

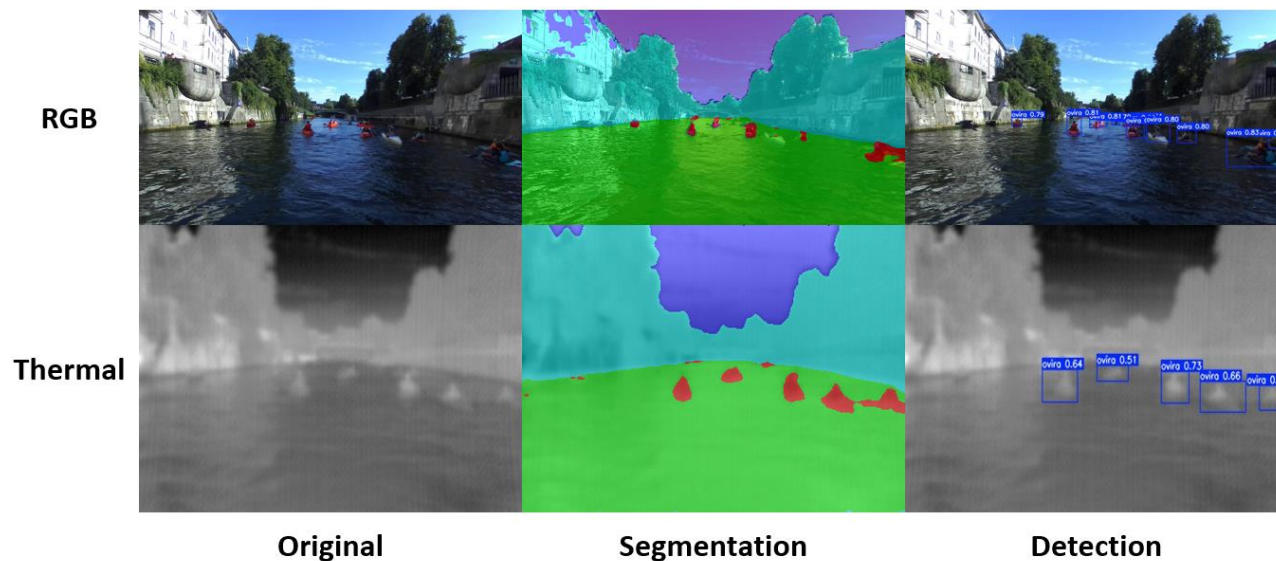
Večmodalna detekcija ovir na vodni površini

Napredne metode računalniškega vida

Tilen Tinta, maj 2025

Cilji seminarja:

- ▶ Osnovni namen:
 - ▶ Semantična segmentacija prizorov iz slik,
 - ▶ Detekcija over na vodni gladini
- ▶ Primerjati uspešnost modelov na barvnih in termalnih slikah,
- ▶ Oceniti kakovost rezultatov s pomočjo metrik mIoU (za segmentacijo) in F1 (za detekcijo),
- ▶ Preizkus različnih modelov nevronske mreže za isto nalogo.



Metodologija:

► Semantična segmentacija:

- Cilj: razvrstiti vsak piksel slike (voda, ovira, ozadje, nebo),
- Namenjena razvrstitvi objektov na sliki v razrede.

► Uporabljeni modeli:

► U-Net:

- Konvolucijska mreža z U-obliko (enkoder, dekoder),
- Uporablja skip povezave za združevanje konteksta in podrobnosti,
- Pogosto uporabljena (tudi razvita) na medicinskem področju.

► SegFormer:

- Transformerju podoben model (encoder del) z MLP dekoderjem,
- Učinkovit in lahek, združuje lokalne in globalne značilnosti,
- Dosega visoko natančnost segmentacije na različnih vhodnih podatkih.

► Metrika za oceno uspešnosti:

► mIoU (angl. mean Intersection over Union):

- Povprečna vrednost preseka in unije med napovedanimi in resničnimi območji,
- Višja vrednost pomeni boljše prekrivanje segmentacije.

Metodologija:

► Detekcija:

- Cilj: Prepoznati in lokalizirati ovire na vodni gladini,
- Vsaka ovira je označena z omejitvenim (angl. bounding box) okvirom.

► Uporabljen model:

► YOLO11

- Zadnja različica serije YOLO (angl. You Only Look Once),
- Izvaja detekcijo objektov v enem prehodu (enostopenjska detekcija),
- Velika hitrost in dobra natančnost.

► Metrika za oceno uspešnosti:

► F1-metrika:

- Kombinacija natančnosti (angl. precision) in priklica (angl. recall),
- Primerna pri neuravnoteženih podatkih,
- $$F1 = 2 * \left(\frac{precision \times recall}{precision + recall} \right)$$

Eksperiment - semantična segmentacija:

- Barvne slike

► Podatki:

- Vhodne slike: barvne slike vodne gladine z okolico,
- Zajem z običajno kamero v naravnem okolju.
- Učna zbirka razdeljena na 80:20 (učna:validacijska)
- Slike resolucije 640x360 (pomanjšane slike)

► Postopek:

- Vhodne barvne slike so bile podane modeloma **U-Net** in **SegFormer**.
- Izhod: semantične maske (vsak piksel dobi oznako - ovira / voda / okolica / nebo).
- Ocenjevanje natančnosti: primerjava napovedane maske z ročno izdelano masko (angl. ground truth) vsako iteracijo
- Na osnovi rezultata validacije se je spremljalo učenje in po potrebi to ustavilo

Eksperiment - semantična segmentacija:

- Termalne slike

► Podatki:

- Vhodne slike: termalne slike vodne gladine z okolico,
- Zajem s termalno kamero,
- Učna zbirka razdeljena na 80:20 (učna:validacijska),
- Slike resolucije 384x288 (nespremenjena resolucija).

► Postopek:

- Vhodne termalne slike so bile podane modeloma **U-Net** in **SegFormer**,
- Izhod: semantične maske (vsak piksel dobi oznako - ovira / voda / okolica / nebo),
- Ocenjevanje natančnosti: primerjava napovedane maske z ročno izdelano masko (ground truth) vsako iteracijo,
- Na osnovi rezultata validacije se je spremljalo učenje in po potrebi to ustavilo.

Eksperiment - detekcija ovir:

- Barvne slike

► Podatki:

- Vhodne slike: barvne slike vodne gladine z ovirami (enake kot za segmentacijo),
- Učna zbirka razdeljena na 80:20 (učna:validacijska),
- Slike resolucije 640x360 (yolo jih dodatno obdela za so te kvadratne),
- Shranjene v ustrezno hierarhijo map,
- Detekcije na slikah pretvorjene iz *.json* zapisa v "YOLO" zapis (*.txt* datoteke za vsako sliko posebej)

► Postopek:

- Uporabljen model YOLO11-s, YOLO11-m in YOLO11-x za **detekcijo** objektov.
- Vhod: barvna slika,
- Izhod: omejitveni okviri z razredom objekta (le ena kategorija - ovira),
- Ocenjevanje uspešnosti detekcije na validacijskem delu zbirke,
- Sprotno spremljanje vrednosti F1-metrike skozi učenje.

Eksperiment - detekcija ovir:

- Termalne slike

► Podatki:

- Vhodne slike: termalne slike vodne gladine z ovirami (enake kot za segmentacijo),
- Učna zbirka razdeljena na 80:20 (učna:validacijska),
- Slike resolucije 384x288,
- Shranjene v ustrezno hierarhijo map,
- Detekcije na slikah pretvorjene iz *.json* zapisa v "YOLO" zapis (*.txt* datoteke za vsako sliko posebej).

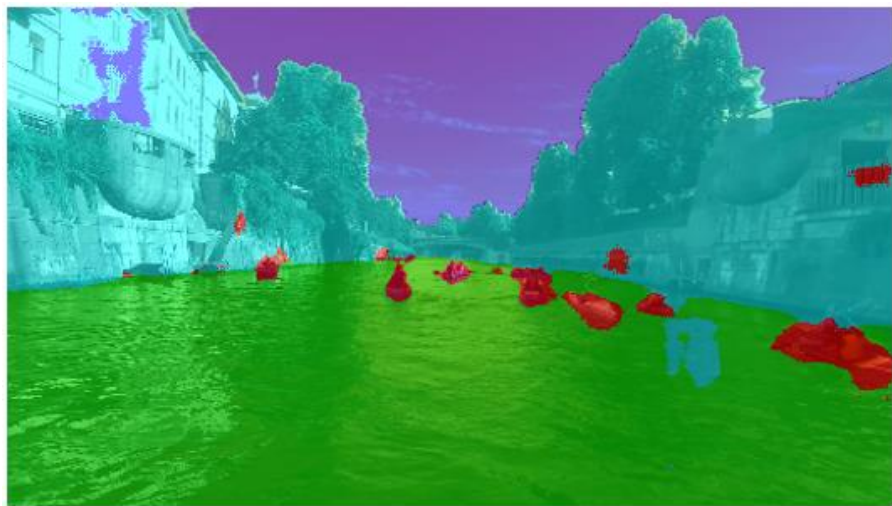
► Postopek:

- Uporabljen model YOLO11-s, YOLO11-m in YOLO11-x za **detekcijo** objektov,
- Vhod: termalna slika,
- Izhod: omejitveni okviri z razredom objekta (le ena kategorija - ovira),
- Ocenjevanje uspešnosti detekcije na validacijskem delu zbirke,
- Sprotno spremljanje vrednosti F1-metrike skozi učenje.

Rezultati - segmentacija:

Arhitektura	mIoU
U-Net	0.757838
SegFormer	0.772018

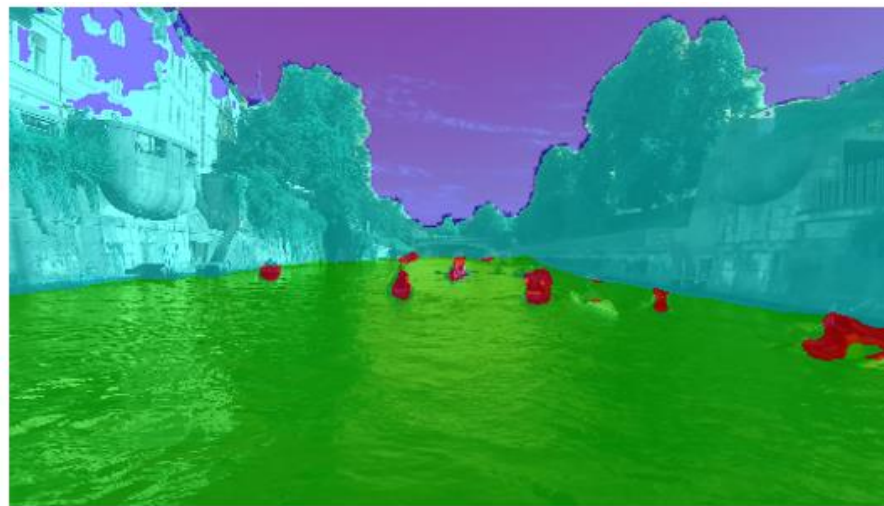
Barvne slike



U-Net

Arhitektura	mIoU
U-Net	0.681646
SegFormer	0.793875

Termalne slike



SegFormer

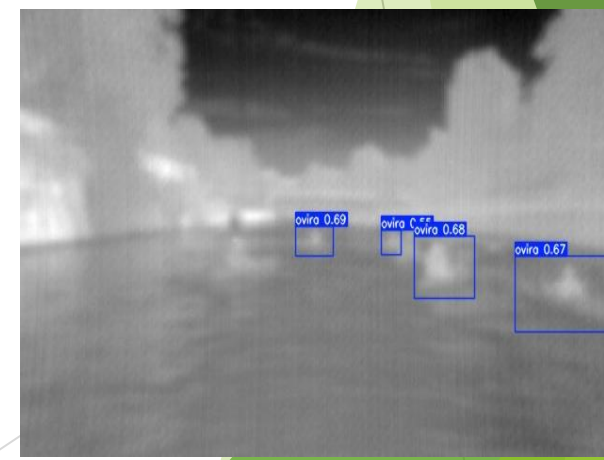
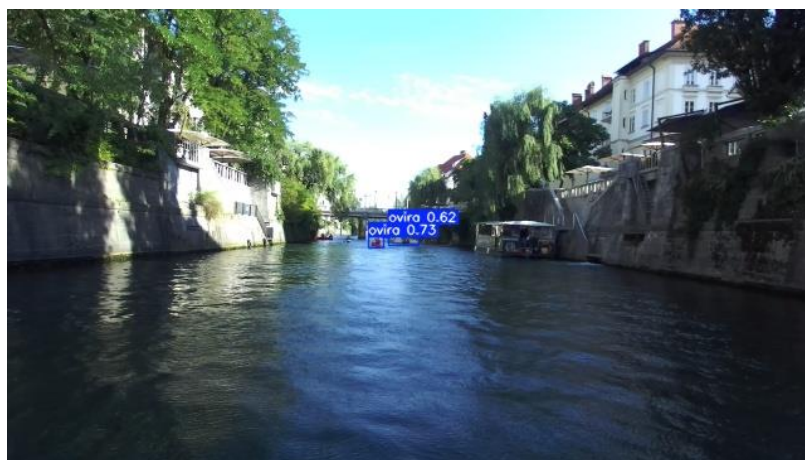
Rezultati - detekcija:

Barvne slike

Mera	YOLO11-s	YOLO11-m	YOLO11-x
Precision	0.682	0.782	0.722
Recall	0.410	0.449	0.477
F1 Score	0.512	0.570	0.574

Termalne slike

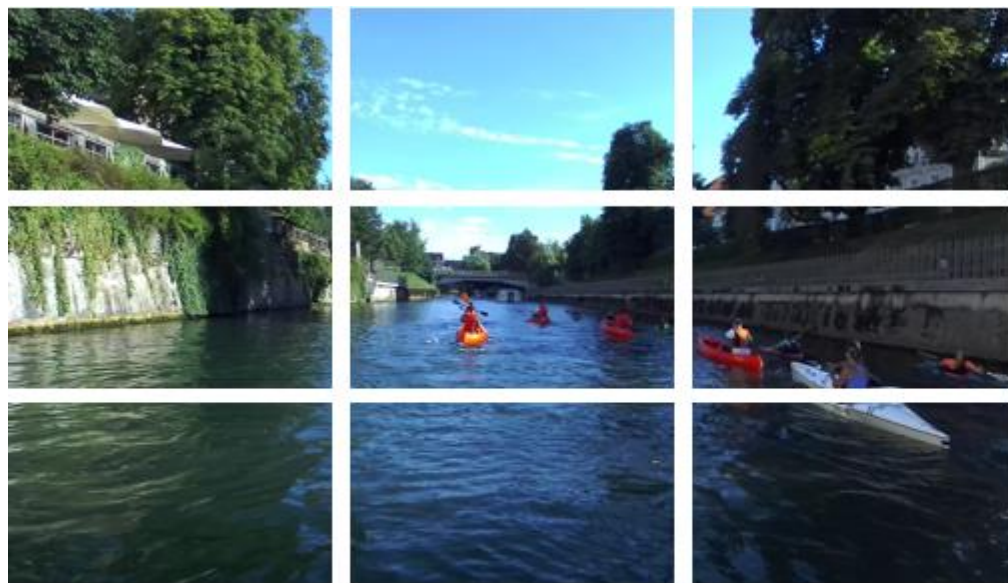
Mera	YOLO11-s	YOLO11-m	YOLO11-x
Precision	0.500	0.562	0.578
Recall	0.297	0.387	0.279
F1 Score	0.373	0.380	0.376



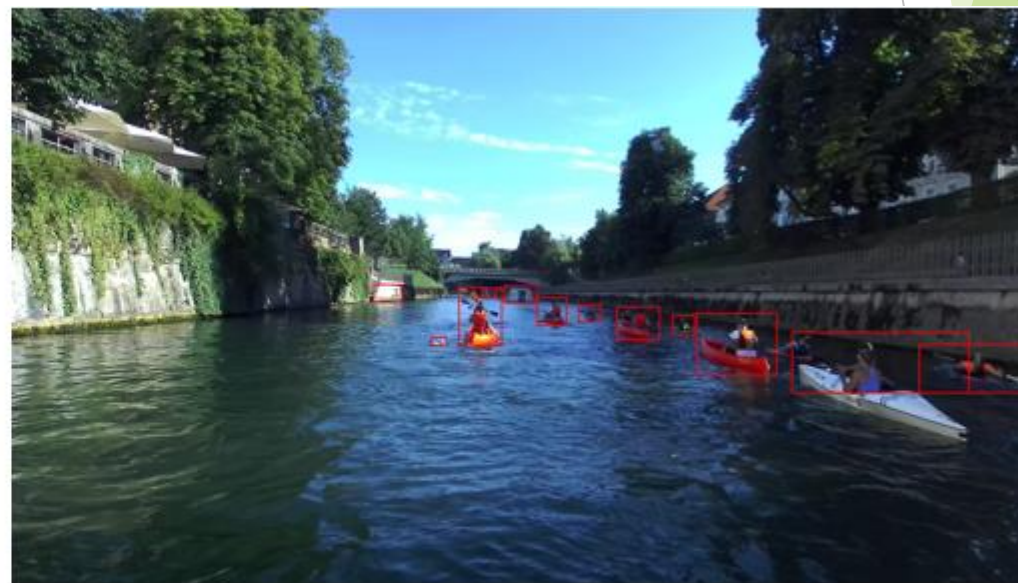
Rezultati - detekcija (razrezane slike):

Barvne slike

Mera	YOLO11-s	YOLO11-m	YOLO11-x
Precision	0.685	0.814	0.774
Recall	0.540	0.561	0.560
F1 Score	0.604	0.664	0.650



Razrezana slika



Slika z detekcijami

Zaključek:

- ▶ Segmentacija in detekcija sta bili uspešno izvedeni na barvnih in termalnih slikah,
- ▶ Rezultati kažejo, da barvne slike omogočajo bolj zanesljivo prepoznavo ovir v pogojih, kjer so slike bile zajete
- ▶ U-Net in SegFormer sta pri segmentaciji oba pokazala dobro sposobnost segmentacije tako na termalnih kot barvnih slikah - SegFormer je bil natančnejši pri robovih,
- ▶ YOLOv11 je učinkovito zaznaval ovire, vendar so rezultati zelo slabi (težave pri detekciji manjših objektov),
- ▶ Najboljše je deloval YOLO11-m
- ▶ Razrez slik na dele izboljša detekcijo manjših ovir a krši osnovno idejo YOLO-ta (single pass)

Hvala za pozornost!