdje ruka ulazi u vodu, izlaze li noge previše iznad površine vode), a detektirat ćemo tehnike plivanja (je li to što pliva delfin, leđno, prsno ili kraul).

5. Snimanje i prikupljanje uzoraka

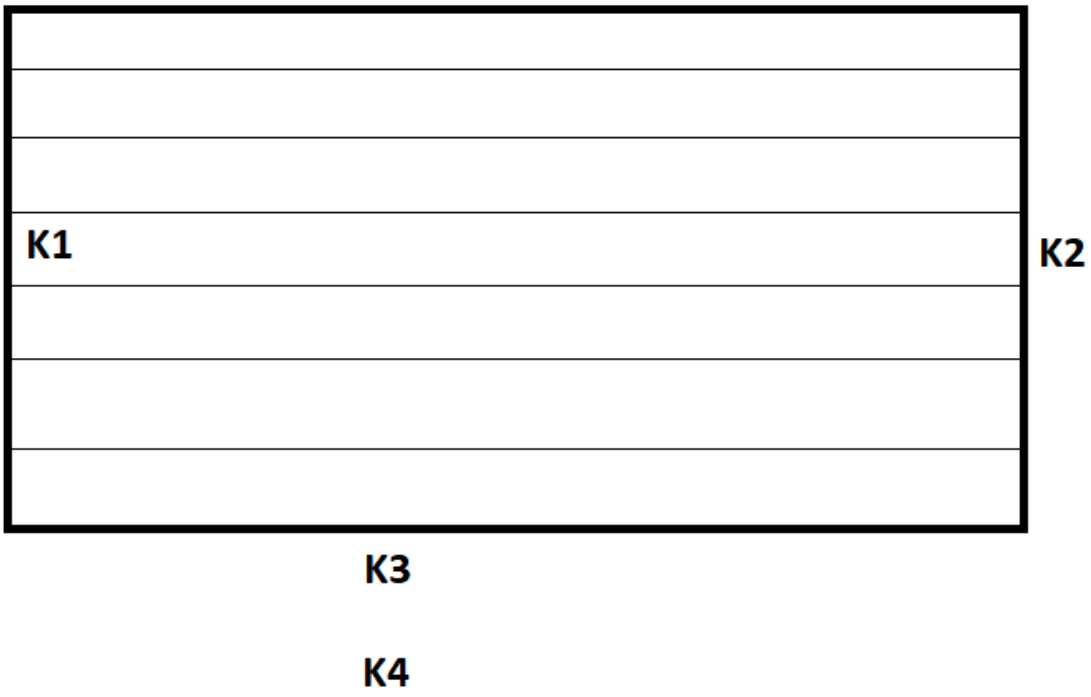
Ključan korak za strojno učenje i analizu bilo kakvih podataka je njihovo prikupljanje, odnosno u slučaju video materijala njihovo snimanje. Nakon ishodovanja potrebnih dozvola, snimanje plivača se provodilo na Bazenima Kantrida u Rijeci. To je kompleks pet bazena gdje veliki broj djece trenira plivanje i gdje se veliki broj odraslih ljudi rekreira.

Od pet bazena, za snimanje je odabran bazen naziva Zdravko Ćiro Kovačić, to jest „glavni bazen“ (slika 7).



Slika 7. Izgled jednog od bazena kompleksa (Zdravko Ćiro Kovačić) gdje je vršena glavnina snimanja

Duljina bazena 50 metara, širina 25 metara, dubina 2,20 metara uz 10 plivačkih staza. U jednom dijelu snimki snimanje je odrađeno na manjem bazenu duljine 25 metara, širine 10 metara i dubine 1,40 metara s četiri plivače staze.



Slika 8. Skica bazena s ucrtanim položajem kamera

Kamera na položaju K1 bila je uronjena u vodu na 2,20 metara dubine. Kamera na položaju K2 nalazila se pored startnog bloka (na suhome) na visini 30 centimetara od površine vode. Kamera na poziciji K3 također se nalazila na suhome na udaljenosti od 15 metara od početka bazena na stalku visine 2 metra. Kamera K4 nalazila se na stropu bazena na visini od 13 metara.

Kao modeli, za jedan dio snimanja poslužili su plivači Plivačkog kluba Kantrida iz Rijeke te treneri istog kluba. Drugi dio snimanja je tempiran za vrijeme natjecanja, jednog od većih natjecanja na Bazenima Kantrida. To nam je jamčilo da će tehnika plivanja i stilovi kojima plivaju snimani plivači biti što sličniji, odnosno da će odstupanja biti što manja. Naravno, nije moguće da odstupanja nema.

Snimljeno je puno materijala. Točnije nešto više od četiri sata video materijala, ukupne „težine“ 170 GB.

5.1. Tijek snimanja

Za snimanje su korištene četiri GoPro kamere.

Slika 9. Korištena GoPro kamera

Svaka je kamera imala svoje postavke. Od broja fotografija u sekundi (fps[3]), širine kuta snimanja do, naravno, samog položaja kamere. Kamere mogu bez problema ići pod vodu, ali imaju dva važna ograničenja. Jedno ograničenje je vijek trajanja baterije koje se u slučaju snimanja na visokoj rezoluciji od 4K isprazni za 30-tak minuta. FPS je bio postavljen na 60 sličica u sekundi što nam je osiguralo da ne dobijemo zamućene dijelove s obzirom da snimamo akcijske slike.

Drugo ograničenje jest to što nema povezanosti između mobilnog telefona i same kamere. Naime, za razliku od drugih tehnoloških uređaja kao što su računala, mobiteli, kamere koji nisu koji s vodom nisu dobri „prijatelji“. GoPro je vodootporan i nema problema da će voda naštetiti sklopovlju, međutim upravljanje kamerom i snimanjem kada je kamera pod vodom nije jednostavno. GoPro kamera ima ekran na kojem možete gledati što snimate. Ukoliko kameru stavite na neki drugi položaj s kojeg ne možete direktno gledati u ekran, postoji

[3] engl. **F**rames **P**er **S**econd – broj sličica u sekundi koje opisuju broj sličica koje kamera snima kako bismo dobili dojam neprekinutosti

mogućnost da GoPro kameru i mobilni telefon povežete Bluetooth-om ili bežičnom Wi-Fi mrežom te putem aplikacije instalirane na vašem mobilnom uređaju pratite što se snima. Voda tu vezu prekida i nemate prijenos u realnom vremenu. Ne možete vidjeti kada se baterija istrošila, je li dobar kut snimanja te je li neki od plivača, odnosno vodena struja stvorena samim plivanjem pomakla kameru koja sada snima zid umjesto plivača. Upravo su to i bili glavni problemi u počecima snimanja. Te smo probleme razriješili boljim fiksiranjem kamera za podlogu te dobrim i plućima punim kisikom kako bismo utvrdili je li kut snimanja povoljan. Iz tih razloga navodimo kako još uvijek, čak i za vrlo sofisticiranu tehnologiju (GoPro ima tako javno mnijenje) voda nije prijateljsko, a ni dobro istraženo okruženje.

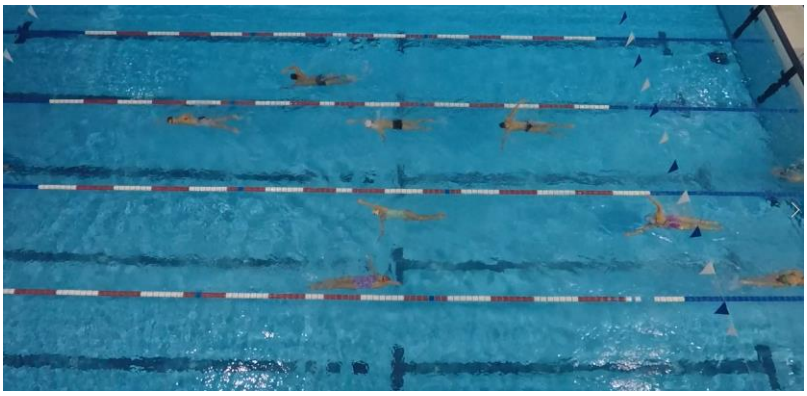
5.2. Pozicije kamere

Plivače smo snimali iz nekoliko pozicija. Snimali smo pod vodom i nad vodom (na suhom) jer se potpuno različite stvari odvijaju pod i nad vodom. Zapravo ono što je pod vodom i ono što se događa nad vodom nema nikakve poveznice, osim položaja tijela. Slikovno ćemo prikazati kako smo postavili kamere (slika 10., slika 11., slika 12., slika 13., slika 14., slika 15.).

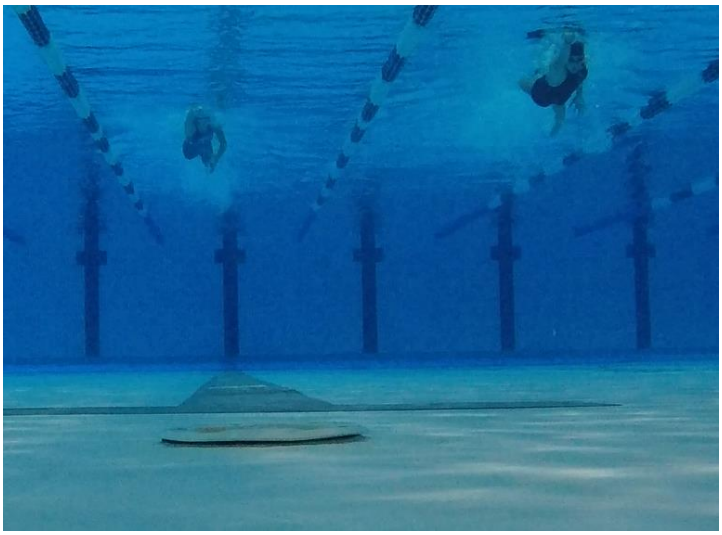
Slika 10. Kamera postavljena bočno od plivača pod vodom

Slika 11. Kamera postavljena ispred plivača točno ispod površine vode u ravnini s plivačem

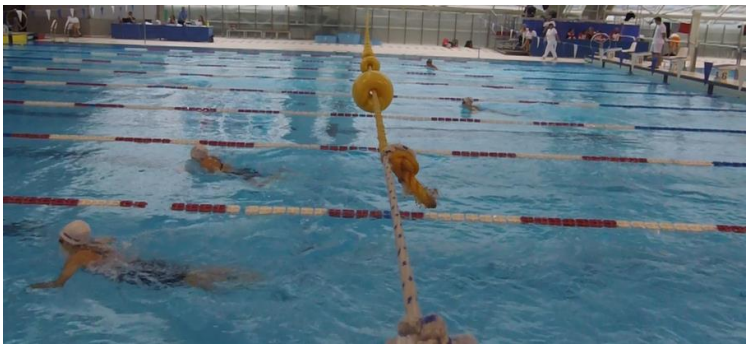
Slika 12. Kamera postavljena ispred plivača iznad vode



Slika 13. Kamera postavljena 13 metara iznad plivača



Slika 14. Kamera postavljena na dno bazena ispred plivača



Slika 15. Kamera postavljena bočno iznad vode

5.3. Problemi kod snimanja na bazenu

Snimke snimane pod vodom su vrlo često mutne (gustoća vode, razni spojevi klora koji vodu zamućuju, nečistoće). Snimke koje su snimane iznad vode često su bile „mokre“, točnije, objektiv kamere bi bio pošprican dolaskom plivača te daljnja snimka ne bi bila čista kao što bismo očekivali (slika 16., slika 17.). Sve su to bile neizvjesnosti i nismo znali kako će sve to utjecati na daljnju obradu snimaka i postupak detekcije.

Slika 16. Prilaz plivača kameri

Slika 17. Izgled objektiva kamere nakon što je plivač došao do kamere

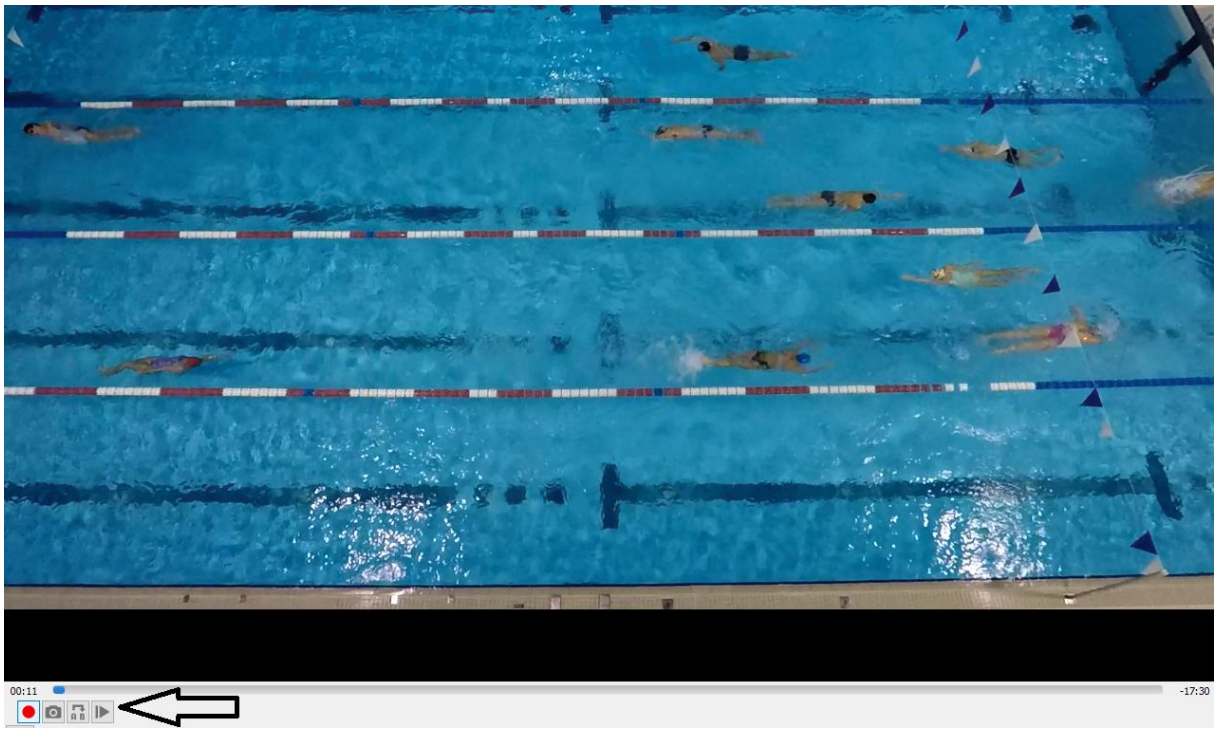
Slika 18. Izgled objektiva kamere nakon što se plivač odmaknuo od kamere i odbljesak na vodi od sunca

5.4. Predobrada podataka

Nakon što je snimanje završilo, slijedi faza pregleda što je snimljeno te faza dosnimavanja i pripreme videozapisa za strojno učenje kako bismo naučili neuronsku mrežu da raspoznaje plivačke tehnike.

Iako je plivanje ciklički sport što znači da se čitavo vrijeme ponavlja jednaka radnja (svaki je zaveslaj jednak, ruke se vrte u istom smjeru, na isti način, a noge lupaju bez prekida), snimke smo rezali upravo nakon svakog zaveslaja kako bismo dobili čim manju jedinicu pokreta.

To nije intelektualno zahtjevan posao, ali je mukotrpan. Svakako želite pronaći alat za obradu video zapisa koji pritiskom na tipku automatski izrezuje i sprema video. Na sreću, takvih je (besplatnih) alata podosta. Za pregledavanje videa koristili smo VLC media player [11] koji među svojim naprednijim opcijama ima i opciju snimanja i spremanja snimljenog. Snima se ono što je na ekranu pritiskom miša na crvenu oznaku. Ponovnim pritiskom, snimanje se zaustavlja, a taj dio videozapisa sprema u zadani direktorij. Taj smo posao sveli na minimum klikanja mišem (slika 19., označeno strelicom).



Slika 19. Dodatne opcije u VLC-u

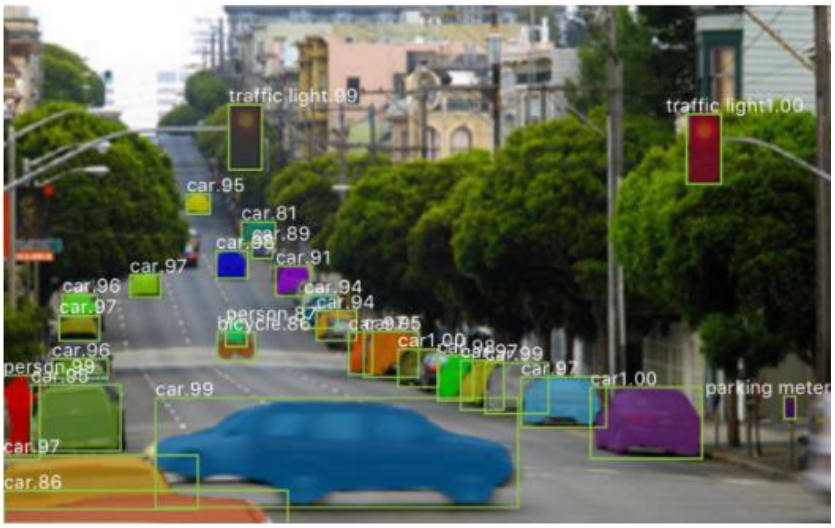
Potom, svaki je od novonastalih videozapisa dobio svoju oznaku, točnije jednu od četiri klase. Svaka klasa označava jednu od tehnika. Ranije smo spomenuli kako imamo četiri tehnike, i to delfin, leđno, prsno i kraul tehniku tako da ćemo prema tome imati četiri klase. Snimke za svaku od tehnika bile su imenovane po uzoru za klasu delfin kao *delfin1*, *delfin2*, …

Za svaku smo tehniku pripremili između 150 i 220 snimki.

6. YOLO i Mask R-CNN

YOLO[4] [12] je detektor kojeg smo koristili prilikom izrade ovog rada. Način rada prilikom detekcije objekata je sljedeći. Na danoj fotografiji YOLO iscrtava mrežu ćelija koja je veličine A x A. Fotografija se obrađuje kao cjelina. Svaka ta ćelija predviđa točno jedan objekt u koji se smješta predviđeni centar objekta. Tada se oko tog objekta iscrtava pravokutnik koji objekt zatvara („okružuje ga“). Tada za svaki pojedini pravokutnik, YOLO pokušava odrediti klasu kojoj taj objekt u pravokutniku najvjerojatnije pripada koristeći naučene modele za klase kao što su kao što je osoba, automobil, cvijeće, pas, …. Modeli YOLO mreže naučeni su na COCO skupu podataka.

Mask R-CNN najnovija je vrsta neuronske mreže za klasifikaciju i detekciju elemenata. Na određenoj slici radi tako da pronađene objekte stavlja u pravokutnike te ih imenuje (klasificira). Radi u dva postupka. Prvo, pretpostavlja gdje bi određeni objekt na slici mogao biti. Drugo, predviđa klasu objekta, oko njega crta siluetu kako bi ga označio te na objekt stavlja „masku“ odnosno ispunjava ga bojom [13].



Slika 20. Mask R-CNN

[4] **Y**ou **O**nly **L**ook **O**nce

Kako Yolo i Mask R-CNN detekciju rade na slikama, odabrane i selektirane videozapise smo trebali transformirati u sličice (frame-ove). Obzirom da je prilikom snimanja korišten visoki fps od 60 fps, za testiranje smo uzeli svaki 20 - i frame.

Kako je provođenje ovog eksperimenta i testiranje već naučenih modela detekcije na našim slikama za računalo kućnih performansi zahtjevan posao, taj je dio odrađen na računalima s GPU jedinicom u laboratoriju Odjela za informatiku Sveučilišta u Rijeci.

7. Analiza rezultata

Dobiveni rezultati nisu dovoljno dobri za zadatak detekcije akcija. Zapravo, kako na situaciju pogledamo. S obzirom da prije pisanja ovog rada nismo znali što nas očekuje i kakve ćemo rezultate dobiti u uvjetima rada s vodom te samo testiranje snimki s modelima koji su naučeni na drugim bazama slika, rezultate možemo dvojno i protumačiti.

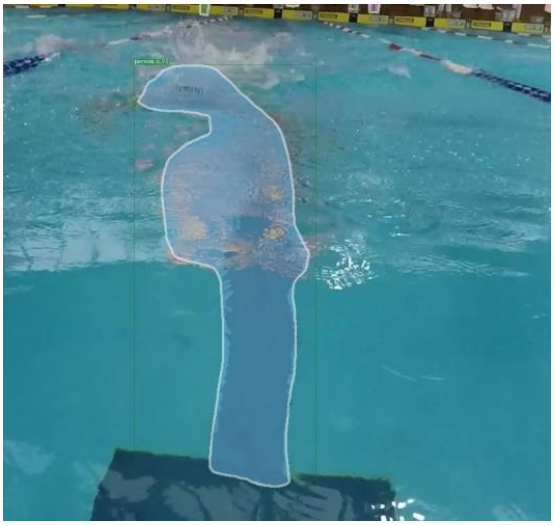
Iz fotografije koje je detektor Mask R-CNN obradio vidljivo je da snimke nisu pogrešno snimljene (potpuno pogrešan kut, mutno) i da su modeli za objekte koji su postojali u skupu za učenje dobro naučeni jer se osobe koje su na suhom detektirane vrlo točno (engl. true positive – TP detekcija), iako su prilično udaljene od same kamere i pored nejednolične pozadine. Riječ je o sucima, osobama koje šeću po rubu bazena, (slika 21., slika 22.).

Slika 21. Mask R-CNN detektira osobe na suhom

Slika 22. Bliži pogled na detektirane osobe i pouzdanost s kojom ih je model klasificirao

Za razliku od sudaca, plivače koji prilaze kameri, dakle nalaze se bliže kameri, detektor uopće ne prepoznaje.

Točnije, u velikom broju slučajeva je detektirao da postoji osoba tamo gdje ona uopće nije bila (to se naziva engl. false positive – FP detekcija). Položaj crne oznake na bazenu i vrtlog vode za njega je stvorio siluetu osobe koja je lažno prepoznata (slika 23.).

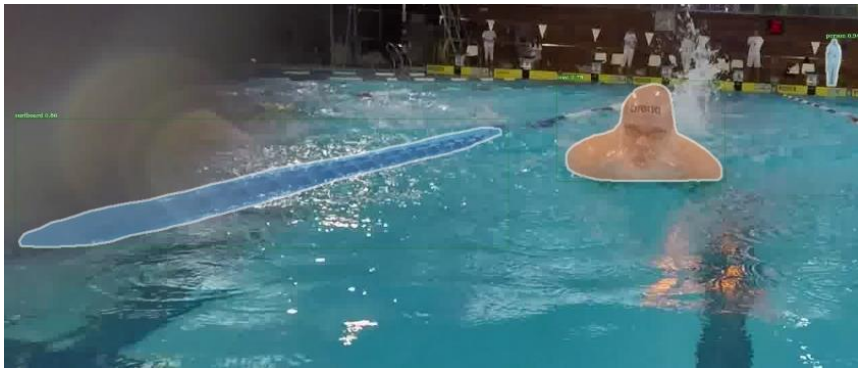


Slika 23. Lažna klasifikacija osobe

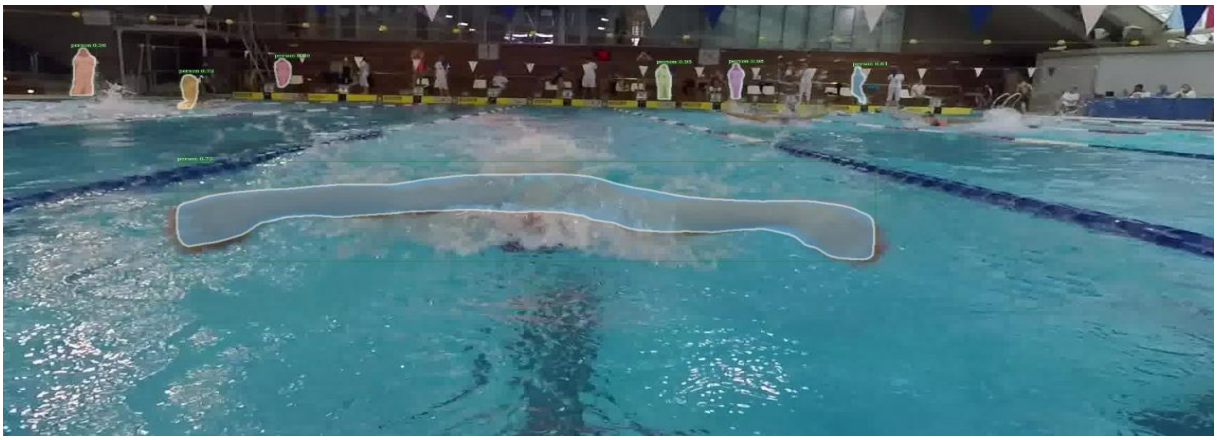
Na ovaj slici detektor nije prepoznao glavu plivača kada je plivač bio prilično blizu kameri, dok na nekim slikama prepoznaje osobu samo na osnovu dijela ruke koja izviruje iz vode dok plivač pliva kraul. Vjerojatno je u ovom slučaju problem „šum“ pjena koju stvara plivač prilikom plivanja. Pojam „šum“ podrazumijeva sve ono što remeti pravilno poimanje fotografije. Primijetimo kako je osobu u daljini prepoznao (slika 24.).

Slika 24. Osobu u daljini prepoznaje, osobu koja izlazi iz vode ne prepoznaje

U nastavku su primjeri krive detekcije kada je plivač detektiran kao čamac i to s prilično visokom pouzdanošću, a plivačka traka kao daska za surfanje (slika 25.).



Slika 25. Lažna detekcija čamca i daske za surfanje

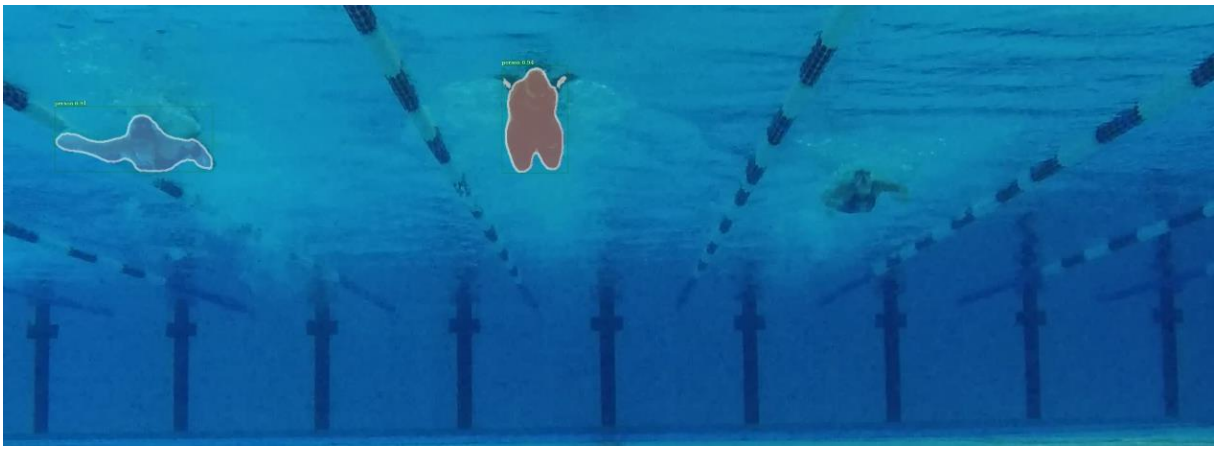


Slika 26. Detekcija svih osoba u daljini

Razne promjene na vodi tipa mreškanje vode i zrcaljenje zbog utjecaja svjetla također su pogrešno detektirane kao klasa (slika 27.) .

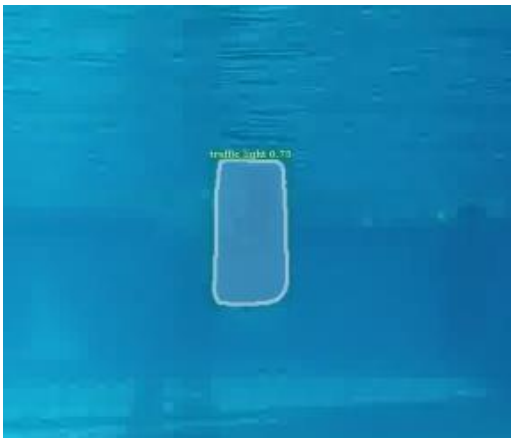
Slika 27. Lažna detekcija čamca

Pokazalo se da je pozicija kamere vrlo važna za uspješnost detekcije. To su situacije kada je kamera postavljena ispred plivača. Pogledajmo što Mask R-CNN detektira kada kameru stavimo na dno bazena (slika 28.).

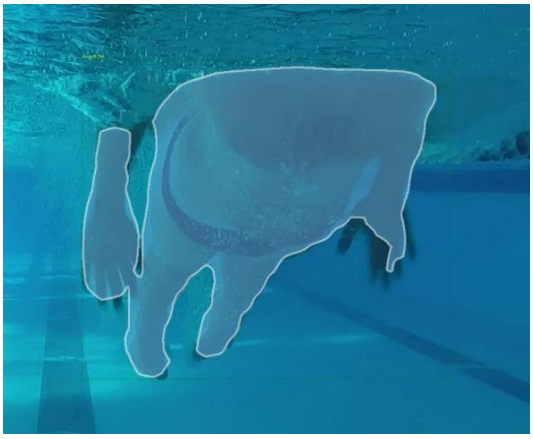


Slika 27. Puno bolja osoba detekcija s kamerom postavljenom na dno bazena

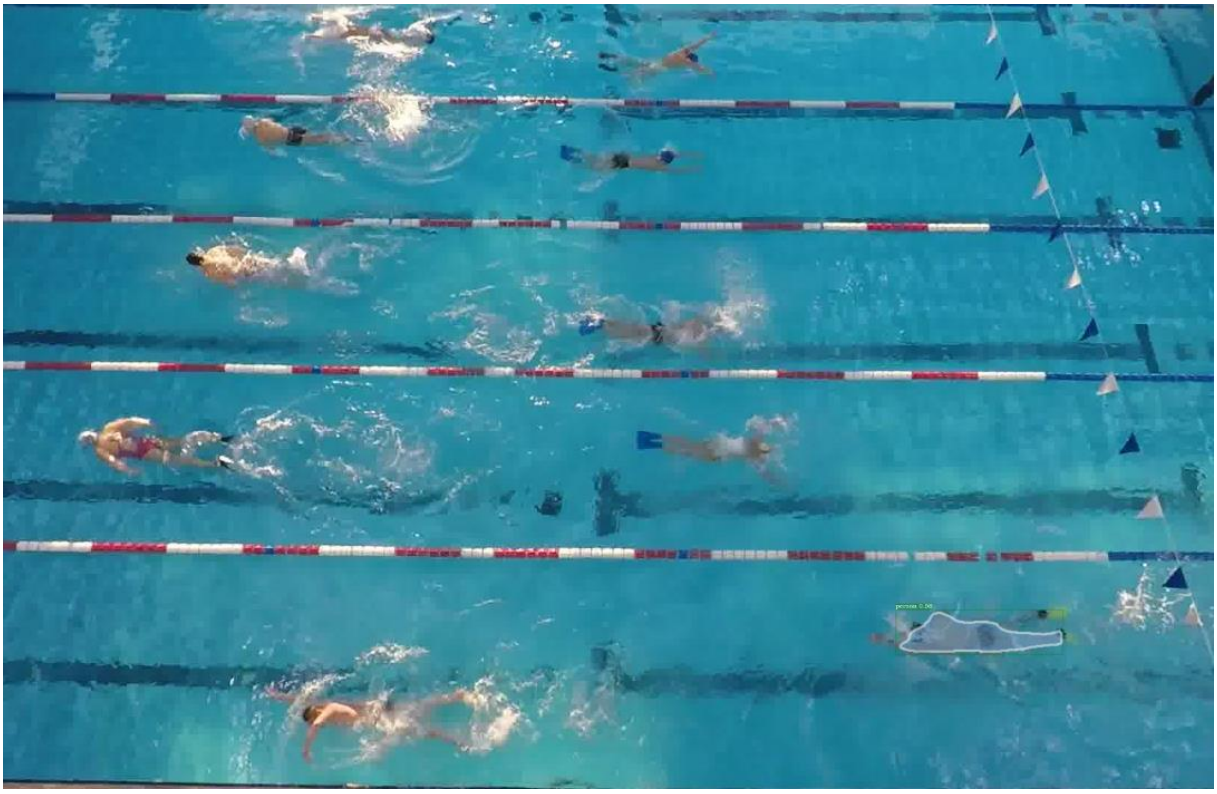
U ovom je slučaju detekcija jako dobra. Uspio je detektirati dvije od tri osobe, iako ne da plivaju.



Slika 29. Detekcija plivača kao semafora



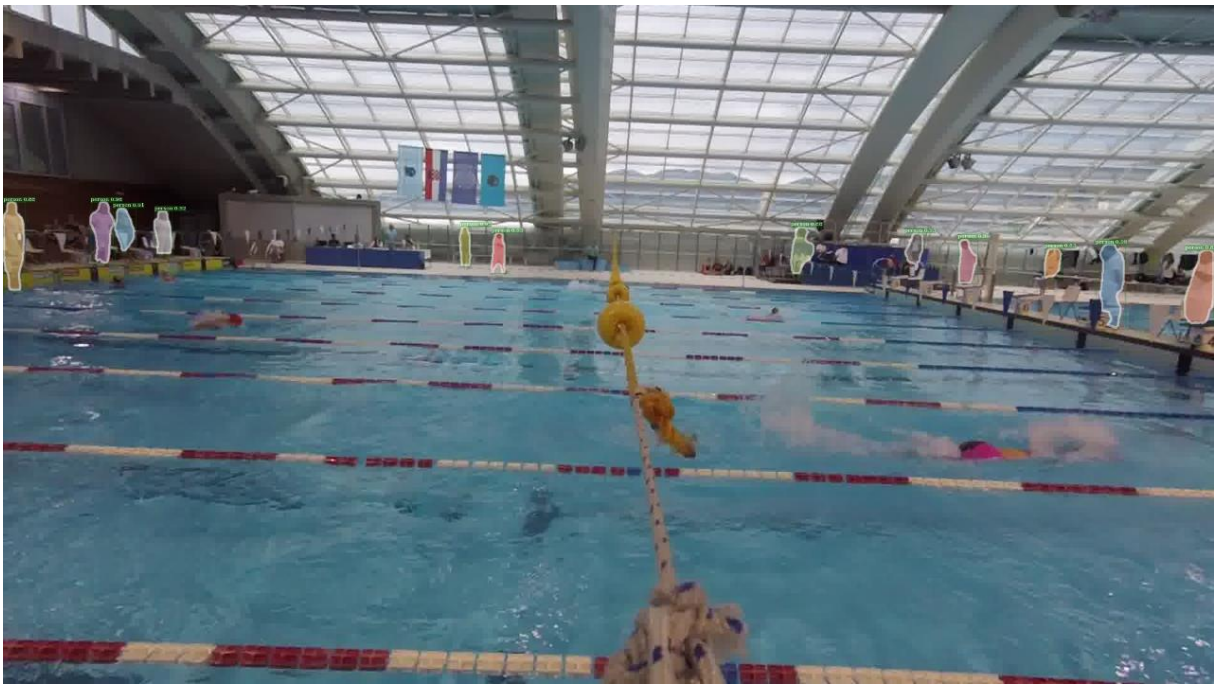
Slika 30. Detekcija psa umjesto osobe (vjerojatno zbog zanimljivog položaja četiriju udova)



Slika 31. Vrlo slaba detekcija osoba

Slika 32. Napokon detektirana osoba

Slika 33. Detektirana osoba, pravilno omeđena



Slika 34. Dobra detekcija osoba izvan vode

Dobra detekcija plivača bila je na bazenu kada nije bilo smetnji u vidu dodatnih osoba ili objekata sa strane. Tu smo dobili najbolje rezultate (slika 35., slika 36., slika 37., slika 38., slika 39.).

Slika 35. Iako je izronila samo glava, detektor je uspio detektirati osobu

Slika 28. Polovična detekcija osobe

Slika 37. Potpuna, pravila detekcija osobe pod vodom

Slika 38. Potpuna, pravilna detekcija plivača

Slika 39. Potpuna, pravilna detekcija plivača, ali ne i tehnike

Dobivene rezultate prikazat ćemo i matematički.

Mjere koje nas zanimaju jesu [14]:

- True positive (TP) – broj slika na kojima su objekti točno detektirani (u radu, klasifikacija plivača pod klasu osobe)
- True negative (TN) – broj slika na kojima su objekt nije detektiran a trebao je biti (u radu, plivač koji nije klasificiran kao osoba)
- False positive (FP) – broj slika netočno detektiranih (u radu, nije plivač, a označen kao osoba)
- False negative (FN) – broj slika pogrešno nedetektiranih (u radu, plivač koji nije označen kao plivač)
- Preciznost – broj ispravnih detekcija s obzirom na sve detekcije

  Prec = $\frac{TP}{TP+FP}$
- Odziv – omjer broja ispravnih detekcija s broj detekcija koje su morale biti prijavljene Rec = $\frac{TP}{TP+FN}$
- F1- mjera - harmonijska sredina preciznosti i opoziva, često se koristi kao mjera koja pokazuje pravu točnost modela F1 = 2 x $\frac{Prec\ x\ Rec}{Prec+Rec}$

Dobivene su sljedeće vrijednosti (na skupu 500 slika):

TP	402
TN	0
FP	30
FN	198
Preciznost	93%
Odziv	67%
F1	77,8%

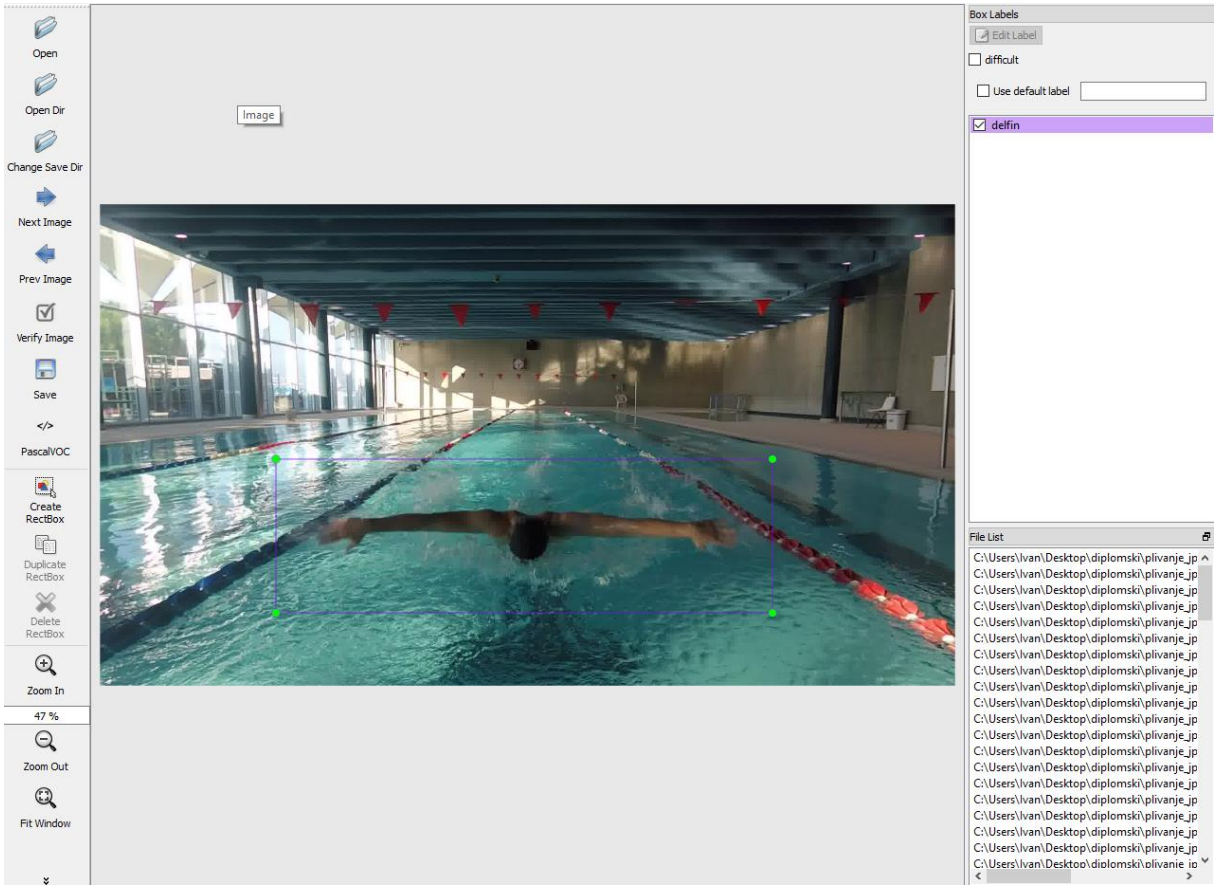
Tablica 1. Matematički prikaz preciznosti sustava

Istaknimo kako je preciznost zaista visoka, 93% pravilnih detekcija. Napomenimo kako je ovdje mreža detektirala plivače kao osobe ali nije detektirala kojom tehnikom plivači plivaju.

8. Daljnje učenje mreže

Pokazalo se da za detekciju plivača ne možemo koristiti modele koji su naučeni na drugoj bazi podataka već da trebamo naučiti mrežu s dodatnim primjerima plivača u bazenu iz naše baze. Za tu svrhu je potrebno preciznije označiti plivača na slici kako bi model mogao što bolje odabrati odgovarajuće značajke koje dobro opisuju plivača u bazenu.

Kao što smo ranije naveli, detektor s kojim smo radili, Mask R-CNN, radi na slikama, tako da smo nakon obrade imali oko 2 000 slika. Alat *labelImg* [15] omogućava da se na prikazanoj slici iscrta pravokutnik koji omeđuje objekt. Tako označene slike spremili smo u već ranije određene klase (njih četiri – delfin, leđno, prsno, kraul). Na slici mišem označimo objekt koji je tada omeđen pravokutnikom te mu dodamo klasu, vidljivo desno na slici 40.



Slika 40. Izgled ekrana alata labelImg

```
<annotation>
  <folder>jpg_subset - oznake</folder>
  <filename>delfin12_00014.jpg</filename>
  <path>
    C:\Users\Ivan\Desktop\diplomski\plivanje_jpg\jpg_subset - oznake\delfin12_00014.jpg
  </path>
  <source>
    <database>Unknown</database>
  </source>
  <size>
    <width>1920</width>
    <height>1080</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>delfin_backout</name>
    <pose>Unspecified</pose>
    <truncated>0</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>833</xmin>
      <ymin>463</ymin>
      <xmax>1114</xmax>
      <ymax>575</ymax>
    </bndbox>
  </object>
</annotation>
```

Slika 41. Kôd koji se generira nakon označavanja slike

Moguće je da će trebati dodatni primjeri slika da mreža prepozna i plivačke tehnike.

Uočene su značajne razlike u detekciji s obzirom na poziciju kamere. Razlog tome je i to što određena tehnika izgleda potpuno drugačije obzirom na kut snimanja (kamera pod vodom i kamera nad vodom vide sasvim suprotne akcije – pokrete). Iz tog je razloga napravljena detaljnija podjela na podskupove (tehnika snimana sprijeda, straga, bočno, pod vodom, iznad vode, sa stropa). U svakom od tih skupova nalazi se između 100 i 200 slika (ovisno o kvaliteti same početne snimke). Smatramo kako će ta podjela doprinijeti boljem učenju modela.

9. Zaključak

Iako nismo uspjeli „istrenirati“ mrežu da pravilno klasificira pojedinu tehniku, puno se toga može zaključiti. Detektor na ovim snimkama radi (detektira osobe u daljini), no na ovom skupu podataka prije nije bio učen. Bez dodatnog učenja i bez još veće količine podataka s plivačima koji plivaju određenim stilom koji demonstrira određenu tehniku nećemo biti u mogućnosti dobiti kvalitetne rezultate, odnosno nećemo biti u mogućnosti definirati odgovarajući model koji će čovjeku taj jednostavan posao obaviti bez problema.

Jedna od mogućnosti poboljšanja je svakako bolje biranje pozicije kamere. Iz rada je vidljivo kako kamera koja se nalazi na stropu visokom 13 metara ne uspijeva detektirati čak niti plivače koje bi klasificirao kao osobe. Najbolji su rezultati bili postignuti kada je kamera bila na dnu bazena ispred plivača i pri maloj udaljenosti. Također, kamera koja se nalazi metar iznad površine vode u okolišu koji nema distraktora i dobro je osvijetljen, također pruža dobru podlogu za detekciju i klasificiranje.

Nadalje, pored optimizacije položaja kamere, svakako treba poraditi i na ručnoj klasifikaciji seta podataka za učenje. Potrebno je na svakoj fotografiji preciznije ručno označiti plivača i označiti s odgovarajućom klasom plivačke tehnike. Točnije, ukoliko je na fotografiji plivač koji pliva tehnikom delfin, treba ga omeđiti pravokutnikom i fotografiju spremiti pod klasu delfin. Tim bismo postupcima osigurali kvalitetnije rezultate u detekciji odnosno raspoznavanju.

10. Literatura

[1] »https://www.bug.hr/vijesti/preminuo-marvin-minsky-pionir-umjetne-inteligenci/149997.aspx,« [Mrežno]. [Pokušaj pristupa rujan 2018.].

[2] »https://hr.wikipedia.org/wiki/Tri_zakona_robotike,« [Mrežno]. [Pokušaj pristupa rujan 2018.].

[3] C. C. Aggarwal, »http://neuralnetworksanddeeplearning.com/,« rujan 2018.. [Mrežno].

[4] D. Božić-Štulić, »https://data.fesb.unist.hr/public/news/Dunja%20Bo%C5%BEi%C4%87-%C5%A0tuli%C4%87-a3b4859ad0.pdf,« rujan 2018.. [Mrežno].

[5] G. L. J. S. Klara Šiljeg, »BIOMEHANIČKE KARAKTERISTIKE ZAVESLAJA U KRAUL TEHNICI,« p. https://hrcak.srce.hr/file/243096, 2016..

[6] »http://www.swimminguae.com/Butterfly.aspx,« [Mrežno]. [Pokušaj pristupa rujan. 2018.].

[7] »http://www.swimminguae.com/Backstroke.aspx,« [Mrežno]. [Pokušaj pristupa rujan 2018.].

[8] »http://www.swimminguae.com/swim/breaststroke.aspx,« [Mrežno]. [Pokušaj pristupa rujan 2018.].

[9] »https://swimmingtechnology.com/get-a-cue-three-freestyle-technique-elements-you-cant-swim-well-without/,« [Mrežno]. [Pokušaj pristupa rujan 2018.].

[10] G. Leko, »Kraul : slobodni način plivanja : sveučilišni priručnik,« Zagreb, Kineziološki fakultet, Sveučilište u Zagrebu, 2010., pp. 10-15.

[11] »https://www.videolan.org/vlc/index.hr.html,« [Mrežno]. [Pokušaj pristupa rujan 2018.].

[12] »https://pjreddie.com/darknet/yolo/,« [Mrežno]. [Pokušaj pristupa rujan 2018.].

[13] »https://medium.com/@alittlepain833/simple-understanding-of-mask-rcnn-134b5b330e95,« [Mrežno]. [Pokušaj pristupa rujan 2018.].

[14] »https://repozitorij.unizd.hr/islandora/object/unizd%3A2298/datastream/PDF/view,« rujan 2018.. [Mrežno].

[15] »https://github.com/tzutalin/labelImg,« [Mrežno]. [Pokušaj pristupa rujan 2018.].