

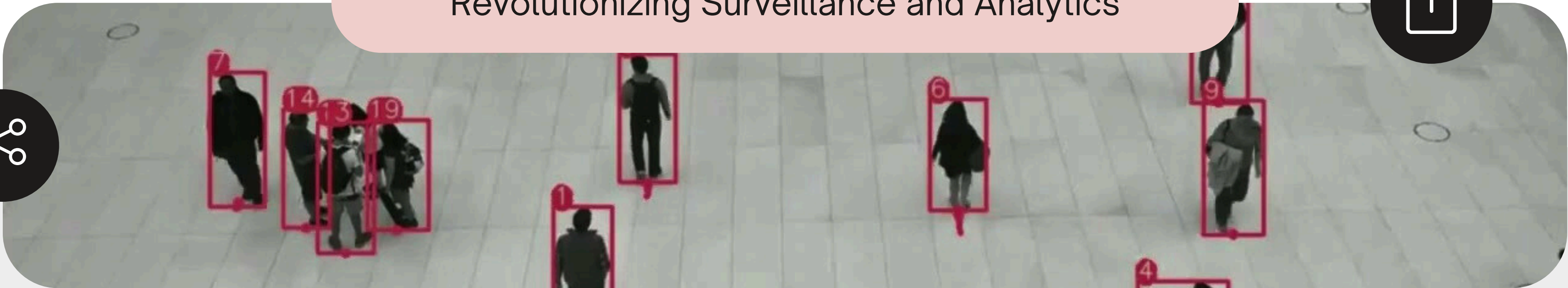


Person Tracking

Team Alan Turing (Team CV B)



“Revolutionizing Surveillance and Analytics”



Team Members

Andy
Ezar Resha
Khusnul Muchlisin
Melvin Tan



The Purpose of the Project

Membuat solusi *Person Tracking* yang memanfaatkan teknik *Computer Vision* untuk mentrack secara akurat dan efisien dalam berbagai lingkungan & kondisi.





Market Potential



Enhanced Security : Pengawasan secara real-time di tempat umum.

Retail Analytics : Analisis perilaku pelanggan untuk mengoptimalkan tata letak toko dan meningkatkan penjualan.

Smart City: Manajemen lalu lintas yang lebih baik dan keselamatan publik yang ditingkatkan

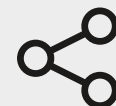
Market Potential

Transportation : Penggunaan untuk Pengaturan lalu lintas

Healthcare : Monitoring jumlah orang di area tertentu untuk mengatur kapasitas dan mencegah penyebaran penyakit



Problem to be Solved



Akurasi Model
yang Rendah

Waktu Latensi
yang tinggi

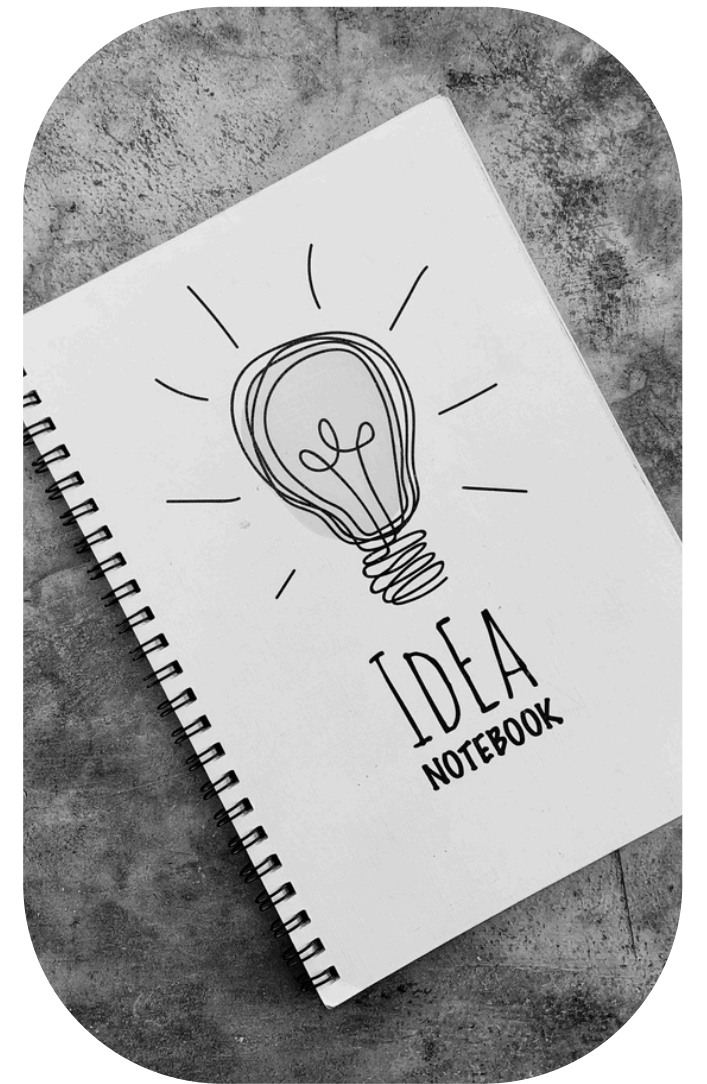
Kinerja buruk dalam kondisi
lingkungan yang buruk

Proposed Solution

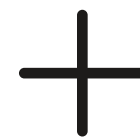
Faster R-CNN

YOLO

ResNet



Dataset



COCO 2017

Common Objects in Context (COCO) adalah dataset deteksi objek, segmentasi, dan captioning skala besar yang berisi lebih dari 200.000 gambar berlabel dengan 80 kategori.

Dataset

COCO 2017

Menggunakan module fiftyone,
di Filter untuk data "Human".

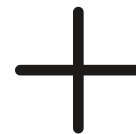
Train : (118.000) -> 64.115

Validation: 5.000

Test : 40.670



Dataset



COCO 2017

Dataset diatur ke dalam folder masing-masing dengan folder label yang berisi data anotasi untuk setiap gambar individual untuk YOLO, dan JSON gabungan untuk Faster R-CNN.

Data Augmentation

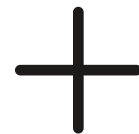


Image Augmentation

- Random Horizontal Flip
 - Color Jitter (brightness, contrast, saturation, and hue)
- Untuk meningkatkan kemampuan generalisasi model.

Model – YOLO



- You Only Look Once - model YOLO dirancang untuk deteksi objek secara real-time dengan menggunakan jaringan saraf tunggal.
- Model ini digunakan dalam berbagai ukuran: (n)ano, (s)mall, (m)edium, (x)tra large.
- Model ini diinisialisasi dengan pre-trained, dan dilatih selama 10 Epoch, ukuran gambar 640, AdamW LR= 0.002, momentum = 0.9 (melalui optimizer) dan menggunakan GPU.

Model – Faster RCNN

- Backbone yang kami gunakan adalah resnet50. Lapisan 'res'idual membantu dalam melatih jaringan yang sangat dalam yang memungkinkan pelatihan model yang lebih akurat dengan kinerja yang terbukti.
- Model Faster-RCNN lebih akurat, tetapi lebih lambat daripada YOLO.
- Model ini diinisialisasi dengan bobot pra-latih, dan dilatih selama 10 Epoch, ukuran gambar 640, dan menggunakan GPU.

Model – Faster RCNN (Extra)

- Pengujian tambahan untuk membandingkan kinerja kecepatan menggunakan mobileNet sebagai backbone untuk FasterRCNN.
- Menggunakan lebih sedikit parameter dan persyaratan komputasi yang lebih rendah, membuatnya lebih cepat dan ringan untuk digunakan.
- Model ini **tidak kami train, hanya untuk komparasi kecepatan**

Training Parameter Faster RCNN

- Epoch : 10
- Batch Size : 1 (mencegah masalah memory)
- Optimizer : SGD
- LR: 0.001
- Momentum: 0.9
- Training fokus ke deteksi person dan background

Home

Content

Sign Out



7

Use Case (Demo)



Berikut adalah Use Case dari Model yang kami buat



Model Evaluation



MAP

Rata2 Inference
Time

Rata2 Person
Tracked

		YOLO				Faster RCNN	
		Nano	Small	Medium	Varying	RESNET50	MobileNet
Various Images	MAP	1	1	1	1	1	1
	Average Person Detected						
	Inference Time	0.2198	0.2122	0.3177	0.5491	3.1772	2.344
	Inference Time Ratio	1.035815269	1	1.497172479	2.587653157	14.9726673	11.04618285

Model: yolov8m.pt
Person (amount): 8
Frame: 92



Model: yolov8s.pt
Person (amount): 8
Frame: 92



Model: yolov8n.pt
Person (amount): 6
Frame: 92

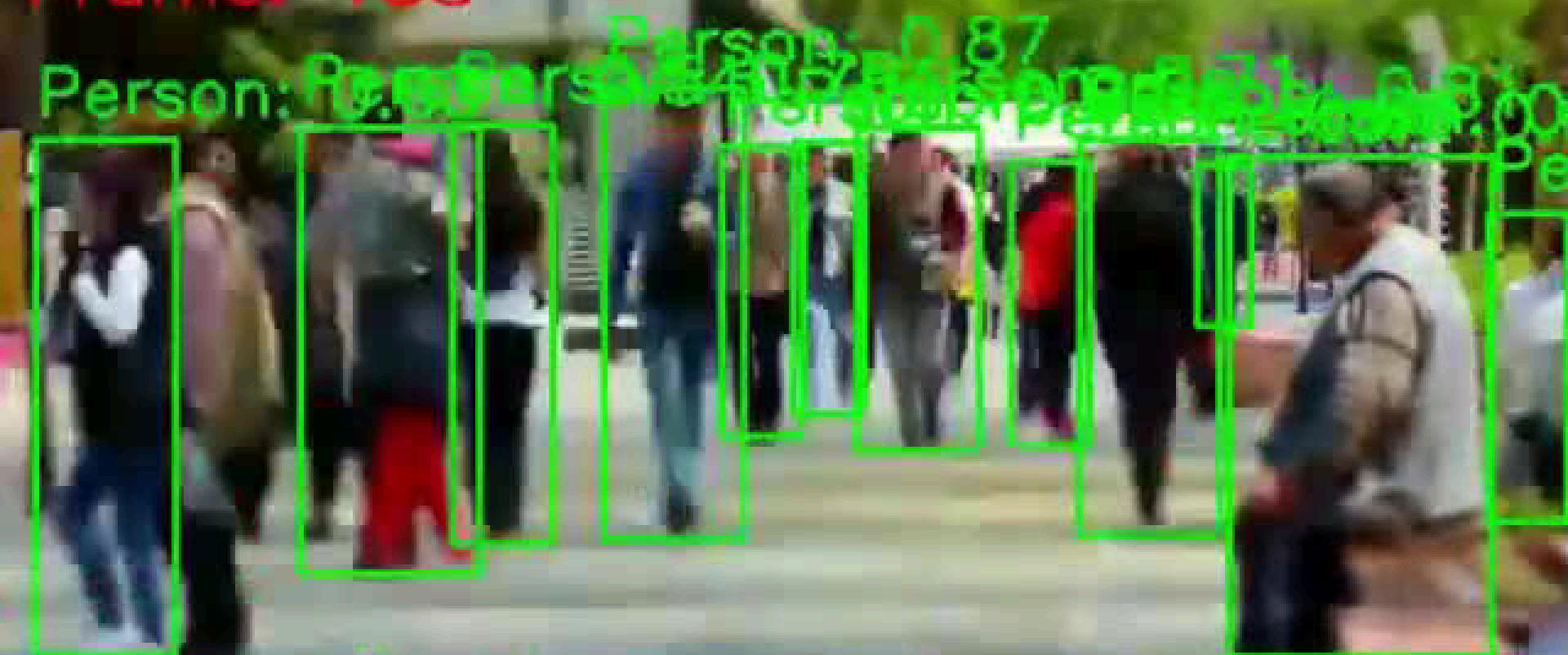


Model: yolov8x.pt
Person (amount): 11
Frame: 92

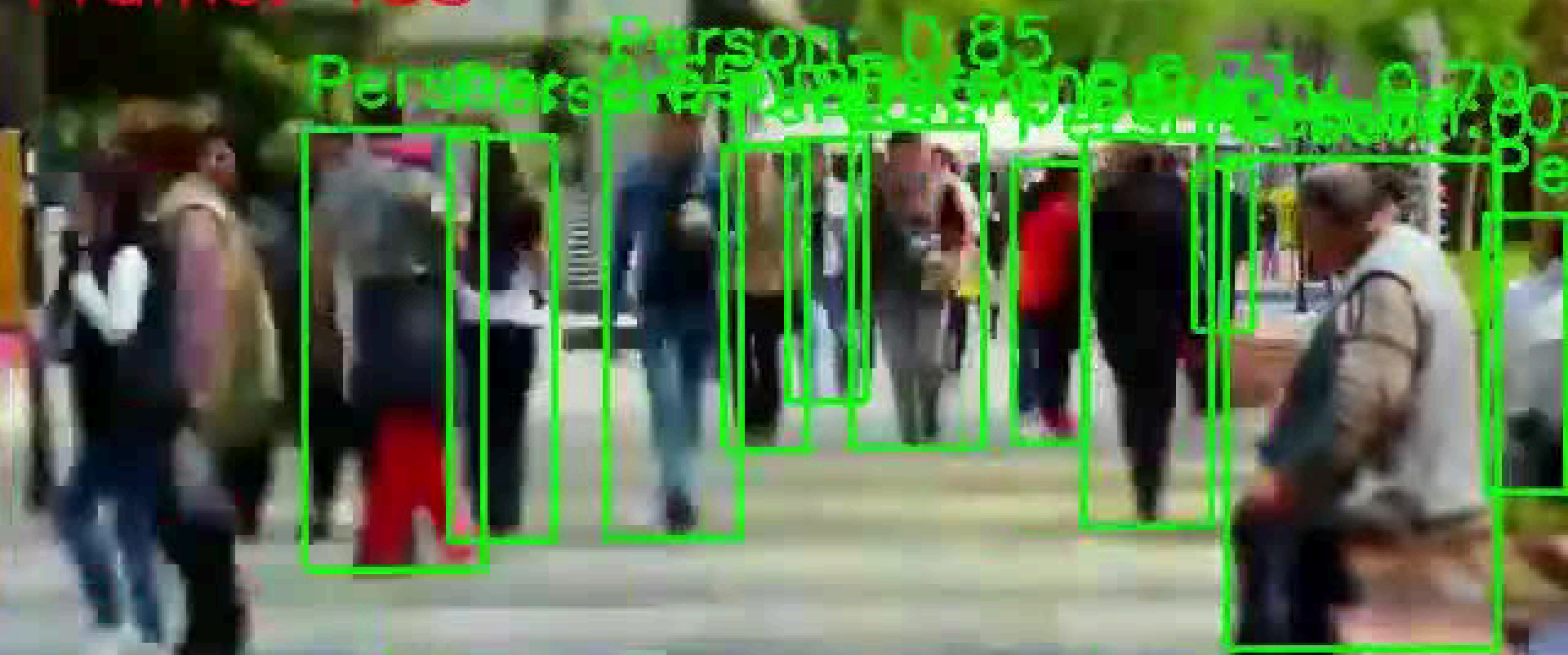


		YOLO				Faster RCNN	
		Nano	Small	Medium	Varying	RESNET50	MobileNet
Video 1 - Times Square Crowd	MAP	1	1	1	1	1	0.908
	Average Person Detected	7.9591	8.8112	9.2806	0.0220	27.0638	100
	Inference Time	0.0137	0.0148	0.0415	0.022	3.8164	1.5023
	Inference Time Ratio	1	1.080291971	3.02919708	1.605839416	278.5693431	109.6678832

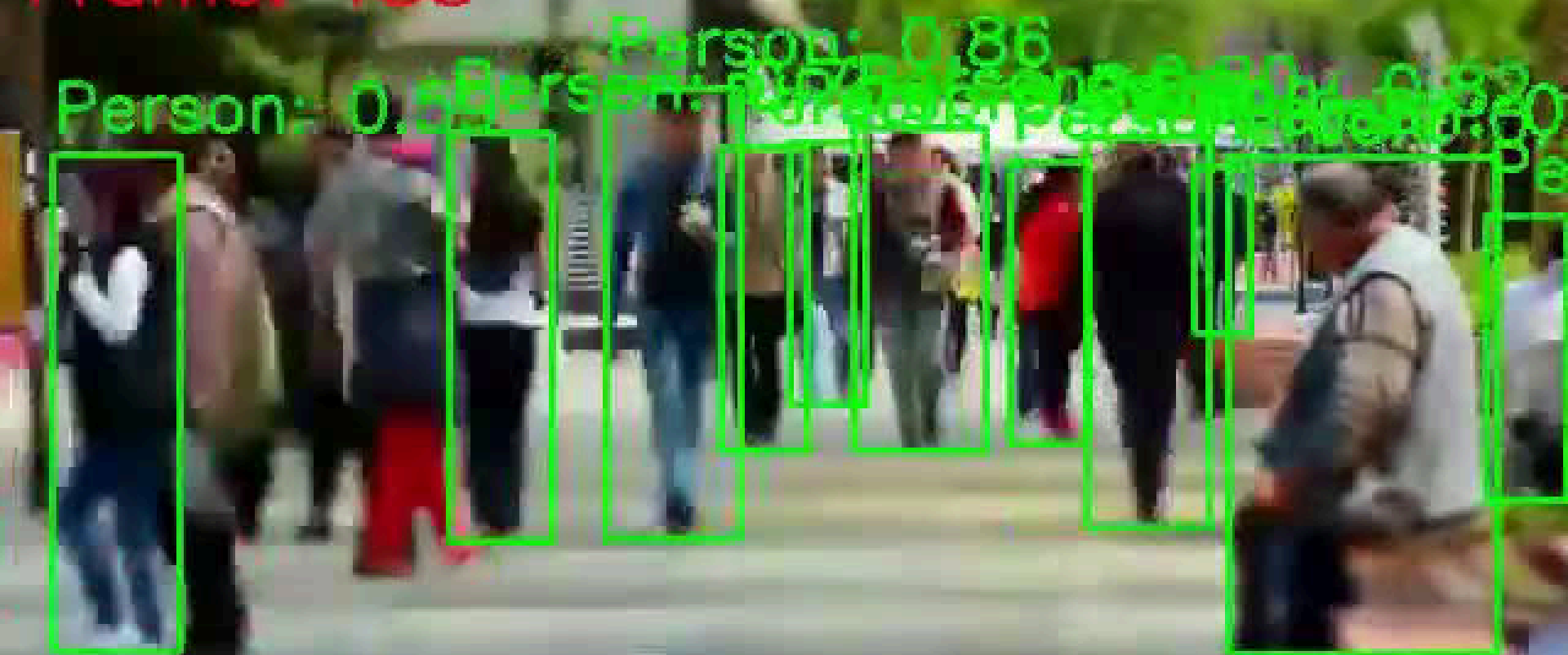
Model: yolov8n.pt
Person (amount): 12
Frame: 158



Model: yolov8s.pt
Person (amount): 11
Frame: 158



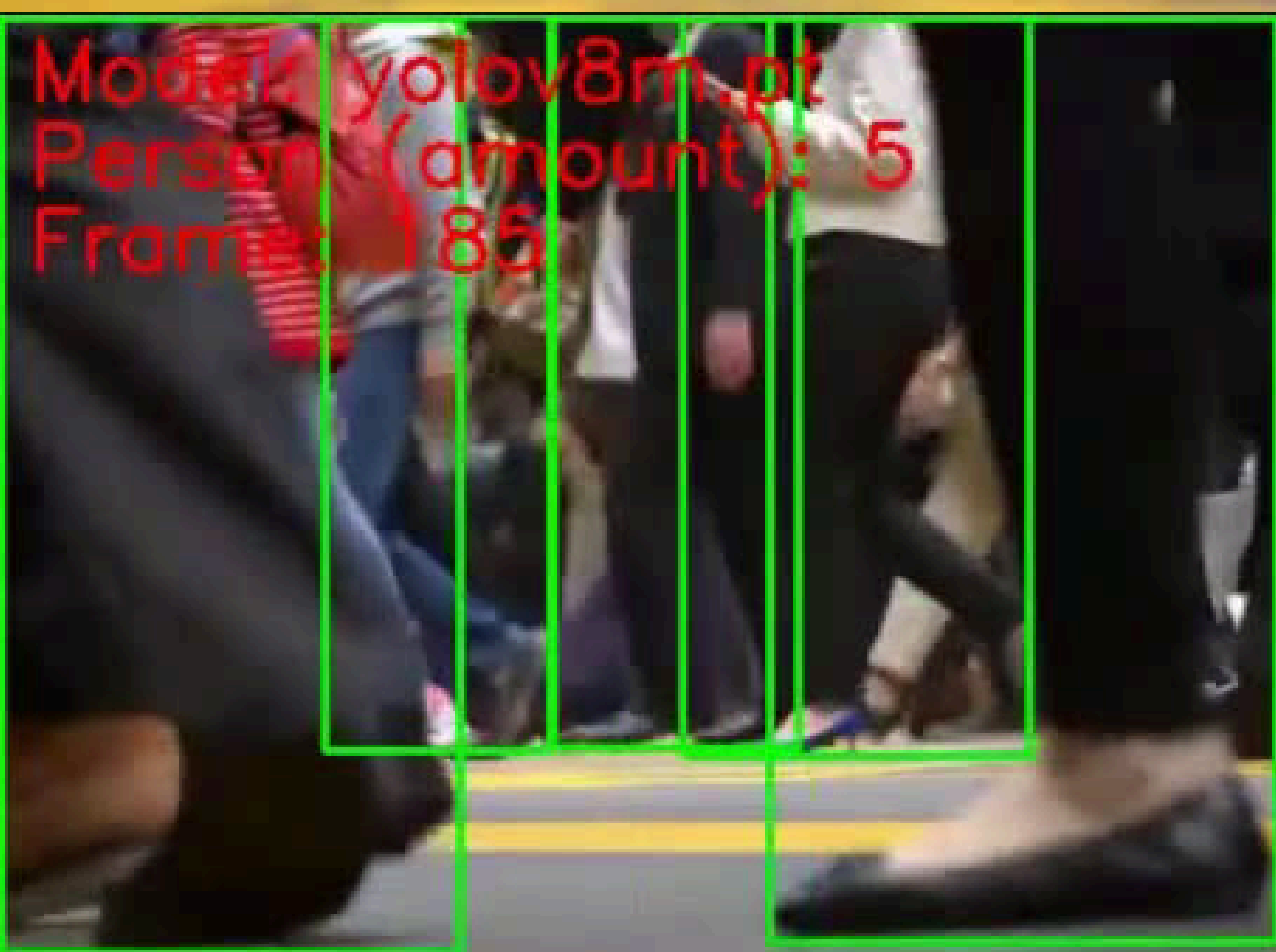
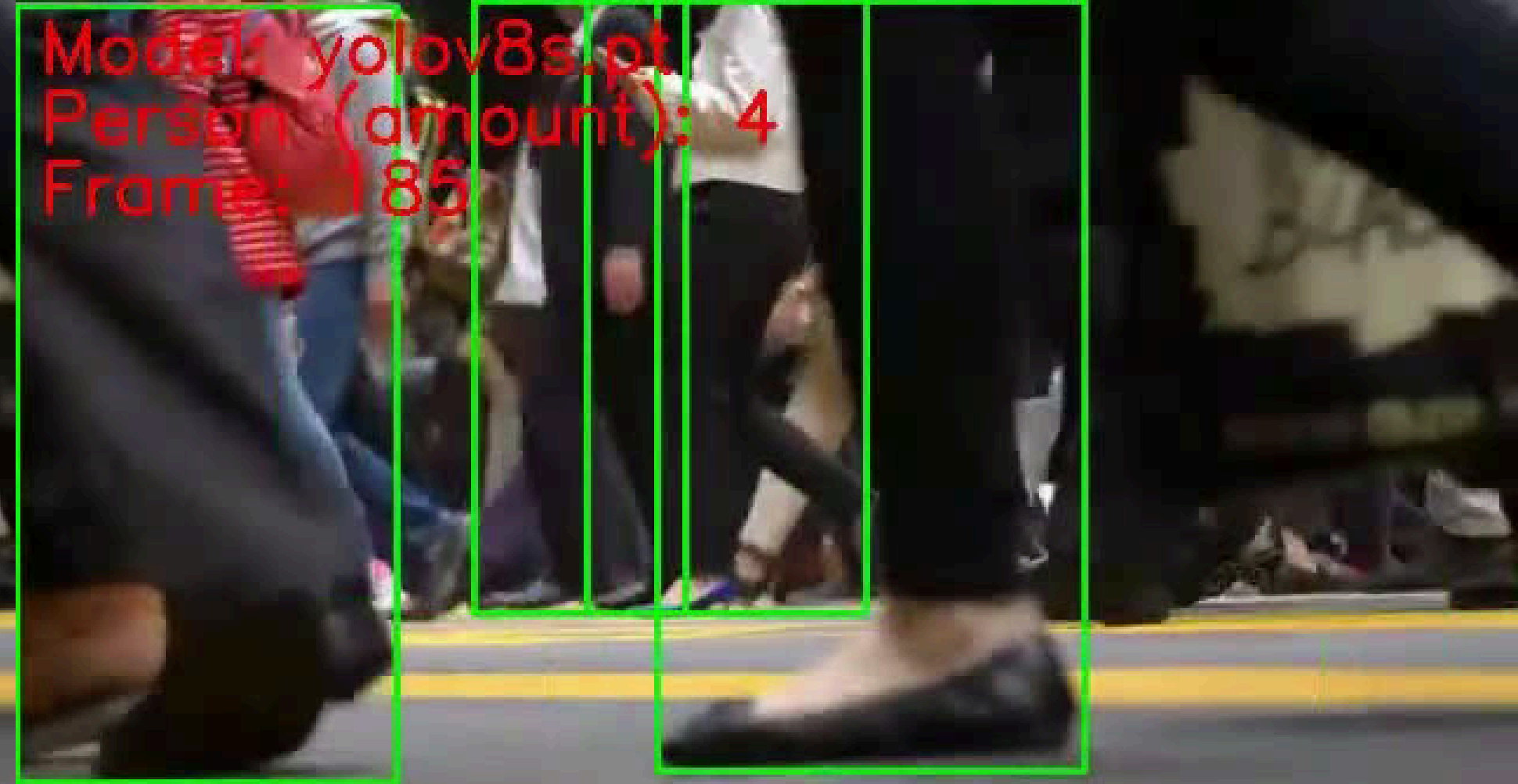
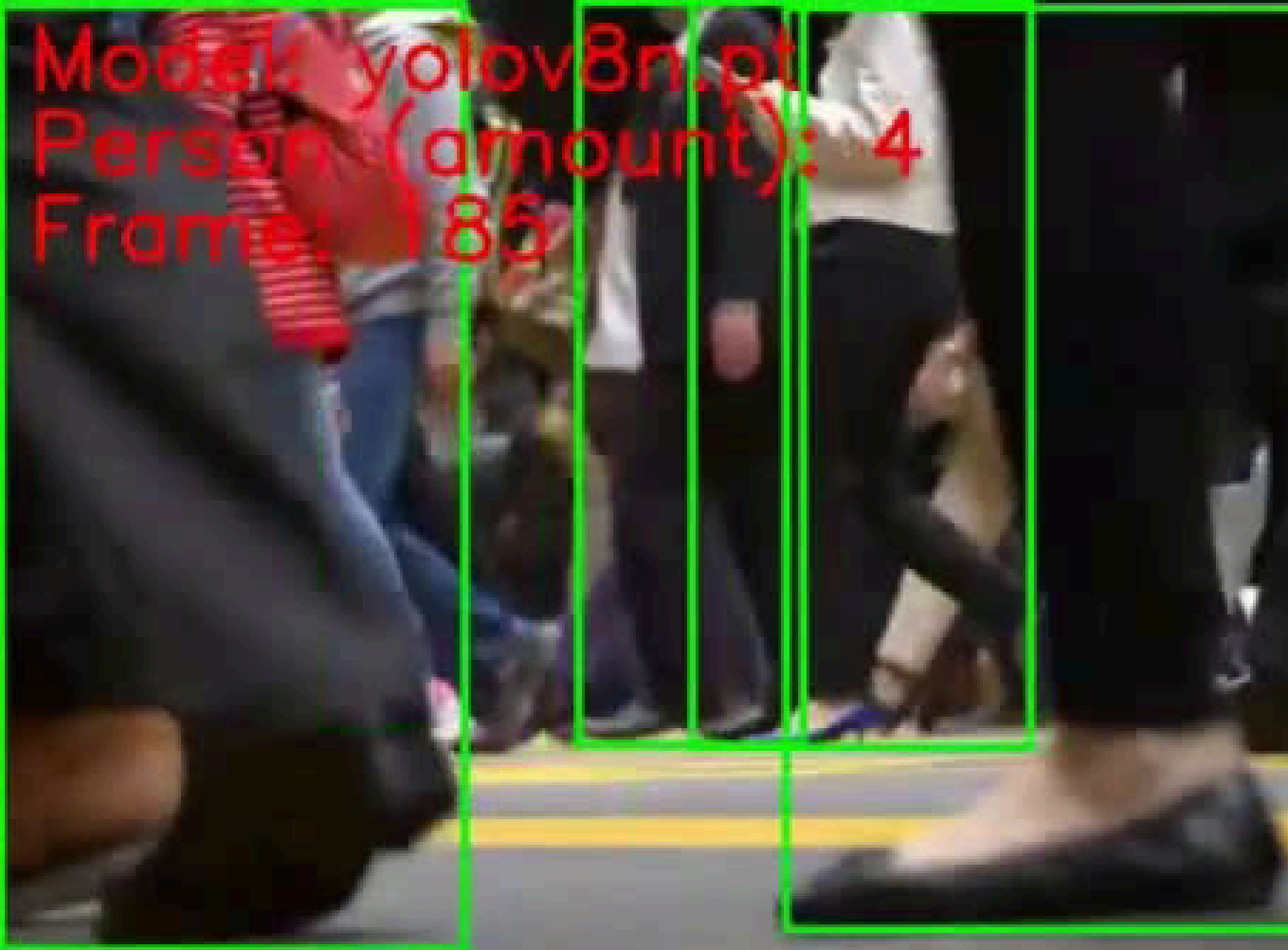
Model: yolov8m.pt
Person (amount): 11
Frame: 158



Model: yolov8x.pt
Person (amount): 11
Frame: 158

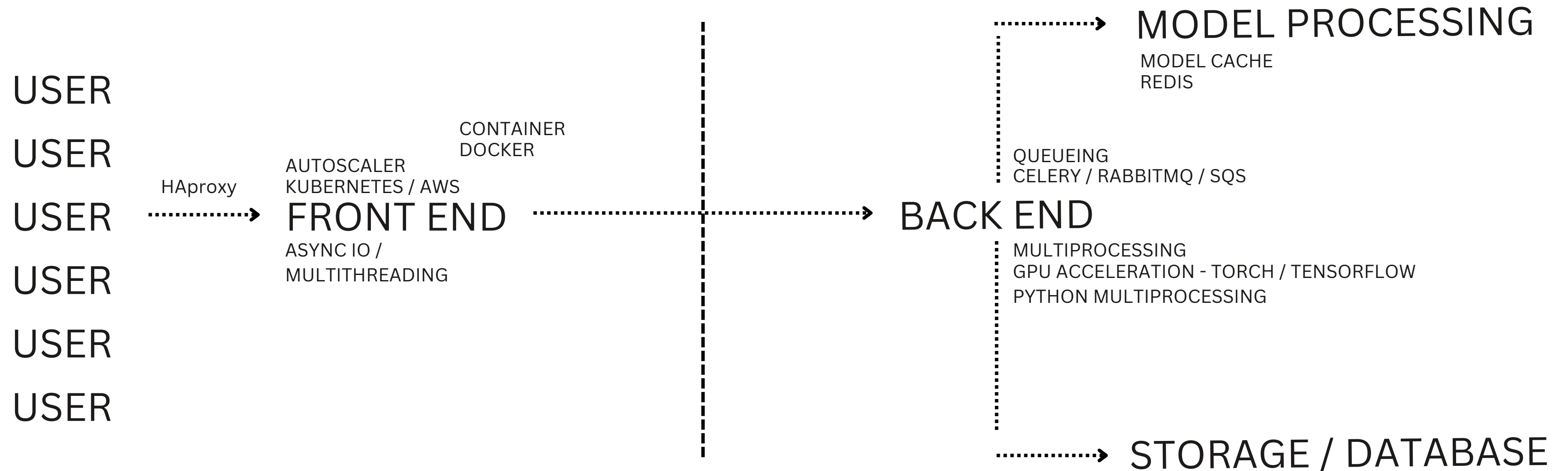


		YOLO				Faster RCNN	
		Nano	Small	Medium	Varying	RESNET50	MobileNet
Video 2 - People walking fast	MAP	1	1	1	1	1	1
	Average Person Detected	7.935	8.685	8.361	8.024	15.3143	100
	Inference Time	0.0514	0.0126	0.02041	0.0406	3.7059	1.78
	Inference Time Ratio	4.079365079	1	1.61984127	3.222222222	294.1190476	141.2698413



		YOLO				Faster RCNN	
		Nano	Small	Medium	Varying	RESNET50	MobileNet
Video 2 - People walking fast	MAP	1	1	1	1	1	0.9081
	Average Person Detected	4.427	5.6098	5.332	5.5144	12.985	100
	Inference Time	0.0107	0.0119	0.0191	0.0404	3.703	1.496
	Inference Time Ratio	1	1.112149533	1.785046729	3.775700935	346.0747664	139.8130841

High Availability Implementation



Kesimpulan

YOLO (You Only Look Once):

- Kecepatan: Sangat cepat dan cocok untuk aplikasi real-time.
- Kelebihan: Penggunaan sumber daya yang efisien dengan proses pelatihan dan struktur model yang lebih sederhana.
- Keterbatasan: Kurang akurat saat mendeteksi objek kecil atau padat, dan lemah saat mengidentifikasi objek dengan variasi tinggi.

Faster-RCNN (Region-based Convolutional Neural Networks):

- Kecepatan: Lebih lambat dibandingkan dengan YOLO.
- Kelebihan: Presisi dan konsistensi tinggi, terutama efektif dalam mendeteksi objek dengan variasi. Secara keseluruhan, menawarkan deteksi yang lebih akurat dan andal.
- Keterbatasan: Proses pelatihan yang lebih kompleks dan memakan waktu, dengan konsumsi sumber daya yang lebih tinggi.

**keterbatasan model-model ini bisa dikurangi dengan penggunaan tools lain, augmentasi data (baik training ataupun di penggunaan) !*

[Home](#)

[Content](#)

[Sign Out](#)



10

Thank You!