VIDEO ANALYSIS FOR VIOLENCE DETECTION

Generative AI Track

TEAM

Eman Arafa

Heba Mohsen

Toka khaled

Shaimaa Mohamed

Yossef Ramadan

IMPORTANCE

- AI-powered system to detect violent incidents captured by surveillance cameras.

- Efficient Frame Retrieval based on Questions.

Business Domains

- Corporate Security in offices, and buildings
- Retail and Commercial Spaces like malls, shops
- Transportation and Airports
- Universities

OBJECTIVES

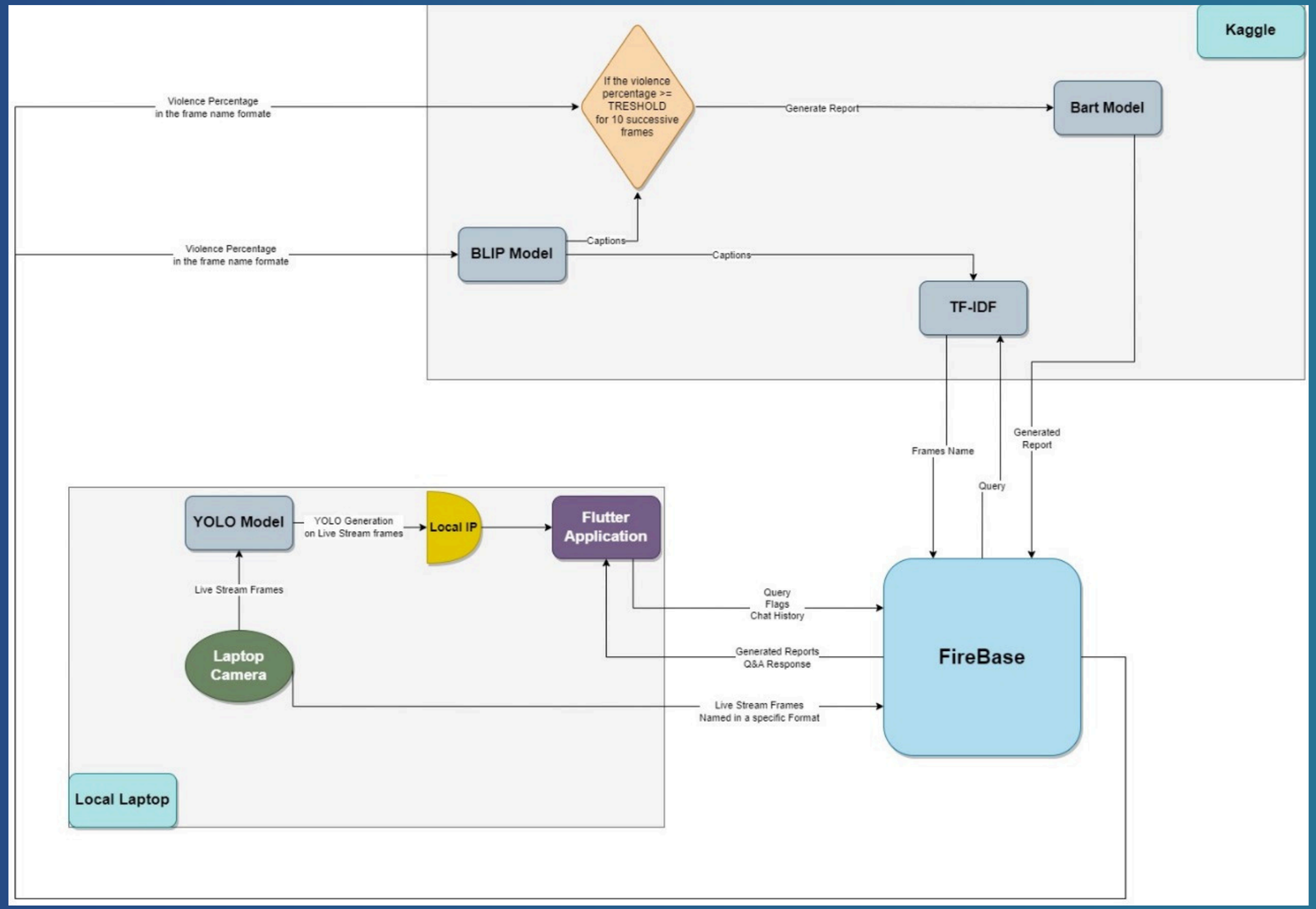
- Process video frames in real-time
- Generate accurate captions
- Detect violence levels from captions and generate refined report
- Q & A and retrieve relevant Frames

SOLUTION OVERVIEW

- Yolov8 for violence detection
- BLIP-2 : Caption generation
- BART: Incident report generation
- Transformers and cosine similarity
- Firebase: Storage and real-time updates

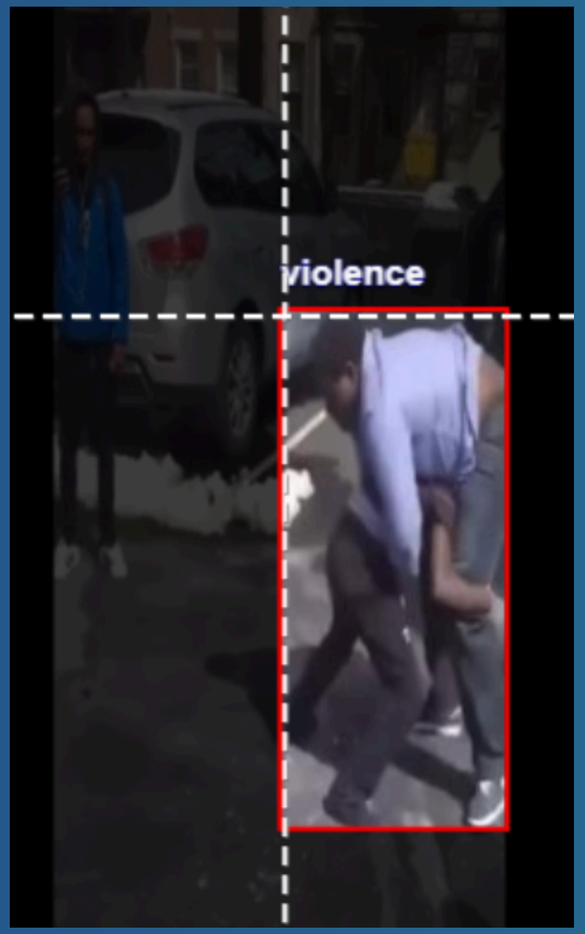
Demo

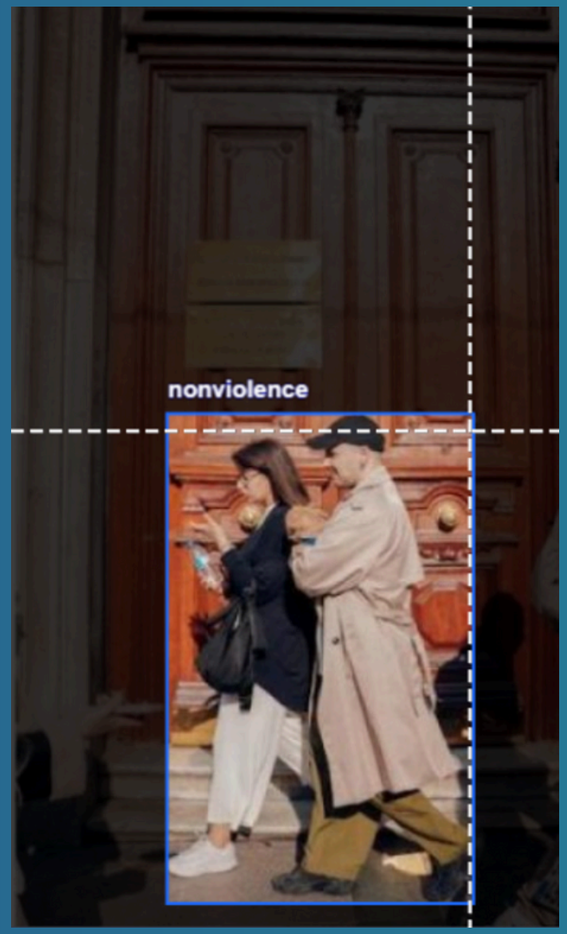
FLOW


Kaggle
Violence Percentage
in the frame name formate
If the violence
percentage >=
TRESHOLD
for 10 successive
frames
Generate Report
Bart Model
Violence Percentage
in the frame name formate
BLIP Model
Captions
Captions
TF-IDF
Frames Name
Generated
Report
Query
YOLO Model
YOLO Generation
on Live Stream frames
Local IP
Flutter
Application
Live Stream Frames
Laptop
Camera
Query
Flags
Chat History
Generated Reports
Q&A Response
Live Stream Frames
Named in a specific Format
FireBase
Local Laptop

YOLO DATASET

Dataset Samples

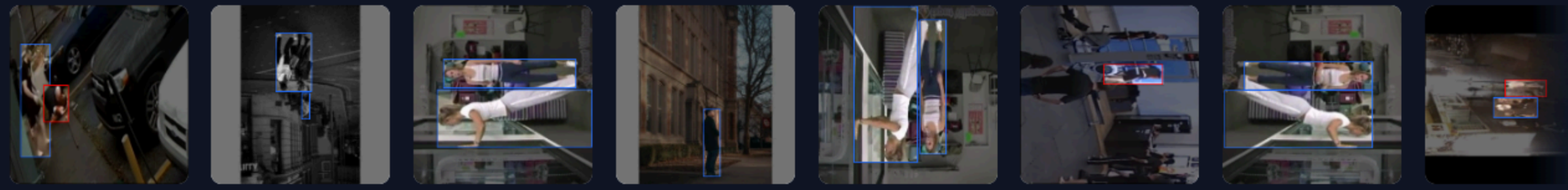




YOLO DATASET

6265 Total Images

View All Images →



Dataset Split

TRAIN SET	VALID SET	TEST SET
81%	10%	10%
5070 Images	**599** Images	**596** Images

Preprocessing

Auto-Orient: `Applied`

Resize: `Stretch to 640x640`

Augmentations

Outputs per training example: `2`

Flip: `Horizontal, Vertical`

90° Rotate: `Clockwise, Counter-Clockwise`

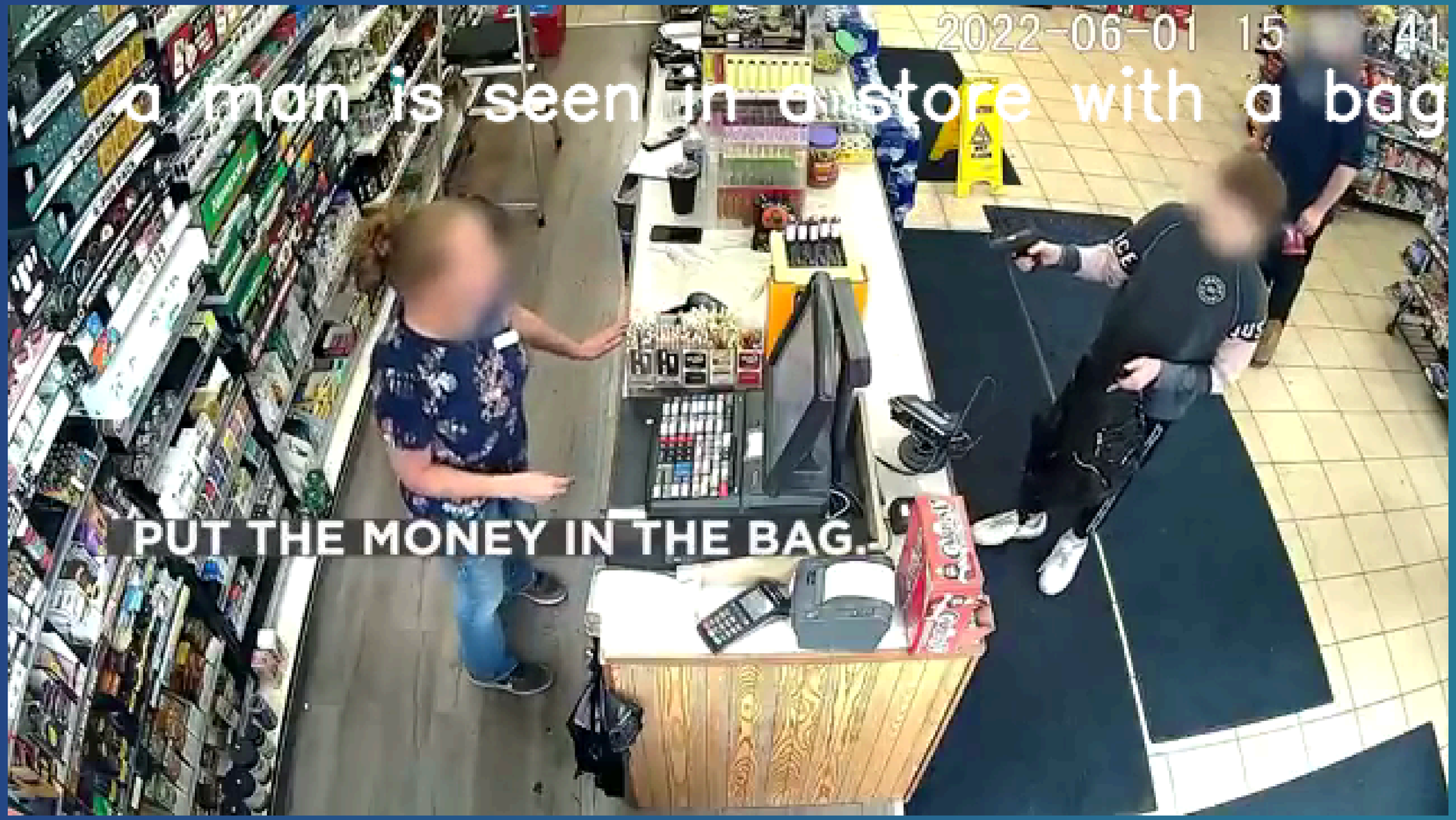
ACCURACY

Reaching 84 % mAP

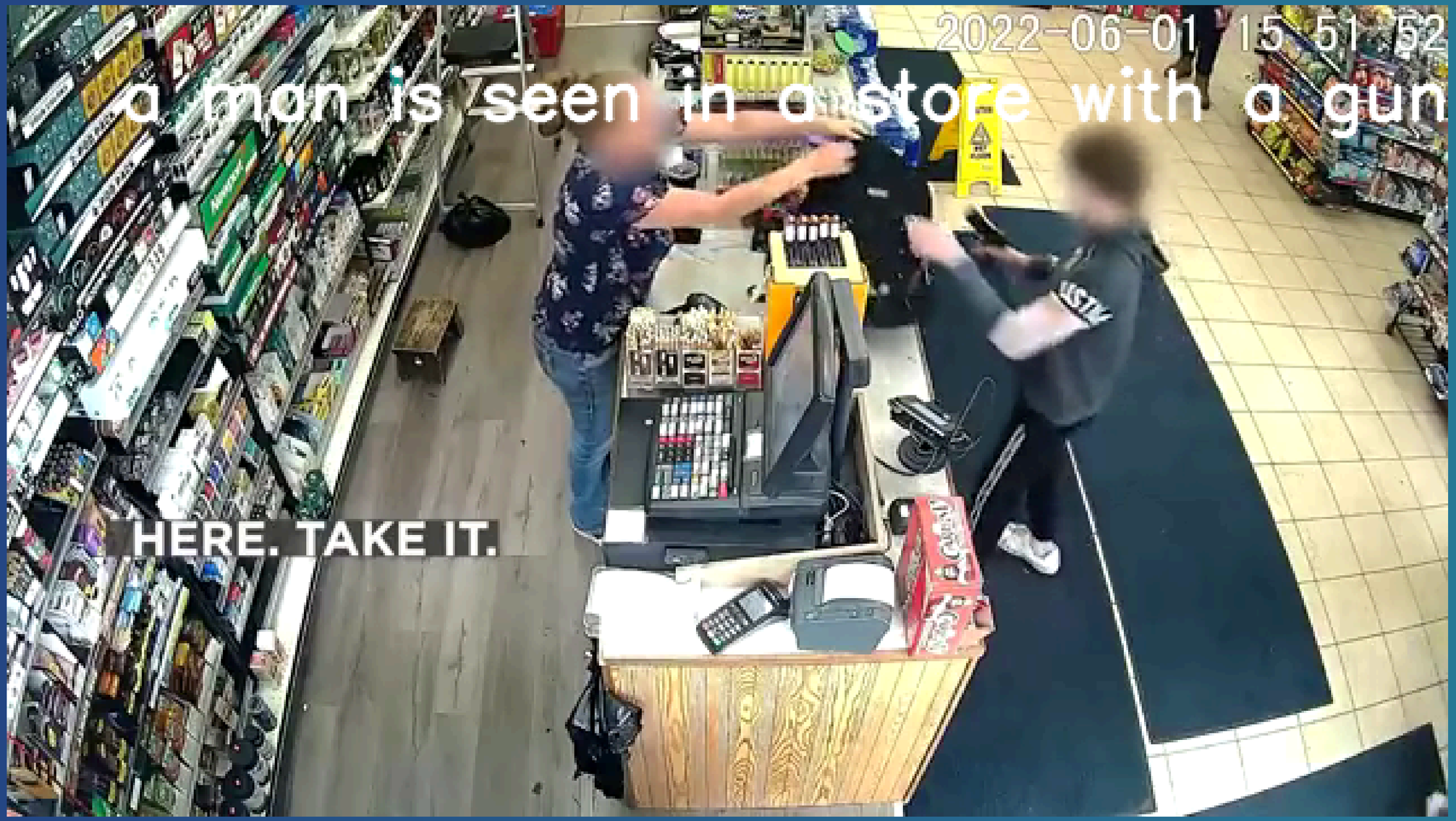
```
Validating runs/detect/train/weights/best.pt...
Ultralytics YOLOv8.2.96 🚀 Python-3.10.12 torch-2.4.0+cu121 CUDA:0 (Tesla T4, 15102MiB)
Model summary (fused): 168 layers, 3,006,038 parameters, 0 gradients, 8.1 GFLOPs
                 Class     Images  Instances      Box(P          R      mAP50  mAP50-95):
                   all        599        942      0.794       0.77      0.844      0.635
           nonviolence        340        592      0.723      0.679      0.774      0.572
              violence        336        350      0.865       0.86      0.915      0.698
Speed: 0.3ms preprocess, 2.4ms inference, 0.0ms loss, 3.0ms postprocess per image
Results saved to runs/detect/train
ultralytics.utils.metrics.DetMetrics object with attributes:
```

YOLO RESULTS

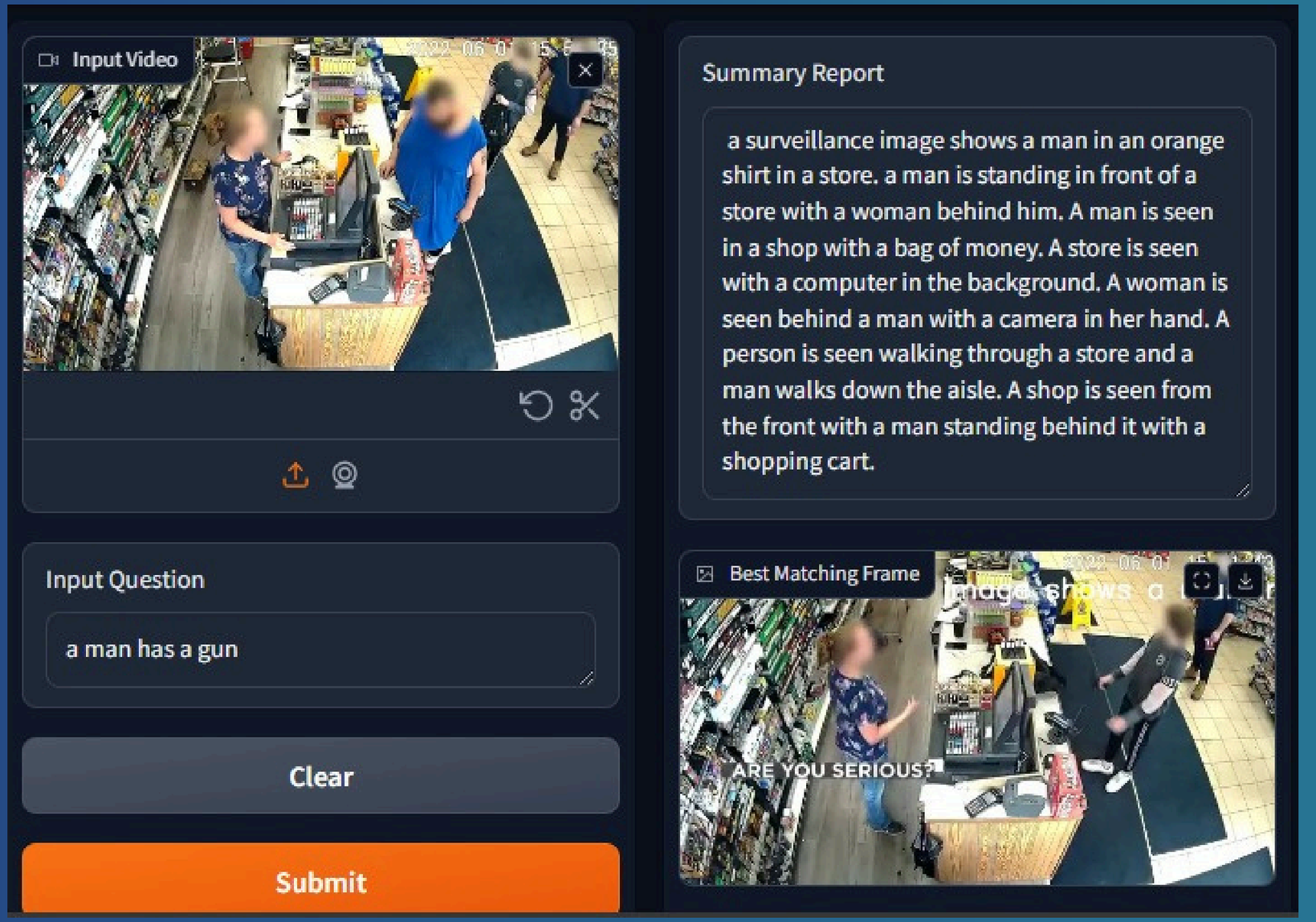
RESULTS


2022-06-01 15 41
a man is seen in a store with a bag
PUT THE MONEY IN THE BAG.

RESULTS


2022-06-01 15:51:52
a man is seen in a store with a gun
HERE. TAKE IT.

APP


Input Video
Summary Report
a surveillance image shows a man in an orange shirt in a store. a man is standing in front of a store with a woman behind him. A man is seen in a shop with a bag of money. A store is seen with a computer in the background. A woman is seen behind a man with a camera in her hand. A person is seen walking through a store and a man walks down the aisle. A shop is seen from the front with a man standing behind it with a shopping cart.
Input Question
a man has a gun
Clear
Submit
Best Matching Frame
ARE YOU SERIOUS?

KEY CHALLENGES

Real-time Model Performance

Large extracted frames

Model Integration and Optimization

ADVANTAGES

Scalable and real time speed

Accurate detection results

Enhanced Report using Prompt Engineering

THANK YOU!