

Sveučilište u Rijeci
Fakultet Informatike i Digitalnih Tehnologija



Projektni zadatak iz kolegija

Meko računarstvo

Detekcija prtljage

Autor: Ivan Vrsalović

Rijeka, 14.02.2024

Sadržaj

Uvod	3
YOLOv8	4
Test YOLOv8m modela	5
Fine-tuning YOLOv8.....	7
Priprema podataka.....	7
Inicijalizacija modela	8
Treniranje na novim podacima	8
Validacija i evaluacija	9
Opis rezultata treniranja modela na pronađenim/gotovim skupovima podataka.....	10
Označavanje vlastitih skupova podataka	11
Usporedba modela detekcije	12
Gotovi skupovi podataka.....	12
YOLOv8-S.....	12
YOLOv8-M	13
YOLOv8 retrainiran na slikama COCO skupu podataka samo za prtljagu i ljude	13
YOLOv8 retrainiran na luggage image skupu podataka	14
Vlastiti skupovi podataka	14
Model treniran na slikama sa korza	14
Model treniran na aerodromu	15
Zaključak.....	16
Skupovi podataka	16

Uvod

Detekcija prtljage predstavlja važan segment sigurnosnih sustava u različitim javnim prostorima, poput aerodroma. Razvoj učinkovitih tehnika detekcije prtljage ključan je za osiguranje sigurnosti i sprječavanje incidenata. U ovom istraživanju fokusiramo se na analizu i evaluaciju dostupnih datasetova s ciljem unapređenja detekcije prtljage korištenjem dubokog učenja.

Kroz raznolike izvore podataka poput online platformi i istraživačkih zajednica, pružaju se različiti datasetovi koji obuhvaćaju različite aspekte detekcije prtljage. Ovaj rad usredotočuje se na pregled dostupnih datasetova, analizu njihove raznolikosti, kvalitete i uporabljivosti za detekciju prtljage.

Cilj ovog istraživanja je identificirati najprikladnije datasetove za detekciju prtljage te istaknuti njihove prednosti i nedostatke. Na temelju analize dostupnih datasetova, planiramo donijeti preporuke za korištenje odabranih datasetova u razvoju i evaluaciji detekcijskih modela prtljage.

Detekcije će biti implementirane pomoću YOLOv8 modela.

YOLOv8

YOLOv8, skraćenica za "You Only Look Once version 8", predstavlja najnoviju iteraciju popularne arhitekture dubokog učenja koja se koristi za detekciju objekata u slikama. Ova arhitektura ističe se visokom brzinom izvođenja i iznimnom preciznošću u detekciji, čineći je idealnim izborom za zadatke detekcije objekata u stvarnom vremenu.

YOLOv8 koristi tehnologiju tzv. "single neural network", što znači da jedna neuronska mreža obrađuje cijelu sliku odjednom i izvodi detekciju objekata u realnom vremenu. Ova tehnika omogućuje izuzetno brzo i efikasno pretraživanje slike, što rezultira minimalnim vremenom obrade i visokom stopom detekcije.

Model YOLOv8 je pretreniran na COCO (*Common Objects in Context*) datasetu. COCO dataset je jedan od najpoznatijih i najkorištenijih skupova podataka u području detekcije objekata. Sadrži više od 200 tisuća slika s preko milijun označenih objekata u 80 različitih kategorija, među kojima su i nama zanimljive „Backpack“, „Suitcase“, „Handbag“ i „Person“ koje opisuju prtljagu i osobe. Prije započinjanja postupka fine-tunninga YOLOv8 modela, potrebno je provjeriti performanse već pretreniranog modela kroz testiranje njegove sposobnosti prepoznavanja i detekcije objekata u različitim scenarijima.

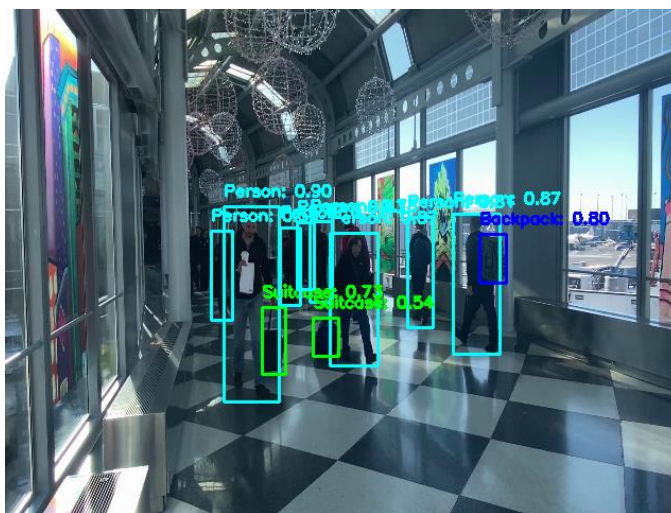
YOLOv8 dolazi u nekoliko varijanti, uključujući YOLOv8-S, YOLOv8-M, YOLOv8-L, i YOLOv8-XL, gdje svaka verzija označava različite veličine i složenost modela. Glavne razlike među ovim varijantama su veličina modela, brzina izvršavanja i preciznost detekcije.

- **YOLOv8-S:** Ovo je najmanja verzija modela, koja se ističe brzim izvođenjem i manjim brojem parametara. Iako je brza, može imati ograničenu preciznost u odnosu na ostale verzije.
- **YOLOv8-M:** Ova srednja varijanta nudi balans između brzine izvršavanja i preciznosti detekcije. Sadrži više parametara i složenosti od YOLOv8-S, što rezultira boljom preciznošću, ali može biti sporija u odnosu na YOLOv8-S.
- **YOLOv8-L:** Ova verzija ima veći broj parametara i složenosti u odnosu na prethodne dvije. To rezultira još većom preciznošću u detekciji, ali može biti sporija u izvršavanju zbog većeg broja parametara.
- **YOLOv8-XL:** Ova varijanta je najveća i najkompleksnija verzija modela, s najviše parametara i složenosti. To omogućava najbolju preciznost u detekciji objekata, ali može biti najsporija u izvršavanju.

U ovom radu korišteni su YOLOv8-s i YOLOv8-m modeli.

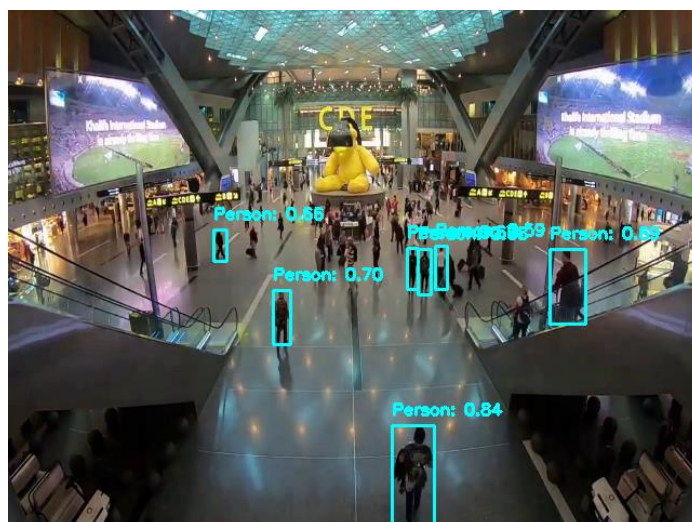
Test YOLOv8m modela

Već preuzeti model pokazuje solidnu sposobnost detekcije ljudi i prtljage kada su objekti snimljeni iz normalne perspektive ili iz blizine. Detekcija objekata, kao što su ljudi i prtljaga, izvodi se s visokom preciznošću kada su slike snimljene pod kutom koji je uobičajen za svakodnevne situacije. Međutim, kako se kut gledanja slike mijenja i postaje sve više sličan ptičjoj perspektivi, model počinje pokazivati sve veće izazove u detekciji objekata.



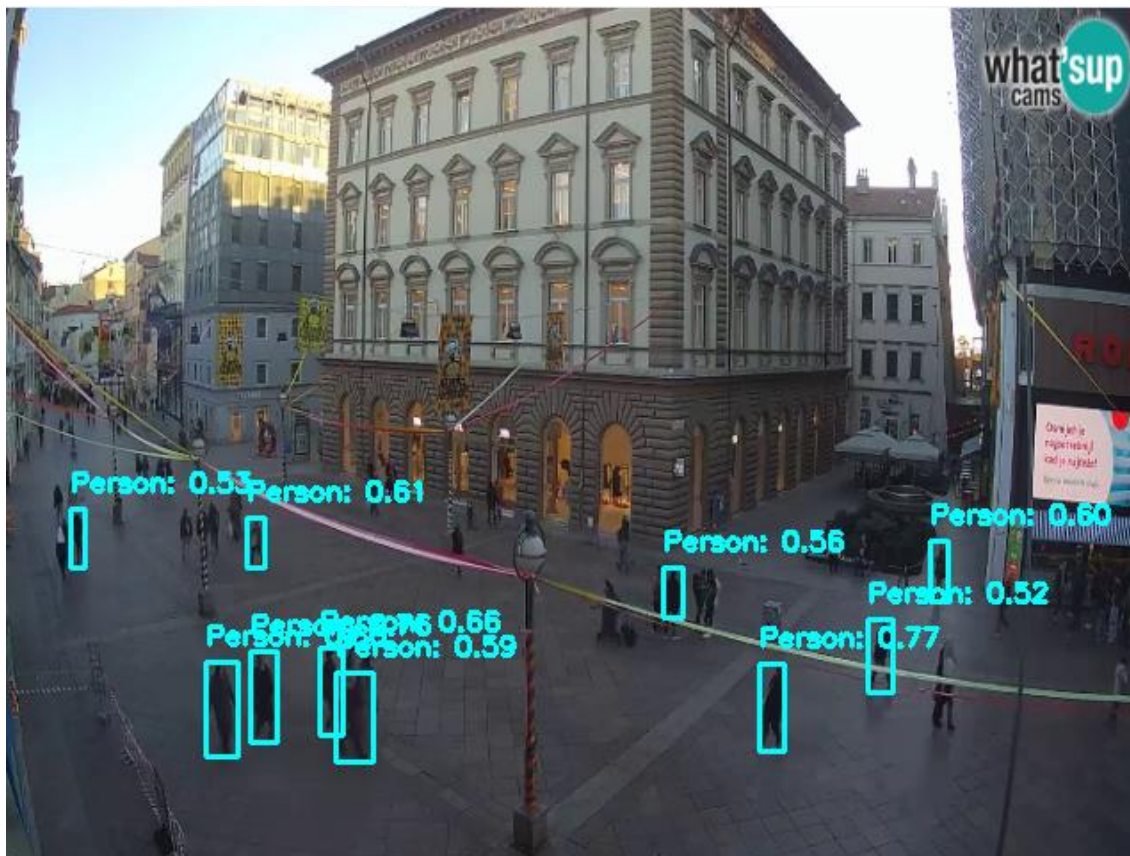
Slika 1 YOLOv8-M model - izravni pogled

U takvim situacijama, gdje je perspektiva naglašeno visoka, dolazi do izraženijih problema u prepoznavanju ljudi i prtljage, što može rezultirati nepreciznim ili čak izostavljenim detekcijama. Ovi problemi mogu biti posljedica nedostatka raznolikosti u treningu na takvim perspektivama slika. Stoga je važno razmotriti ove aspekte prilikom daljnjeg finog podešavanja modela kako bi se poboljšala njegova sposobnost detekcije objekata iz različitih kutova gledanja.



Slika 2 YOLOv8-M model - ptičja perspektiva

Već sada primjećujemo izazove u detekciji objekata iz daljine. Tokom snimanja s web kamere, primijetili smo da niti jedna prtljaga nije detektirana, ali i veliki broj ljudi na korzu. Ovaj nedostatak detekcije može biti rezultat nedostatka sličnih situacija u trening skupu podataka. Model treniran na skupu podataka koji uglavnom sadrži slike s objektima snimljenim iz blizine ili normalne perspektive možda nije u potpunosti sposoban za prepoznavanje objekata u situacijama u kojima su objekti snimljeni iz daljine ili ptičje perspektive. Stoga je važno razmotriti ove nedostatke u trenutnom modelu i razmotriti dodavanje više raznolikosti u trening skupu podataka kako bi model postao sposoban za prepoznavanje objekata u različitim perspektivama i udaljenostima.



Slika 3 YOLOv8-M Detekcija ljudi i prtljage na web kameri korza

Fine-tuning YOLOv8

Proces „fino podešavanja“ (fine-tuning) YOLOv8 modela je tehnika koja se koristi kako bi se prilagodila postojeća arhitektura modela za specifičan zadatak detekcije objekata u novom okruženju ili scenariju. Ovaj proces obično uključuje sljedeće korake:

1. **Priprema podataka:** Prvo, pripremamo novi skup podataka koji sadrži slike s označenim objektima koje želimo detektirati. Ove slike moraju biti reprezentativne za stvarne scenarije u kojima će model biti korišten.
2. **Inicijalizacija modela:** Koristimo prethodno treniran YOLOv8 model kao početnu točku za fino podešavanje. Ovaj inicijalni model može biti pretreniran na velikom skupu podataka poput COCO dataset-a.
3. **Treniranje na novim podacima:** Koristimo novi skup podataka kako bismo nastavili trenirati YOLOv8 model. Tijekom treninga, model se prilagođava novom skupu podataka, učeći prepoznavati specifične objekte i obrasce prisutne u tim slikama.
4. **Validacija i evaluacija:** Nakon završetka treniranja, model se evaluira na odvojenom skupu validacijskih podataka kako bi se procijenila njegova sposobnost detekcije objekata u stvarnom svijetu.

Jedna od glavnih prednosti ovog procesa je povećana preciznost u detekciji objekata. Finim podešavanjem, model se prilagođava specifičnim karakteristikama novog skupa podataka, što rezultira većom preciznošću u identifikaciji objekata.

Priprema podataka

Za pripremu podataka u ovom slučaju koristit ćemo gotove specifične datasetove koji se sastoje samo od slika prtljage i osoba kako bismo model naučili prepoznavati te specifične objekte. Takvi datasetovi omogućavaju fokusiranje modela isključivo na detekciju prtljage i osoba, što može rezultirati boljom preciznošću u tim područjima.

Osim toga, izradit ćemo i vlastiti skup podataka koji će sadržavati slike prtljage i osoba kako bi vidjeli kako model uči i koliko kvalitetno radi predikcije na nad našim oznakama.

Inicijalizacija modela

Inicijalizacija s pretreniranim YOLOv8 modelom koristi već postojeći model treniran na bogatom skupu podataka poput COCO dataset-a kao polazišnu točku za daljnje prilagodbe. Ovaj pristup omogućuje brže prilagođavanje modela s obzirom da već ima osnovno razumijevanje objekata iz prethodnog treniranja. Za inicijalizaciju pretreniranog modela koristimo ultralytics biblioteku, specificirajući željenu verziju YOLO modela.

```
from ultralytics import YOLO
model = YOLO('yolov8m.pt')
```

Kod 1 Inicijalizacija YOLO modela

Treniranje na novim podacima

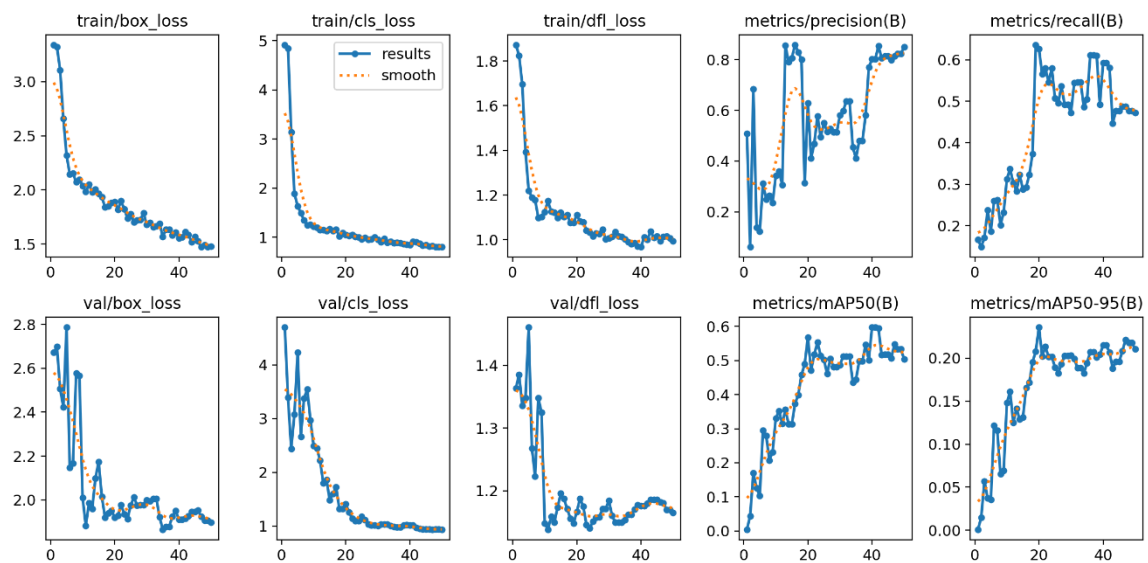
Nakon pripreme podataka, sljedeći korak je fino podešavanje inicijaliziranog modela kako bi se prilagodio specifičnom zadatku detekcije prtljage i osoba. Korištenjem fine tuning tehnike, model se dalje trenira na novim podacima radi poboljšanja preciznosti detekcije i prilagodbe raznolikosti novog skupa podataka, postižući optimizaciju performansi za stvarni svijet. U data.yaml datoteci sadržane su osnovne informacije o skupu podataka korištenom za treniranje modela, kao što su broj klasa objekata, putanje do podataka te ostali relevantni parametri potrebni za konfiguraciju i treniranje modela.

```
results = model.train(data = r'luggage.v1i.yolov8/data.yaml', epochs = 20)
```

Kod 2 Fine-tuning modela

Validacija i evaluacija

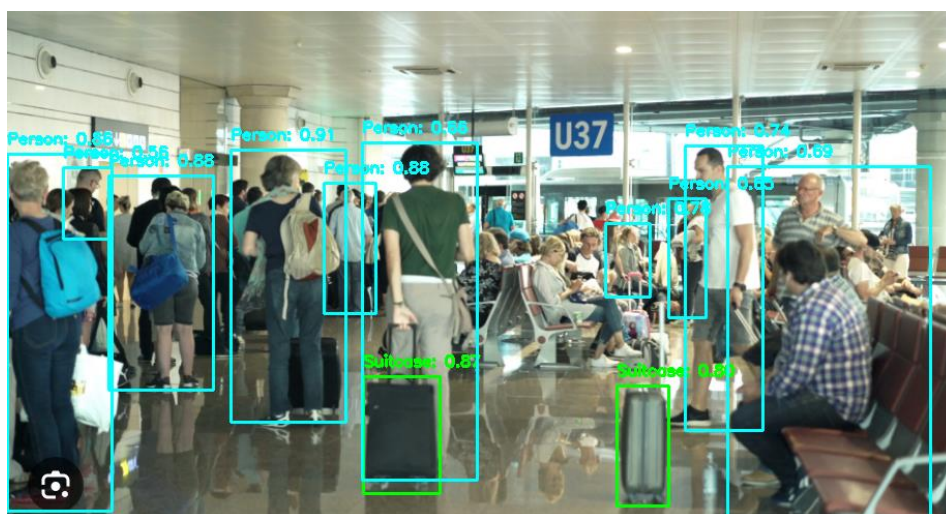
Validacija i evaluacija YOLOv8 modela su ključni koraci u ocjeni performansi i pouzdanosti modela detekcije objekata. Koristeći odvojeni skup podataka za validaciju, provjeravamo sposobnost modela da precizno identificira objekte izvan skupa podataka korištenog za treniranje. Nakon toga slijedi evaluacija, gdje se koriste metrike poput matrice konfuzije, preciznosti, gubitka itd. kako bi se kvantitativno ocijenila preciznost i osjetljivost modela. Ove faze omogućuju procjenu učinkovitosti modela i identificiranje područja za daljnje poboljšanje kako bi se postigle optimalne performanse u stvarnom svijetu.



Slika 4 Primjer metrike modela

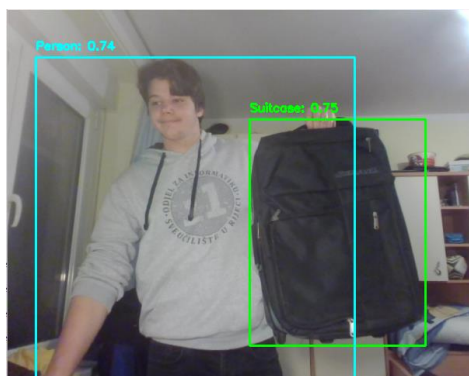
Opis rezultata treniranja modela na pronađenim/gotovim skupovima podataka

Rezultati treniranja modela na pronađenim ili već dostupnim skupovima podataka mogu pokazati visoku preciznost detekcije objekata, posebno prtljage i osoba. Ovi skupovi podataka često su obilježeni i sadrže raznolike primjere objekata, što pruža dobar temelj za učenje modela. Međutim, jedan od glavnih izazova s ovim skupovima podataka je njihova uska specifičnost - često se slike snimaju iz jedne perspektive, što može ograničiti sposobnost modela da detektira objekte iz drugačije perspektive, što je ključno za detekciju prtljage na većim javnim prostorima.



Slika 5 Detekcija prtljage i ljudi izravna perspektiva

Ova ograničenja mogu rezultirati manjom učinkovitošću modela u stvarnom svijetu gdje su kamere postavljene pod različitim kutovima i visinama, stvarajući raznolike perspektive. Stoga, iako su ovi skupovi podataka korisni za osnovno treniranje modela, daljnje prilagodbe i fino podešavanje su nužni kako bi se modeli bolje prilagodili raznolikim perspektivama i okruženjima na stvarnim lokacijama.



Slika 6 Primjer detekcije na vlastitoj kameri

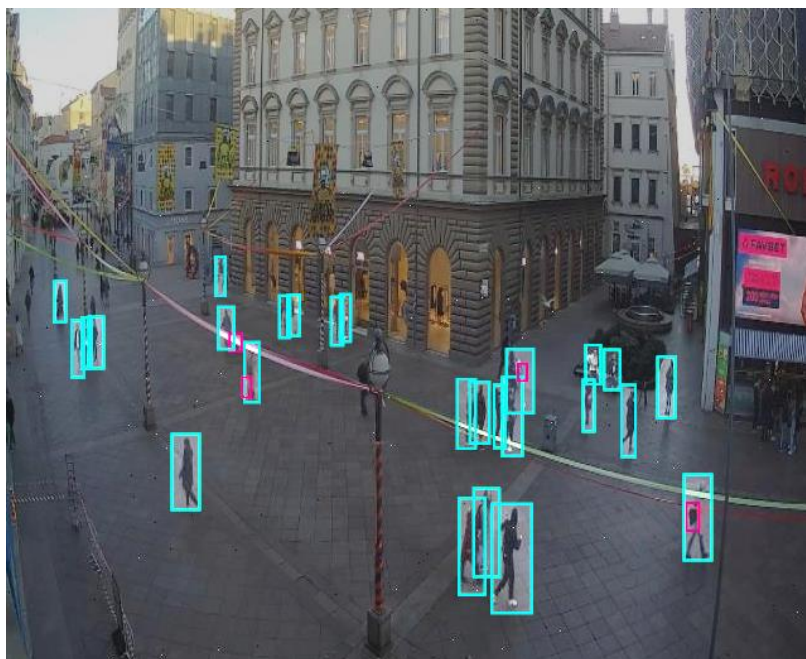
Ukratko, iako su pronađeni ili već dostupni skupovi podataka veoma kvalitetni i opsežni, njihova ograničenja u perspektivi nadzornih kamera mogu zahtijevati dodatne korake u procesu treniranja kako bi modeli bili učinkoviti u detekciji prtljage na većim javnim prostorima.



Slika 7 Loša detekcija prtljage iz ptičje perspektive

Označavanje vlastitih skupova podataka

Vlastiti skupovi podataka su označeni pomoću alata za označavanje *Roboflow* kako bismo nadomjestili gotove skupove podataka koji ne obuhvaćaju perspektivu nadzorne kamere. Jedan od tih skupova podataka sadrži slike prtljage i osoba s riječkog Korza, iako je broj slika s prtljagom u tom skupu relativno mali.



Slika 8 Označavanje skupa podataka Korzo

Drugi skup podataka sastoji se od snimke web kamere s aerodroma u Düsseldorfu. Korištenjem vlastitih označenih skupova podataka, omogućujemo modelu da se bolje prilagodi različitim

perspektivama nadzornih kamera, što bi trebalo rezultirati poboljšanom sposobnošću detekcije prtljage i osoba na tim lokacijama.



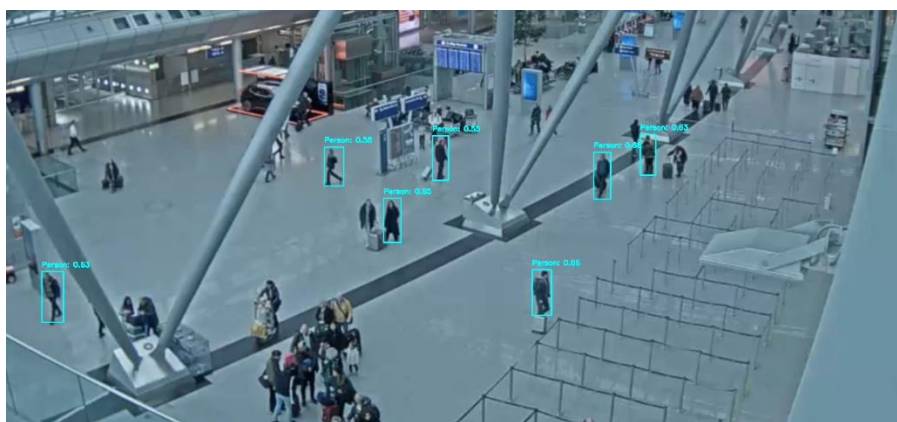
Slika 9 Označavanje skupa podataka aerodrom

Usporedba modela detekcije

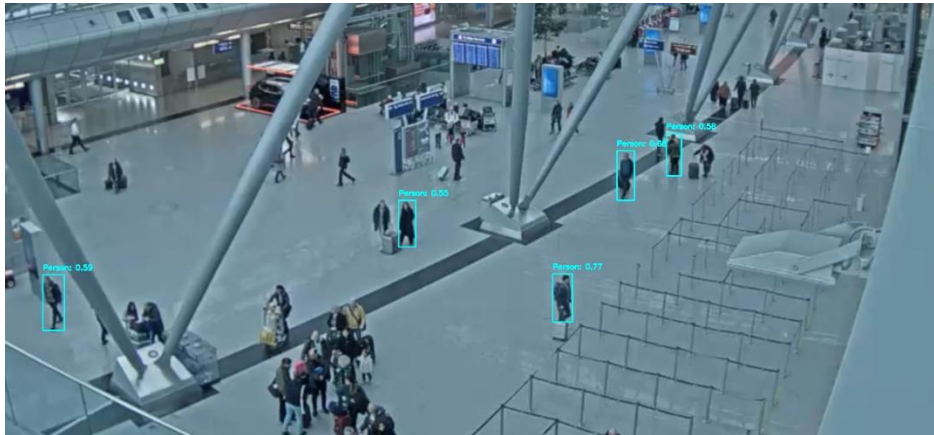
Gotovi skupovi podataka

Svakom modelu ćemo pružiti istu sliku s aerodroma Düsseldorf kako bismo detaljno analizirali razlike u njihovim sposobnostima detekcije objekata, posebno prtljage. Korištenjem iste slike za sve modele omogućit ćemo preciznu usporedbu njihovih performansi na istom skupu podataka. Ova metodologija će nam omogućiti da identificiramo prednosti i nedostatke svakog modela u detekciji prtljage te da donesemo informirane zaključke o najučinkovitijem pristupu za naše specifične potrebe u kontekstu aerodromske sigurnosti.

YOLOv8-S



YOLOv8-M



Iznenadujuće je primijetiti da YOLOv8s model pokazuje bolje detekcije od YOLOv8m. Iako niti jedan od modela nema veliki uspjeh detekcije ljudi, a prtljagu ne prepoznaje niti na jednoj slici.

YOLOv8 retrainiran na slikama COCO skupu podataka samo za prtljagu i ljude

<https://universe.roboflow.com/playground-v6gip/person-and-luggage>



YOLOv8 retrainiran na luggage skupu podataka

<https://universe.roboflow.com/luggage-7rqr6/luggage-kcuiy/dataset/1>

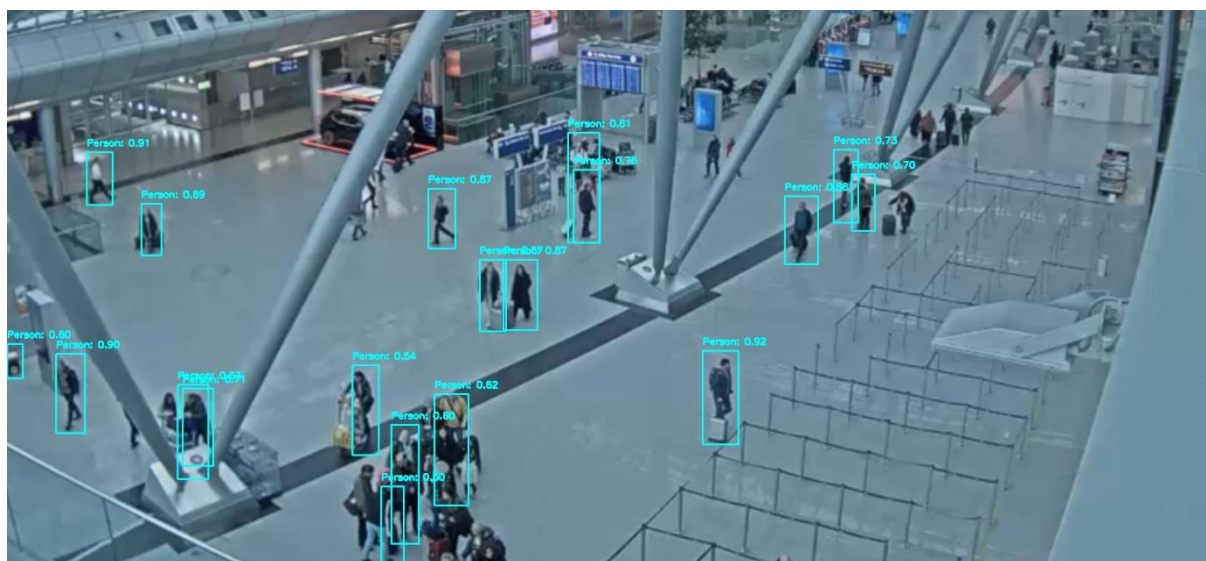


Očekivano je da će rezultati biti lošiji zbog nedostatka prikladnih primjera prtljage i različitih perspektiva u treniranju modela. Većina slika u ovim skupovima podataka sadrži objekte koji su izravno slikani, dok mi ovdje promatramo drugačiju perspektivu, što može biti izazov za detekciju objekata. Ovo naglašava važnost korištenja raznolikih skupova podataka koji odražavaju stvarne scenarije i perspektive nadzornih kamera kako bi se osigurala adekvatna generalizacija modela detekcije objekata.

Vlastiti skupovi podataka

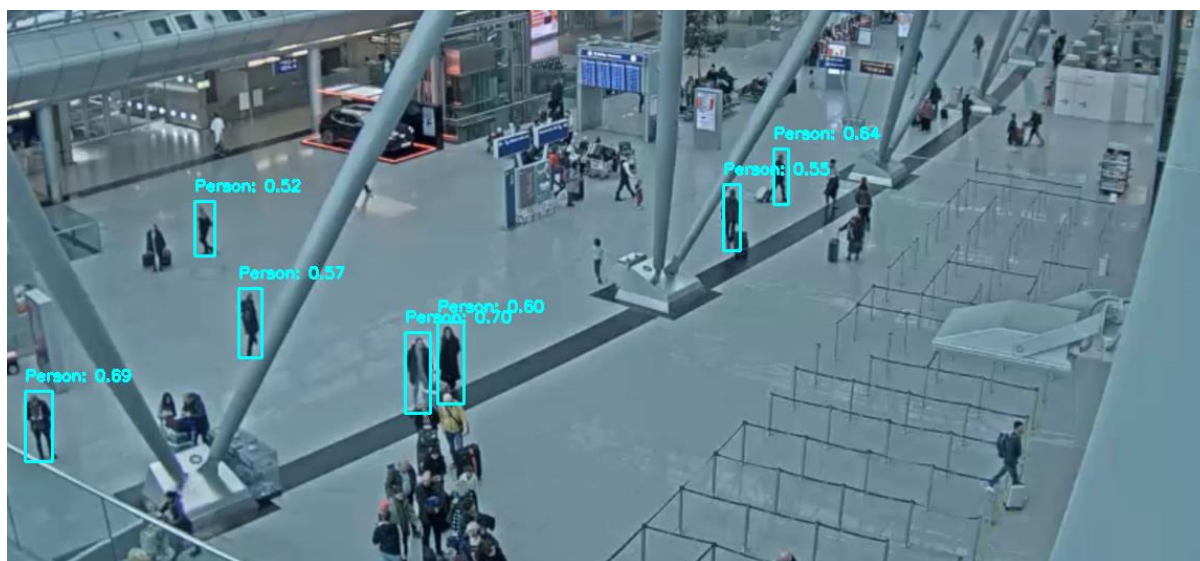
Model treniran na slikama sa korza

<https://app.roboflow.com/cars-0jbgu/luggage-detection-axdmv/1>



Model treniran na aerodromu

<https://app.roboflow.com/cars-0jbgu/luggage-person-detection-airport/2>



Primjećujemo značajne razlike u performansama modela treniranog na skupu podataka s riječkog Korza i modela treniranog na snimci s aerodroma. Model treniran na Korzu pokazuje znatno bolje rezultate u detekciji osoba, dok model aerodroma nije uspio prepoznati prtljagu ni u jednom slučaju. Ovo pokazuje kako modeli mogu varirati u učinkovitosti ovisno o skupu podataka na kojem su trenirani te naglašava potrebu za raznolikim i reprezentativnim skupovima podataka kako bi se postigla pouzdana generalizacija modela u stvarnim uvjetima. Također, primjećujemo da iako modeli ponekad uspješno detektiraju prtljagu tijekom testiranja videa, ta detekcija je nepouzdana i ne konstantna, što ukazuje na potrebu za daljnjim poboljšanjima kako bi se osigurala dosljedna i precizna detekcija prtljage u različitim scenarijima.



Slika 10 Primjer rada modela na snimku aerodroma

Zaključak

U zaključku, istaknuli smo složenost u pronalaženju kvalitetnih skupova podataka za detekciju prtljage iz perspektive nadzorne kamere. Iako su pronađeni skupovi podataka veći u broju fotografija, njihova učinkovitost u detekciji prtljage značajno je umanjena zbog nedostatka reprezentativnih primjera i raznolikih perspektiva. Ovo jasno ističe važnost daljnjeg istraživanja i razvoja kako bismo osigurali da modeli detekcije objekata budu adekvatno prilagođeni stvarnim scenarijima nadzora.

Najvjerojatnije rješenje za postizanje kvalitetne detekcije prtljage leži u stvaranju vlastitog skupa podataka koji će biti veći, sa više fotografija i primjera željenih detekcija kako bi bolje odražavali raznolike perspektive i uvjete s kojima se susrećemo u stvarnom svijetu. Ovaj pristup omogućit će nam da izgradimo modele detekcije prtljage koji su bolje prilagođeni specifičnim potrebama i zahtjevima našeg sustava nadzora.

Skupovi podataka

1. <https://universe.roboflow.com/luggage-7rqr6/luggage-kcuiy>
2. <https://app.roboflow.com/cars-0jbgu/luggage-detection-axdmv/1>
3. <https://www.kaggle.com/datasets/dataclusterlabs/suitcaseluggage-dataset>
4. <https://app.roboflow.com/cars-0jbgu/luggage-person-detection-airport/1>
5. <https://app.roboflow.com/cars-0jbgu/luggage-person-detection-airport/2>