

# Usporedna analiza YOLOv5, YOLOv8 i YOLOv11 modela u detekciji naslovnica stripova

Vedran Dejanović, Ivan Šimac, Ana Vrcelj Božić

Veleučilište u Rijeci  
CASE 2026

Rijeka, 10. ožujka 2026.

# Sažetak projekta

- **Problem:** Detekcija naslovnica stripova (Zagor, Tex, Dragonero) u različitim uvjetima
- **Metoda:** Usporedba YOLOv5s, YOLOv8s i YOLOv11s na datasetu od 822 slike
- **Rezultati:** YOLOv11s najbolji ( $mAP@0.5=0.78$ , Precision=0.82)

# Kontekst i motivacija

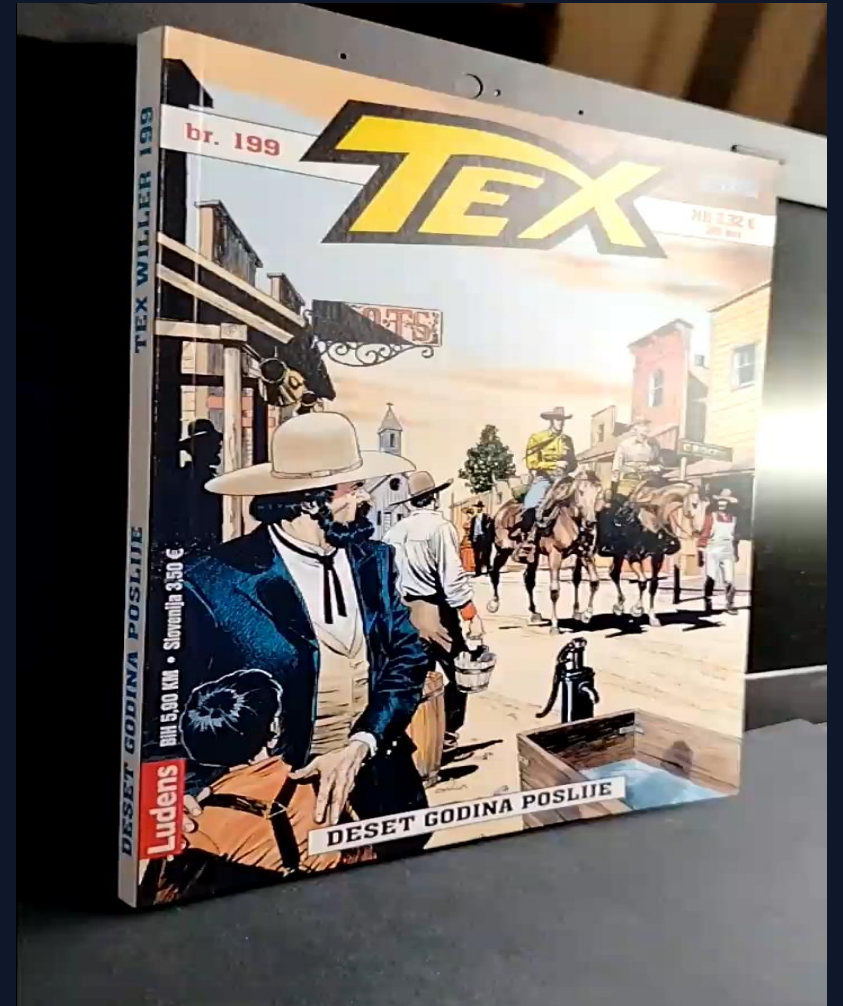
- **Računalni vid:** Detekcija objekata jedan od najraširenijih zadataka
- **YOLO modeli:** Iznimna brzina i mogućnost rada u stvarnom vremenu
- **Osobna potreba:** Organizacija strip kolekcije - automatska identifikacija naslovnica iz fotografija

# Cilj projekta

- **Analiza i usporedba** tri verzije YOLO modela: YOLOv5s, YOLOv8s i YOLOv11s
- **Dataset:** Tri klase naslovnica stripova (Zagor, Tex, Dragonero)
- **Alati:** Label Studio za anotaciju, Google Colab za treniranje, Anaconda za testiranje
- **Evaluacija:** Kvantitativne metrike i kvalitativne analize robustnosti

# Priprema dataseta: Snimanje

- **Inovativni pristup:** Video materijal umjesto pojedinačnih fotografija
- **36 videa** (720p, 30fps) s naslovnicama Zagora, Texa i Dragonera
- **Različiti uvjeti snimanja:** Kutovi (ravno, nagore, sa strane), osvjetljenja (prirodno, umjetno, mješovito)
- **Dinamički pokreti kamere** tijekom snimanja



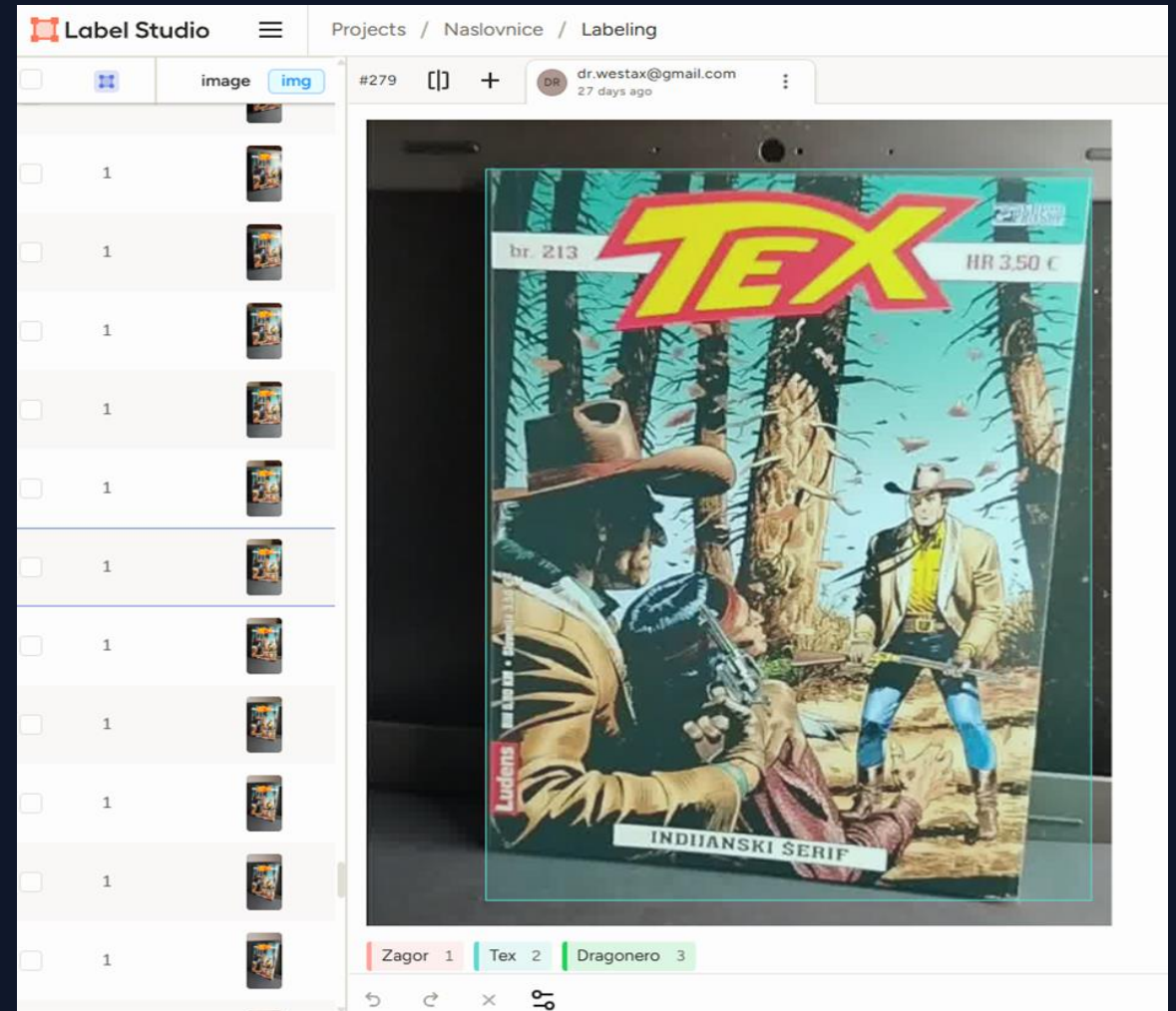
# Ekstrakcija frameova

- **FFmpeg alat:** Ekstrakcija svakog 25. framea (~1 frame/sekundi)
- **Rezultat:** 822 frameova iz 36 videa
- Raznovrsni uglovi i osvjetljenja za dobru generalizaciju

```
ffmpeg -i Zagor1.mp4 -vf "select=not(mod(n\\,25))" -vsync vfr -start_number 1  
Zagor_frame_%03d.jpg
```

# Anotacija u Label Studio

- **Anaconda okruženje:**  
Python 3.12, Label Studio instalacija
- **Projekt "Naslovnice":** 3 klase (Zagor, Tex, Dragonero)
- **Proces:** Import slika, crtanje bounding boxova, klasifikacija, review
- **822 frameova** ručno anotirano s YOLO formatom



# Distribucija dataseta

## Klase

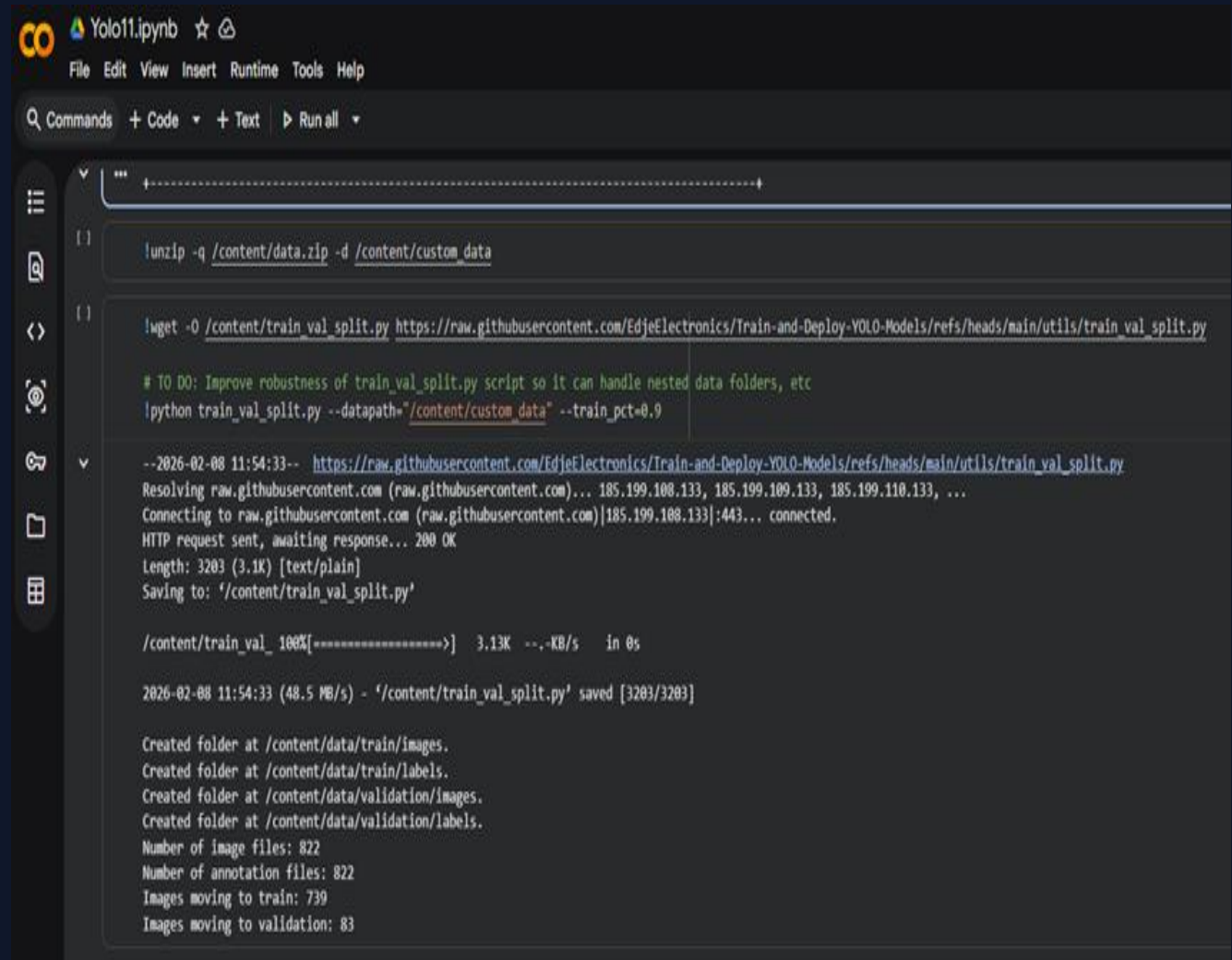
- Dragonero: 349 (42.8%)
- Tex: 239 (28.8%)
- Zagor: 234 (28.4%)

## Struktura

- 822 JPG slika (images/)
- 822 TXT labela (labels/)
- classes.txt datoteka

# Google Colab: Upload i priprema

- **Export:** ZIP arhiva iz Label Studio (YOLO with Images format)
- **Upload:** Prebacivanje u Google Colab custom\_data mapu
- **Podjela:** 90% train (739 slika), 10% validation (83 slike)



```
Yolo11.ipynb
File Edit View Insert Runtime Tools Help

Q Commands + Code + Text ▶ Run all ▼

[ ]
!unzip -q /content/data.zip -d /content/custom_data

[ ]
!wget -O /content/train_val_split.py https://raw.githubusercontent.com/EdgeElectronics/Train-and-Deploy-YOLO-Models/refs/heads/main/Utils/train_val_split.py

# TO DO: Improve robustness of train_val_split.py script so it can handle nested data folders, etc
!python train_val_split.py --datapath="/content/custom_data" --train_pct=0.9

--2026-02-08 11:54:33-- https://raw.githubusercontent.com/EdgeElectronics/Train-and-Deploy-YOLO-Models/refs/heads/main/Utils/train_val_split.py
Resolving raw.githubusercontent.com (raw.githubusercontent.com)... 185.199.108.133, 185.199.109.133, 185.199.110.133, ...
Connecting to raw.githubusercontent.com (raw.githubusercontent.com)|185.199.108.133|:443... connected.
HTTP request sent, awaiting response... 200 OK
Length: 3203 (3.1K) [text/plain]
Saving to: '/content/train_val_split.py'

/content/train_val_100%[=====] 3.13K --KB/s in 0s

2026-02-08 11:54:33 (48.5 MB/s) - '/content/train_val_split.py' saved [3203/3203]

Created folder at /content/data/train/images.
Created folder at /content/data/train/labels.
Created folder at /content/data/validation/images.
Created folder at /content/data/validation/labels.
Number of image files: 822
Number of annotation files: 822
Images moving to train: 739
Images moving to validation: 83
```

# Instalacija i konfiguracija

## Ultralytics instalacija

Omogućuje učitavanje, detekciju, validaciju i treniranje svih YOLO modela (v5/v8/v11)

### data.yaml konfiguracija

- Putanje do train/val foldera
- Broj klasa: 3
- Nazivi: Dragonero, Tex, Zagor



Created config file at /content/data.yaml

File contents:

```
path: /content/data
train: train/images
val: validation/images
nc: 3
names:
  - Dragonero
  - Tex
  - Zagor
```

# Parametri treninga

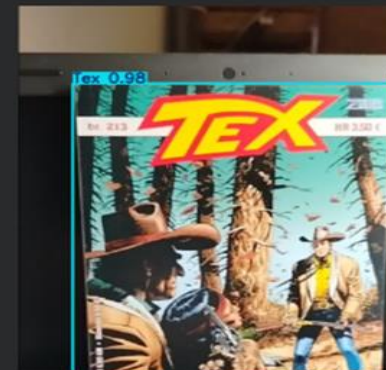
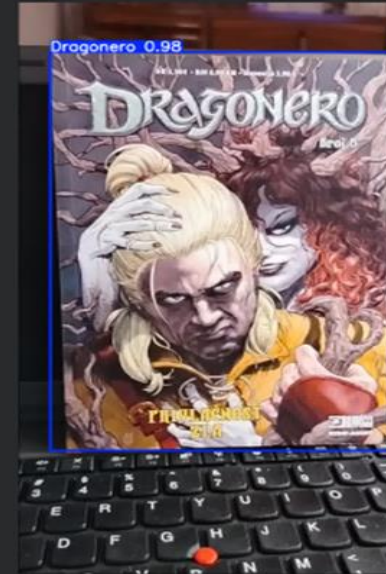
- **60 epoha** za sve tri modela (YOLOv5s, YOLOv8s, YOLOv11s)
- **Batch size:** 16 slika (automatsko prilagođavanje GPU memoriji)
- **Rezolucija:** 640x640 piksela
- **Early stopping:** 25 epoha bez poboljšanja metrika

```
!yolo detect train data=/content/data.yaml model=yolov5s.pt epochs=60 imgsz=640  
!yolo detect train data=/content/data.yaml model=yolov8s.pt epochs=60 imgsz=640  
!yolo detect train data=/content/data.yaml model=yolo11s.pt epochs=60 imgsz=640
```

# Treniranje i validacija

- **Epoha:** Jedan kompletni prolaz kroz training set od 739 slika
- **Validacija:** Automatska nakon svake epohe na 83 slike
- **best.pt model:** Spremanje najboljeg modela prema mAP metrici
- **Vizualizacija:** Predikcije s bounding boxovima za praćenje napretka

```
[ ] import glob
from IPython.display import Image, display
for image_path in glob.glob(f'/content/runs/detect/predict/*.jpg')[:10]:
    display(Image(filename=image_path, height=400))
    print('\n')
```



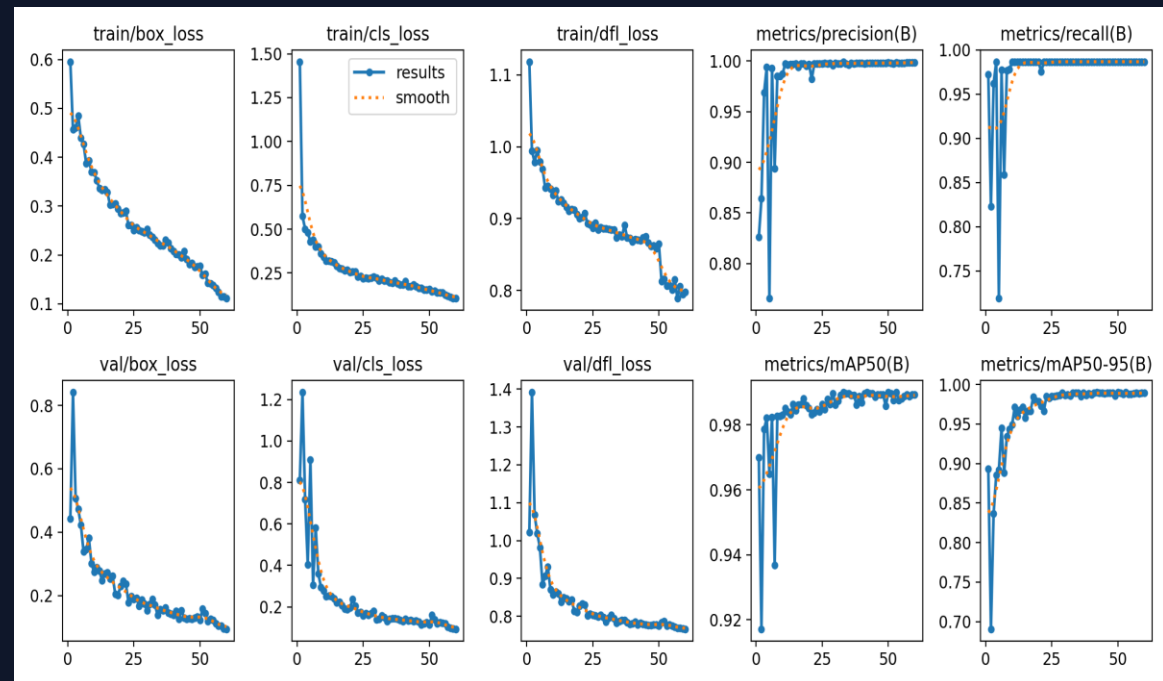
# YOLOv5s - Rezultati

## Metrike

- Precision: ~0.70
- Recall: ~0.72
- mAP@0.5: ~0.68
- mAP@0.5:0.95: ~0.45

## Karakteristike

- Sporiji, ali stabilan napredak
- Box loss: 1.6 → 0.55
- Bez naglog overfittinga
- Baseline referenca



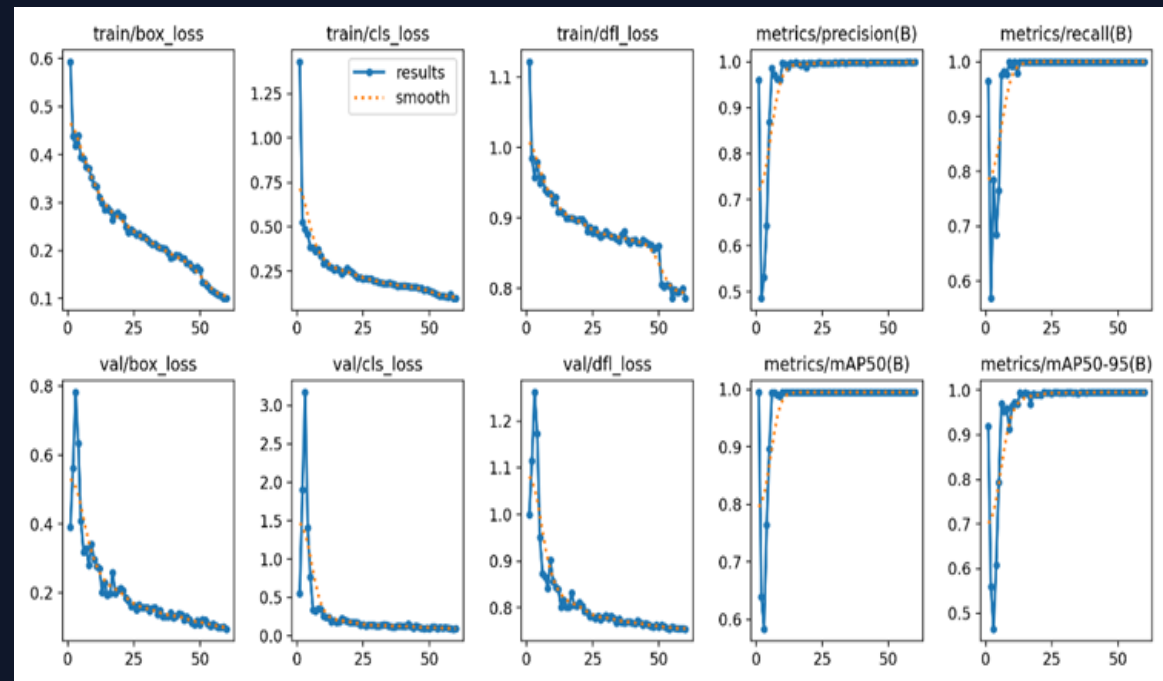
# YOLOv8s - Rezultati

## Metrike

- Precision: 0.78
- Recall: 0.75
- mAP@0.5: 0.74
- mAP@0.5:0.95: 0.51

## Karakteristike

- Brža konvergencija
- Box loss: 1.5 → 0.45
- C2f moduli - efikasna ekstrakcija
- Optimalan balans



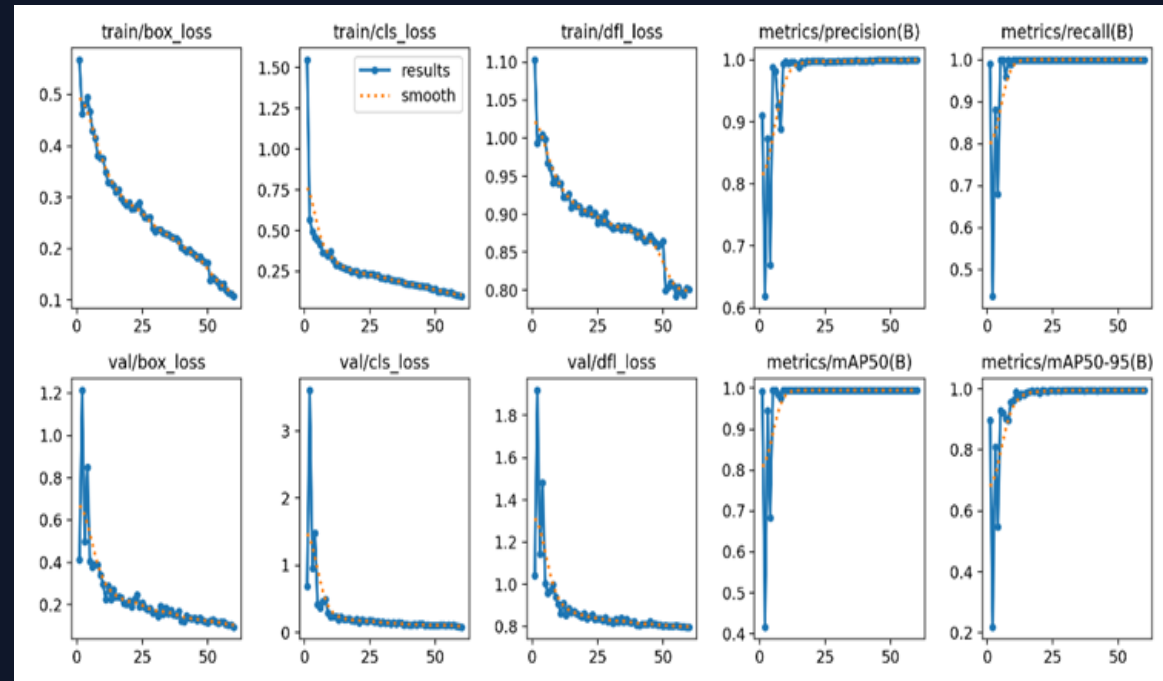
# YOLOv11s - Rezultati

## Metrike

- Precision: 0.82
- Recall: 0.77
- mAP@0.5: 0.78
- mAP@0.5:0.95: 0.55

## Karakteristike

- Najbrža konvergencija
- Box loss: 1.3 → 0.40
- C2PSA neck arhitektura
- Vrhunac YOLO evolucije

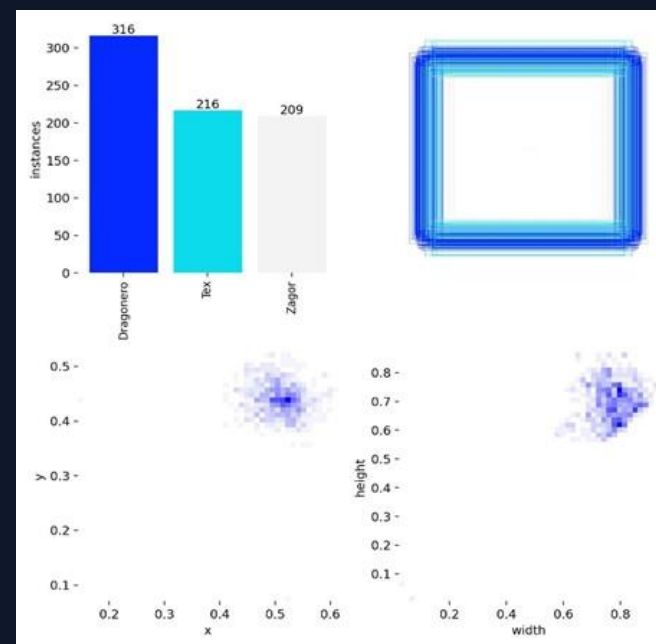
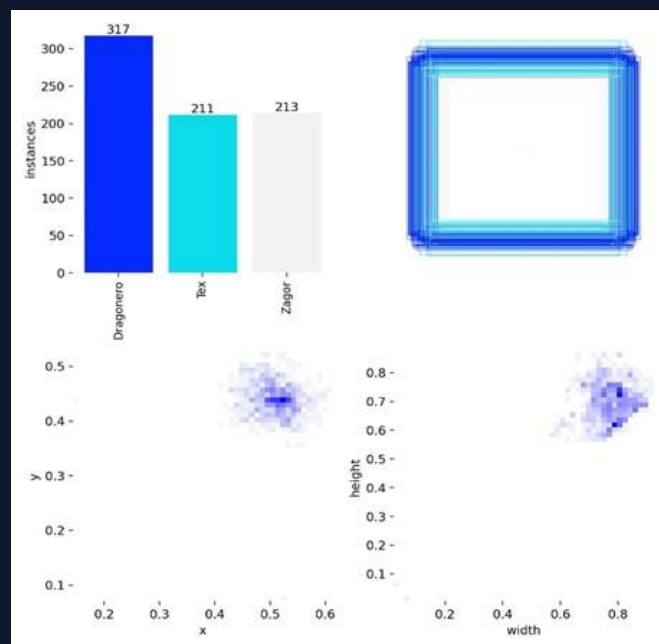
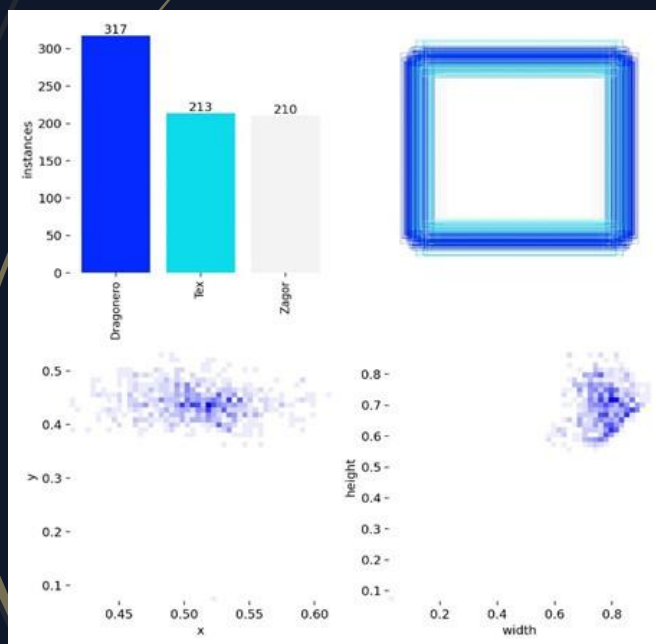


# Usporedba performansi modela

| Metrika   | YOLOv5s | YOLOv8s | YOLOv11s |
|-----------|---------|---------|----------|
| Precision | 0.70    | 0.78    | 0.82     |
| Recall    | 0.72    | 0.75    | 0.77     |
| mAP@0.5   | 0.68    | 0.74    | 0.78     |

# Usporedba prepoznavanja klasa

- **Konzistentnost:** Svi modeli registriraju sličnu distribuciju
- **Dragonero:** 316-317 instanci (~38%) - dominantna klasa
- **Tex i Zagor:** 209-213 instanci (~25% svaki)
- **Bounding box karakteristike:** 70-80% površine slike, centri ~0.5

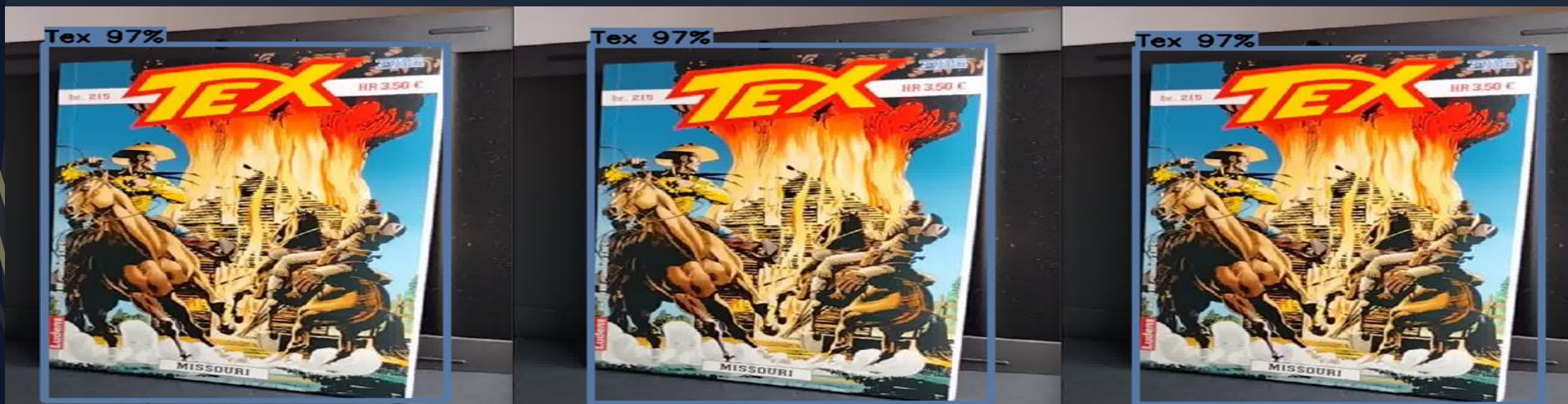


# Desktop testiranje - Poznate naslovnice

yolo\_detect.py skripta za lokalno testiranje u Anaconda okruženju

## Rezultati

- Svi modeli: **96-97% confidence**
- Precizni bounding boxovi na svim klasama
- Bez lažnih pozitiva



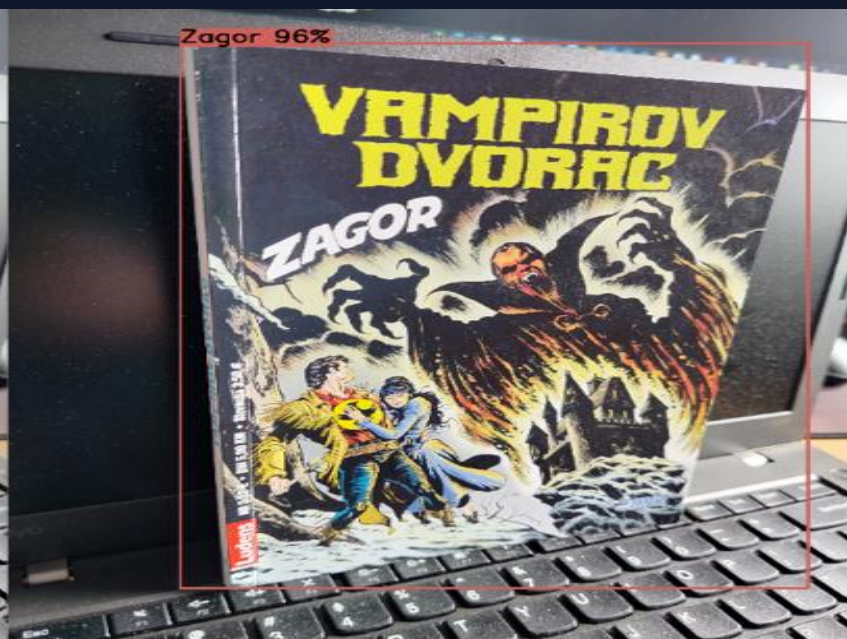
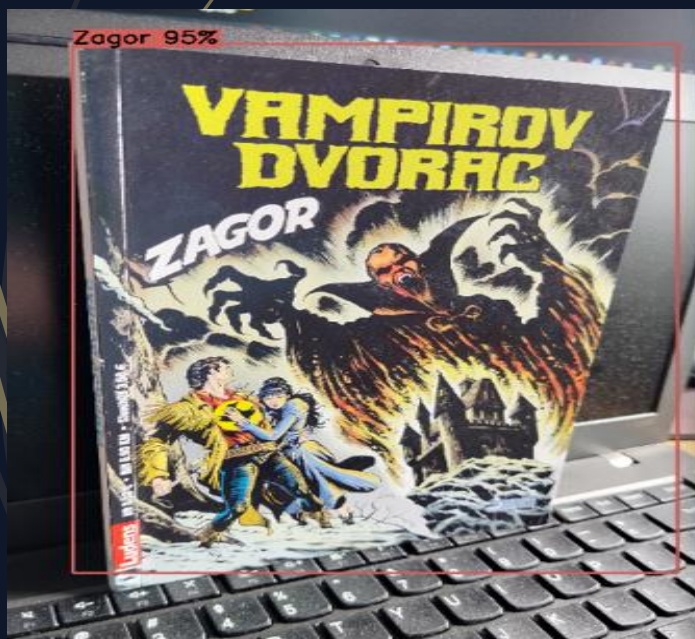
# Desktop testiranje - Nove naslovnice

## Robusnost

- YOLOv5s: 94.7%
- YOLOv8s: 95%
- YOLOv11s: 92.3%

## Zaključak

YOLOv8s najstabilniji na neviđenim slikama



# Real-time testiranje: Zagor

Web kamera testiranje u stvarnim uvjetima - simulacija praktične upotrebe

## Rezultati - Zagor naslovnice

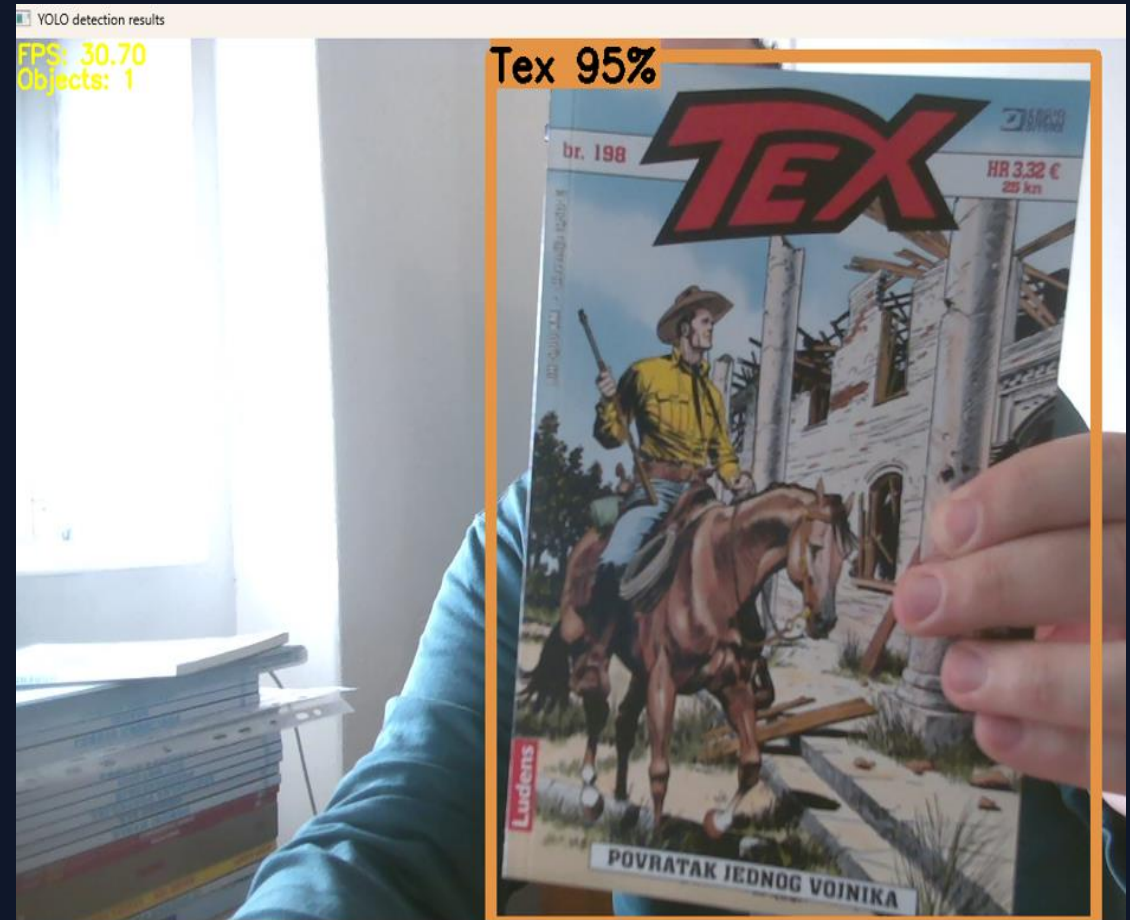
- YOLOv5s, v8s, v11s: **>95% confidence**
- Točni bounding boxovi pri ručnom držanju



# Real-time testiranje: Tex

## Rezultati - Tex naslovnice

- YOLOv8s: 97% confidence
- YOLOv11s: 95-97%
- YOLOv5s: 97%
- Bez lažnih pozitiva, precizna lokalizacija



# Real-time testiranje: Dragonero

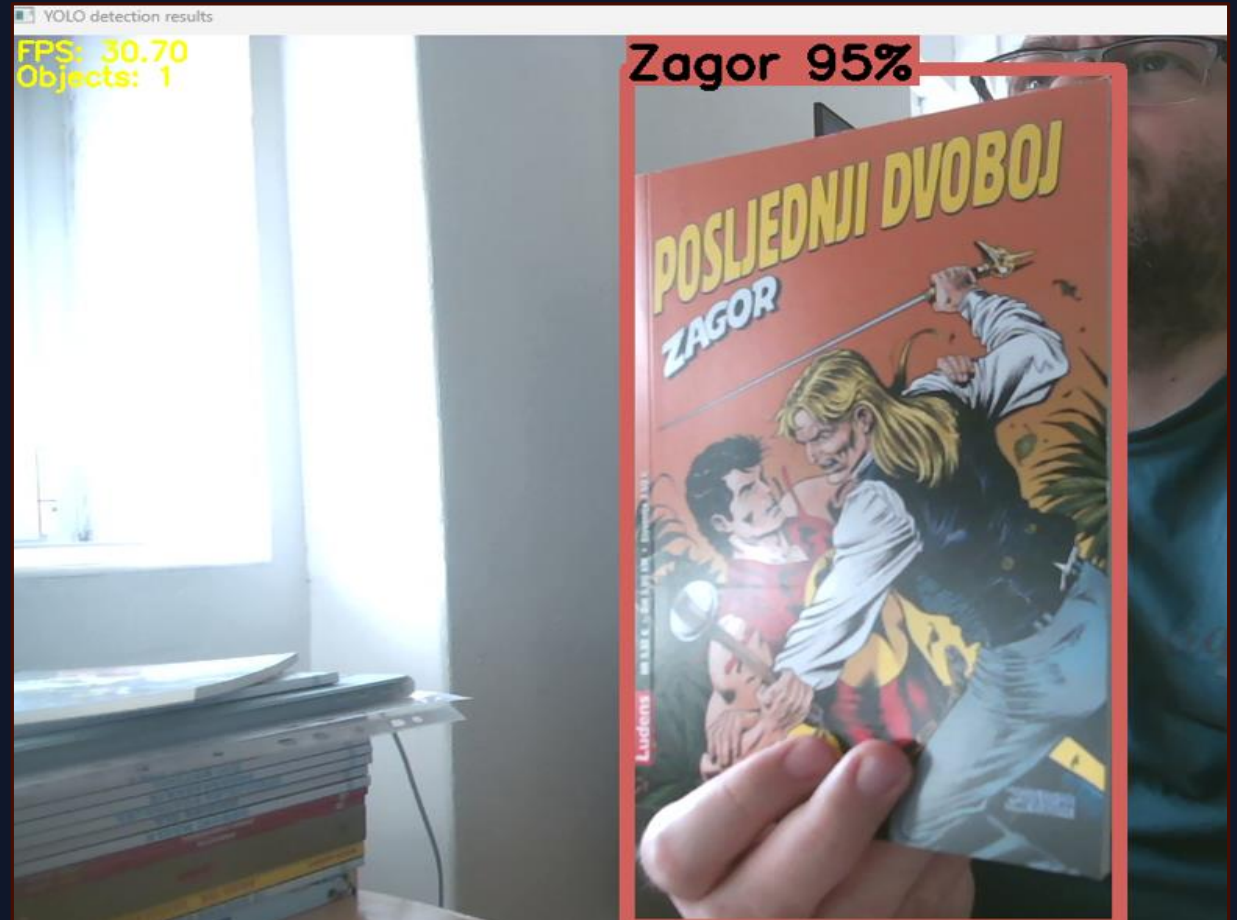
## Rezultati - Dragonero naslovnice

- YOLOv8s: **95% confidence**
- Svi modeli: **>94%**
- Konzistentne performanse preko svih klasa



# Real-time performanse

- **30+ FPS** na svim modelima (produksijska brzina)
- **Robusnost:** Različita osvjetljenja, kutovi, ručno držanje

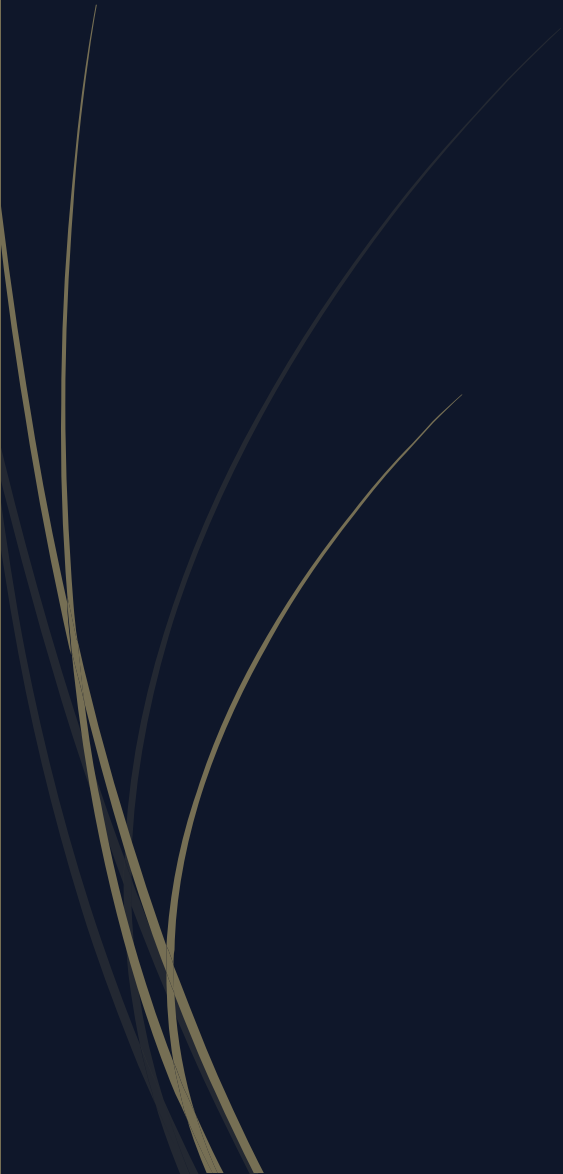


# Daljnja istraživanja

- **Proširenje dataseta:** 1000+ slika za još bolju generalizaciju
- **Fino podešavanje:** Dodavanje više strip junaka i izdavača

# Zaključak

- **YOLOv11s superioran:**  $mAP@0.5=0.78$ , Precision=0.82, najbrža konvergencija
- **YOLOv8s optimalan balans:**  $mAP@0.5=0.74$ , najstabilniji na neviđenim slikama
- **Praktična primjena:** Real-time katalogizacija strip kolekcije s 30+ FPS
- **Kvalitetan dataset:** 822 slike, uravnotežena distribucija klasa



# Hvala na pažnji!

Pitanja?