

SVEUČILIŠTE U ZAGREBU
FAKULTET ELEKTROTEHNIKE I RAČUNARSTVA

SEMINAR

Praćenje nogometaša primjenom YOLOv5 modela i DeepSORT algoritma

Marko Tutić

Voditelj: Izv. prof. dr. sc. Zoran Kalafatić

Zagreb, svibanj 2022.

SADRŽAJ

1. Uvod	1
2. Skup podataka	3
3. Arhitektura sustava	5
3.1. Detekcija nogometaša	5
3.2. Praćenje objekata	6
3.2.1. Kalmanov filter	7
3.2.2. Dodjela trajektorija detekcijama	7
3.3. Izrada mape aktivnosti	8
4. Implementacija sustava	11
4.1. Podsustav za detekciju nogometaša	11
4.2. Podsustav za praćenje nogometaša	12
4.3. Podsustav za izradu mape aktivnosti nogometaša	12
4.4. Povezivanje podsustava u cijelinu	13
4.5. Podsustav za vrednovanje praćenja i detekcije objekata	13
5. Analiza rezultata	14
6. Zaključak	15
7. Literatura	16

1. Uvod

Praćenje objekata je grana računalnog vida u kojoj se detekcije zadane u određenom okviru nastoje pratiti u nadolazećim okvirima. Metode praćenja objekata se najčešće razvrstavaju u metode praćenja jednog objekta i metode praćenja više objekata. U metode praćenja jednog objekta ubrajaju se metode koje ne zahtjevaju zadavanje detekcija u svakom okviru videozapisa. U takvim metodama dovoljno je zadati poziciju objekta u nekom početnom okviru, a metoda zatim može pratiti objekt u daljnjim okvirima isključivo na temelju karakteristika objekata i promjena koje se događaju u slici.

Praćenje više objekata se realizira kao detektiranje objekata i dodijeljivanje identifikatora detekcijama na temelju prethodnih detekcija te je cilj očuvati ispravne identifikacije objekata kroz više okvira videozapisa. Ove metode zahtjevaju detekcije u svakom okviru videozapisa. Praćenje više objekata često se vodi pretpostavkom da objekti neće naglo mijenjati brzinu i smjer kretanja.

Praćenje više objekata predstavlja veliki izazov u području računalnog vida jer se redovito javljaju situacije koje mogu izazvati propuštanje dodijeljivanja identifikatora ili pogrešno dodijeljivanje identifikatora poput zaklanjanja objekta i zamjene identiteta. Zaklanjanje objekta je pojava kada objekt nije vidljiv na slici, ali se pozicija objekta i dalje nalazi unutar slike. Do zamjene identiteta dolazi kada se objekti mimoilaze ili se nalaze vrlo blizu određeno vrijeme te se kao posljedica toga identitet jednog objekta dodijeli drugom i obrnuto.

Praćenje nogometaša predstavlja situaciju u kojoj je potrebno pratiti više objekata pri čemu su zamjene identiteta česte s obzirom da se igrači kroz trajanje cijele utakmice mimoilaze i zaklanjaju na kameri. Dodatan izazov pri praćenju nogometaša predstavlja činjenica da igrači izgledaju vrlo slično na velikoj udaljenosti zbog identičnih boja dresova.

Originalni SORT algoritam za praćenje objekata imao je problem sa velikim brojem zamjena identita s obzirom da je uzimao u obzir samo pozicije objekata prilikom dodijeljivanja identiteta. DeepSORT algoritam nastoji riješiti taj problem uvođenjem duboke metrike asocijacije koja se temelji na opisnicima izgleda objekata koji gene-

riraju vektor karakteristika koji opisuje izgled za svaki objekt. Taj vektor se također razmatra prilikom izračuna udaljenosti detekcija i predviđenih trajektorija.

DeepSORT algoritam zahtjeva detekcije objekata u svakom okviru kako bi koristio svoja interna stanja koja koristi za predviđanje novih trajektorija. Za detekciju nogometaša koristi se YOLOv5s model koji je naučen na specifičnom skupu podataka. Dobivene detekcije se prosljeđuju algoritmu za praćenje objekata.

Pozicije nogometaša u nekom okviru bilježe se za svakog nogometaša zasebno te se na temelju tih pozicija izrađuje mapa aktivnosti igrača. Pozicija nogometaša se tretiraju kao slučajna varijabla pri čemu je mapa aktivnosti vizualizacija funkcije gustoće te varijable. Za izradu mape aktivnosti koristi se procjena parametara gustoće jezgrom (kernel density estimate) koja koristi pozicije igrača za procjenu funkcije gustoće.

2. Skup podataka

Za skup podataka na kojem se vrši učenje i testiranje sustava iskorišten je videozapis duljine 5 minuta koji sadrži snimku jedne nogometne utakmice. Kamera kojom je utakmica snimana je statična. Svi okviri u videozapisu sadrže pregled cijelog nogometnog terena. Jedna sekunda navedenog videozapisa sadrži 25 slikovnih okvira. Širina videozapisa iznosi 3260 piksela dok visina okvira videozapisa iznosi 570 piksela. Slika 2.1 prikazuje primjer jednog okvira iz videozapisa

Skup podataka dodatno sadrži tekstualni dokument koji sadrži pozicije i identifikatore igrača i sudca. Oznake su zapisane u obliku propisanom MOT16 specifikacijom. Jedan redak tekstualne datoteke ima sljedeći format:

<redni broj okvira>, <identifikator igrača>, <x koordinata>, <y koordinata>, <širina>, <visina>, -1, -1, -1, -1.

Slika 2.2 prikazuje primjer oznaka pozicija i identifikatora igrača za prvi okvir videozapisa.

Koordinate x i y te širina i visina odnose se na pravokutnik koji omeđuje igrača na slici. Koordinate x i y označuju poziciju gornje lijeve točke takvog pravokutnika. Na temelju širine i visine mogu se izvesti i ostale točke omeđujućeg pravokutnika. Niz brojeva -1 se ne koristi, no služe kao nadopuna za zadovoljavanje MOT specifikacije.

Skup podataka dijeli se na skup za učenje, skup za provjeru i skup za testiranje. Okviri videozapisa se dijele u navedene skupove slijedno kako bi se očuvala mogućnost praćenja objekata u navedenim skupovima podataka. Za potrebe učenja detektora objekata oznake koje su zadane u MOT16 formatu potrebno je pretvoriti u oznake koje



Slika 2.1: Primjer okvira iz videozapisa

```

1,0,841,175,13,35,-1,-1,-1,-1
1,1,1551,405,24,54,-1,-1,-1,-1
1,2,1341,217,15,40,-1,-1,-1,-1
1,3,1374,154,14,34,-1,-1,-1,-1
1,4,1457,103,12,27,-1,-1,-1,-1
1,5,1448,186,13,36,-1,-1,-1,-1
1,6,1466,168,12,33,-1,-1,-1,-1
1,7,1621,269,15,44,-1,-1,-1,-1
1,8,1624,175,16,37,-1,-1,-1,-1
1,9,1613,142,13,28,-1,-1,-1,-1
1,10,1627,105,11,26,-1,-1,-1,-1
1,11,1528,133,16,27,-1,-1,-1,-1
1,12,2429,176,15,34,-1,-1,-1,-1
1,13,2056,195,18,39,-1,-1,-1,-1
1,14,1997,263,16,42,-1,-1,-1,-1
1,15,1964,151,14,33,-1,-1,-1,-1
1,16,1925,137,13,29,-1,-1,-1,-1
1,17,1812,283,16,47,-1,-1,-1,-1
1,18,1846,197,15,37,-1,-1,-1,-1
1,19,1758,200,14,35,-1,-1,-1,-1
1,20,1758,155,12,33,-1,-1,-1,-1
1,21,1756,119,12,29,-1,-1,-1,-1
1,22,1618,223,19,37,-1,-1,-1,-1

```

Slika 2.2: Primjer oznaka skupa podataka

zadovoljavaju YOLOv5 format oznaka. YOLOv5 format propisuje izradu tekstualne datoteke s oznakama za svaki okvir te se omeđujući pravokutnik opisuje središtem tog pravokutnika i njegovom širinom i visinom. Identifikatori igrača se ne uzimaju u obzir prilikom pretvorbe jer detektor objekata samo treba lokalizirati objekte te ih klasificirati. Pretvorba gornje lijeve točke u središnju točku pravokutnika vrši se primjenom sljedećih formula:

$$x' = x + \frac{width}{2}$$

$$y' = y + \frac{height}{2}$$

Skup podataka se dijeli tako da skup za učenje sadrži prvih 4500 okvira videozapisa, skup za validaciju sadrži idućih 1500 okvira, dok skup za testiranje sadrži preostalih 1515 okvira.

Skup za učenje koristi se za učenje modela detekcije objekata. Skup za validaciju koristi se za provjeru svojstva generalizacije tijekom različitih epoha prilikom učenja modela detekcije objekata. Konačno, skup za testiranje koristi za evaluaciju modela detektora objekata nakon što je završen proces učenja.

Iako se u DeepSORT algoritmu za praćenje objekata može naučiti opisnik izgleda objekata pomoću istog skupa podataka, taj postupak će biti preskočen s obzirom da izvorni algoritam već sadrži opisnik izgleda osoba koji je prethodno naučen na MARS (Motion Analysis and Re-Identification Set) skupu podataka koji sadrži slike različitih osoba. Takav opisnik izgleda naučen je za izlučivanje značajki koje opisuju izgled jedne osobe. Sustav za praćenje objekata će se samo testirati na skupovima za učenje i testiranje.

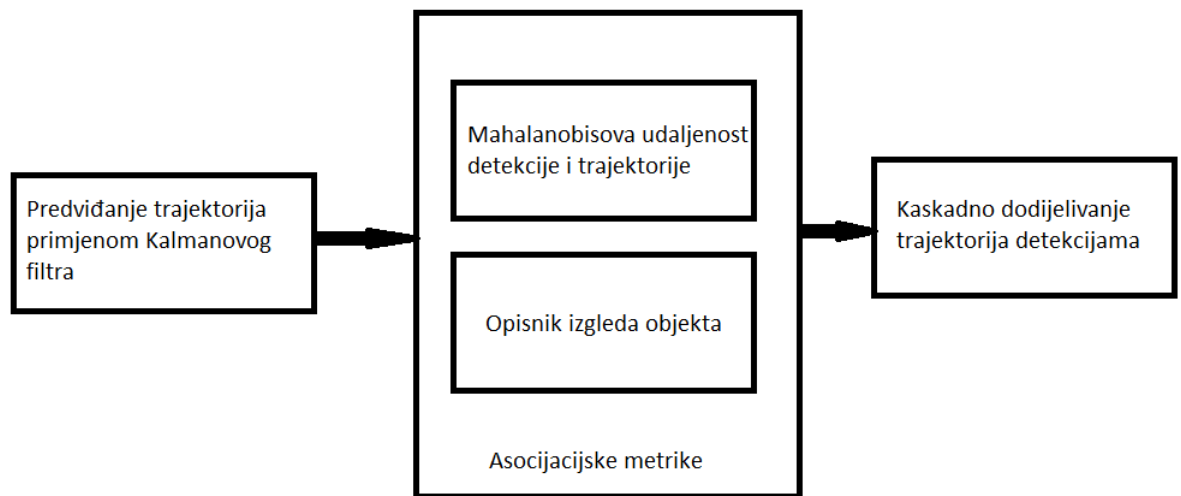
3. Arhitektura sustava

Sustav za praćenje nogometaša i generiranje mape aktivnosti sastoji se od 3 komponente:

- Podsustav za detekciju pozicija nogometaša u okviru
- Podsustav za praćenje nogometaša između okvira
- Podsustav za generiranje mape aktivnosti na temelju rezultata dobivenih praćenjem nogometaša

3.1. Detekcija nogometaša

Za detekciju nogometaša na slici koristi se YOLOv5s model iz skupa modela YOLOv5. Karakteristike YOLOv5 modela je da sliku podijeli na ćelije te je svaka ćelija zadužena za detektiranje objekata kojima se središte nalazi u toj ćeliji. Sustav na ulazu prima slike koje predstavljaju okvire videozapisa te na izlazu vraća skup detekcija objekata na slici. Model vraća uvijek isti broj detekcija te je potrebno ukloniti redundantne detekcije primjenom algoritma potiskivanja nemaksimalnih odziva (engl. non-maximum suppression). S obzirom da se na terenu većinu vremena nalazi 22 nogometaša i jedan sudac, očekivani broj detekcija u svakom okviru iznosi 23. Model za detekciju objekata služi isključivo za određivanje pozicija nogometaša u trenutnom okviru, odnosno ne može identificirati pojedine objekte jer ne uzima u obzir detekcije iz prethodnih okvira.



Slika 3.1: Arhitektura DeepSORT algoritma

3.2. Praćenje objekata

Za praćenje objekata koristi se algoritam DeepSORT.

Algoritam DeepSORT nastao je kao proširenje SORT algoritma koji se bazirao na korištenju Kalmanovog filtra za predviđanje trajektorija te mađarskog algoritma za dodjelu detekcija trajektorijama na temelju njihovog omjera presjeka i unije. Iako je takav algoritam imao veliku brzinu izvođenja, zamjena identiteta je bila česta pojava. Algoritam DeepSORT nastoji riješiti taj problem uvođenjem opisnika izgleda objekta koji se također uzima u obzir prilikom dodjele detekcija trajektorijama. Arhitektura mreže za dobivanje opisnika izgleda objekta prikazana je tablicom na Slici 3.2.

Algoritam također koristi Kalmanov filtar u svakoj trajektoriji za praćenje trenutnog stanja i predviđanje idućeg stanja, odnosno pozicije trajektorije u sljedećem okviru videozapisa. Predviđene trajektorije se uspoređuju sa detekcijama sa izlaza detektora objekata te se na temelju međusobne udaljenosti uparuju. Udaljenost se definira kao težinska suma Mahalanobisove udaljenosti trajektorije i detekcije te kosinusne udaljenosti opisnika izgleda detekcije i trajektorije.

Uparivanje se vrši primjenom kaskadnog uparivanja koje se bazira na iterativnoj primjeni Mađarskog algoritma pri čemu se prilikom uparivanja daje prioritet češće ažuriranim trajektorijama. Rezultat algoritma su uparene detekcije i trajektorije. Arhitektura sustava prikazana je na Slici 3.1.

Naziv	Vel. jezgre/Korak	Vel. izlaza
Conv 1	$3 \times 3/1$	$32 \times 128 \times 64$
Conv 2	$3 \times 3/1$	$32 \times 128 \times 64$
Max Pool 3	$3 \times 3/2$	$32 \times 64 \times 32$
Residual 4	$3 \times 3/1$	$32 \times 64 \times 32$
Residual 5	$3 \times 3/1$	$32 \times 64 \times 32$
Residual 6	$3 \times 3/2$	$64 \times 32 \times 16$
Residual 7	$3 \times 3/1$	$64 \times 32 \times 16$
Residual 8	$3 \times 3/2$	$128 \times 16 \times 8$
Residual 9	$3 \times 3/1$	$128 \times 16 \times 8$
Dense 10		128
Batch and ℓ_2 normalization		128

Slika 3.2: Arhitektura opisnika izgleda

3.2.1. Kalmanov filtar

Kalmanov filtar se u algoritmu DeepSORT koristi za praćenje internog stanja trajektorije te predviđanje pozicije objekta ukoliko detektor objekta ne uspije detektirati praćeni objekt u novom okviru. Interno stanje trajektorije se ažurira kada god je dostupan vektor mjerenja. Vektor mjerenja u kontekstu ovog algoritma jednak je jednoj detekciji objekta.

Algoritma DeepSORT stanje trajektorije definira kao 8-dimenzionalni vektor koji se sastoji od: x koordinate objekta, y koordinate središta objekta, omjera visine i širine, visine te brzina promjena za svaku od prethodno navedenih komponenti. Mjerenje se definira kao 4-dimenzionalni vektor koji se sastoji od: x koordinate pozicije objekta, y koordinate pozicije objekta, omjera visine i širine omeđujućeg pravokutnika detekcije te visine tog pravokutnika.

Za detekcije koje nisu uparene s postojećim trajektorijama prilikom procesa uparivanja, stvaraju se nove trajektorije koje su na početku u stanju pripravnosti, odnosno još se ne tretiraju kao rezultat praćenja objekata. Tek ukoliko se novostvorenoj trajektoriji dodijeli mjerenje kroz određen broj uzastopnih okvira, trajektorija prelazi u aktivno stanje te se tek tada tretira kao ispravna trajektorija. Na taj način se sprječava praćenje nepostojećih objekata [4].

Kada je trajektorija u aktivnom stanju te vektor mjerenja nije dostupan, odnosno detekcija nije dodijeljena, sljedeće stanje se predviđa pomoću internog stanja trajektorije [1]. Ukoliko se nekoj aktivnoj trajektoriji ne dodijeli mjerenje unutar određenog broja uzastopnih okvira, trajektorija i njeno stanje se briše. [4].

3.2.2. Dodjela trajektorija detekcijama

U svakom okviru potrebno je dodijeliti detekcije trajektorijama tako da parovi detekcija i trajektorija budu minimalno udaljeni. Algoritam DeepSORT ne pokušava upariti

sve detekcije i trajektorije odjednom već daje prednost trajektorijama kojima se interno stanje češće ažurira pomoću novih mjerenja, odnosno detekcija. Algoritam kreće od dodjelivanja nedodijeljenih detekcija najčešće ažuriranim trajektorijama.

Kriterij koji se razmatra prilikom uparivanja trajektorija i detekcija sastoji se od dvije komponente: kvadrirane mahalanobisove udaljenost između predviđene pozicije objekta (dobiva se primjenom Kalmanovog filtra) i detekcije te minimalne kosinusne udaljenosti između opisnika izgleda detektiranog objekta i prethodnih opisnika objekta za određenu trajektoriju. Za jednu trajektoriju se pamti određen broj opisnika izgleda detekcija koje su bile prethodno uparene s tom trajektorijom. Prethodno navedene udaljenosti se linearno kombiniraju u konačnu udaljenost koja se razmatra prilikom dodjele detekcija trajektorijama.

Trajektorije i detekcije s minimalnom udaljenošću uparuju uzimajući u obzir ažurnost trajektorija i prag udaljenosti koji mora biti zadovoljen. Rezultat uparivanja s ovim ograničenjima je da će neke trajektorije biti uparene s nekim detekcijama, ali i dalje mogu ostati neuparene detekcije i neuparene trajektorije. U idućem koraku se ponovo uzimaju nove neuparene trajektorije koje su rijeđe ažurirane od prethodno uparenih te se pokušavaju upariti s preostalim neuparenim detekcijama.

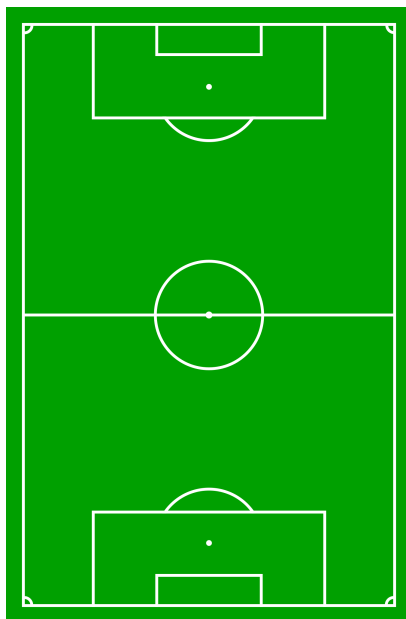
Kao parametar algoritma se postavlja maksimalni broj okvira, odnosno maksimalno vrijeme koje je prošlo od ažuriranja trajektorije te se ono koristi za određivanje koliko puta će pokušati dodijeliti neuparene detekcije neuparenim trajektorijama.

Ako nakon ovakvog iterativnog dodjeljivanja i dalje postoje neuparene detekcije i trajektorije, one se uparuju primjenom Mađarskog algoritma gdje se kao kriterij optimizacije prilikom dodjeljivanja koristi omjer presjeka i unije između područja detekcije i predviđenog područja trajektorije. Za ovo dodjeljivanje se također uvodi prag koji mora biti zadovoljen kako bi se uparivanje smatralo uspješnim[4].

Ukoliko i dalje postoje neuparene detekcije, za njih se stvaraju nove trajektorije koje zatim moraju proći fazu pripravnosti. Neuparene trajektorije ne dobivaju mjerenje u tom okviru te ukoliko su u aktivnom stanje, same predviđaju svoju poziciju primjenom kalmanovog filtra.

3.3. Izrada mape aktivnosti

Mape aktivnosti izrađuju se nakon analize cijelog videozapisa. Pozicije praćenih objekata zapisane su u tekstualnoj datoteci te se one zapisuju u skladu s MOT16 formatom. Mape aktivnosti se izrađuju za svakog nogometaša zasebno te se pohranjuju kao slike.



Slika 3.3: Nogometni teren iz ptičje perspektive

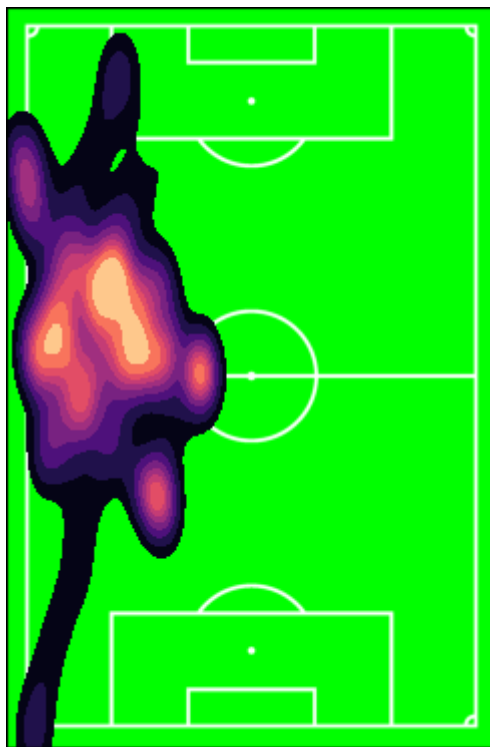
Za izradu mape aktivnosti, koristi se slika igrališta iz ptičje perspektive kao pozadina. Navedena slika prikazana je na Slici 3.3.

Na početku se za svakog igrača izrađuju se liste x i y koordinata njihovih pozicija na okvirima videozapisa. Pozicije igrača na okvirima videozapisa ne moraju odgovarati pozicijama na slici terena iz ptičje perspektive te je stoga potrebno provesti perspektivnu transformaciju. Za poziciju igrača odabrana je točka koja se nalazi na polovištu donje lijeve i donje desne točke omeđujućeg pravokutnika detekcije igrača.

Za dobivanje matrice perspektivne transformacije potrebno je zadati četiri točke koje predstavljaju pozicije rubova igrališta u videozapisu te četiri točke koje predstavljaju rubove igrališta na slici terena iz ptičje perspektive.

Na temelju tih točaka izračunava se transformacijska matrica pomoću koje se sve točke koje predstavljaju poziciju igrača na okviru videozapisa preslikavaju u poziciju igrača na terenu snimljenom iz ptičje perspektive. S obzirom da se pozicija kamere može razlikovati u videozapisima, potrebno je za svaki videozapis precizno zadati točke koje predstavljaju rubove terena, jer će u protivnom rezultirajuće mape aktivnosti biti netočne.

Jednom kada su sve točke pozicija u okvirima preslikane u odgovarajuće točke pozicija na terenu iz ptičje perspektive, izračunava se mapa aktivnosti. Pozicije nogometaša tretiraju se kao slučajne varijable koje imaju nepoznatu funkciju gustoće. Za procjenu te razdiobe koristi se procjena pomoću jezgre. Konačan rezultat je procjenjena funkcija gustoće pozicija koja ujedno predstavlja i mapu aktivnosti nekog igrača.



Slika 3.4: Mapa aktivnosti za igrača

Primjer jedne mape aktivnosti prikazan je na Slici 3.4

4. Implementacija sustava

4.1. Podsustav za detekciju nogometaša

Podsustav za detekciju nogometaša sastoji se YOLOv5 detektora objekata koji na ulazu prima slikovne okvire videozapisa, a na izlazu vraća detekcije sa prikladnim podacima o detekciji. Implementacija YOLOv5 detektora preuzeta je sa izvornog repozitorija [2]. U repozitoriju se nalaze skripte za učenje modela detekcije kao i za vrednovanje modela dobivenog učenjem. Dodatno se prilikom implementacije sustava praćenja nogometaša iz repozitorija iskorištavaju funkcije za potiskivanje nemaksimalnih odziva, označavanje izlaznih detekcija na slikama, dohvaćanje treniranog modela te provođenje detektiranja objekata u slikovnom okviru. Potiskivanje nemaksimalnih odziva (engl. non-maximum suppression) potrebno je provesti kako bi se uklonile redundantne detekcije, tj. detekcije niske pouzdanosti i višestruke detekcije.

Detektor objekata YOLOv5 zahtjeva skup podataka za učenje, provjeru i testiranje u specifičnom formatu koji se sastoji od slika te istoimenih tekstualnih datoteka za svaku sliku koje sadrže oznake pozicija objekata kao i njihove razrede. Pozicije objekata zadane su relativnim koordinatama tako da budu invarijantne na veličinu slike koja se dovodi na ulaz detektora. Omeđujući pravokutnik jednog objekta zadan je središtem tog pravokutnika, širinom i visinom tog pravokutnika. Postojeće oznake su zadane u MOT16 formatu koji definira omeđujuće pravokutnike definira pomoću lijeve gornje točke pravokutnika, širine i visine pri čemu su navedene komponente zadane u apsolutnim koordinatama.

Za potrebe izrade skupa podataka koji odgovara YOLOv5 formatu izrađena je skripta *dataset_builder.py*. Navedena skripta slijedno čita okvire u videozapisu i za njih stvara slike i prikladnu tekstualnu datoteku sa oznakama. Rezultat izvođenja ove skripte je skup slika koje predstavljaju okvire videozapisa i tekstualne datoteke za svaku sliku koje sadrže oznake objekata u pojedinim slikama.

Dodatno je izrađena skripta *dataset_splitter.py* koja slike i oznake dobivene skriptom *dataset_builder.py* dijeli u skup za učenje, provjeru i testiranje. Za potrebe trani-

ranja modela odabrano je prvih 4500 slika iz skupa podataka, za provjeru je korišteno idućih 1500 slika, a za testiranje je korišteno posljednjih 1515 slika.

4.2. Podsustav za praćenje nogometaša

Za praćenje nogometaša koristi se algoritam DeepSORT. Implementacija navedenog algoritma također je preuzeta s izvornog repozitorija [3][4]. Algoritam DeepSORT koristi značajke izgleda objekata prilikom izračuna udaljenosti detekcija i trajektorija.

Za izlučivanje značajki može se koristiti bilo koji model konvolucijske neuronske mreže koji na izlazu daje vektor značajki. Treniranje navedenog opisnika u implementaciji nije napravljeno, već je iskorišten izvorni model koji je naučen izlučivati značajke osoba.

U implementaciji sustava potrebne su 3 uzastopne detekcije dodijeljene trajektoriji prije nego što trajektorija postane aktivna. Najveća dozvoljena kosinusna udaljenost prilikom dodjele detekcija trajektorijama iznosi 0.4. Ako se nekoj trajektoriji ne dodijeli detekcija kroz uzastopnih 50 okvira, ta se detekcija briše. Jedna trajektorija prema posljednjih 50 vektora značajki izgleda objekata s kojima uspoređuje vektore značajki detekcija prilikom dodijeljivanja detekcija trajektorijama. Najmanji dozvoljeni omjer presjeka i unije između detekcije i trajektorije prilikom dodijeljivanja iznosi 0.1.

4.3. Podsustav za izradu mape aktivnosti nogometaša

Dio sustava zadužen za generiranje mapi aktivnosti implementiran je pomoću skripte *heatmap_builder.py*. Skripta za izradu mape aktivnosti koristi tekstualnu datoteku sa pozicijama igrača u MOT16 formatu iz koje se za svakog igrača stvara lista njegovih pozicija u videozapisu. Pozicija igrača se tretira kao slučajna varijabla, odnosno mapa aktivnosti predstavlja funkciju gustoće distribucije pozicije igrača. Navedena funkcija gustoće aproksimira se pomoću jezgara pri čemu pozicije igrača predstavljaju uzorke funkcije gustoće. Procjena funkcije gustoće definirana je formulom:

$$f_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right)$$

Za aproksimaciju funkcije i njen slikovni prikaz koristi se funkcija *kdeplot* iz biblioteke *Seaborn*. Slika 3.3 koja prikazuje nogometni teren iskorištena je kao pozadina. Na izlazu se generiraju slike koje sadrže mape aktivnosti pojedinih igrača na terenu.

4.4. Povezivanje podsustava u cijelinu

Glavna skripta *system.py* na početku inicijalizira posustav za praćenje objekata. Za inicijalizaciju podsustava potrebno je navesti putanju do datoteke s informacijama o modelu za izlučivanje značajki te je potrebno navesti parametre algoritma DeepSort navedene u potpoglavlju 4.2.

Nakon inicijalizacije sustava za praćenje, inicijalizira se sustav za detekciju objekata. Za inicijalizaciju navedenog sustava potrebno je navesti putanju do datoteke koja sadrži informacije o treniranom modelu te sklopovlje na kojem će se izvršavati detekcija nogometaša (u implementaciji se detekcija izvršavala na grafičkoj procesnoj jedinici).

Sustav na izlazu daje videozapis sa označenim nogometašima te je stoga potrebno stvoriti razrede za izradu videozapisa te se oni stvaraju nakon inicijalizacije sustava za detekciju objekata.

Nakon inicijalizacija podsustava iterativno se obrađuju okviri videozapisa tako da se za svaki okvir provodi detekcija objekata, potiskivanje nemaksimalnih odziva, dodjeljivanje detekcija trajektorijama i dodavanje okvira s rezultatima u videozapis.

4.5. Podsustav za vrednovanje praćenja i detekcije objekata

Za evaluaciju detekcije objekata korištena je skripta *test.py* iz implementacije YOLOv5 detektora. Navedena skripta na ulazu prima model i skup podataka nad kojima vrši vrednovanje. Skripta na temelju ulaza generira podatke koji opisuju kvalitetu modela.

Za evaluaciju praćenja objekata iskorišten je program TrackEval koji kao ulaz zahtjeva datoteku sa pozicijama pojedinih objekata koje je dao sustav te datoteku sa stvarnim pozicijama objekata. Program zahtjeva da obje datoteke sadrže pozicije i identitete objekata u MOT16 formatu.

5. Analiza rezultata

6. Zaključak

7. Literatura

- [1] Alex Bewley, Zongyuan Ge, Lionel Ott, Fabio Ramos, i Ben Upcroft. Simple online and realtime tracking. *2016 IEEE International Conference on Image Processing (ICIP)*, Sep 2016. doi: 10.1109/icip.2016.7533003. URL <http://dx.doi.org/10.1109/ICIP.2016.7533003>.
- [2] Arne Hoffhues Jonathon Luiten. Yolov5. <https://github.com/ultralytics/yolov5>, 2020.
- [3] Nicolai Wojke i Alex Bewley. Deep cosine metric learning for person re-identification. U *2018 IEEE Winter Conference on Applications of Computer Vision (WACV)*, stranice 748–756. IEEE, 2018. doi: 10.1109/WACV.2018.00087.
- [4] Nicolai Wojke, Alex Bewley, i Dietrich Paulus. Simple online and realtime tracking with a deep association metric, 2017.