

Senior Projects

JubJones

Intelligent Multi-Camera Person
Tracking and Analytics System

Next Slide



Overview

The Intelligent Multi-Camera Person Tracking and Analytics System is an AI-powered solution designed to identify, track, and analyze individuals or objects across multiple camera views in real time.

Next Slide



Features

Core

LLM-Powered Person Selection

Real-Time Person Detection and Re-Identification

Path Tracking and Visualization

Optional

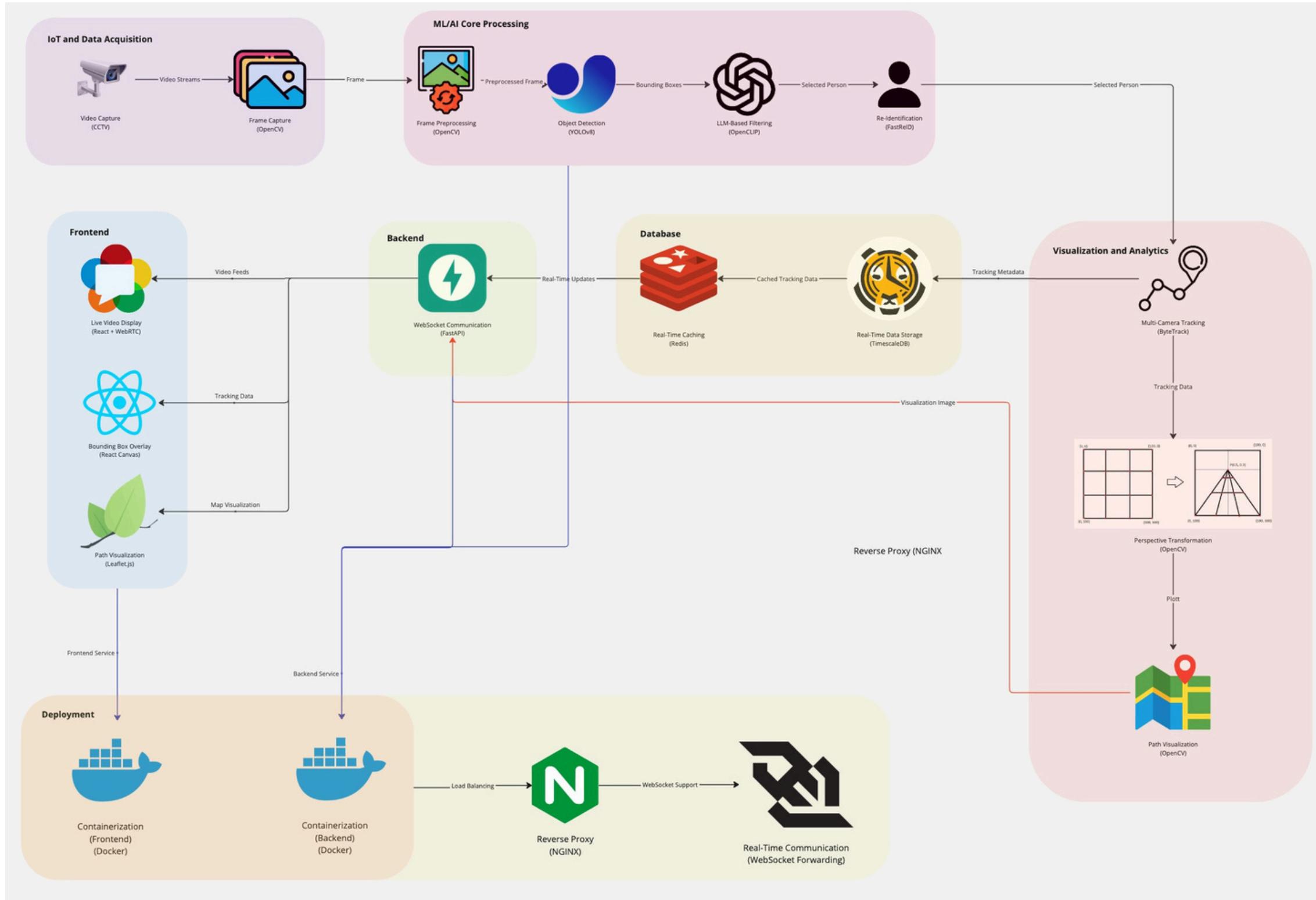
Multi-Object Tracking

Voice Command Integration

Dynamic Analytics and Alerts

Automatically generate map

Tech Stacks & Architecture



IoT and Data Acquisition



Video Capture
(CCTV)

Video Streams



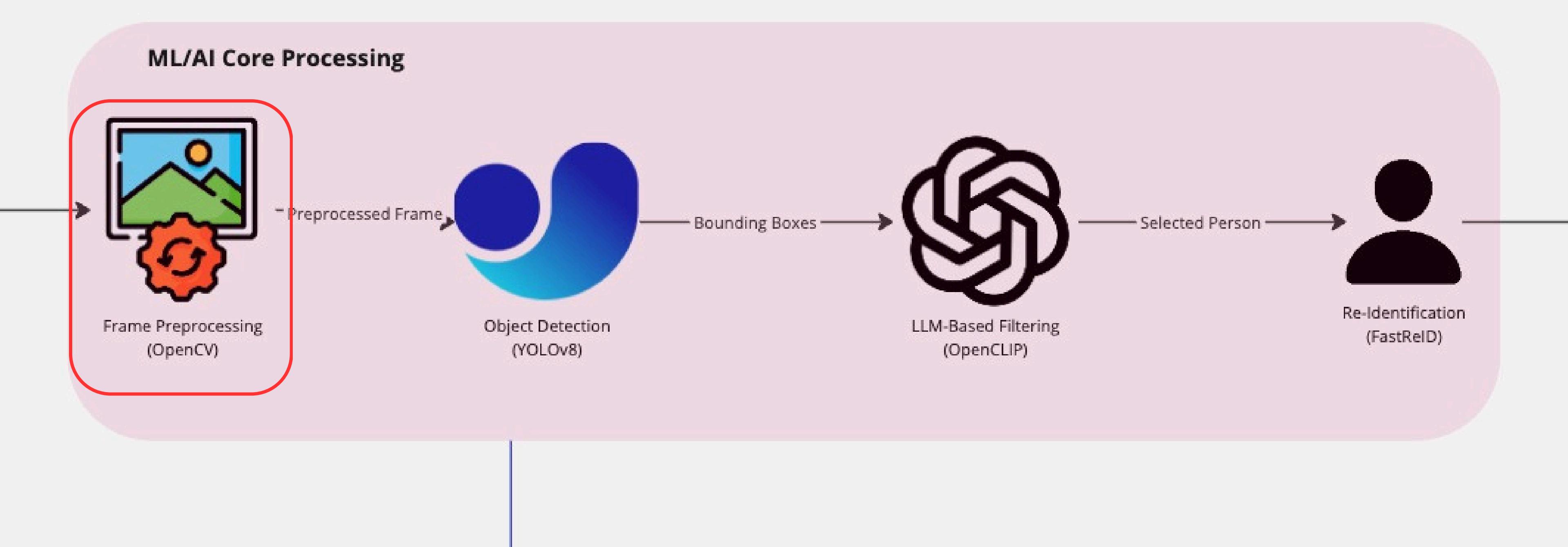
Frame Capture
(OpenCV)

Frame

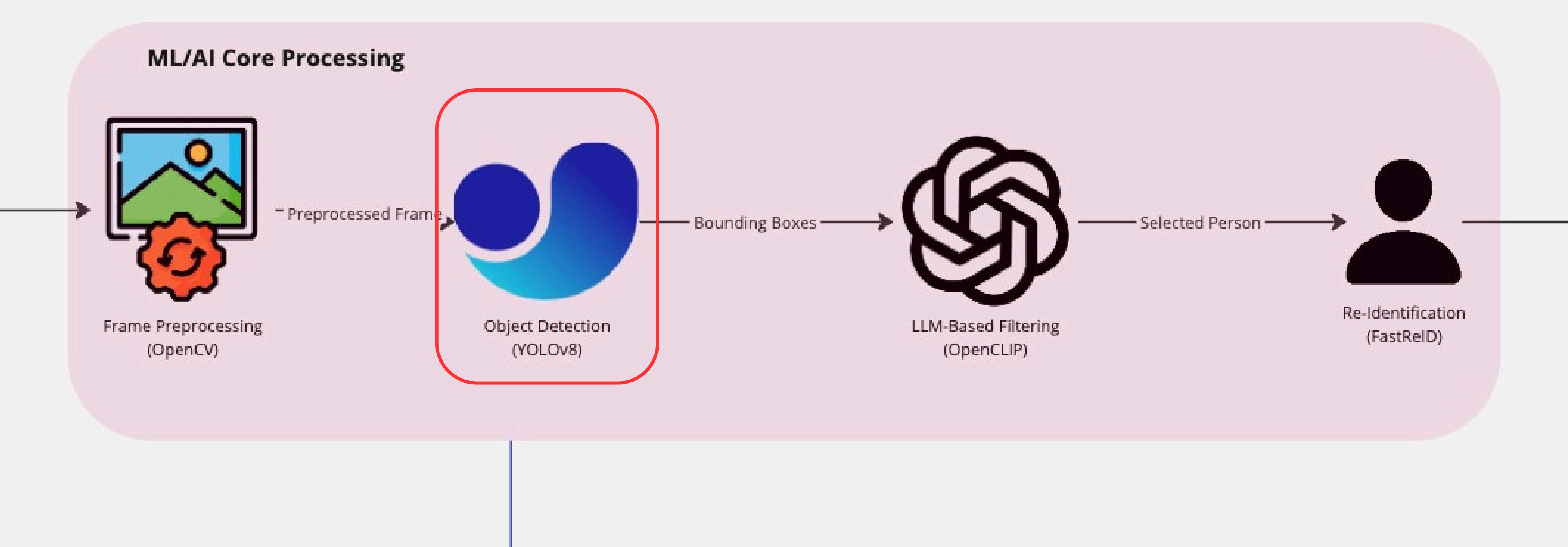


Frame Preprocess
(OpenCV)

- Connect to CCTV cameras (via RTSP or WebRTC) and retrieve live video streams.
- Use OpenCV (Python) to capture frames at a target FPS (e.g., 10-30 FPS).
- Ensure timestamps are synchronized for multi-camera scenarios.

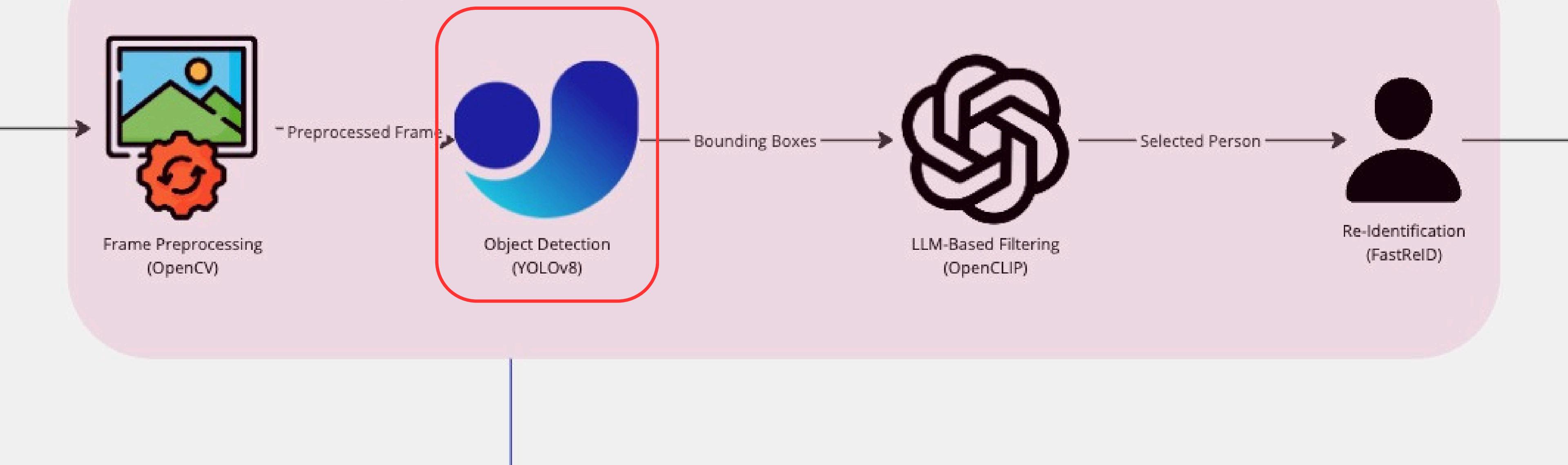


- Preprocess frames for consistent input to detection models.
 - Normalize and resize frames (OpenCV).
 - Optionally apply image enhancements for low-light or noisy footage.

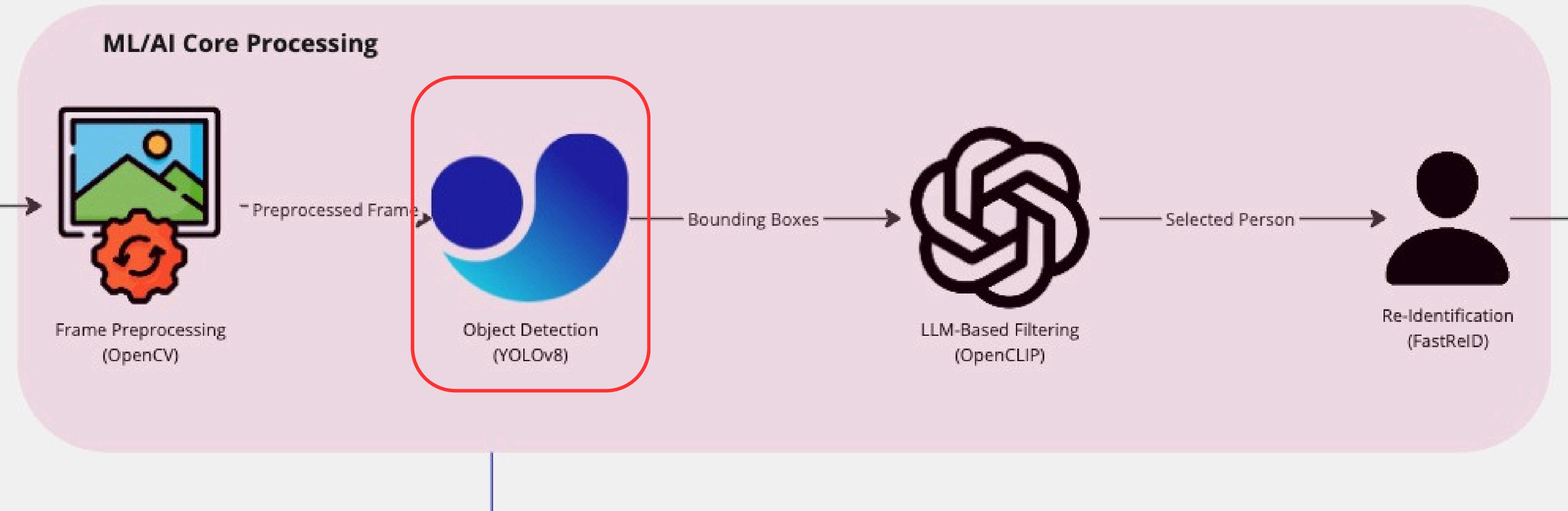


- Detects all individuals (or objects) in each video frame.
 - Use YOLOv8 for real-time object detection.
 - Generate bounding boxes, confidence scores, and class labels (just focus on “person”).

ML/AI Core Processing

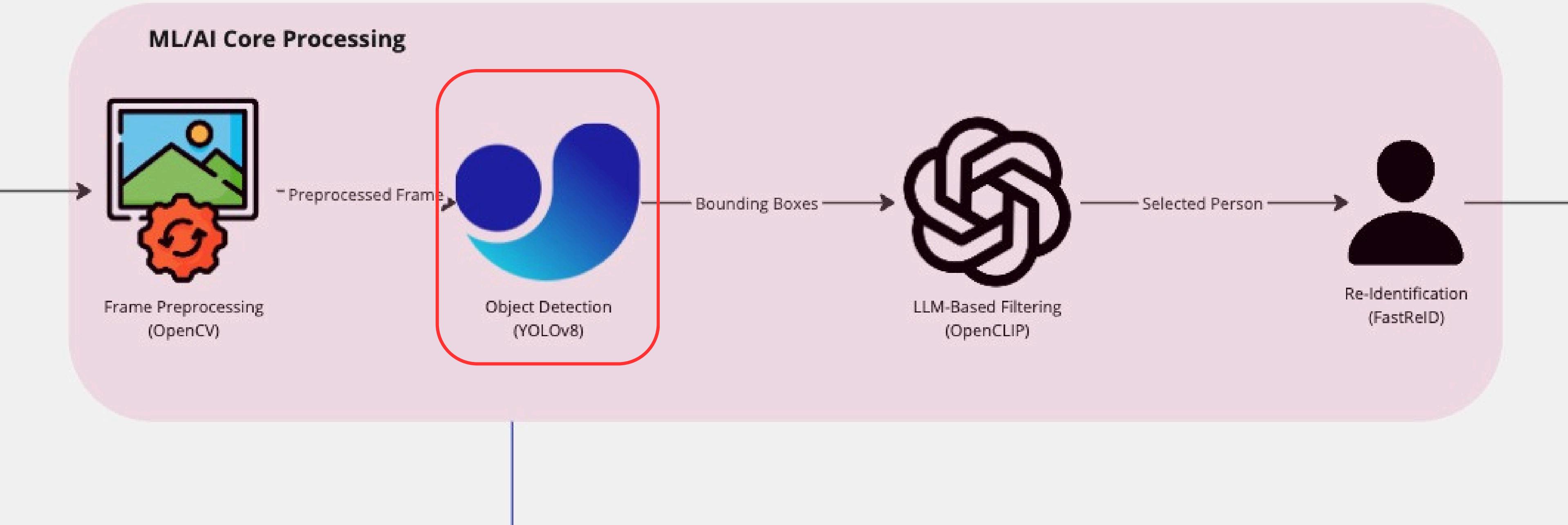


- Any other models?
- Why YOLOV8?



- Any other models?
 - **YOLOv11**: An unofficial iteration with anchor-free detection and improved small-object performance, but offers only **modest** accuracy gains over YOLOv8 and **lacks official support/documentation**.
 - **YOLO-NAS (Neural Architecture Search)**: Excels in small-object detection, localization accuracy, and edge-device efficiency with post-training quantization, but has **limited adoption and ecosystem support** compared to YOLOv8.
 - **RT-DETR**: Transformer-based model with NMS-free architecture, offering competitive speed and accuracy but **requiring more computational resources** and being **less mature** than YOLO.
 - Non-Maximum Suppression (NMS) is traditionally used to remove redundant bounding boxes by selecting the one with the highest confidence score while suppressing others that overlap significantly

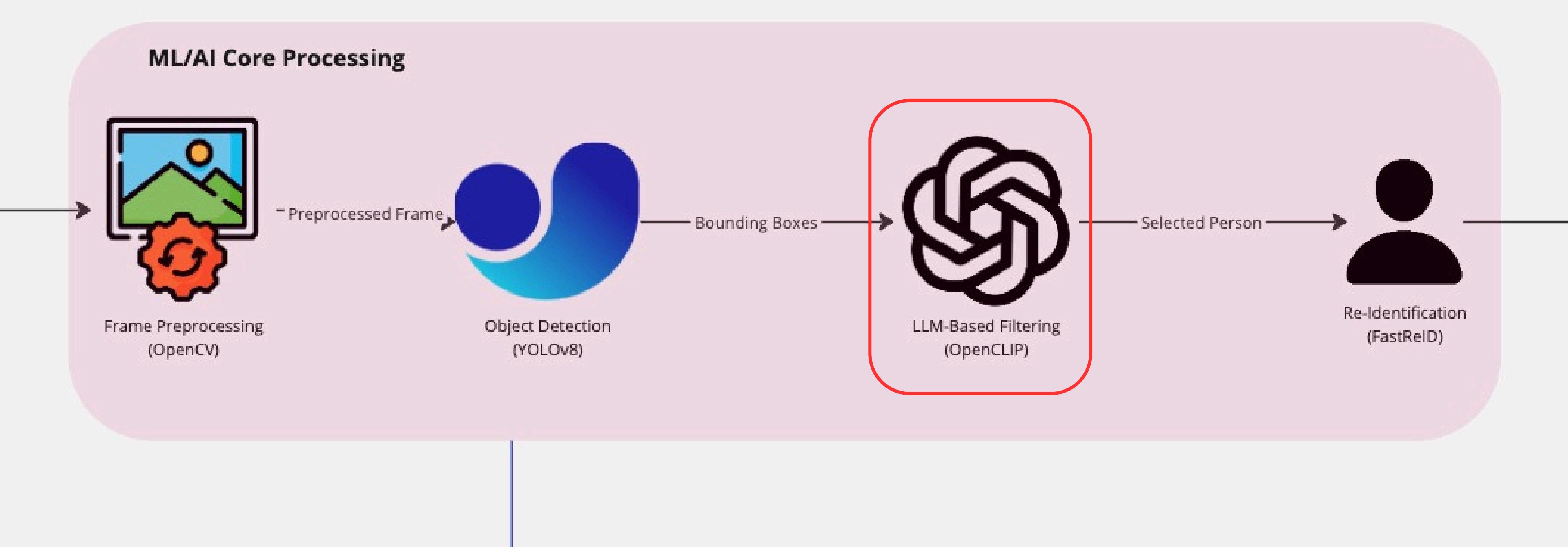
ML/AI Core Processing



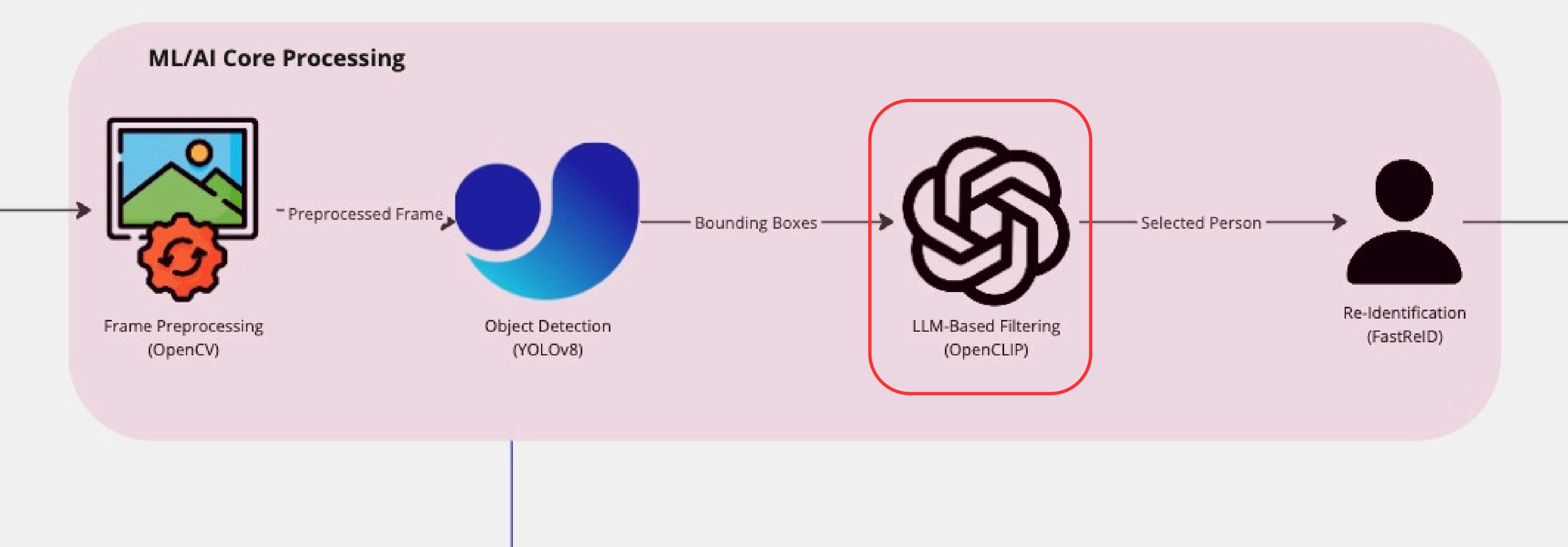
- Why YOLOV8?
 - YOLOv8 provides a perfect balance of speed, accuracy, and ease of integration, backed by a mature ecosystem, extensive documentation, and reliable support for real-time applications.

ML/AI Core Processing

Model	Accuracy (mAP)	Speed (FPS)	Ease of Use
YOLOv8	mAP@0.50:0.95 ~ 50-55% (varies by size)	120-150 FPS on RTX 4090 (YOLOv8n)	Very easy to use; integrated with Ultralytics, supports PyTorch, and has good documentation.
YOLOv11	mAP@0.50:0.95 ~ 55-60% (improved over YOLOv8)	130-160 FPS on RTX 4090 (YOLOv11n)	Easy to use; similar to YOLOv8 but with better modularity and scalability.
YOLO- NAS	mAP@0.50:0.95 ~ 55-62% (high accuracy)	100-120 FPS on RTX 4090	Moderate; requires Daci's platform for optimal use, but offers automated architecture search.
RT-DETR	mAP@0.50:0.95 ~ 50-58% (varies by dataset)	80-100 FPS on RTX 4090	Moderate; requires more setup compared to YOLO models, but offers transformer-based flexibility.

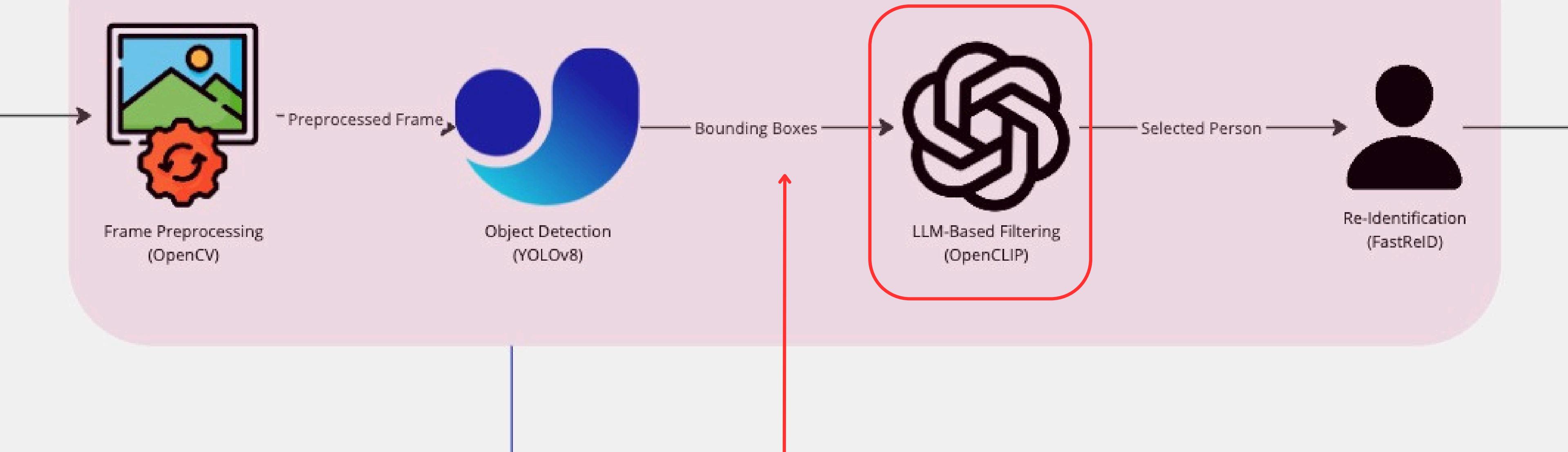


- OpenCLIP: An open-source implementation of the CLIP (Contrastive Language-Image Pretraining) model
 - The goal is to make the embeddings of matching image-text pairs closer together while pushing apart the embeddings of non-matching pairs. (**Embed text, Embed image, compare similarity**)
 - **Zero-shot learning**



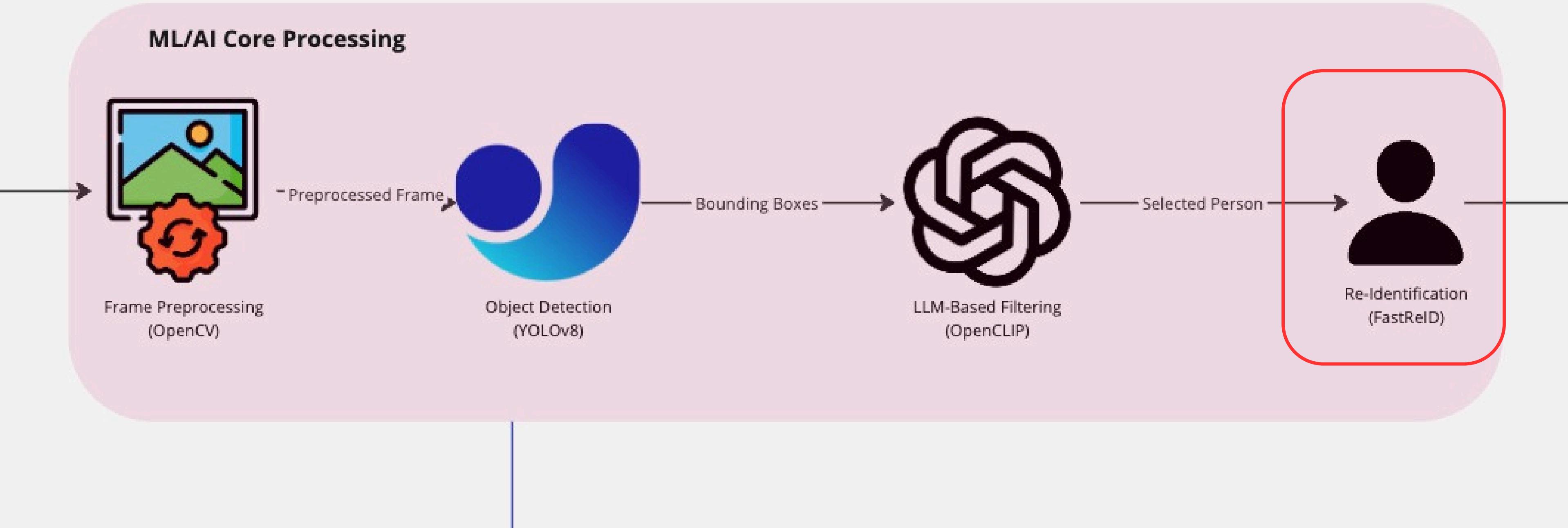
- Alternatives
 - Google's Flamingo
 - Bootstrapped Language-Image Pretraining (BLIP)

ML/AI Core Processing



- What if we got a complex prompt from the user?
 - Consider hooking up something like Llama 2 or GPT-4 for query expansion.

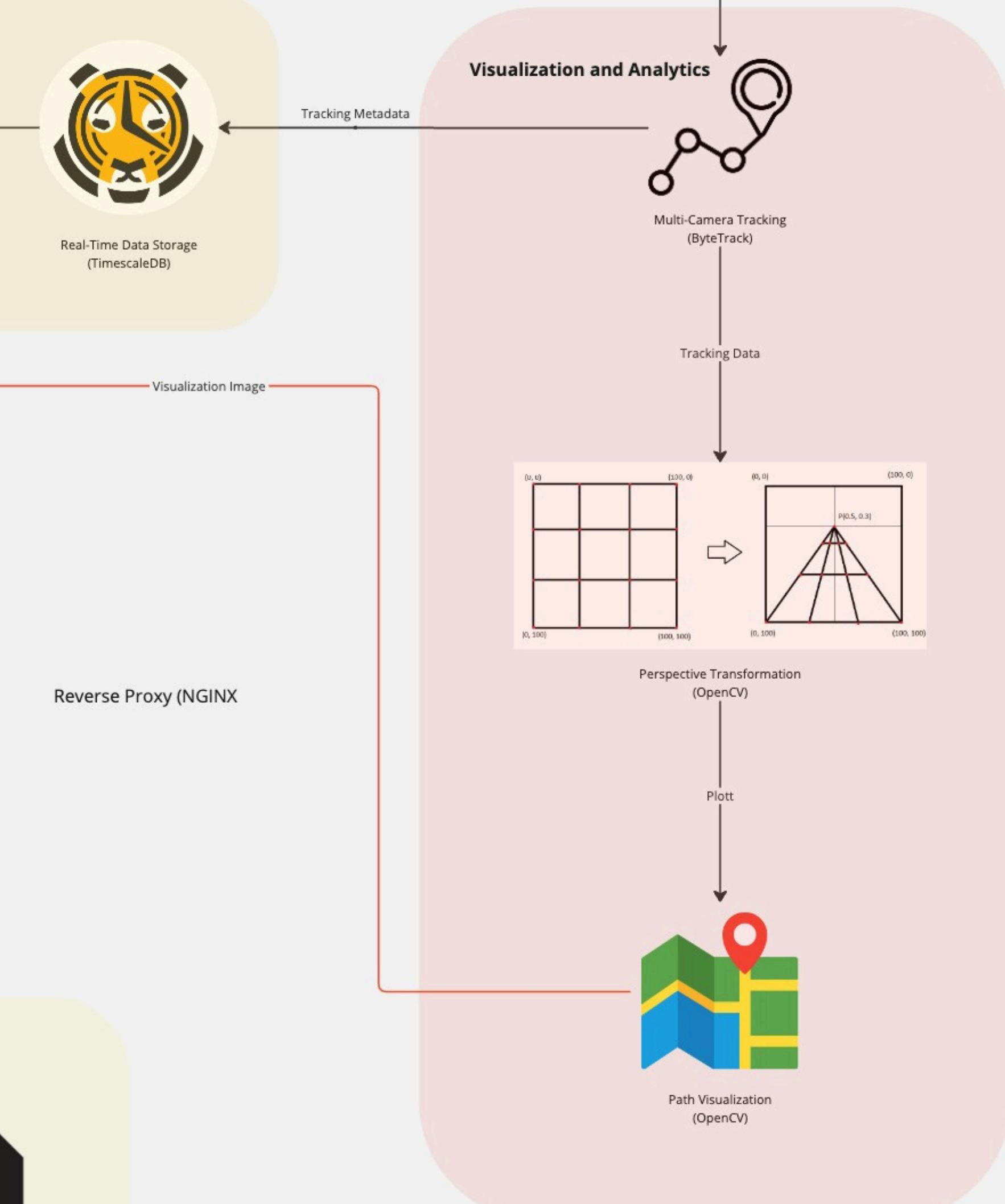
ML/AI Core Processing



- FastReID remain state-of-the-art toolkits for well-known pipeline re-ID.
- Might have to investigate the open-set re-ID method.

ML/AI Core Processing





ByteTrack: A **state-of-the-art** tracker that improves on StrongSORT by handling low-confidence detections more effectively. It is highly recommended for real-world scenarios with occlusions and noisy detections.

FairMOT: Combines detection and tracking in a single network. It's more robust for crowded scenes but **computationally heavier**.

CenterTrack: Best for **end-to-end detection and tracking** in dynamic environments, such as autonomous driving or sports.



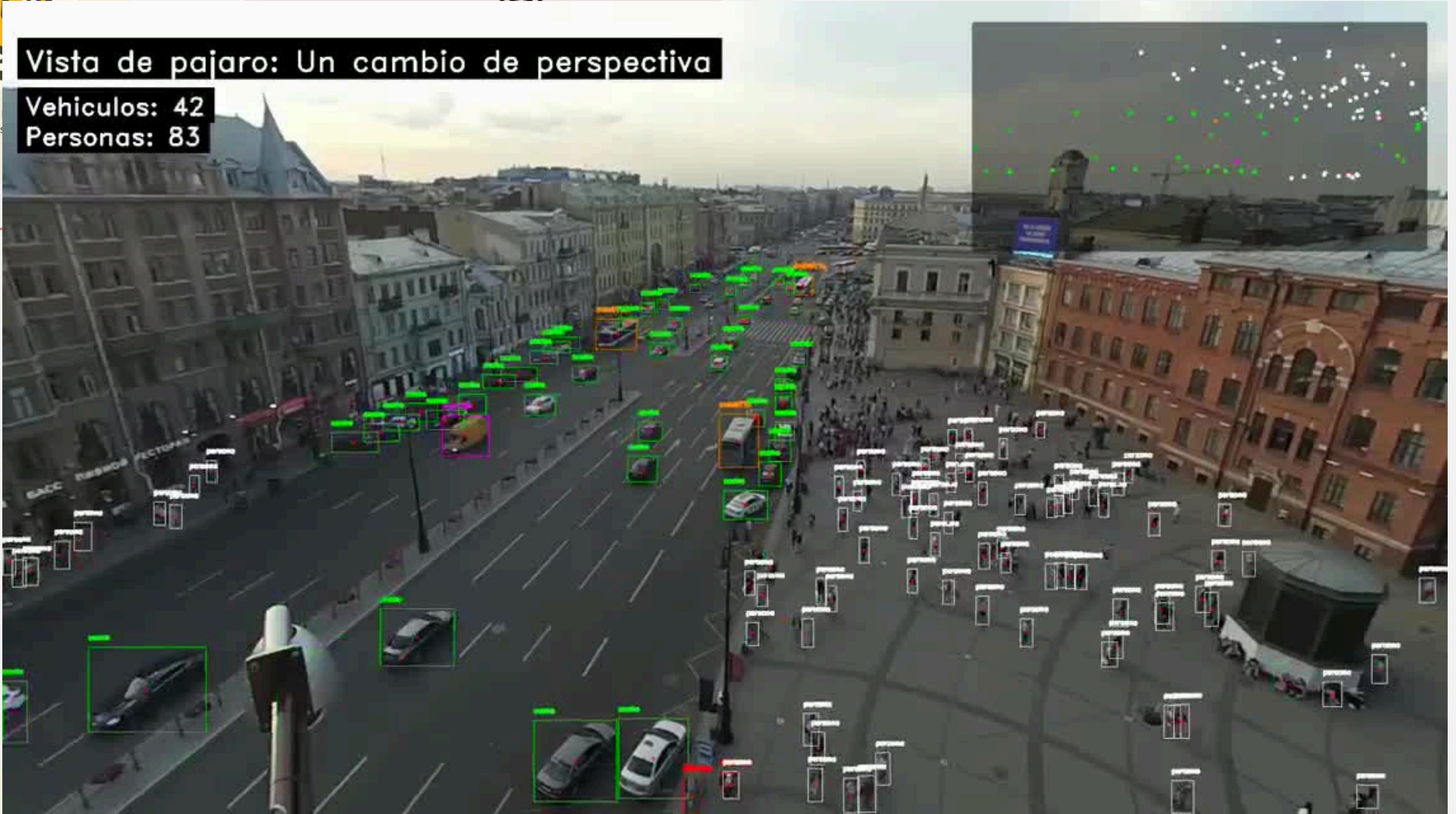
Vista de pajaro: Un cambio de perspectiva

Vehiculos: 42

Personas: 83

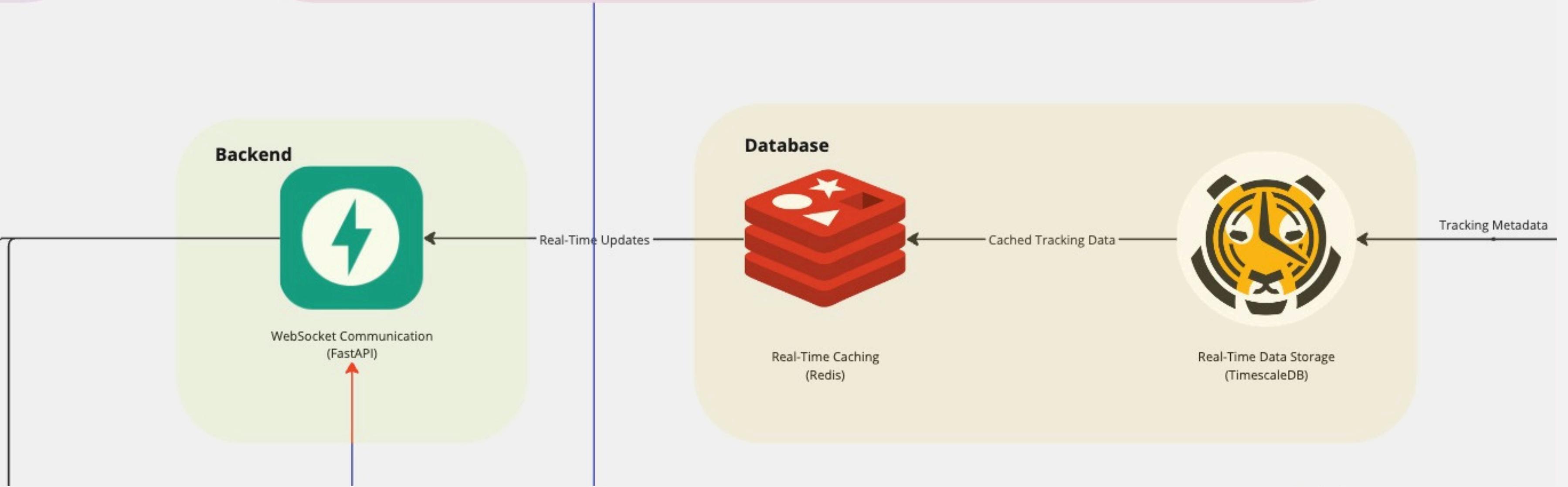
Real-Time
(Times)

Reverse



(OpenCV)



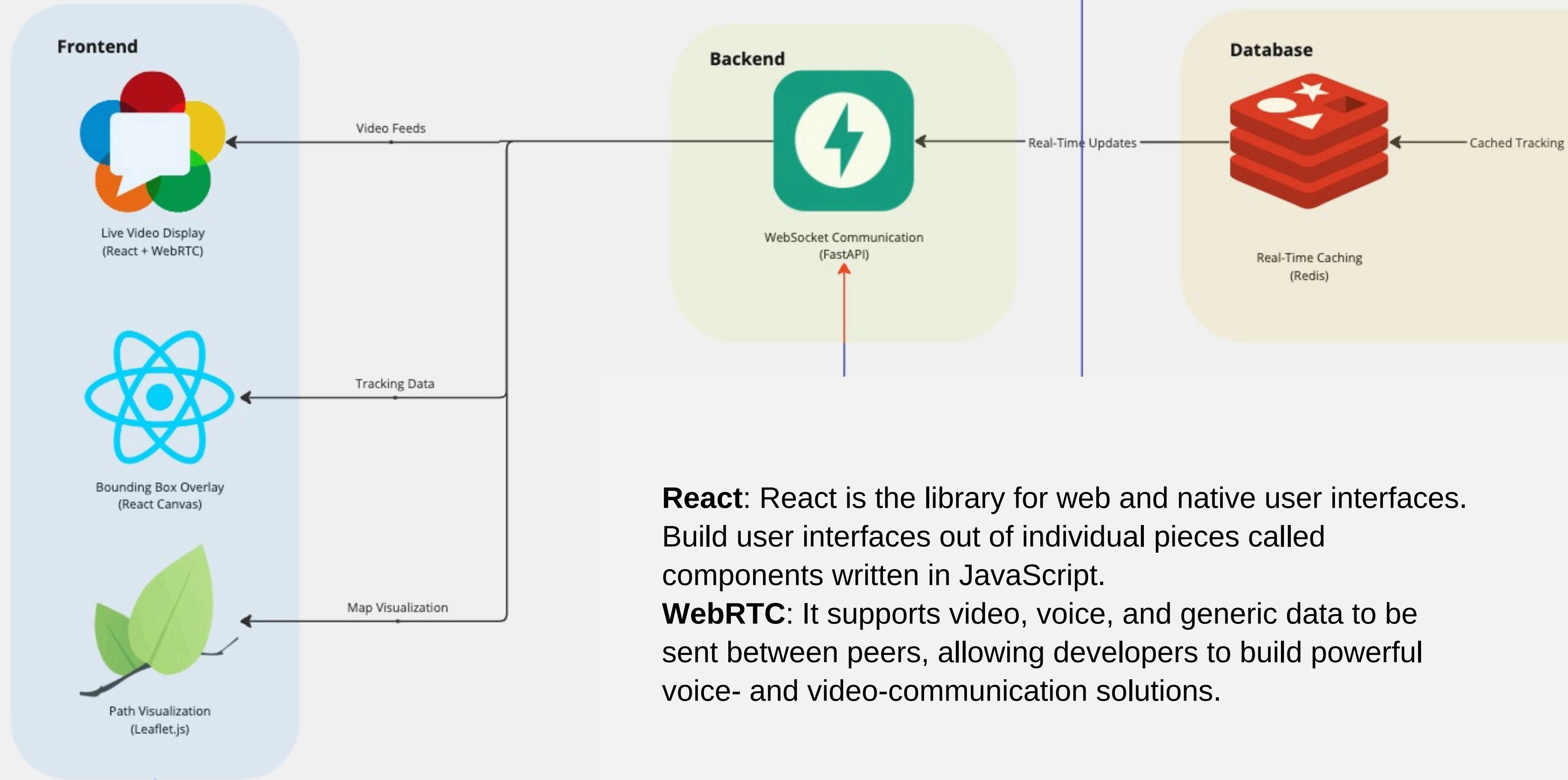


FastAPI: web framework for building APIs with Python based on standard Python type hints.

TimescaleDB: PostgreSQL extension designed for improved handling of time-series data.

Redis: Redis is a high-performance in-memory database that excels at real-time caching and pub/sub messaging.

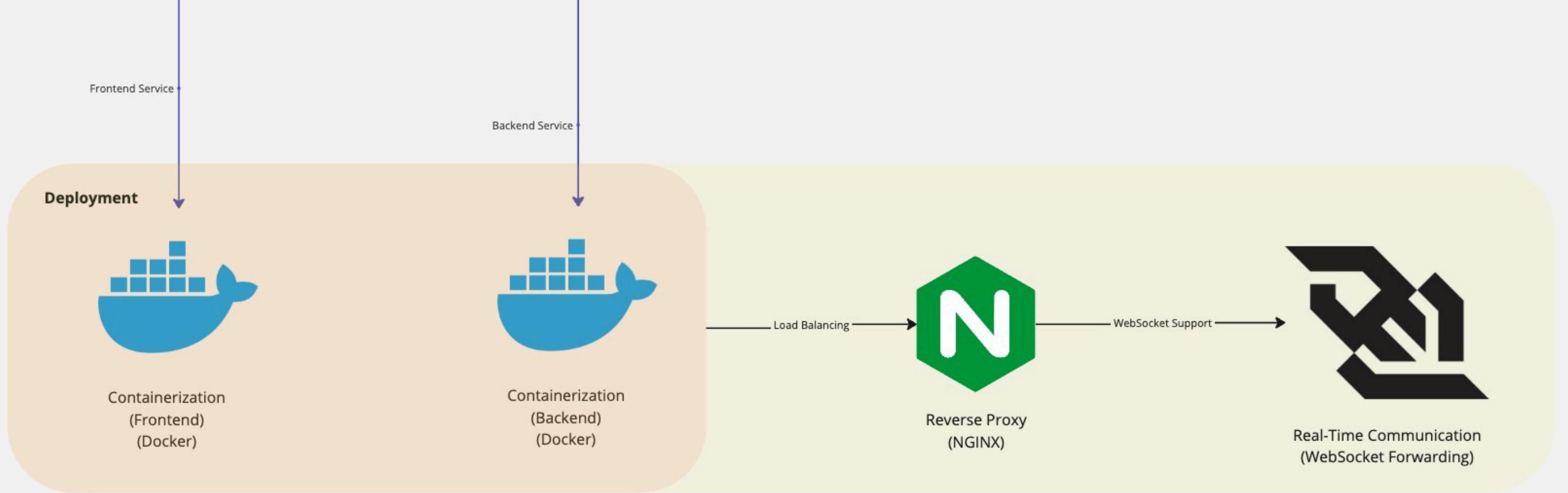
Websocket: WebSocket enables the backend to send real-time updates to the frontend without requiring the frontend to continuously poll the server.



React: React is the library for web and native user interfaces.

Build user interfaces out of individual pieces called components written in JavaScript.

WebRTC: It supports video, voice, and generic data to be sent between peers, allowing developers to build powerful voice- and video-communication solutions.



Docker: Docker is used to containerize both the frontend and backend services.

Reverse Proxy (NGINX): acts as a reverse proxy to manage incoming traffic and distribute it to the appropriate services.

WebSocket Forwarding: WebSocket forwarding ensures low-latency, persistent communication between the backend and frontend.

Thank You

