

Detekcia objektov z nočných scén

Tomáš Koštrna

2023

1 Tutoriál

Cieľom bolo absolvovať tutoriál zo stránky <https://docs.ultralytics.com/quickstart/>. Verzia programovacieho jazyka Python je 3.9.10, verzia kompilátora pre grafickú kartu Cuda je 11.2. Modelom grafickej karty je Nvidia GTX 1660 Super 6GB. Pamäť RAM je 16GB. Spracovávame dataset Coco s 128 obrázkami a anotáciami. Model s ktorým pracujeme je už predtrénovaný **yolov8n.pt**.

1.1 Problém 1

Znenie: AssertionError: Torch not compiled with CUDA enabled in spite upgrading to CUDA version.

Riešenie: Inštalácia python package ultralytics nainštaluje verziu pytorch pre CPU a nie pre GPU, preto treba túto verziu pytorchu vymazať a nainštalovať podľa stránky: <https://pytorch.org/get-started/locally/>. (1)

1.2 Problém 2

Znenie: RuntimeError: An attempt has been made to start a new process before the current process has finished its bootstrapping phase.

Riešenie: Kód modelového procesu musí byť pod názvom prostredia najvyššej úrovne (if name == 'main:'). Samotná inicializácia modelu môže byť nad mainom ale práca s modelom musí byť v rámci mainu. (2)

1.3 Problém 3

Znenie: Default hodnota argumentu amp do funkcie model.train() kompletne pokazila celé tréovanie.

Riešenie: Nastaviť argument amp na false. (3)

2 Získavanie datasetu

Po konzultáciách sme sa rozhodli použiť dataset Night City tento dataset však neobsahoval potrebné anotácie vo formáte .txt alebo .json ale iba ako re-kolorizované obrázky čiže sa v podstate jednalo o segmentáciu ktorú využijeme neskôr. Nad'alej som sa pokúšal pre-konvertovať ExDark dataset ktorý využíva zastaralú a málo používanú formu anotácii bbgt. Konverzia tohto datasetu vyžaduje pre každý obrázok re-škálovanie t.j získať aktuálnu veľkosť a následne skonvertovať na 640x640 pričom treba prepočítať vertikálny aj horizontálny škálovací faktor. Ak máme prepočítané faktory vieme upraviť bbgt anotáciu tak aby sedela pre

640x640 (kompatibilné s YOLOv8) z pôvodných napr. 500x375. Následne potrebujeme bbox (bounding box) ktorý tvorí ľavý horný roh a pravý horný roh prekonvertovať do yolo formátu. Táto časť bola problematická pretože dochádzalo k degenerovaným prípadom kedy tieto hodnoty boli zamenené. Mojou snahou bolo konverziu realizovať manuálnymi výpočtami avšak dostával som záporné hodnoty pre súradnice čo znamená chybu.

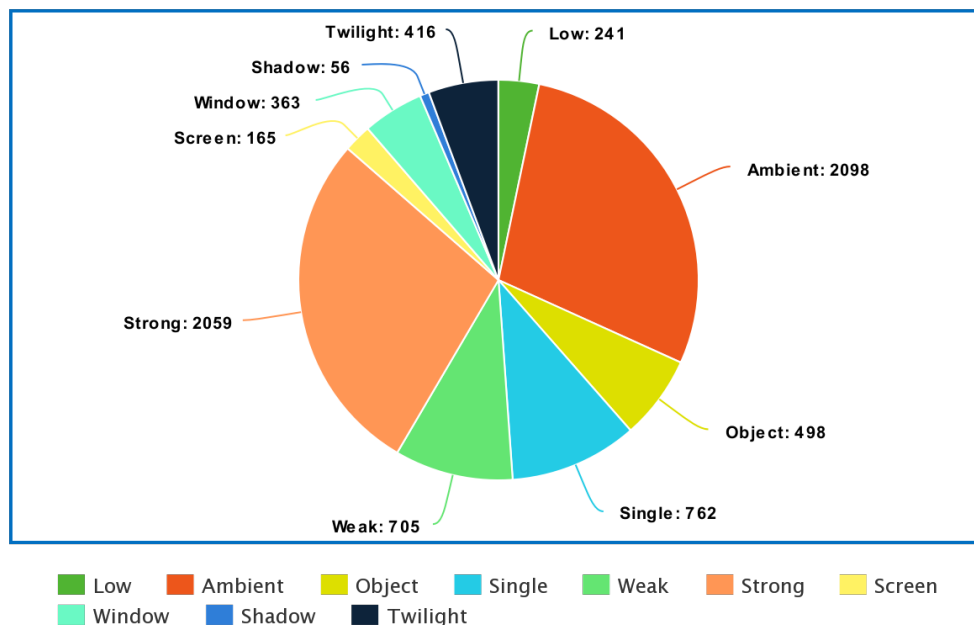
2.1 Získanie datasetu

Zistil som že daný bbgt formát anotácií datasetu ExDark je použiteľný, formát v akom boli zapísané detekcie objektu boli vo formáte ľavý horný roh, šírka a výška. Tento formát je známy aj pod pojmom Coco. Pomocou knižnice ktorú som využil je veľmi jednoduché tieto formáty prekonvertovať. Prekonvertovaný dataset som nahral na stránku Roboflow pomocou ktorej je možné export do YOLOv8 formátu spolu s rozdelením datasetu na tréningovú, testovaciu a validačnú množinu.

3 YOLOv8 EDA na ExDark datasete

3.1 Popis datasetu

Tento dataset pozostáva z 7,363 snímok so slabým osvetlením. Ďalej už len obrázky, tabuľky a grafy.



meta-chart.com

3.2 Porovnanie modelov YOLOv8

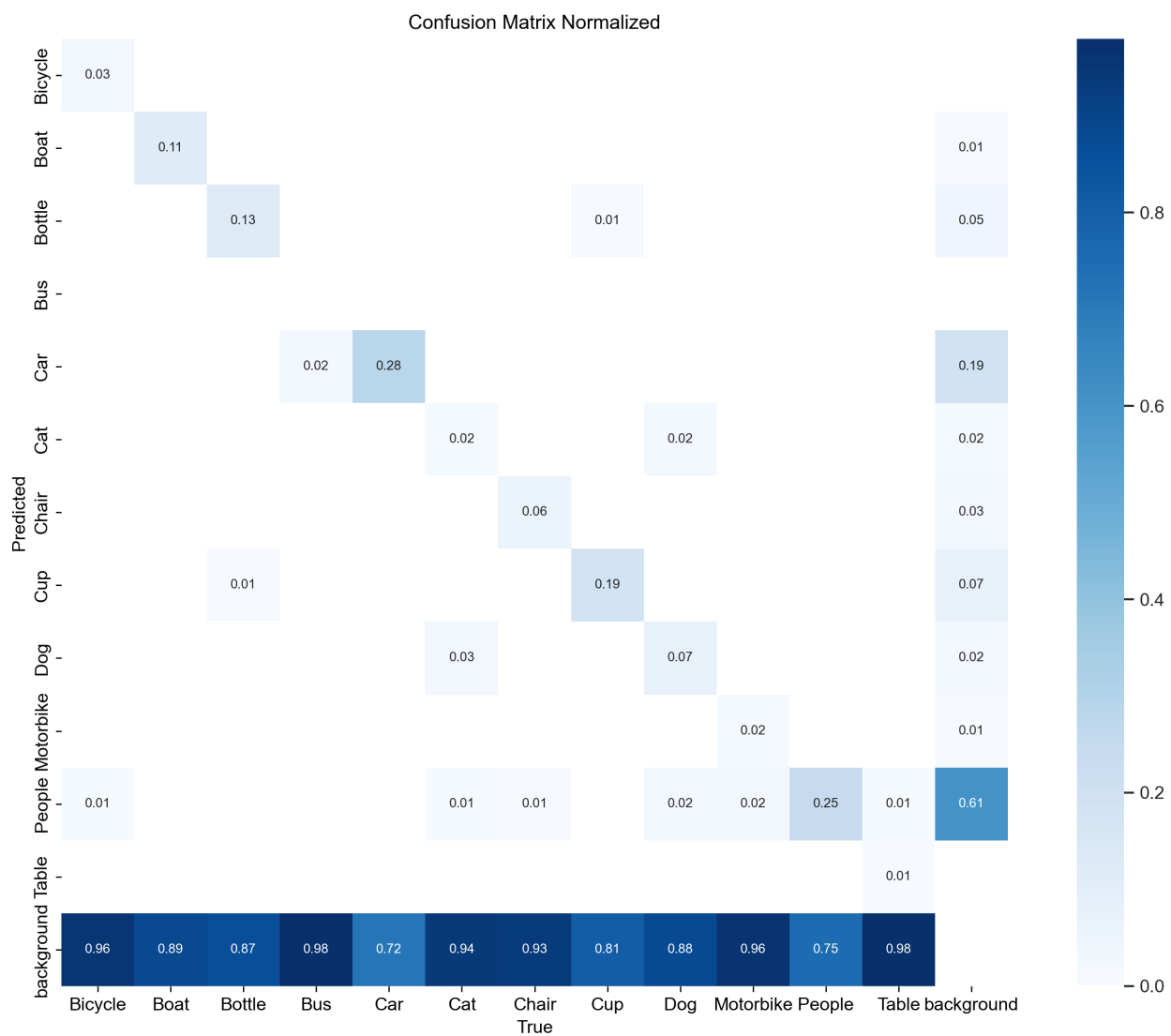
V tomto pokuse som porovnával 3 rôzne modely YOLOv8.

Model	size (pixels)	mAP ^{val} 50-95	Speed CPU ONNX (ms)	Speed A100 TensorRT (ms)	params (M)	FLOPs (B)
YOLOv8n	640	37.3	80.4	0.99	3.2	8.7
YOLOv8s	640	44.9	128.4	1.20	11.2	28.6
YOLOv8m	640	50.2	234.7	1.83	25.9	78.9

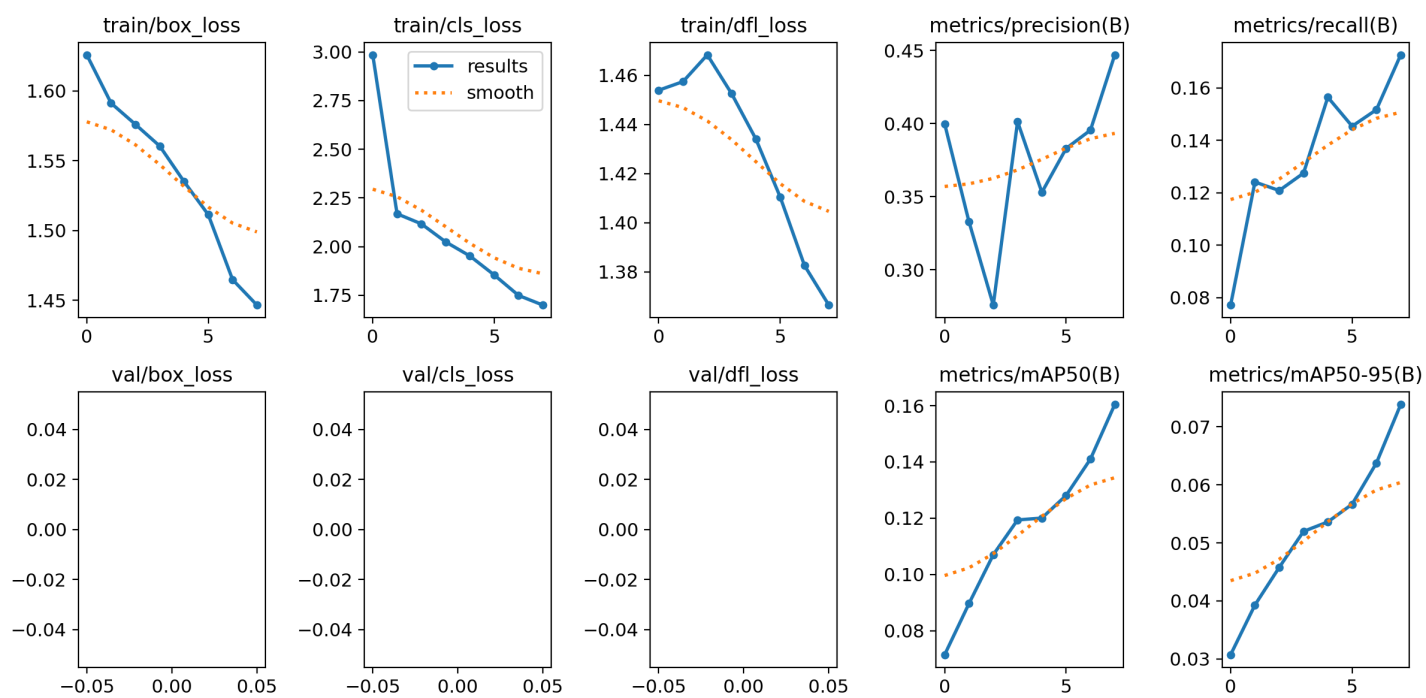
Tieto modely sú odstupňované od najmenšieho nano až po medium. Existujú aj modely s väčším počtom parametrov avšak tieto by sa mi nezmestili do VRAM. Všetky modely boli trénované na datasete ExDark. Vzhľadom na rôzne veľkosti jednotlivých modelov som musel upraviť aj parametre a to hlavne veľkosť dávky a rozmery obrázka. Modely sú schopné automaticky upraviť rozmery podľa vstupného parametra a to na 640x640 alebo 416x416.

3.2.1 YOLOv8n Prvé (krátke) tréningy

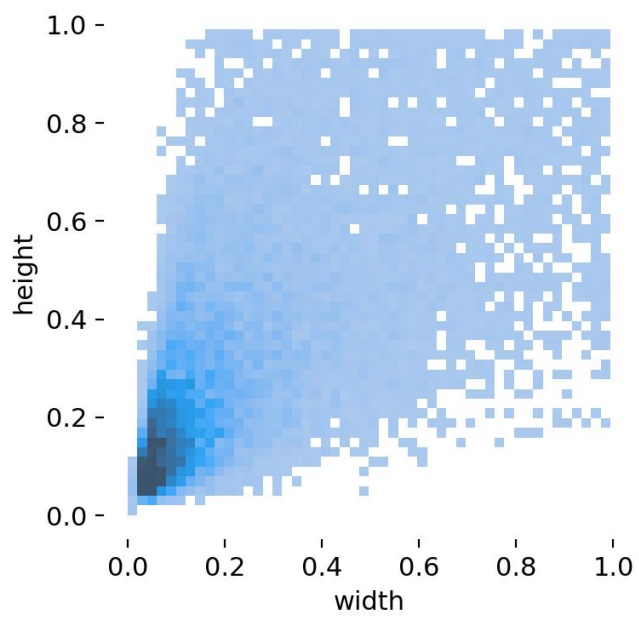
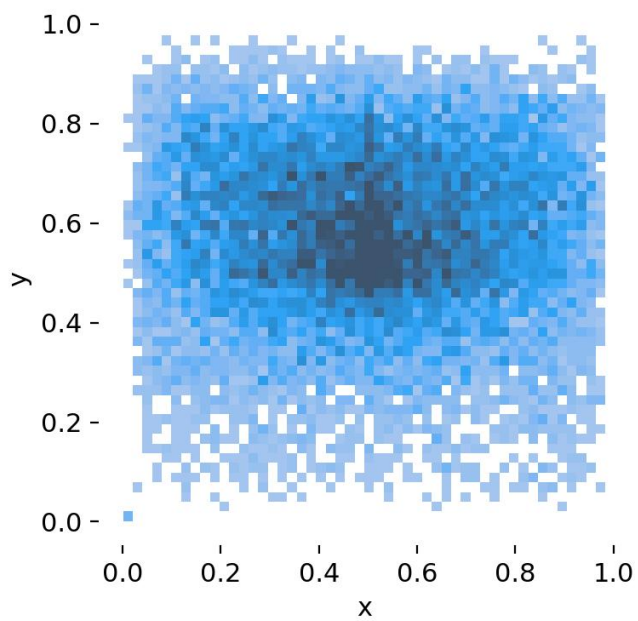
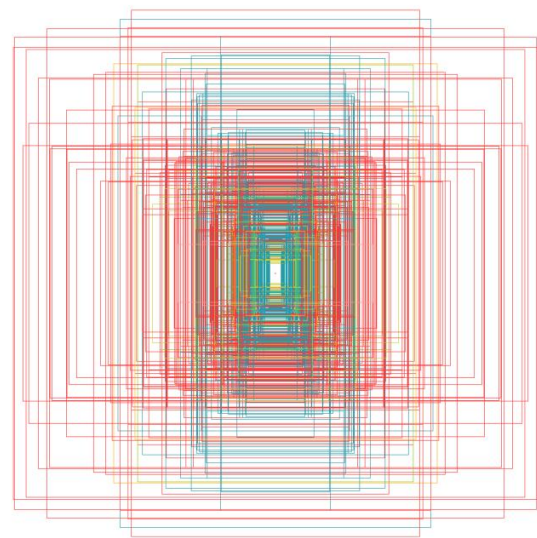
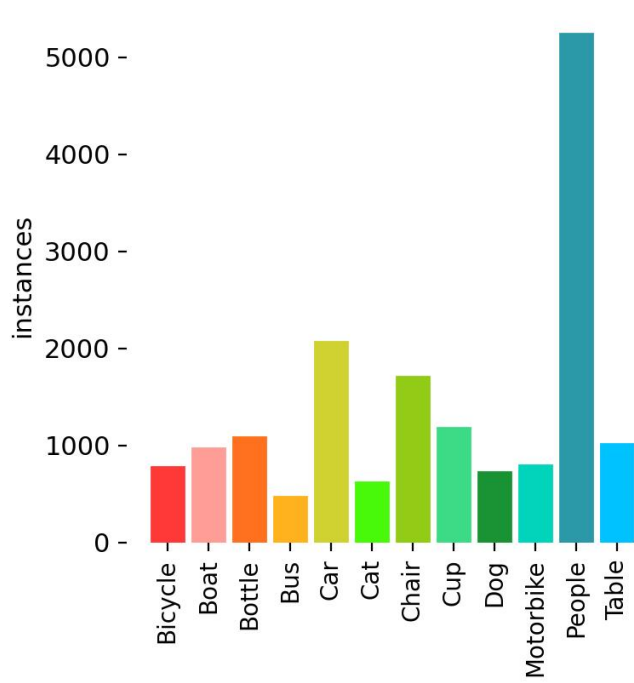
Parameter	Hodnota
Epochy	8
Dávka	32
Rozmery	416x416
lr-0	0.005
lr-f	0.01



Obr. 1: Normalizovaná konfúzna matica



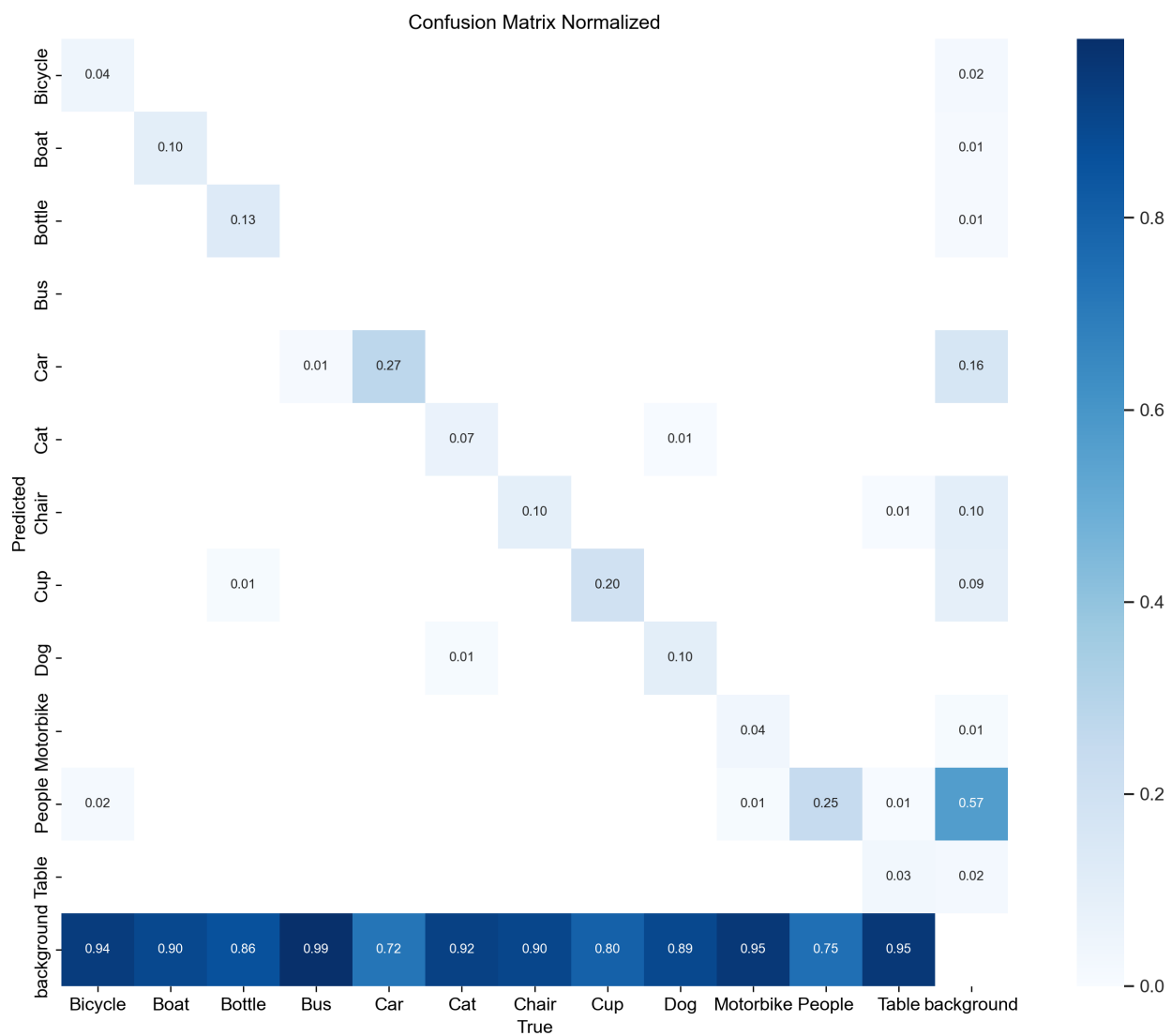
Obr. 2: Výsledky tréningu



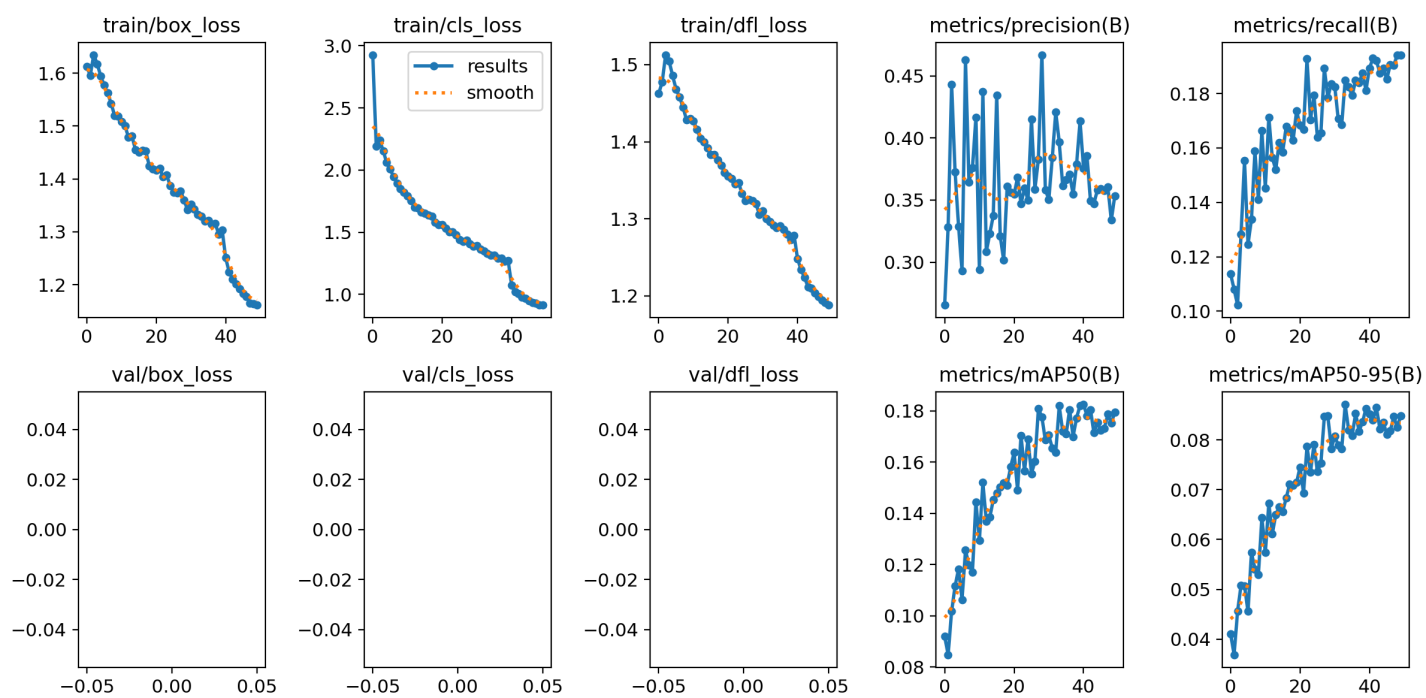
Obr. 3: Anotácie

3.2.2 YOLOv8n Druhé (dlhé) tréovanie

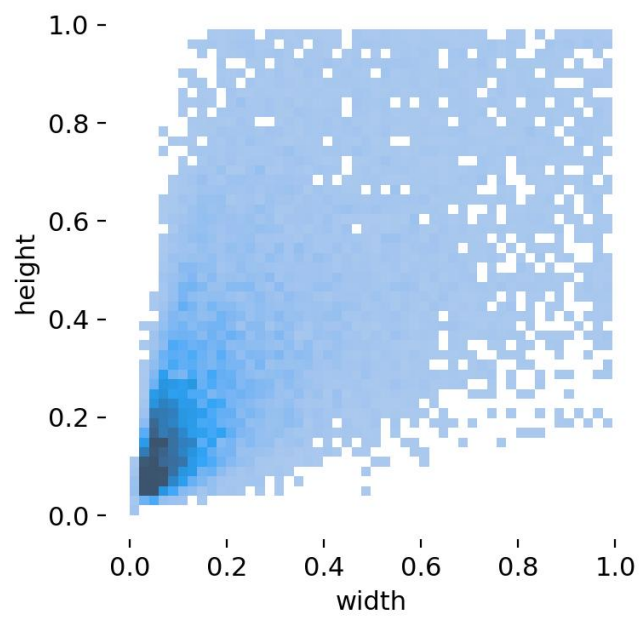
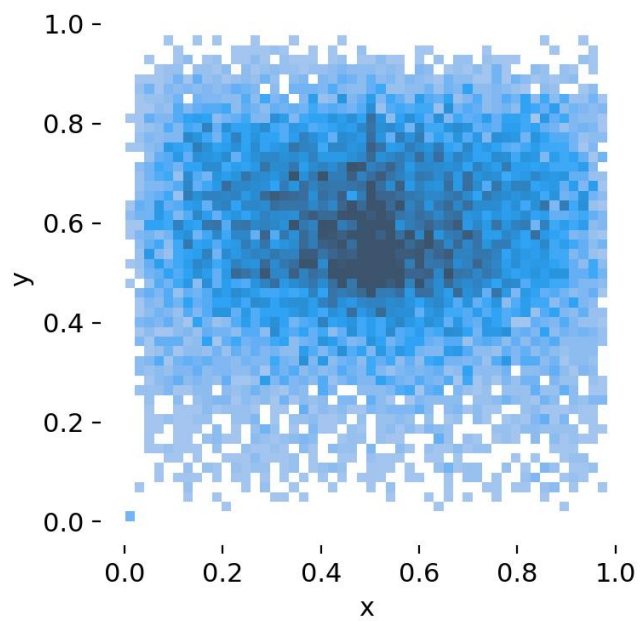
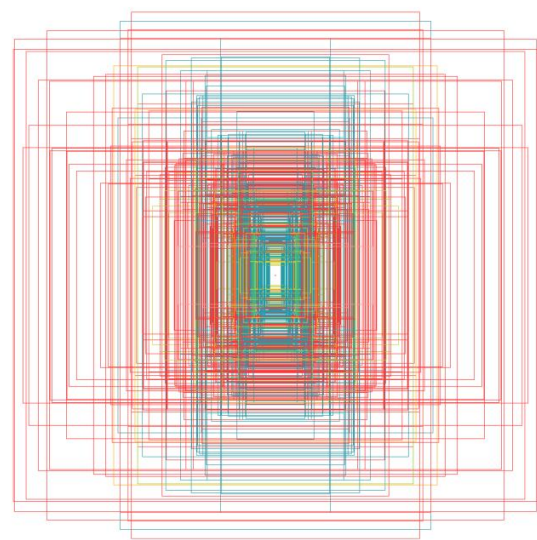
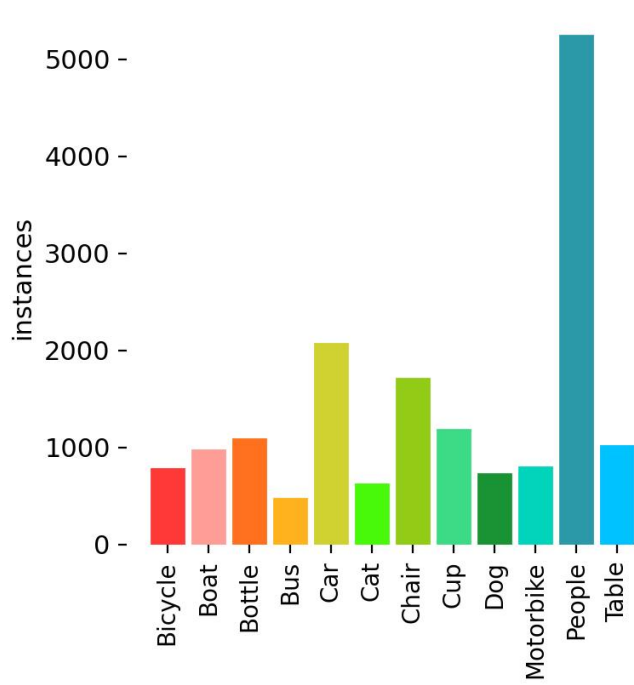
Parameter	Hodnota
Epochy	50
Dávka	35
Rozmery	416x416
lr-0	0.005
lr-f	0.01



Obr. 4: Normalizovaná konfúzna matica



Obr. 5: Výsledky tréovania

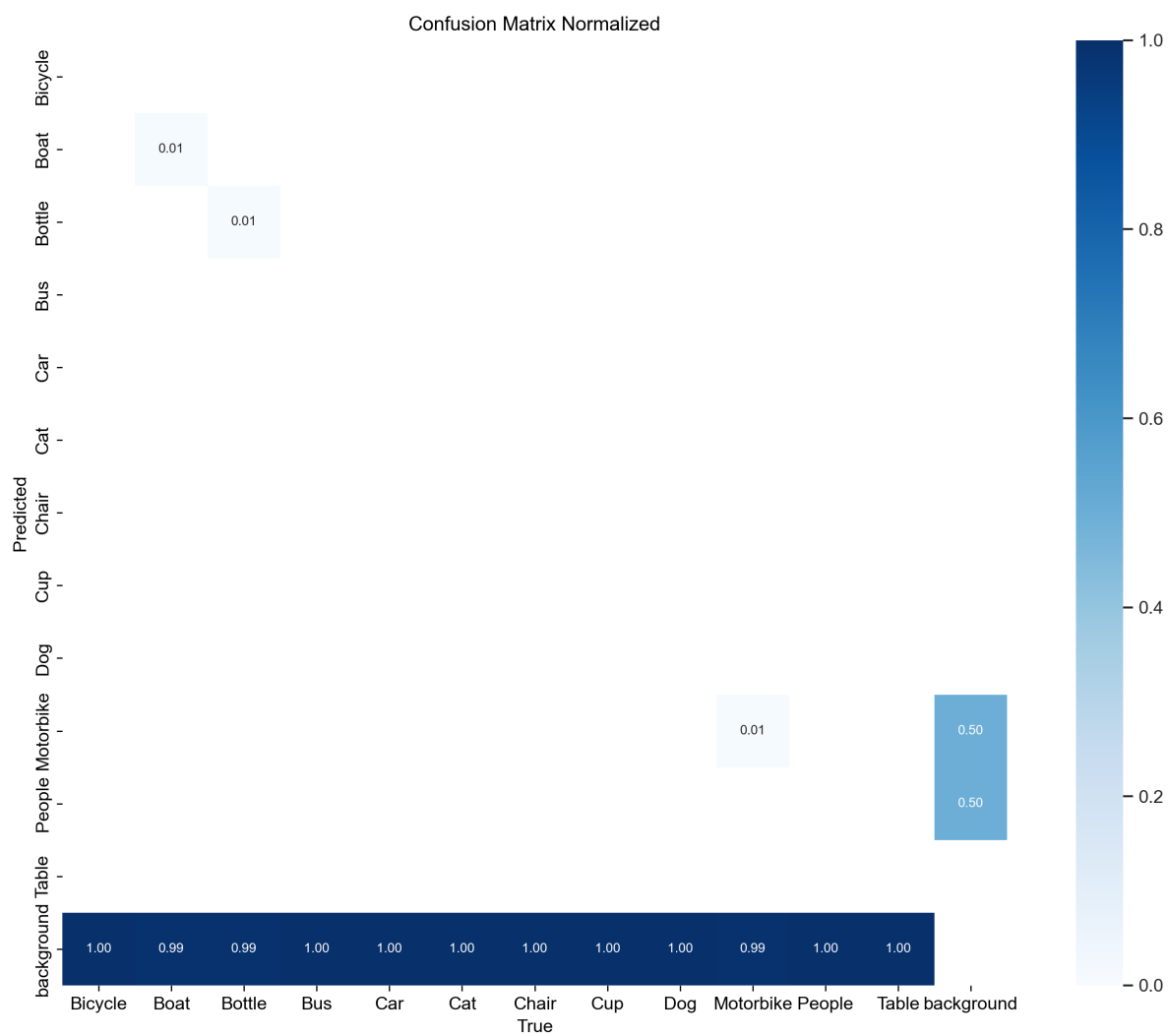


Obr. 6: Anotácie

3.2.3 YOLOv8s a YOLOv8m

Parameter	Hodnota
Epochy	8
Dávka	20
Rozmery	416x416
lr-0	0.005
lr-f	0.01

Detektor sa počas trénovania nič nenaučil. Tabuľka pre YOLOv8m je totožná až na to že bolo potrebné zmenšiť veľkosť dávky na 10 čiže o polovicu čo vlastne aj odpovedá pomeru trénovacích parametrov. Záverom je teda že grafická krata NVIDIA GTX 1660 Super 6GB nepostačuje na to aby bola schopná pri rozumných parametroch trénovať väčšie modely detektorov.



Obr. 7: Výsledky trénovania

4 Zdroje

Referencie

- [1] Loh, Yuen Peng and Chan, Chee Seng, (2019), Getting to Know Low-light Images with The Exclusively Dark Dataset.