Sveučilište u Rijeci

Fakultet informatike i digitalnih tehnologija

Sveučilišni diplomski studij informatike

**PRAĆENJE OBJEKTA**

**izvještaj**

IZRADIO: Marko Ribarić

siječanj, 2025. godine

Sadržaj

[1. UVOD 4](#_Toc189170318)

[2. DEEPSORT + YOLO 5](#_Toc189170319)

[3. CSRT 8](#_Toc189170320)

[4. FairMOT 10](#_Toc189170321)

[5. ZAKLJUČAK 12](#_Toc189170322)

[Popis slika 13](#_Toc189170323)

# UVOD

U ovom eksperimentalnom zadatku provedeno je istraživanje i usporedba triju različitih algoritama za praćenje objekata u video materijalu. Odabrana su četiri videozapisa (zbog raznovrsnosti) iz različitih okruženja, koji prikazuju aktivnosti skijanja, nogometa, plesanja i interakcije pasa. Cilj eksperimenta bio je testirati učinkovitost algoritama DeepSORT, CSRT i FairMOT u praćenju više objekata istovremeno pri čemu su se pratili objekti različitih vrsta, kao što su ljudi (skijaši, nogometaši, plesači) i životinje (psi).,

Odabrani algoritmi uključuju: DeepSORT, koji koristi duboke neuronske mreže za praćenje objekata uz pomoć detektora poput YOLO-a, CSRT (Channel and Spatial Reliability Tracking) koji je specifičan za praćenje u složenim uvjetima sa stabilnim performansama u raznim okruženjima, te FairMOT koji kombinira detekciju i praćenje u jedinstvenoj mreži.

# DEEPSORT + YOLO

YOLO (You Only Look Once) jedan je od najpoznatijih i najbržih algoritama za detekciju objekata, koji se temelji na dubokim neuronskim mrežama. Njegova glavna prednost je sposobnost da u samo jednom prolazu kroz mrežu istovremeno predvidi i detektira sve objekte u kadru, što omogućuje izuzetno brzu analizu videa u stvarnom vremenu. YOLO mreža dijeli sliku na mrežu manjih ćelija i za svaku od njih predviđa okvire (bounding boxes) i klase objekata, omogućujući prepoznavanje više objekata istovremeno.

DeepSORT (Deep Learning-based SORT) je prošireni verzija SORT algoritma koji koristi duboko učenje za poboljšanje preciznosti i robusnosti praćenja objekata u video materijalu. DeepSORT se oslanja na duboku neuronsku mrežu za dodatno prepoznavanje i identifikaciju objekata na temelju njihove vizualne osobitosti, što omogućava bolje praćenje objekata u izazovnim uvjetima.

DeepSORT je testiran na četiri različita videozapisa, a rezultati su pokazali varijacije u performansama. Najbolje rezultate postigao je u videozapisu nogometne utakmice, gdje je model uspješno pratio više igrača, rijetko gubeći pojedine subjekte. Također se vrlo dobro pokazao u videozapisu plesa, gdje je uspješno prepoznavao svih sedam plesača, ali se povremeno zbunio kada bi neka osoba bila zaklonjena drugom, dodjeljujući joj novi ID nakon što bi se ponovno pojavila. U videozapisu sa psom model je radio vrlo dobro, no postupno je gubio praćenje i pogrešno određivao dimenzije objekata. Međutim, dobro je prepoznavao pse jer je bio podešen da označava samo njih. Najslabije rezultate dao je u videozapisu skijanja, gdje je model često netočno predviđao kretanje skijaša. Budući da se u skijanju smjer kretanja često i naglo mijenja, model bi pri svakom skretanju nastavio pratiti prethodni pravac, što je dovodilo do učestalih pogrešaka.

Na slici 1 prikazan je jedan kadar DeepSORT praćenja tijekom nogometne utakmice.

A football game with green lines and dots

Description automatically generated with medium confidence

Slika DeepSort Nogometna utakmica

Na drugoj slici je prikazan jedan frame DeepSort tracking na videu plesanja.

A group of people on a stage

Description automatically generated

Slika DeepSort Plesanje

Na trećoj slici je prikazan jedan frame DeepSort tracking na videu skijanja.

A person skiing on a snowy hill

Description automatically generated

Slika DeepSort Skiing

Na četvrtoj slici je prikazan jedan frame DeepSort tracking-a na videu psi.

A person and dogs on a bed

Description automatically generated

Slika Psi - DeepSort

# CSRT

CSRT (Channel and Spatial Reliability Tracker) je algoritam za praćenje objekata u OpenCV-u. CSRT radi u tri glavne faze. Prvo, tijekom inicijalizacije, korisnik označava objekt koji želi pratiti ručno pomoću funkcije cv2.selectROI. Zatim, u fazi praćenja, algoritam analizira boje i teksture odabranog objekta te predviđa njegovu novu poziciju u sljedećem kadru videa. Na kraju, tijekom ažuriranja, ako se objekt pomakne ili promijeni oblik, CSRT prilagođava svoj model kako bi osigurao precizno i stabilno praćenje. CSRT je odličan za scenarije gdje je potrebna preciznost, ali nije najbolji izbor za stvarnovremenske (real-time) primjene zbog većeg opterećenja procesora.

U programskom kodu mijenjamo koliko objekta želimo označiti. U slučaju videa plesa bilo je 7 objekta. Označen je jedan po jedan plesač kao što je prikazano na slici 5.

A group of people on a stage

Description automatically generated

Slika Označavanje plesača

Nakon određenih brojeva frame-ova možemo vidjeti kako je uspio i dalje pratiti plesače na sljedećoj slici.

A group of men walking on a stage

Description automatically generated

Slika Plesanje – CSRT

Međutim nakon nekog vremena algoritam prestane dobro pratiti objekte kada plesači idu iza plesača te radi previše nagle pokrete.

Na videu nogometne akcije možemo vidjeti kako nakon par sekundi algoritam i dalje uspješno prati sva 3 igrača koje smo označili.

A football field with people in the background

Description automatically generated

Slika Nogomet - CSRT

Na primjeru psi i skijaša se CSRT jako brzi izgubi zbog čestih promjena scena.

# FairMOT

FairMOT je napredni algoritam za praćenje više objekata u videozapisima, koji koristi jednu mrežu za simultano detektiranje objekata i praćenje njihovog pomaka. Glavna prednost FairMOT-a je što istovremeno detektira i prati objekte u realnom vremenu, što ga čini brzim i preciznim, čak i u zagušenim scenama s mnogo objekata. FairMOT je izuzetno učinkovit za primjene poput video-nadzora, autonomnih vozila i analize sportskih događanja.

Za FairMOT nije napisan nikakav kod nego je praćen tutorial na FairMOT githubu (https://github.com/ifzhang/FairMOT) kako bi procesirao video kroz algoritam.

Algoritam je uglavnom točno prepoznavao igrače i pratio njihovo kretanje, ali je ponekad prestao pratiti igrača te neke igrače uopće nije detektirao. Jedan frame je prikazan na sljedećoj slici.

A group of people on a football field

Description automatically generated

Slika Nogomet – FairMOT

Na videu plesača je model radio jako dobro gubeći plesača samo kada se plesač jedva vidi golim okom kao što je prikazano na sljedećoj slici.

A group of people on stage

Description automatically generated

Slika Plesanje – FairMOT

Za videje psa i skijaša je model bio vrlo loš te nije dobro prepoznavao objekte te nije bio uspješan u praćenju njih. Za skiajenj je detektirao svako malo, ali i dalje dosta gubio.

# ZAKLJUČAK

DeepSORT je dao najbolje rezultate na videu nogometne utakmice, gdje je uspješno pratio više objekata, dok je na videu skijanja imao poteškoća zbog čestih promjena smjera kretanja. CSRT se dobro ponašao u uvjetima sa stabilnim objektima, ali je imao poteškoće u dinamičnim scenama, poput skijanja i plesanja. FairMOT je bio vrlo učinkovit na videu plesa, gdje je uspješno pratio objekte, no imao je problema u praćenju objekata na videu skijanja i sa psima.

Ovi rezultati ukazuju na to da izbor algoritma ovisi o specifičnostima zadatka. Algoritmi poput DeepSORT i FairMOT pokazuju bolje rezultate u dinamičnim scenama i kada je potrebno pratiti više objekata u stvarnom vremenu, dok CSRT može biti bolji za precizno praćenje u stabilnijim uvjetima.

# Popis slika

[Slika 1 DeepSort Nogometna utakmica 6](#_Toc189170301)

[Slika 2 DeepSort Plesanje 6](#_Toc189170302)

[Slika 3 DeepSort Skiing 6](#_Toc189170303)

[Slika 4 Psi - DeepSort 7](#_Toc189170304)

[Slika 5 Označavanje plesača 8](#_Toc189170305)

[Slika 6 Plesanje – CSRT 9](#_Toc189170306)

[Slika 7 Nogomet - CSRT 9](#_Toc189170307)

[Slika 8 Nogomet – FairMOT 10](#_Toc189170308)

[Slika 9 Plesanje – FairMOT 11](#_Toc189170309)