



Sveučilište u Rijeci
University of Rijeka
<http://www.uniri.hr>

Polytechnica: Journal of Technology Education, Volume 8, Number 1 (2024)
Politehnika: Časopis za tehnički odgoj i obrazovanje, Svezak 8, Broj 1 (2024)



Politehnika
Polytechnica
<http://www.politehnika.hr/journal>
cte@uniri.hr

DOI: <https://doi.org/10.36978/cte.8.1.4>

Stručni članak
Professional paper
UDK: 681.783.322

Detekcija objekata s pomorskih nadzornih kamera

Miran Pobar

Sveučilište u Rijeci, Fakultet informatike i digitalnih tehnologija

Radmile Matejčić 2, 51000 Rijeka

mpobar@uniri.hr

Sažetak

Automatska detekcija objekata u moru na slikama nadzornih ili panoramskih kamera otvara mogućnost automatskog praćenja prometa, detekcije neovlaštenoga kretanja, opasnosti ili onečišćenja. U ovom radu analiziraju se performanse modela temeljenih na arhitekturi YOLOv7 za zadatak detekcije plovila i plutača na takvim slikama. Modeli su naučeni na vlastitom skupu podataka različitih pomorskih scena izrađenom za tu svrhu, korištenjem prijenosa učenja s modela naučenih na općenitim slikama. Također, ispitane su dvije varijante rukovanja ulazom u mrežu, te je korištenje strategije rezanja ulazne slike značajno poboljšalo rezultate detekcije posebno malih objekata u odnosu na osnovni model.

Ključne riječi: računalni vid; neuronske mreže; mali objekti; detekcija brodova.

1 Uvod

Detekcija objekata na slikama predstavlja ključni korak u mnogim zadacima gdje se primjenjuje računalni vid, kao što su kontrola kvalitete u proizvodnoj industriji (Yang i sur., 2020), autonomna vožnja (Arnold i dr., 2019) ili nadzor (Zou i sur., 2023).

Iako se u posljednjim istraživanjima često koriste veliki skupovi podataka poput MS COCO (Lin i sur. 2014), koji sadrže označene svakodnevne objekte, specifični zadaci zahtijevaju prilagođene skupove podataka. Ti skupovi trebaju sadržavati označene dodatne kategorije objekata relevantnih za specifične zadatke, u realističnom kontekstu kakav se očekuje u primjeni.

Fokus ovog rada je na detekciji pomorskih plovila i malih objekata na slikama s nadzornih i panoramskih kamera. Pomorske nadzorne kamere već su instalirane na mnogim lokacijama, ali bi se njihova

funkcionalnost mogla značajno unaprijediti kroz automatiziranu detekciju objekata. Automatskom detekcijom neovlaštenih plovila ili drugih objekata korištenjem računalnog vida moglo bi se pridonijeti povećanju sigurnosti (Gianinetto i sur., 2016), (Shao i sur., 2020), ali i pomoći u navigaciji u situacijama gdje se ti objekti teško detektirani drugim sredstvima poput radara ili sustava automatske identifikacije plovila (AIS) (Petković i sur., 2021; Lyu i sur., 2023).

U radu se za detekciju objekata razmatra više varijanti modela temeljenih na konvolucijskoj neuronskoj mreži. Za učenje i evaluaciju modela koristi se izvorni skup podataka s označenim slikama pomorskih scena s različitih lokacija. Modeli se uče uz prijenos znanja iz postojećeg modela obučenog na velikom skupu podataka s općenitim objektima kako bi se prilagodili specifičnoj domeni interesa i kako bi učenje na relativno malenom specijaliziranom skupu podataka bilo moguće.

Iako je posljednjih godina postignut veliki napredak u detekciji objekata na slikama korištenjem metoda temeljenih na neuronskim mrežama, detekcija malih ili sitnih objekata ostaje izazov (Tong i Wu, 2022), posebno kada je pozadina pretrpana ili kada su objekti mali i slične boje kao pozadina. U radu se stoga pristupa detekciji najmanjih objekata kroz jednostavan pristup rezanja ulazne slike na manje dijelove, čime se osigurava da se najmanji objekti ne izgube prilikom prilagodbe slike ulazu u mrežu. Ovaj pristup omogućuje detekciju objekata i kada su oni prisutni samo kao nekoliko piksela, što je često slučaj s plutačama ili udaljenim brodovima.

Doprinosi rada uključuju usporedbu performansi detekcije dviju varijanti modela YOLOv7 na novom skupu podataka pomorskih slika. Također, istražuje se kako dvije strategije rukovanja s ulazom, odnosno promjena veličine ili rezanje, utječu na performanse detekcije, posebno malih objekata.

Ostatak rada organiziran je na sljedeći način: u 2. odjeljku opisana su povezana istraživanja, 3. odjeljak opisuje predloženi skup podataka i modele za detekciju pomorskih objekata s detaljima postavki eksperimenta. Rezultati su predstavljeni u 4. poglavlju. Naposljetku, zaključak i prijedlog za daljnji rad navedeni su u 5. odjeljku.

2 Povezana istraživanja

Većina današnjih metoda detekcije objekata kao što su SSD (Liu i sur., 2016), Yolo obitelj detektora (Redmon i sur., 2016; Redmon i Farhadi, 2018; Bochkovskiy, Wang i Liao, 2020; Bochkovskiy i Liao, 2022) i EfficientDet (Tan, Pang i Le, 2020) temelji se na konvolucijskim neuronskim mrežama (engl. convolutional neural network, CNN). Iako univerzalni detektori objekata, naučeni na skupovima podataka općenitih objekata, dobro funkcioniraju za mnoge zadatke, nove domene primjene često zahtijevaju prilagodbe i specijalizirane skupove podataka kako bi se modeli prilagodili novim zadacima, poput detekcije u termalnom snimanju (Krišto, Ivašić-Kos i Pobar, 2020) ili kada se radi o malim objektima, poput detekcije pčela (Stojnić i sur., 2021.) ili osoba (Sambolik i Ivašić-Kos, 2021) u videozapisima s bespilotnim letjelicama.

Detekcija malih objekata aktivna je tema te su neka istraživanja, primjerice (Stojnić i sur., 2021), usmjerena na metode specifične za zadatke, dok su druga usmjerena na općenitija rješenja s modifikacijom postojećih arhitektura kao što je YOLO-Z (Benjumea i sur. 2023). Iako je u mnogim praktičnim primjenama detekcije objekata pravi izvor slika videozapis, a ne pojedinačne neovisne slike, relativno malo metoda se usredotočuje na iskorištavanje informacija iz više filmskih slika za

poboljšanje otkrivanja objekata. Na primjer, u (Broad, Jones i Lee, 2018) arhitektura detektora objekata temeljenog na CNN-u koji radi s jednom slikom modificirana je tako da se podaci iz više slika spajaju na razini izlučenih značajki i dalje obrađuju u mreži.

U ovom radu istražuje se jednostavna strategija rezanja ulaznih slika kako bi se izbjeglo smanjenje veličine slike i sačuvali mali objekti na ulaznim slikama, slično pristupu u radu (Unel, Ozkalayci i Cigla, 2019).

Nedavno su postali dostupni skupovi podataka za detekciju malih objekata u moru kao što je SeaDroneSee (Varga i sur., 2022), usmjeren na scenarije potrage i spašavanja.

SPSCD skup podataka predložen u (Petković i sur. 2023) sadrži slike sa stacionarne kamere u luci i usmjeren je na detekciju nekoliko klasa brodova koji mogu ući u luku. Obuhvaća različita doba dana i vremenske uvjete, ali se posebno ne odnosi na male objekte i ograničen je na jednu luku.

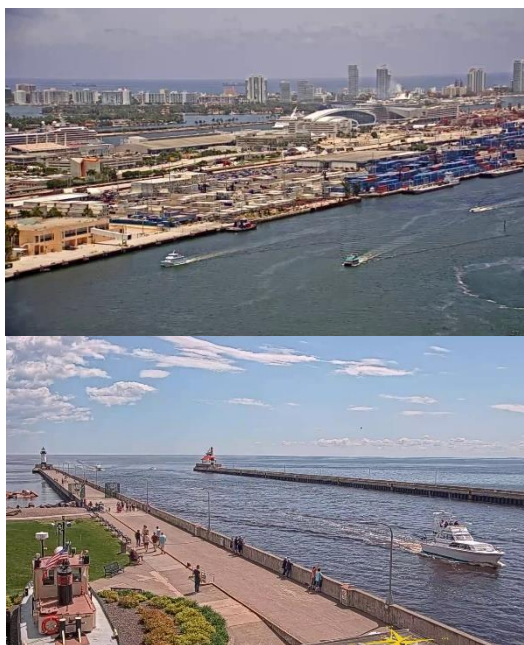
Skup podataka predložen u ovom radu ima sličan cilj kao SPSCD, no naglasak je na raznolikim lokacijama i detekciji malih objekata poput plutača. Unatoč manjem broju slika i klasa objekata, ovaj skup podataka pruža izazov zbog komprimirane kvalitete videozapisa, što čini otkrivanje malih objekata još složenijim.

3 Podaci i metode

U ovom eksperimentu su za otkrivanje pomorskih objekata na slikama naučena i evaluirana na prilagođenom skupu podataka dva različita modela temeljena na arhitekturi YOLOv7 (Wang, Bochkovskiy i Liao, 2023). YOLOv7 je iteracija Yolo (Redmon i sur., 2016) obitelji detektora objekata temeljenih na konvolucijskoj neuronskoj mreži koji u jednom prolazu detektiraju pravokutne granice objekata (granične okvire) na slikama i njihove oznake klase. Na skupu podataka za detekciju objekata MS COCO postiže izvrsne rezultate u pogledu brzine i točnosti (Wang, Bochkovskiy i Liao, 2023). Nedavno je primijenjen i na problem otkrivanja pomorskih objekata (Zhu i sur., 2023, a autori bilježe srednju prosječnu preciznost (engl. mean average precision - mAP) od 75,66% pri pragu 50% omjera presjeka i unije (engl. intersection over union - IoU) na skupu podataka za zadatak traganja i spašavanja na moru (Gasienica-Jozkowsky, Knapik i Cyganek, 2021). Autori su predložili modificiranu arhitekturu YOLOv7 i postigli povećanje mjere mAP pri detekciji malih objekata od 4,2%, te poboljšanje točnosti detekcije s 0,578 na 0,754 za posebno za klasu plovila i s 0,615 na 0,643 za klasu plutače u jednoj testiranoj varijanti mreže.

3.1 Skup podataka

Skup podataka za detekciju objekata priređen je iz 20 video prijenosa koji su dostupni na internetu i emitiraju se sa stacionarnih kamera postavljenih u različitim lukama, kanalima, marinama i drugim značajnim lokacijama (slika 1). Ovi video prijenosi obuhvaćaju širok spektar scena, pružajući raznolikost u analizi objekata. Preciznije, 10 videa potječe iz manjih luka ili marina, gdje se često pojavljuje više malih brodova i plutača istovremeno, stvarajući preklapanja s obzirom na njihovu blizinu. Pet videa prikazuje turističke panorame, gdje je uobičajeno prisutno manje brodova, ali su oni često u pokretu. Preostalih pet videa snimljeno je u velikim lukama ili pomorskim kanalima gdje dominiraju veliki komercijalni brodovi. Većina videozapisa emitirana je u rezoluciji 1920x1080, a manjina u 1280x720, s različitim stupnjevima kvalitete i jačine kompresije. Neke od postavljenih kamera su nepokretne, dok druge koriste rotaciju oko svoje osi ili zumiranje kako bi pokrile veće područje. Videozapisi su prikupljeni tijekom razdoblja od travnja do svibnja 2022. godine.

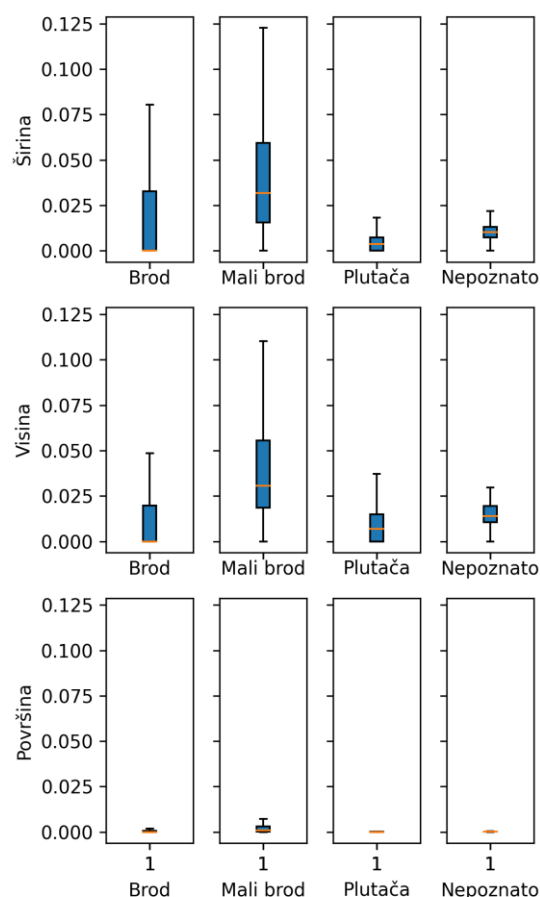


Slika 1. Primjeri scena u izvornim videozapisima.

Iz videozapisa su izdvojeni pojedinačne sličice u trenucima kada se scena subjektivno značajno promijenila. U nekim slučajevima, npr. kada se kamera rotira ili zumira, interval između izdvojenih slika može biti samo nekoliko sekundi, dok u nekim slučajevima, npr. u luci koja nije prometna, interval može biti nekoliko sati ili dana.

Objekti na izdvojenim sličicama su zatim ručno označeni graničnim okvirima i oznakama u pet klasa: *brod*, *mali brod*, *plutača*, *plivač* i *nepoznato* za

neidentificirane objekte u moru. Za potrebe označavanja, klasa *mali brod* je definirana kao jedrilice, mali ribarski brodovi, mali putnički brodovi, male i srednje jahte, dok su teretni brodovi, trajekti, srednji i veliki putnički brodovi, velike jahte itd. označeni oznakom klase *brod*. S obzirom na mali broj pojavljivanja klase *plivač*, ta je klasa konačno spojena s klasom *nepoznato*. Konačni skup podataka sadrži 135 označenih slika s 2755 instanci objekta. Promatrajući veličine objekata, klasa *mali brod* obično je zauzimala najveće područje slike, zbog položaja nekoliko kamera koje su se nalazile u marinama gdje su brodovi prolazili u blizini. S druge strane, klasa *brod* je obično vizualno manja jer su veliki brodovi u pravilu bili mnogo udaljeniji od kamere. Klase *plutača* i *nepoznato* očekivano su bile najmanje bez obzira na položaj kamere, te su najčešće zauzimale manje od 0,01% ukupne površine slike, ponekad samo nekoliko piksela.



Slika 2. Rasponi relativnih veličina objekata u odnosu na cijelu sliku u skupu podataka.

3.2 Postavke eksperimenta

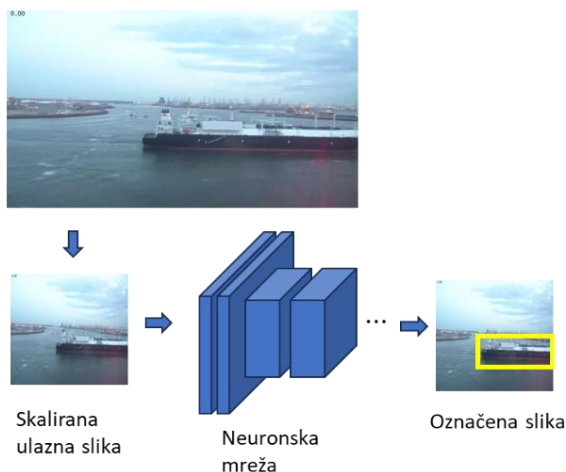
Za zadatak detekcije objekata u moru naučeni su modeli temeljeni na dvije varijante arhitekture YOLOv7 (Wang, Bochkovskiy i Liao, 2023), pri čemu je

izlazni broj klasa prilagođen korištenom skupu podataka.

Prva korištena varijanta je osnovna YOLOv7 arhitektura s 415 slojeva i oko 37 milijuna parametara, dok je druga mnogo jednostavnija i računalno manje zahtjevna varijanta YOLOv7-tiny s 263 sloja i oko 6 milijuna parametara. Osnovna YOLOv7 arhitektura je prvenstveno namijenjena upotrebi na grafičkim procesorima (GPU) kakvi su dostupni za desktop računala i za zaključivanje zahtijeva oko 104,7 milijardi izračuna s pomičnim zarezom. Varijanta YOLOv7-tiny znatno je manje računalno zahtjevna s obzirom da za zaključivanje zahtijeva oko 13,8 milijardi operacija s pomičnim zarezom te je stoga prikladnija za upotrebu na uređajima s manje resursa poput ugradbenih sustava. Veličina ulazne slike za oba naučena modela je 640x640 piksela.

Zadani zadatak je na slikama detektirati objekte klase *broda*, *mali brod*, *plutača* i *nepoznato*.

Skup podataka najprije je nasumično podijeljen na skupove za učenje, validaciju i testiranje u omjeru 60:15:25. Skupovi za učenje i validaciju su potom prošireni dijelovima postojećih slika iz istih skupova izrezivanjem dijelova veličine 640x640 piksela s nasumično odabranih pozicija. Konačni skup za učenje sastojao se od mješavine 81 izvorne slike i 276 izrezanih dijelova, ukupno 357 slika. Kako je veličina izrezanih dijelova ista kao veličina ulaza u mrežu, mali objekti poput plutača nisu promijenili veličinu, čime se očuvao njihov izgled i osiguralo da budu vidljivi tijekom učenja, dok se na izvornim slikama zbog smanjenja veličine s 1920x1080, ili 1280x720 na 640x640, mali objekti potencijalno ne vide prilikom učenja.



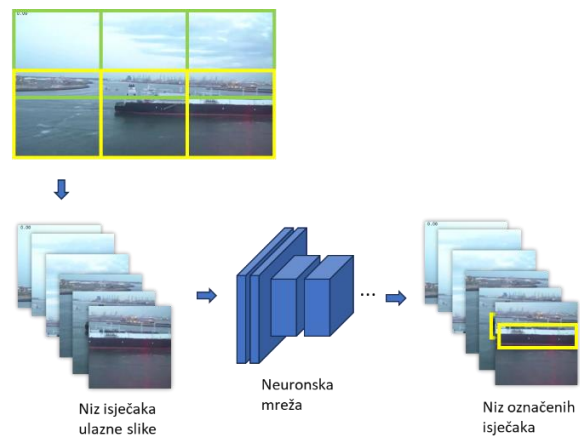
Slika 3. Strategija skaliranja: ulazna slika se smanjuje na veličinu koja odgovara ulaznoj veličini neuronske mreže.

Ostali načini dinamičkog povećanja skupa za učenje poput skaliranja ili pomaka su isključeni prilikom učenja. Učenje je provedeno u 3000 epoha na skupu

za učenje sa stopom učenja postavljenom na 0,01. Za učenje je korišten postupak prijenosa učenja s težina modela prethodno naučenih na skupu podataka MS COCO (Lin i sur., 2014) i dostupnih u repozitoriju YOLOv7 (Wang, 2023).

Nakon 3000 epoha učenja, na testnom skupu su testirane naučene varijante modela YOLOv7 i YOLOv7-tiny koje su postigli najbolje rezultate na validacijskom skupu.

Modeli su prvo procijenjeni testnim slikama koje su na ulazu u mrežu smanjene s izvorne veličine (obično 1920x1080 piksela) na veličinu 640x640 koju mreža očekuje (strategija skaliranja, slika 3.), a zatim na preklapajućim izdvojenim dijelovima veličine 640x640 koje odgovaraju veličini ulaza izvorne mreže (strategija rezanja, slika 4.). Za preračunavanje referentnih graničnih okvira na isječcima korišten je alat yolo-tiling (Neskorozhenyi, 2023).



Slika 4. Strategija rezanja: ulazna slika se reže na dijelove čija veličina odgovara ulaznoj veličini neuronske mreže.

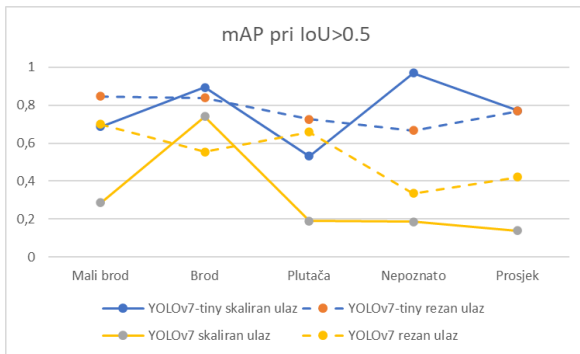
Svi eksperimenti provedeni su na temelju PyTorch implementacije YOLOv7, korištenjem verzije PyTorch 1.10.1, Python 3.8.10 i Ubuntu 20.04 operativnog sustava, na GPU Nvidia RTX 3090.

Za evaluaciju performansi modela korištene su standardne metrike prosječne preciznosti (AP) i srednje prosječne preciznosti (mAP) (Padilla, Neto i Da Silva, 2020) uz prag od 50% indeksa preklapanja (IoU), označene kao mAP@0.5 i u rasponu praga IoU od 50% do 95%, označene kao mAP@.5:.95.

Indeks preklapanja (intersection-over-union, IoU) se računa kao omjer površine presjeka detektiranih i stvarnih objekata i površine unije detektiranih i stvarnih objekata.

4 Rezultati

Rezultati evaluacije modela s obzirom na mjeru prosječne preciznosti uz minimalnu vrijednost IoU 0,5 prikazane su na slici 5.



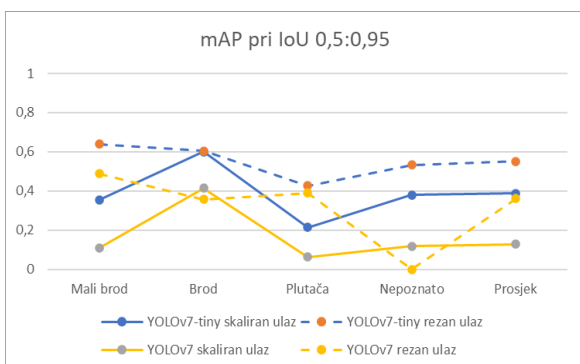
Slika 5. Srednja prosječna preciznost detekcije modela za indeks preklapanja IoU >0.5

U prosjeku, najbolje rezultate ostvaruje model YOLOv7-tiny, bez obzira na korištenu strategiju rezanja ili skaliranja ulaza, uz srednju prosječnu preciznost (mAP) detekcije od 77%.

Gledajući pojedine klase, vidljivo je da korištenje strategije rezanja ulaza znatno poboljšava detekciju za klase *mali brod* i *plutača*, za oba modela. U slučaju modela YOLOv7-tiny za klasu *brod* nema značajnih razlika između strategija rezanja ili skaliranja, dok za model YOLOv7 strategija rezanja daje lošije rezultate.

Klasa *nepoznato* nedovoljno je zastupljena u skupu podataka, no s obzirom da je sveobuhvatna klasa za sve neprepoznate objekte, pojedine instance klase se mogu vizualno značajno razlikovati. Glavna svrha klase je spriječiti klasifikator da dodijeli oznaku drugih klasa dijelovima slike koji imaju kvalitetu "objektnosti", ali se ne mogu jasno svrstati u neku od poznatih klasa. Trebalo bi dodatno istražiti utječe li ova klasa doista na preciznost detekcije drugih klasa.

Slika 6 prikazuje mjeru prosječne preciznosti za pragove indeksa preklapanja detektiranih i stvarnih objekata u rasponu od 0,5:0,95.



Slika 6. Srednja prosječna preciznost detekcije modela za pragove indeksa preklapanja IoU 0,5:0,95

S obzirom na stroži kriterij prihvaćanja detekcije u odnosu na prag IoU od 50%, vrijednosti mAP su očekivano niže, u rasponu od 42,6% za *plutaču* do 63,9% za *mali brod*, kod najbolje kombinacije strategije i modela, tj. YOLOv7-tiny i rezanja ulaza.

U ovakvom režimu, strategija rezanja ulaza daje bolje rezultate od skaliranja ulaza za sve klase u slučaju korištenja modela YOLOv7-tiny, osim za klasu *brod*, za koju se rezultati bitno ne razlikuju. Kod modela YOLOv7, za klase *mali brod* i *plutača* prednost strategije rezanja ulaza je još više izražena, no za klasu *brod* skaliranje daje neznatno bolje rezultate.

Mjere odziva i preciznosti prikazane su u tablici 1. Može se uočiti da se korištenjem strategije rezanja povećava preciznost detekcije u odnosu na skaliranje u svim slučajevima, osim kod klase *nepoznato* za detektor YOLOv7-tiny. Značajno se povećava i broj dobro detektiranih malih objekata poput *plutače*, što se odražava na povećanoj mjeri odziva za tu klasu.

Klasa	Mjera	YOLOv7-tiny		YOLOv7	
		Skaliran ulaz	Rezan ulaz	Skaliran ulaz	Rezan ulaz
Mali brod	Odziv	0,646	0,759	0,402	0,664
	Preciznost	0,926	0,988	0,4	0,715
Brod	Odziv	0,897	0,776	0,759	0,557
	Preciznost	0,907	0,959	0,609	0,677
Plutača	Odziv	0,481	0,609	0,204	0,638
	Preciznost	0,765	0,933	0,752	0,798
Nepoznato	Odziv	0,75	0,667	0,247	0,333
	Preciznost	0,995	0,917	0,497	0,501
Prosjeck	Odziv	0,693	0,703	0,403	0,548
	Preciznost	0,898	0,949	0,564	0,673

Tablica 1. Odziv i preciznost modela po klasama

Unatoč mnogo većem broju parametara i računskoj složenosti modela YOLOv7, u ovom eksperimentu su značajno bolje performanse ostvarene jednostavnijim YOLOv7-tiny modelom. Moguć razlog je skromna veličina skupa podataka i mnogo veći broj parametara punog modela YOLOv7 koji zahtijeva više podataka za učenje.

Eksperiment potvrđuje da je za zadatak detekcije malih objekata primjena strategije rezanja ulazne slike u dijelove korisna, unatoč dodatnoj složenosti postupka i potrebi da se ista slika obradi nekoliko puta (jednaka broju isječaka).

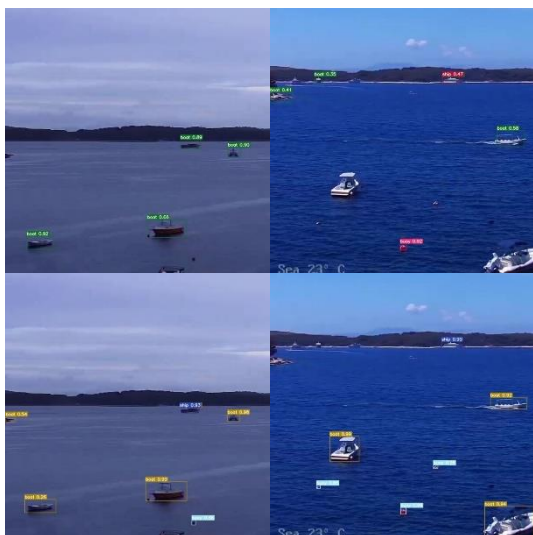
Nekoliko primjera detekcije s oba modela pomoću strategije skaliranja prikazano je na slici 7. i pomoću strategije rezanja na slici 8.

Iz primjera na slikama 7 i 8 vidljivo je kako YOLOv7 model detektira manje stvarnih objekata od tiny modela. Također, model YOLOv7-tiny detektirane objekte češće svrstava u pravu klasu, što se vidi na primjeru lijevo na slici 8, gdje YOLOv7 model sva

detektirana plovila svrstava u klasu *mali brod*, a YOLOv7-tiny udaljeni brod ispravno u klasu *brod*.



Slika 7. Primjeri rezultata detekcije za strategiju skaliranja ulaza. Gornji red: YOLOv7, donji red: YOLOv7-tiny.



Slika 8. Primjeri rezultata detekcije za ulaznu strategiju rezanja. Gornji red: YOLOv7, donji red: YOLOv7-tiny.

5 Zaključak

U ovom je radu razmatran zadatak detekcije objekata u moru na vlastitom skupu podataka prikupljenih s javno dostupnih panoramskih kamera. U provedenom eksperimentu, jednostavniji i brži model baziran na arhitekturi YOLOv7-tiny ostvario je bolje rezultate u usporedbi s modelom baziranim na punoj arhitekturi YOLOv7. Ovi rezultati sugeriraju

prikladnost modela YOLOv7-tiny za predmetni zadatak, ali i na vjerojatnu mogućnost ostvarenja boljih rezultata detekcije uz prikupljanje dodatnih podataka za učenje kako bi se u potpunosti iskoristila arhitektura složenijeg modela.

Arhitektura YOLOv7-tiny izdvaja se zbog manjih računalnih zahtjeva, što je čini idealnom za implementaciju na rubnim GPU uređajima u sklopu ili blizini instaliranih kamera. To je važno kako bi se izbjegao problem gubitka kvalitete videa zbog visoke stope kompresije koja se obično koristi kod streaminga videozapisa putem interneta, za obradu na udaljenom računalu.

Za rješavanje problema detekcije malih objekata, važnih u kontekstu situacija opasnosti ili spašavanja, razmatrana je alternativna ulazna strategija koja uključuje rezanje ulazne slike na segmente točnih dimenzija ulaza mreže. Eksperimentalni rezultati jasno ukazuju na značajno poboljšanje u otkrivanju malih objekata, kao što su plutače, što rezultira povećanjem prosječne preciznosti za tu klasu s 52% na 72% i s 48% na 61% na testnom skupu uz primjenu modela temeljenog na arhitekturi YOLOv7-tiny.

Performanse otkrivanja malih objekata još uvijek nisu savršene, a u budućem radu trebalo bi razmotriti dodatna poboljšanja, na primjer integraciju detekcija iz više uzastopnih slika.

Osim toga, skup podataka proširit će se tako da uključuje više različitih vremenskih uvjeta i lokacija te metodu testiranu na drugim pomorskim skupovima podataka kao što su SPSCD ili SeaDroneSee.

Literatura

- Arnold, E., Al-Jarrah, O. Y., Dianati, M., Fallah, S., Oxtoby, D., Mouzakitis, A. (2019). A Survey on 3D Object Detection Methods for Autonomous Driving Applications. *IEEE Transactions on Intelligent Transportation Systems*, 20(10), 3782–3795. <https://doi.org/10.1109/TITS.2019.2892405>
- Benjumea, A., Teeti, I., Cuzzolin, F., Bradley, A. (2023). *YOLO-Z: Improving small object detection in YOLOv5 for autonomous vehicles* (arXiv:2112.11798). arXiv. <http://arxiv.org/abs/2112.11798>
- Broad, A., Jones, M., Lee, T. Y. (2018, September). Recurrent Multi-frame Single Shot Detector for Video Object Detection. In *BMVC* (p. 94).
- Bochkovskiy, A., Wang, C. Y., Liao, H. Y. M. (2020). Yolov4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934*.
- Gasienica-Jozkowsky, J., Knapik, M., Cyganek, B. (2021). An ensemble deep learning method with optimized weights for drone-based water rescue

- and surveillance. *Integrated Computer-Aided Engineering*, 28(3), 221-235.
- Gianinetto, M., Aiello, M., Marchesi, A., Topputo, F., Massari, M., Lombardi, R., Banda, F., Tebaldini, S. (2016). OBIA ship detection with multispectral and SAR images: A simulation for Copernicus security applications. *2016 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 1229–1232. <https://doi.org/10.1109/IGARSS.2016.7729311>
- Krišto, M., Ivacic-Kos, M., Pobar, M. (2020). Thermal object detection in difficult weather conditions using YOLO. *IEEE access*, 8, 125459-125476.
- Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., ..., Zitnick, C. L. (2014). Microsoft coco: Common objects in context. In *Computer Vision—ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13* (pp. 740-755). Springer International Publishing.
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., Berg, A. C. (2016). Ssd: Single shot multibox detector. In *Computer Vision—ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14* (pp. 21-37). Springer International Publishing.
- Lyu, H., Shao, Z., Cheng, T., Yin, Y., Gao, X. (2023). Sea-Surface Object Detection Based on Electro-Optical Sensors: A Review. *IEEE Intelligent Transportation Systems Magazine*, 15(2), 190–216. <https://doi.org/10.1109/MITS.2022.3198334>
- Neskorozhenyi, R. (2023, June). yolo-tiling [GitHub repository]. GitHub. <https://github.com/slani/yolo-tiling>
- Padilla, R., Netto, S. L., Da Silva, E. A. (2020, July). A survey on performance metrics for object-detection algorithms. In *2020 international conference on systems, signals and image processing (IWSSIP)* (pp. 237-242). IEEE.
- Petković, M., Kezić, D., Vujović, I., Pavić, I. (2021). Target detection for visual collision avoidance system. *Pedagogika*, 93, 159-166.
- Petković, M., Vujović, I., Lušić, Z., Šoda, J. (2023). Image Dataset for Neural Network Performance Estimation with Application to Maritime Ports. *Journal of Marine Science and Engineering*, 11(3), 578.
- Redmon, J., Divvala, S., Girshick, R., Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
- Redmon, J., Farhadi, A. (2018). Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- Sambolek, S., Ivacic-Kos, M. (2021). Automatic person detection in search and rescue operations using deep CNN detectors. *IEEE Access*, 9, 37905-37922.
- Shao, Z., Wang, L., Wang, Z., Du, W., Wu, W. (2020). Saliency-Aware Convolution Neural Network for Ship Detection in Surveillance Video. *IEEE Transactions on Circuits and Systems for Video Technology*, 30(3), 781–794. <https://doi.org/10.1109/TCSVT.2019.2897980>
- Stojnić, V., Risojević, V., Muštra, M., Jovanović, V., Filipi, J., Kezić, N., Babić, Z. (2021). A Method for Detection of Small Moving Objects in UAV Videos. *Remote Sensing*, 13(4), 653. <https://doi.org/10.3390/rs13040653>
- Tan, M., Pang, R., Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 10781-10790).
- Tong, K., Wu, Y. (2022). Deep learning-based detection from the perspective of small or tiny objects: A survey. *Image and Vision Computing*, 123, 104471.
- Unel, F. O., Ozkalayci, B. O., Cigla, C. (2019). The Power of Tiling for Small Object Detection. *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*, 582–591. <https://doi.org/10.1109/CVPRW.2019.00084>
- Varga, L. A., Kiefer, B., Messmer, M., Zell, A. (2022). SeaDronesSee: A Maritime Benchmark for Detecting Humans in Open Water. *2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, 3686–3696. <https://doi.org/10.1109/WACV51458.2022.00374>
- Wang, C. Y., Bochkovskiy, A., Liao, H. Y. M. (2023). YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 7464-7475).
- Wang, C. Y. (2023, June). Yolov7 (Version 0.1) [GitHub repository]. GitHub. <https://github.com/WongKinYiu/yolov7/>
- Yang, J., Li, S., Wang, Z., Dong, H., Wang, J., Tang, S. (2020). Using Deep Learning to Detect Defects in Manufacturing: A Comprehensive Survey and

Current Challenges. *Materials*, 13(24), Article 24.
<https://doi.org/10.3390/ma13245755>

Zhu, Q., Ma, K., Wang, Z., Shi, P. (2023). YOLOv7-CSAW for maritime target detection. *Frontiers in neurorobotics*, 17, 1210470.
<https://doi.org/10.3389/fnbot.2023.1210470>

Zou, Z., Chen, K., Shi, Z., Guo, Y., Ye, J. (2023). Object Detection in 20 Years: A Survey. *Proceedings of the IEEE*, 111(3), 257–276.
<https://doi.org/10.1109/JPROC.2023.3238524>

Object recognition in marine surveillance cameras

Abstract

Automatic object detection in maritime surveillance or panoramic camera images opens up possibilities

for automatic traffic monitoring, unauthorized movement detection, and hazard or pollution identification. This study investigates the performance of models based on the YOLOv7 architecture for the task of detecting vessels and buoys in images captured by panoramic and surveillance cameras. The models are trained on a dedicated dataset comprising diverse maritime scenes created for this purpose, utilizing transfer learning from models trained on generic images. Additionally, two variants of input handling strategies are examined, and the use of the input image cropping strategy significantly improves detection results, especially for small objects, compared to the baseline model.

Keywords: *computer vision; neural networks; small object detection; ship detection.*