Modular Open-Source PTZ AI Camera

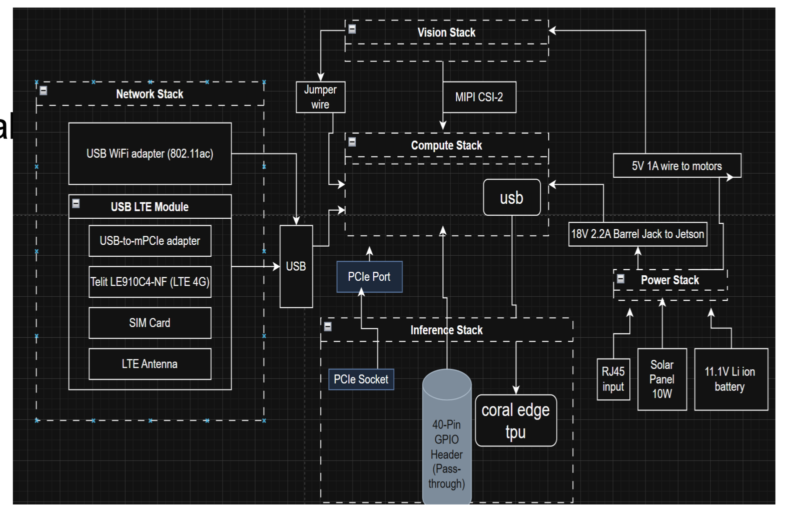
STUDENTS: Derek de la Cruz, Lun Jing, Tyler Braham, Brandon Le, Luke Ferguson

Motivation and Goals

This project builds a prototype modular PTZ camera system with on-device AI detection and tracking. The goal is to create an open system that is easy to modify and extend. Currently available similar systems are typically proprietary and thus are difficult to extend or customize based on the use case. This system supports different power, network, and hardware modules so that each potential user can tailor the device to their specific needs. The compute platform runs AI directly on the device to reduce delay and improve reliability, and the device is powered via either PoE (power over ethernet) and battery and solar support to increase reliability.

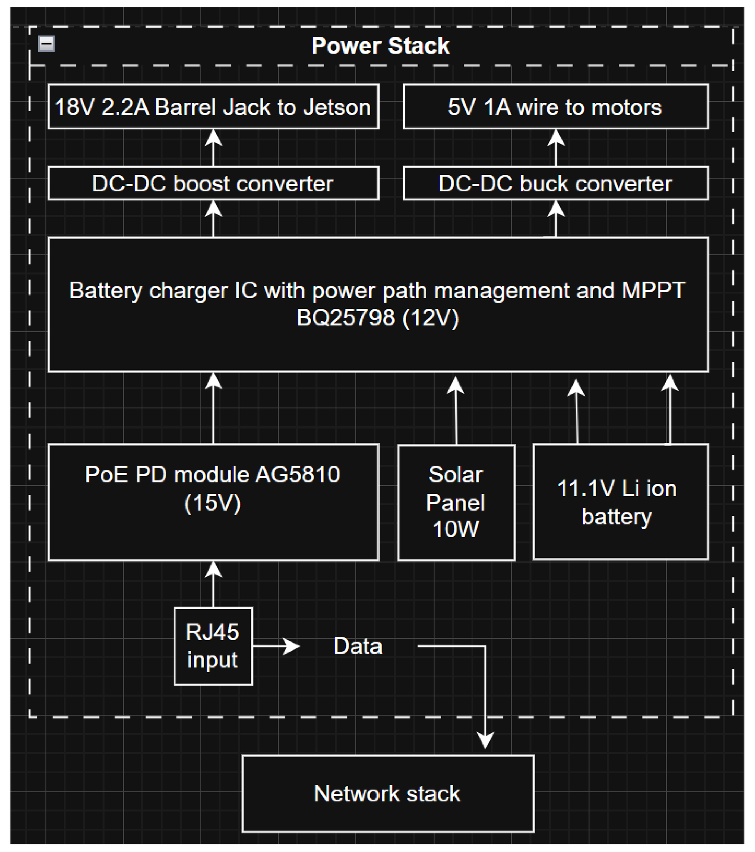
System Diagram and Interfaces

The prioject was broken down into several smaller blocks that each team member was responsible for. The critical interfaces are shown below.
Power to Jetson: 19V Barrel Jack
Jetson to network:
Power to Motors: 5V header

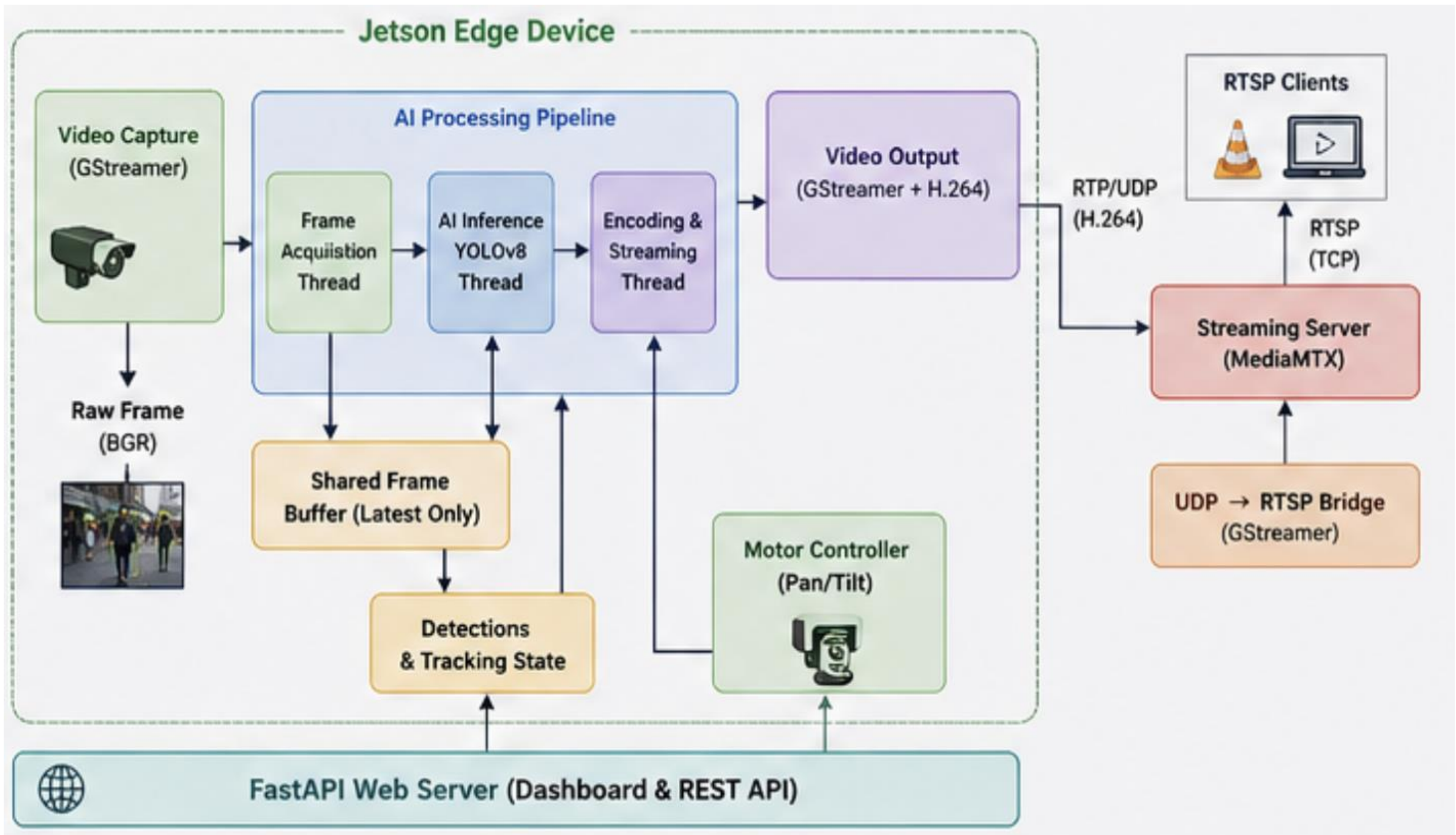


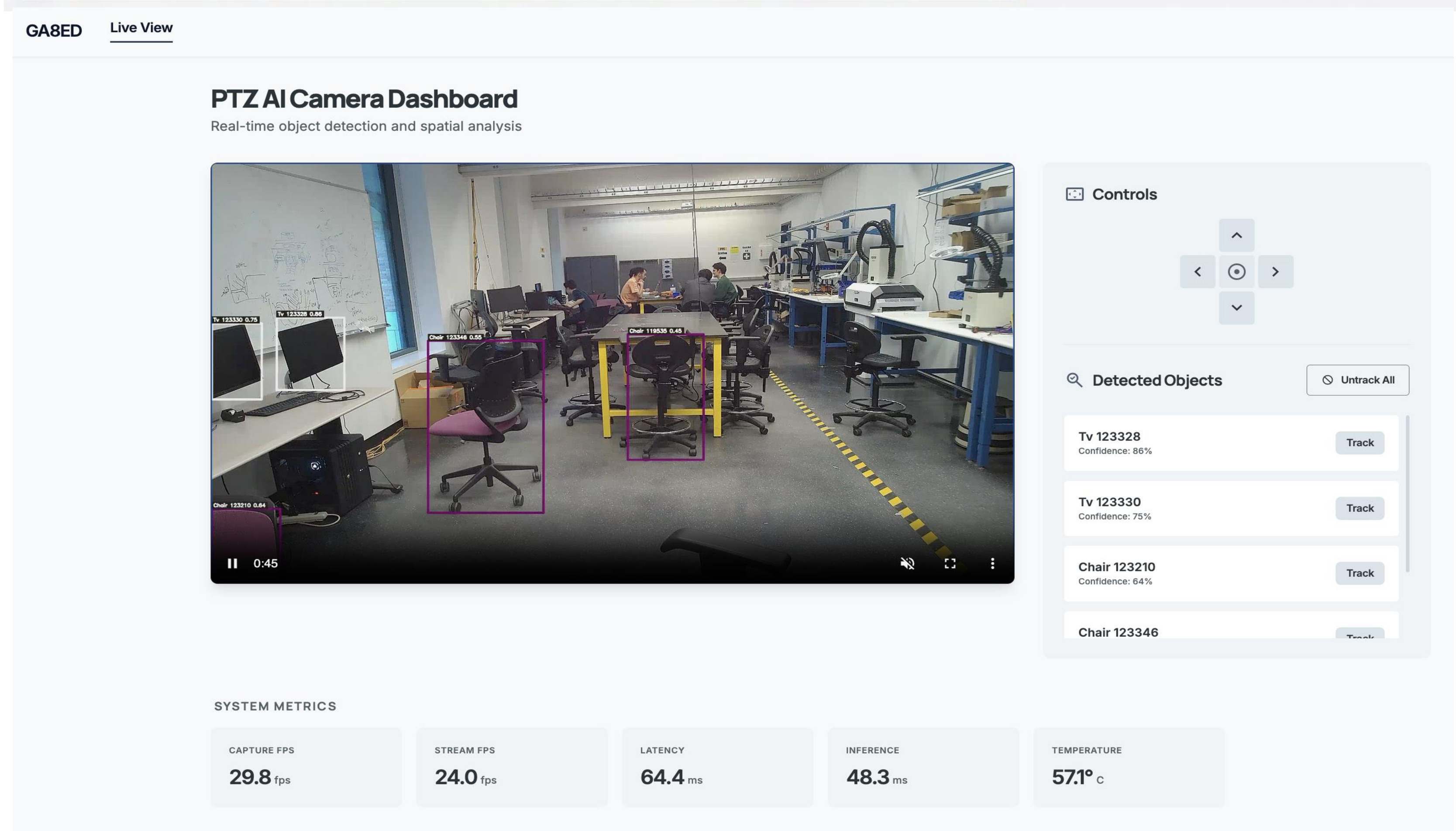
Power Stack

- PoE Module handles IEEE standard power negotiation
- Battery charger contains bidirectional buck-boost converter for charge discharge cycles
- Parallel converters isolate from each other to prevent motor current spikes from affecting compute
- Each unit needs output filtering capacitors to limit voltage ripple
- Battery charger configure I2C
- Compatible with 1-4 cell lithium batteries
- Compatible with solar panels up to 22V open circuit voltage
- High compatibility and multiple power options reflect the desired modularity of the system



IMPLEMENTATION





YOLO

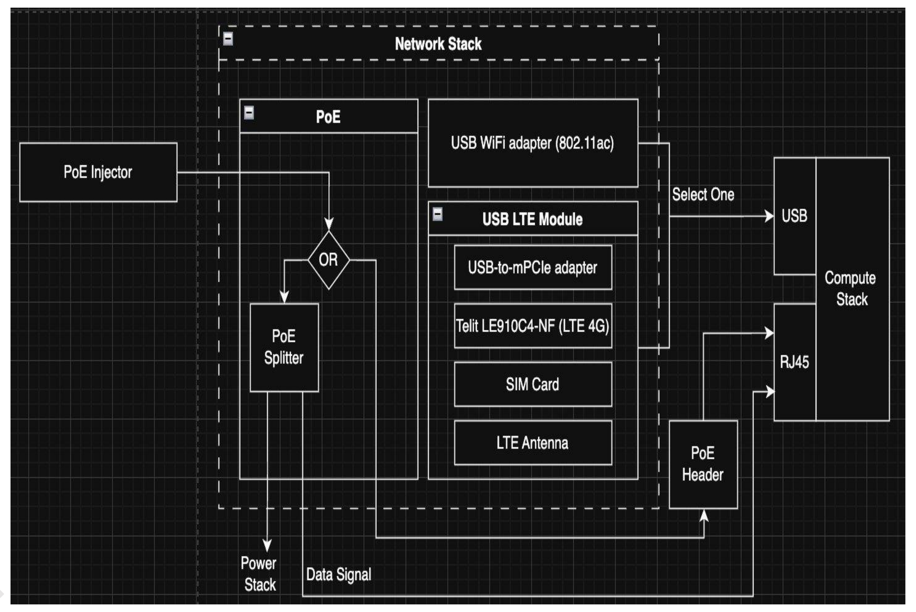
YOLO is a computer vision model that finds and classifies objects in each video frame. It runs locally on the Jetson to detect and track objects in real time from the camera feed. It draws labeled boxes, assigns object IDs, and sends target position data to the PTZ control loop so the camera can follow a selected object.

Vision

- Sony camera sensor with 12MP 60 FPS output capabilities.
- Pan and tilt system utilizes two 9g micro servo motors and PLA 3D printed mounting brackets.
- Camera communicates through MIPI CSI-2 connection directly with the Jetson.
- Motors are driven using an external PWM driver that communicates to the Jetson through I2C.
- Tracking software calculates distance from the center of the screen to the marked object and moves the motors fittingly.

Network Stack

- Edge-to-cloud connectivity using a mix of WiFi and LTE
- Uses FRP (fast reverse proxy) to create a bridge between the Jetson and the internet
- AWS EC2 instance running Caddy hosts the app and handles HTTPS/TLS encryption
- Uses WebRTC and MediaMTX to allow for sub 50ms low latency streaming



Results And Future Work

- Achieved ~30ms latency at 1080p resolution.
- Sustains 30fps using an optimized YOLO and GStreamer pipeline.
- Maintains >90% accuracy for reliable object identification.
- smooth PTZ control for 24/7 operation.
- Better camera sensor.
- Bracket and enclosure design improvements.

Transition to a SMARC SOM and baseboard architecture for connectors and modularity:

- SMARC SOM simplifies mass production integration.
- Dramatic performance-per-watt gains using NPU.
- Lower hardware costs for large-scale deployment.

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON

ADVISERS: Professor Dinuka Sahabandu and Ankit Talluri

SPONSOR: GA8ED: Shaiwal Singh and Gagan Mehta