



OPEN-SOURCE COMPUTER VISION PIPELINE TO TRACK ALASKAN BROWN BEAR ACTIVITY AND SALMON FEEDING EVENTS IN KATMAI FOR ECOLOGICAL RESEARCH AT SCALE

STUDENTS: OORJIT CHOWDHARY, CHLOE ZHONG, DARIAN DING, JINGWEN LI, KAITO YAN



ELECTRICAL & COMPUTER ENGINEERING

Problem & Background

- 2,200+ bears congregate annually each summer at Brooks Falls to feed on spawning salmon
- Monitoring bears relies on slow and error-prone manual observations by researchers and volunteers
- Automating this process unlocks significant ecological research data and insights for wildlife studies and curious public users



Katmai National Park



Relevant Camera Sites

- The region has 5 livestream camera feeds at Explore.org, but no footage is being used for tracking and analysis at scale
- Katmai footage is challenging to generalize due to environmental conditions like water splashes, occlusions, dynamic lighting, and operator led camera zooms

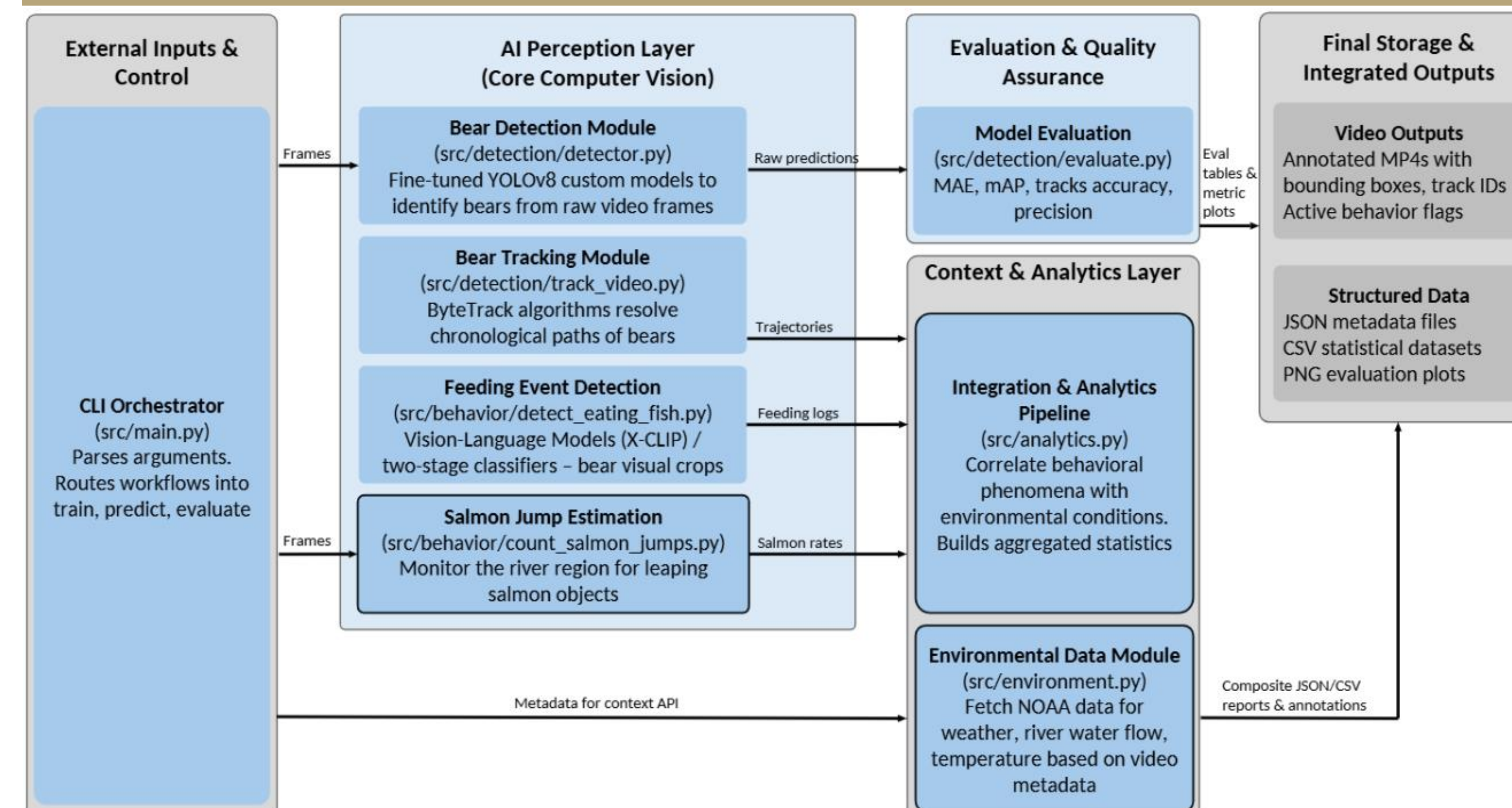
Dataset & Preprocessing

- About 58 GB of trimmed Katmai livestream recordings
- Extracted 24,238 frames at 1 FPS from 5 camera views + 1 multi-view
- Removed YouTube UI and screen-recording distractions
- Used Grounding DINO for zero shot "bear" prompt auto annotations
- Explored CLIP + OWL-ViT + Florence-2 MoE model for salmon annotations
- Created metadata extractor script to fetch recording datetime of inputs
- Organized videos by bear/salmon presence and feeding behavior



Training Data Creation

System Architecture



Repository: <https://github.com/katmai-vision-lab/katmai-cv-pipeline>

Bear Detection Model

- Our data was unlabeled so chose YOLOv8 model fine-tuning to build on top of COCO pretrained weights
- Selected YOLOv8 Nano for speed and efficiency for consumer grade laptops
- Recall improved from 2% --> 91%, mAP@0.5 from 18% --> 95%
- COCO pretrained "bear" class provided anatomy; fine-tuning adapted it to Katmai's unique environment settings



Bear Detection on Brooks River

Multi-Object Tracking

- Integrated ByteTrack algorithm with custom thresholds for scenarios with waterfall occlusions
- Two-stage data association with Kalman-filter and IOU matching: high confidence detections first, then low confidence detections used to recover temporary bears
- Outputs unique bear count, per-frame tracking IDs, and full trajectory data per video



Multi-Bear Recognition and Movement Tracking

Salmon Jump & Feeding Events

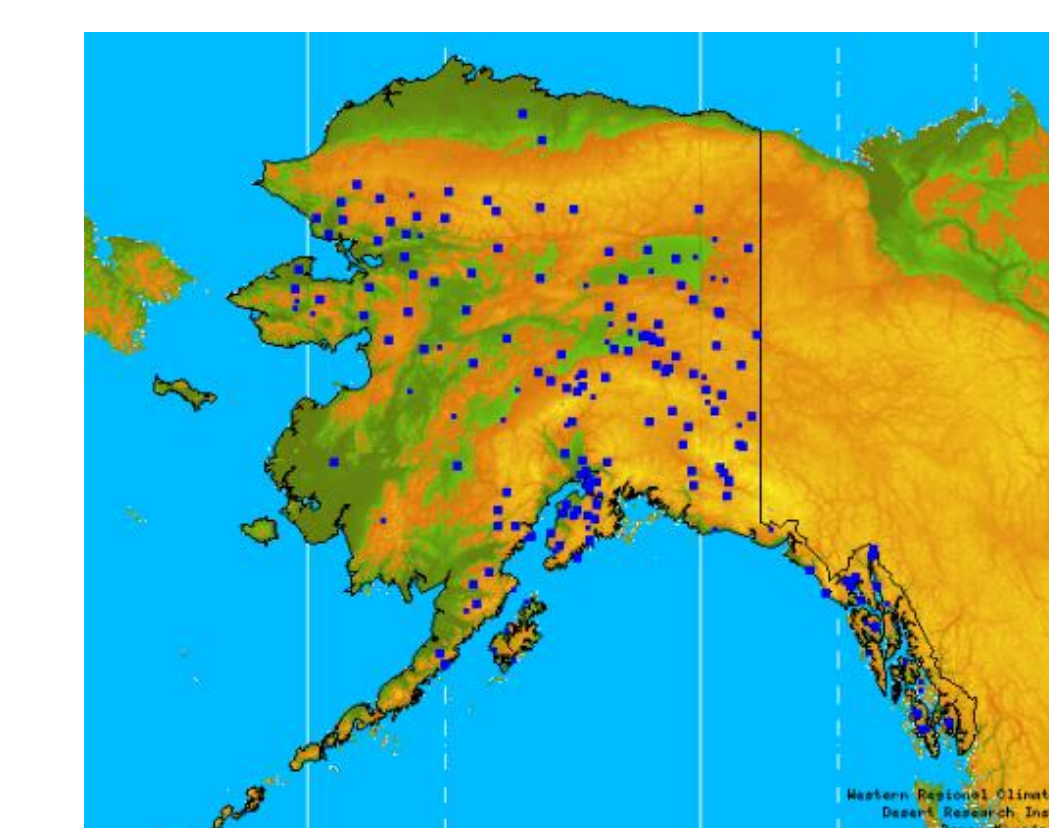
- Salmon jump counting uses MOG2 background subtraction combined with HSV color masking and ROI constraints to isolate jumping fish from turbulent stream
- Blob area peaks across frames are smoothed and peak-detected to count discrete jumps and output per jump times
- Feeding event detection runs YOLO + ByteTrack on each frame, overlays bear IDs, and queries pluggable VLM backend to intelligently caption each frame
- Bear level state timelines are summarized to count confirmed catches, missed attempts, and exact catch timestamps



Salmon Jump Counting

Environmental Data

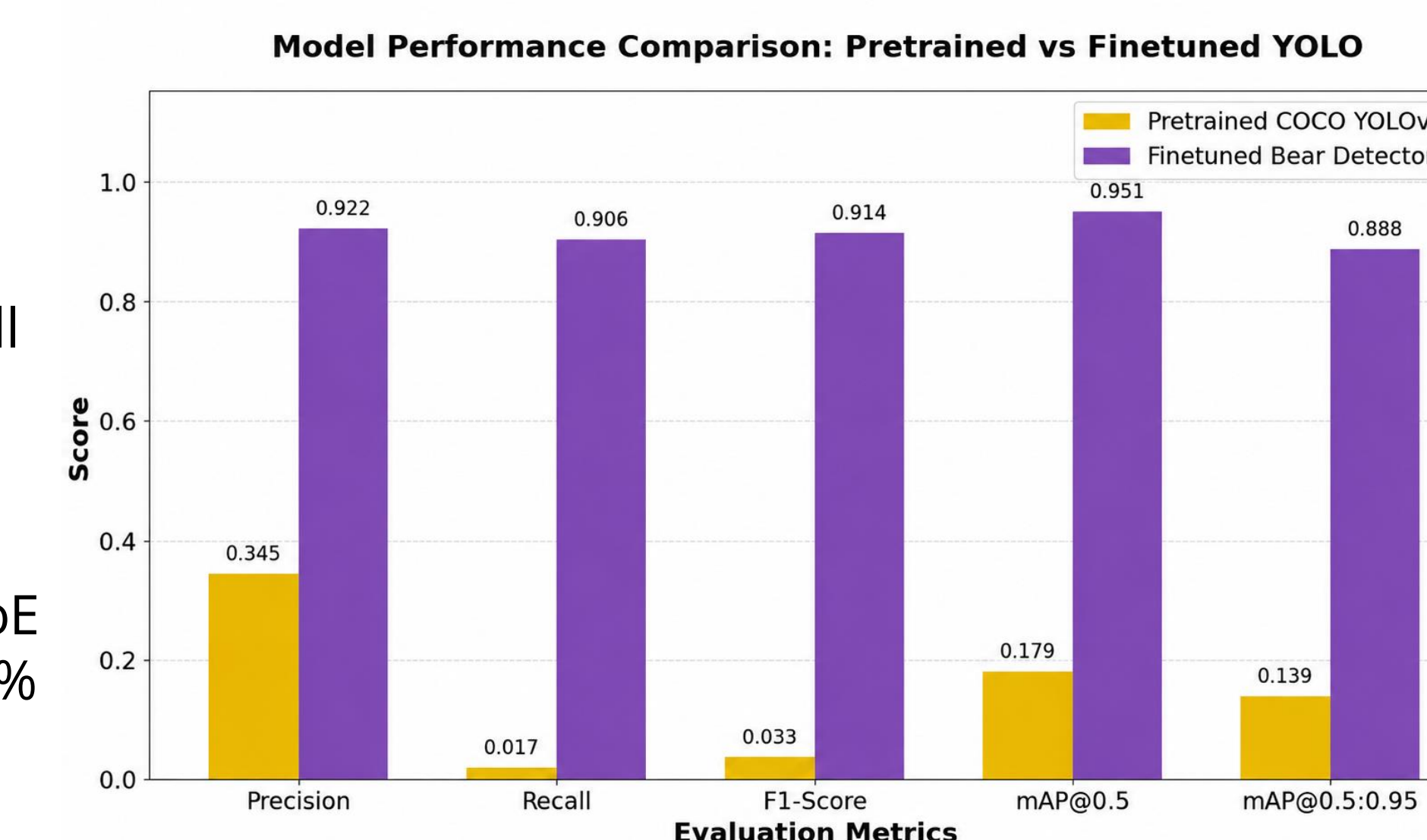
- Weather data is pulled from 3 WRCC RAWs (Coville, Pfaff, and Three Forks) stations within 75 km of Brooks Falls, with hourly data for wind, temperature, humidity, precipitation
- Daily precipitation is supplemented via NADP station AK97 for southwest Alaska
- The nearest active USGS river stream gauge is 150 km away in Kvichak river; not a good proxy but maintained as a research artifact



RAWs Climate Map

Results & Constraints

- Fine-tuned model outperformed pretrained YOLO
- Generalized successfully across all camera angles
- Runs on consumer-grade hardware
- Salmon detection MoE model achieved 97.5% precision and 96.6% recall



Future Work

- Moose detection (and other animals)
- Expansion to use in other locations
- Executable on CPU only
- Bear trajectory over panning videos