|  |  |
| --- | --- |
|  | Zbornik radova Fakulteta tehničkih nauka, Novi Sad |

**UDK:** **621.39**

**DOI:** [**https://doi.org/10.24867/10BE29Mrdjen**](https://doi.org/10.24867/10BE29Mrdjen)

**PRAĆENJE POKRETNIH OBJEKATA U VIDEU KORIŠĆENJEM PASIVNOG FILTRA I DETEKTORA VISOKIH PERFORMANSI**

**TRACKING OF MOVING OBJECTS IN VIDEO USING PASSIVE FILTER AND HIGH PERFORMANCE DETECTOR**

Jelena Mrđen, *Fakultet tehničkih nauka, Novi Sad*

**Oblast – ELEKTROTEHNIČKO I RAČUNARSKO INŽENJERSTVO**

**Kratak sadržaj –** *U ovom radu je razmatran problem praćenja pokretnih objekata u videu i implementirani su elementi sistema za praćenje koji su zasnovani na principu pasivnog filtriranja detekcija u svakom kadru video snimka. Korišćenjem DETRAC protokola kroz numeričke eksperimente određeni su optimalni parametri konfiguracije IOU algoritma sa YOLOv3 detektorom.*

**Ključne reči:** *Praćenje objekata; IOU algoritam; YOLOv3 detekor; DETRAC evaluacija.*

**Abstract** – *The paper describes the problem of tracking moving objects in video and implements system components based on passive filtering of detections in each frame. Through numerical experiments based on DETRAC evaluation protocol the best parameters for the configuration of IOU algorithm and YOLOv3 object detector were determined.*

**Keywords:** *Object tracking; IOU algorithm; YOLOv3 object detector; DETRAC evaluation protocol.*

**1. UVOD**

Praćenje pokretnih objekata u videu predstavlja jedan od klasičnih problema kojim se bavi kompjuterska vizija. Rešenje ovog problema podrazumeva detektovanje objekta od interesa u svakom kadru video sekvence, prepoznavanje njegovih karakteristika, kao i dalje praćenje datog objekta kroz celokupnu sekvencu. Shodno tome se problem praćenja objekata može podeliti na tri zasebna zadatka: detekciju objekata, koja podrazumeva uočavanje objekta od interesa i određivanje njegovog položaja u sceni, prepoznavanje objekta, koje daje uvid u konkretan tip i karakteristike objekta koji je detektovan, a zatim i praćenje objekata čiji je krajnji cilj određivanje putanje kretanja objekta kroz celokupnu video sekvencu.

Cilj ovog rada je razmatranje problema praćenja pokretnih objekata u videu i implementacija elemenata sistema koji omogućavaju praćenje, a koji su zasnovani na pasivnom filtriranju detekcija na nivou pojedinačnih kadrova. U ovom radu je predstavljen sistem za praćenje velikog broja objekata u videu*.* Ovaj sistem upotrebljava IOU (*engl. Intersection over Union)* algoritam za aso-\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

**NAPOMENA:**

**Ovaj rad proistekao je iz master rada čiji mentor je bio dr Branko Brkljač, docent.**

cijaciju detekcija koji je zasnovan na pasivnom filtru dok je u cilju generisanja detekcija, koje mu daju osnovnu funkcionalnost, upotrebljen YOLOv3 (*engl. You Only Look Once, version 3)* detektor. Cilj ovog rada jeste, po­red implementacije elemenata sistema koji omogućavaju praćenje pokretnih objekata, određivanje optimalnih parametara ovog sistema kroz evaluaciju njegovih performansi kao i analiza uticaja YOLOv3 detektora na rad algoritma. Evaluacija performansi sistema je izvršena pomoću DETRAC (*engl. DeTection And TraCking)* protokola za evaluaciju nad 60 zahtevnih video sekvenci koje prikazuju realne scene iz saobraćaja.

**2. YOLOv3 DETEKTOR**

Cilj algoritama za detekciju objekata upotrebom 2D podataka tj. slika u boji je lociranje i klasifikacija objekata od interesa na slikama ili video snimcima koji se mogu posmatrati kao sekvenca pojedinačnih statičnih slika. Kao rezultat algoritama za detekciju se dobija okvir koji obuhvata objekat od interesa, poznatiji pod nazivom *engl. Bounding box*, klasa objekta koji je detektovan kao i verovatnoća da objekat zapravo postoji u datom okviru, koji se takođe naziva i skor sigurnosti.

YOLOv3 je neuronska mreža za detekciju objekata koja omogućava procesovanje detekcija u realnom vremenu. Ova mreža predviđa četiri koordinate za svaki okvir koji definiše objekat. Ukoliko se predlog okvira objekta idealno poklapa sa stvarnim tj. anotiranim okvirom objekta (*engl. Ground Truth – GT)*, skor sigurnosti treba da bude jednak jedinici.

Ukoliko je ovo preklapanje manje od vrednosti 0.5 ta predikcija se zanemaruje. Takođe, mreža definiše klasu objekta koji je zahvaćen okvirom upotrebom klasifikacije u više klasa pomoću logističkog klasifikatora. Predikcija okvira objekata se vrši na tri različite skale i ulazi ove mreže mogu biti rezolucije 320x320, 416x416 i 608x608 piksela

Što je rezolucija ulazne slike veća, detekcije su tačnije tj. imaju veći skor sigurnosti i lokalizacija objekta u okviru je bolja, ali se vreme generisanja detekcija povećava u skladu sa povećanjem rezolucije ulaznih slika.

Za izvlačenje obeležja upotrebljava se hibridna *Darknet-53* mreža koja se sastoji od 53 sloja i trenirana je na *ImageNet* bazi podataka. Iako ova mreža u poređenju sa drugim kompleksnijim mrežama za detekciju objekata daje nižu mAP (*engl. mean Average Precision*)vrednost brzina izvršavanja ove mreže je znatno veća u odnosu na druge metode. Detaljan opis arhitekture i principa rada YOLOv3 mreže se može pronaći u [1].

**3. IOU ALGORITAM ZA PRAĆENJE OBJEKATA U VIDEU**

Ovaj rad je zasnovan na IOU algoritmu za praćenje objekata koji je predstavljen u radu [2]. Za svaki kadar video sekvence se prvobitno generišu detekcije a zatim se primenjuje algoritam za asocijaciju detekcija. Upotreba detektora velike preciznosti i video snimaka koji imaju veliki broj kadrova u sekundiznačajno olakšava proces praćenja objekata. IOU algoritam polazi od pretpostavke da detektor za svaki objekat koji se prati daje odgovarajuće detekcije kao i da se javlja veoma mali broj promašenih detekcija tj. lažnih negativa.

Takođe se pretpostavlja da detekcije istog objekta u su­sednim kadrovima imaju veliki IOU indeks, što je uobi­čajeno za video sekvence koje imaju veliki broj kadrova u sekundi. Ukoliko su sve pretpostavke ispunjene, praćenje objekata u se može izvršiti bez izvlačenja obeležja objekata koji se prate. IOU algoritam ostvaruje praćenje objekata asociranjem detekcija na osnovu IOU indeksa trenutne detekcije i detekcije iz prethodnog kadra, ukoliko je taj indeks veći od praga $σ\_{IOU}$. Sve detekcije koje nisu dode­ljene trenutnoj putanji započinju novu putanju, dok se sve putanje koje nemaju novu asociranu detekciju smatraju završenim. Performanse IOU algoritma su unapređene filtriranjem svih putanja koje su kraće od vrednosti $t\_{min}$ kao i onih putanja koje nemaju nijednu detekciju čiji je skor sigurnosti veći od praga $σ\_{h}$. Kratke putanje se uklanjaju jer dovode do lažnih pozitiva.

Zahtev da svaka putanja ima barem jednu detekciju visokog skora sigurnosti obezbeđuje da putanja pripada pravom objektu od interesa dok istovremeno koristi detekcije malog skora sigurnosti u cilju generisanja kompletnih putanja objekata.

Ulaz u algoritam čini skup detekcija objekata koji se pojavljuju u kadrovima video sekvence, što znači da je pre primene ovog algoritma neophodno prikupiti detekcije svih objekata čije putanje želimo pratiti.

IOU algoritam je predstavljen slikom 1 na kojoj $D^{f}$ označava detekcije u kadru $f$, $d\_{j}$ označava *j*-tu detekciju u tom kadru, $T\_{a}$ označava aktivne putanje, $T\_{f}$ završene putanje dok F označava ukupan broj kadrova u video sekvenci.

Prilikom incijalizacije se vrši filtriranje detekcija čiji je skor sigurnosti manji od $σ\_{l}$ praga čime se, pre samog izvršavanja algoritma, vrši eliminacija detekcija malih skorova sigurnosti koje bi potencijalno dovele do lažnih pozitiva. Sveukupna kompleksnost ovog algoritma je veoma mala u odnosu na druge popularne sisteme za praćenje objekata.

Kako ovaj algoritam ne koristi nikakve vizuelne infor­macije, može se posmatrati kao jednostavan proces filtri­ranja na nivou detekcija. Ovo znači da je pri online upo­trebi IOU algoritma za praćenje, u kombinaciji sa detek­torom visokih performansi, računska zahtevnost ovog algoritma zanemarljiva.

Ukoliko se algoritam izvršava oflajn, sa prethodno proce­sovanim detekcijama, se može postići izvršavanje od 100 kadrova u sekundi (*engl. Frames per second – fps)*.

Takođe je bitno napomenuti da se, zahvaljujući brzini ovog algoritma, na ovu metodu mogu dodati i druge komponente za praćenje objekata, [1].



Slika 1. *IOU algoritam za praćenje objekata u videu [1]*

**4. DETRAC PROTOKOL ZA EVALUACIJU**

Evaluacija performansi sistema za praćenje objekata je neophodna radi procene uspešnosti njegovog rada. Tako­đe je veoma značajna kao metoda koja omogućava pore­đenje različitih sistema za praćenje i njihovu standardiza­ciju. Većina sistema za evaluaciju podrazumeva zasebnu evaluaciju performansi sistema za detekciju objekata i sistema za praćenje objekata. Sistemi za praćenje objekata se obično evaluiraju upotrebom iste metode za detekciju objekata, a samim tim i istih detekcija. Iako je ova strategija opšte prihvaćena u literaturi i daje koristan uvid u sisteme za praćenje velikog broja objekata u video sekvencama, obično je nedovoljna za potpunu analizu ovih sistema. Konkretno, veoma je bitno razumeti uticaj performansi detektora na celokupan sistem za praćenje objekata. Ovo je moguće ostvariti primenom sistema za evaluaciju koji zajedno analizira rad detektora objekata i sistema za praćenje objekata, [3]. Jedan od ovakvih sistema jeste DETRAC protokol za evaluaciju. DETRAC protokol za evaluaciju je zasnovan na DETRAC skupu podataka koji se sastoji od 100 video sekvenci. Sve sekvence su snimane sa frekvencijom od 25fps pri čemu rezolucija svakog pojedinačnog kadra iznosi 960x540 piksela. Postoji više od 140 hiljada kadrova, 8,250 ručno anotiranih motornih vozila i 1.21 miliona anotiranih okvira objekata u DETRAC skupu podataka. DETRAC skup podataka je podeljen na skup za obuku i skup za testiranje, koji se sastoje od 60 i 40 video sekvenci, respektivno.

U osnovne metrike za evaluaciju peformansi algoritama za asocijaciju detekcija spadaju, [3]: primarno praćeni objekti (*engl. Mostly Tracked – MT*), primarno promašeni objekti (*engl. Mostly Lost – ML*), promene identiteta (*engl. Identity Switches - IDS*), fragmenti trajektorija objekata (*engl. Fragmentations of Target trajectories – FM*), lažni pozitivi (*engl. False positives – FP*), lažni negativi (*engl. False negatives – FN*), tačnost (*engl. Multi-Object Tracking Accuracy - MOTA*) i preciznost (*engl. Mult- Object Tracking Precision - MOTP*). FP metrika predstavlja broj izlaza algoritma za asocijaciju detekcija koji predstavljaju lažne pozitive dok FN predstavlja broj objekata koji su promašeni od strane bilo koje trajektorije u svakom kadru tj. broj lažnih negativa. IDS metrika opisuje koliko puta se odgovarajuća oznaka identiteta objekta koji se prati, tj. njegove trajektorije, promenila. FM definiše koliko puta je trajektorija objekta koji se prati prekinuta. IDS i FM metrike definišu tačnost praćenih trajektorija. ML i MT metrike opisuju procenat praćenih trajektorija koji je manji od 20% i veći od 80% ukupne anotirane trajektorije objekta, respektivno.

Na osnovu opisanih osnovnih metrika koje se upotreblja­vaju u cilju evaluacije performansi algoritama za asocija­ciju detekcija se za sve sekvence skupa podataka izraču­navaju DETRAC MOT metrike. DETRAC MOT metrike uzimaju u obzir rad detektora i rad algoritma za asoci­jaciju detekcija i definisane su u zavisnosti od krive pre­ciznosti-odziva detektora (*engl. precision-recal*), u zajed­ničkoj oznaci PR, i uvode PR-MOTA, PR-MOTP, PR-MT, PR-ML, PR-IDS, PR-FM, PR-FP i PR-FN metrike. Performanse algoritama za praćenje objekata se rangiraju spram PR-MOTA vrednosti tj. veća PR-MOTA vrednost nagoveštava bolje performanse algoritma. Ukoliko su PR-MOTA vrednosti dva algoritma jednake, dalje rangiranje se vrši spram PR-MOTP vrednosti.

**5. REZULTATI**

Krajnji cilj ovog istraživanja je određivanje optimalnih parametara konfiguracije IOU algoritma i YOLOv3 de­tektora uz upotrebu DETRAC protokola za evaluaciju, po uzoru na rad [1]. U [1] su upotrebljene četiri konfiguracije IOU algoritma i njegove performanse su testirane uz pomoć četiri različita detektora a to su R-CNN, CompACT, EB i ACF detektori. Svaka od ove četiri konfiguracije je testirana za različite opsege parametara $σ\_{IOU}, σ\_{h}$ i $t\_{min} .$ Po uzoru na [1] je konfiguracija IOU algoritma i YOLOv3 detektora takođe testirana za razli­čite vrednosti ovih parametara u cilju pronalaženja njiho­vih optimalnih vrednosti. Bitno je naglasiti da su svi detektori upotrebljeni u [1] prvobitno obučeni nad DETRAC skupom za obuku nad kojim su takođe odre­đene i vrednosti njihovih optimalnih parametra. Upo­trebljeni YOLOv3 detektor je obučen nad COCO skupom podataka i prepoznaje 80 klasa objekata. Kako je neop­hodno prvobitno prikupiti detekcije svih objekata koje želimo pratiti, a u DETRAC skupu podataka su od inte­resa samo klase koje definišu motorna vozila, prilikom generisanja detekcija za DETRAC skup za obuku upo­trebom YOLOv3 detektora su uzete u obzir samo objekti klasa ‘*car’, ‘bus’* i *‘truck’*. Svaki pojedinačan kadar video sekvence, koji ujedno predstavlja i ulaz YOLOv3 mreže, je smanjen na rezoluciju 608x608 piksela. Primer rada konfiguracije IOU algoritma i YOLOv3 detektora na sekvenci DETRAC skupa za obuku je dat na slici 2.

Nakon generisanja detekcija izvršena je evaluacija perfor­mansi konfiguracije IOU algoritma i YOLOv3 detektora po DETRAC protokolu. DETRAC protkol za evaluaciju integrisan je u *eng. MATLAB* okruženju i može se preuzeti kao celokupan sistem sa [4].



Slika 2. *Vizuelizacija rada IOU algoritma i YOLOv3 detektora*

U Tabeli 1 dati su opsezi parametara koji su upotrebljeni u cilju određivanja optimalnih vrednosti parametara konfiguracije IOU algoritma i YOLOv3 detektora.

Tabela 1*. Opsezi parametara za svaki detektor*

|  |  |  |  |
| --- | --- | --- | --- |
| Detektor | $$σ\_{IOU}$$ | $$σ\_{h}$$ | $$t\_{min}$$ |
| YOLOv3 | 0,4 – 0,7 | 0,4 – 0,9 | 1 - 4 |

Kako bi se izvršila evaluacija sistema je neophodno prvenstveno evaluirati peformanse detektora a zatim izvršiti evaluaciju performansi algoritma za asocijaciju detekcija. Kao rezultat evaluacije performansi detektora se dobija PR kriva i prosečna preciznost detektora za dati skup podataka, dok se kao krajnji rezultat evaluacije performansi algoritma za asocijaciju detekcija dobijaju DETRAC MOT metrike opisane u poglavlju 4. Na Slici 3 je prikazana PR kriva YOLOv3 detektora nad DETRAC skupom za obuku. Njegova prosečna preciznost nad ovim skupom iznosi 70.23%.



Slika 3. *PR kriva YOLOv3 detektora nad DETRAC skupom za obuku*

*Tabela 2. Rezultati evaluacije konfiguracije IOU algoritma i YOLOv3 detektora*

|  |  |
| --- | --- |
| Detektor | YOLOv3 |
| $$σ\_{IOU}$$ | 0,4 |
| $$σ\_{h}$$ | 0,9 |
| $$t\_{min}$$ | 2 |
| PR-MOTA | 22,32% |
| PR-MOTP | 29,06% |
| PR-MT | 15,42% |
| PR-ML | 22,85% |
| PR-IDS | 3108,81 |
| PR-FM | 3224,52 |
| PR-FP | 9798,44 |
| PR-FN | 160200,22 |
| Brzina u FPS | 13,828 |

U Tabeli 2 date su vrednosti optimalnih parametara IOU algoritma za YOLOv3 detektor kao i vrednosti DETRAC MOT metrika koje su ostvarene njihovom upotrebom. Konfiguracija IOU algoritma sa YOLOv3 detekcijama daje najveću, a samim tim i najbolju, PR-MOTA vrednost od 22.32% za kombinaciju parametara $σ\_{IOU}=0.4,$ $σ\_{h}=0.9$ i $ t\_{min}=2$. Na Slici 4 je prikazan grafik koji omogućava poređenje PR-MOTA vrednosti dobijenih za različite kombinacije ovih parametara.

Za svaku kombinaciju pragova $σ\_{IOU}$ i $σ\_{h}$ određena je vrednost $t\_{min}$ koja daje najbolje PR-MOTA vrednosti. Tačke označene crvenom bojom predstavljaju rezultat evaluacije performansi algoritma dok broj iznad njih označava optimalnu vrednost parametra $t\_{min}$. Sa ove slike se može uočiti da se najbolje PR-MOTA vrednosti dobijaju ukoliko je parametar $σ\_{IOU}$ postavljen na vrednost od 0.4. Sa daljim rastom parametra $σ\_{IOU}$ dolazi do pada PR-MOTA vrednosti dok sa rastom parametra $σ\_{h} $dolazi do suptilnog rasta PR-MOTA vrednosti.



Slika 4. *PR-MOTA vrednosti za različite kombinacije parametara* $σ\_{IOU}$ *i* $σ\_{h}$*, predložene konfiguracije IOU algoritma i YOLOv3 detektora*

Na slici 5 prikazane su MOTA vrednosti koje se dobiju upotrebom optimalnih parametara konfiguracije IOU al­goritma i YOLOv3 detektora za različite vrednosti praga $σ\_{l}. $ Radi lakšeg poređenja rezultata su na ovoj slici prikazane i MOTA vrednosti za različite detektore upot­rebljene u [1]. Isprekidane linije predstavljaju odgovara­juće PR-MOTA vrednosti.



Slika 5. *MOTA vrednosti za različite vrednosti praga sigma* $σ\_{l}$ *razmatrane konfiguracije*

Sa Slike 4 može se uočiti da se koleno krive, koja pred­stavlja promenu MOTA vrednosti u skladu sa promenom parametra $σ\_{l}$ za YOLOv3 detektor, javlja kada je parametar $σ\_{l}$ izmedju vrednosti 0.8 i 0.9. Ovo naglašava da sa porastom $σ\_{l}$ preko ovih vrednosti dolazi do pada performansi sistema. U poređenju sa krivama ostalih detektora iz [1] se može uočiti da je PR-MOTA vrednost za konfiguraciju IOU algoritma i YOLOv3 detektora manja u odnosu na konfiguracije iz rada [1]. Manje PR-MOTA vrednosti mogu se pripisati činjenici da YOLOv3 detektor nije prethodno obučen nad DETRAC skupom podataka, nad kojim su takođe određene i vrednosti njegovih optimalnih parametara, dok su svi detektori upotrebljeni u [1] prvobitno obučeni nad ovim podacima.

**6. ZAKLJUČAK**

Ovaj rad kombinuje jednostavan algoritam za praćenje objekata u videu i detektor objekata visokih performansi koji omogućava da se praćenje objekata vrši u realnom vremenu. Iako predstavljeni rezultati ukazuju na mogućnost primene predloženog sistema za praćenje, takođe postoji mogućnost značajnog unapređenja upotrebljene konfiguracije IOU algoritma i YOLOv3 detektora. Naime, kako upotrebljeni YOLOv3 detektor nije prethodno obučavan nad DETRAC skupom podataka, dalje unapređenje predloženog sistema za praćenje bi se moglo ostvariti ukoliko bi se detektor prvobitno obučio nad pomenutim skupom.

Ovo unapređenje bi potencijalno moglo povećati tačnost celokupnog sistema a samim tim dovesti do boljih rezultata. Takođe, u cilju testiranja rada ove konfiguracije i potencijalnog poboljšanja prikazanih rezultata bi trebalo razmotriti optimalnu rezoluciju slika koje se dovode na ulaz YOLOv3 detektora prilikom generisanja detekcija za IOU algoritam.

**4. LITERATURA**

[1] J. Redmon and A. Fahradi, *“YOLOv3: An Incremental Improvement“,* 2018.

[2] E. Bochinski, V. Eiselein and T. Sikora,
*“High-Speed Tracking-by-Detection Without Using Image Information”*, 2017.

[3] L. Wena et al., *“*UA-DETRAC: A *New Benchmark and Protocol for Multi-Object Detection and Tracking”,* 2020.

[4] *"The UA-DETRAC Benchmark Suite*", Detrac-db.rit.albany.edu, 2020. [Online]. Dostupno na: <http://detrac-db.rit.albany.edu/download>.[Pristupljeno: 08- Sep- 2020].

**Kratka biografija:**

|  |  |
| --- | --- |
|  | **Jelena Mrđen** rođena je u Novom Sadu 1996. god. Master rad na Fakultetu tehničkih nauka iz oblasti Energetika, elektronika i telekomunikacije – Obrada signala odbranila je 2020.god. Osnovne akademske studije završila je 2019. godine na studijskom programu Biomedicinsko inženjerstvo.kontakt: jelenamrdjen3@gmail.com |