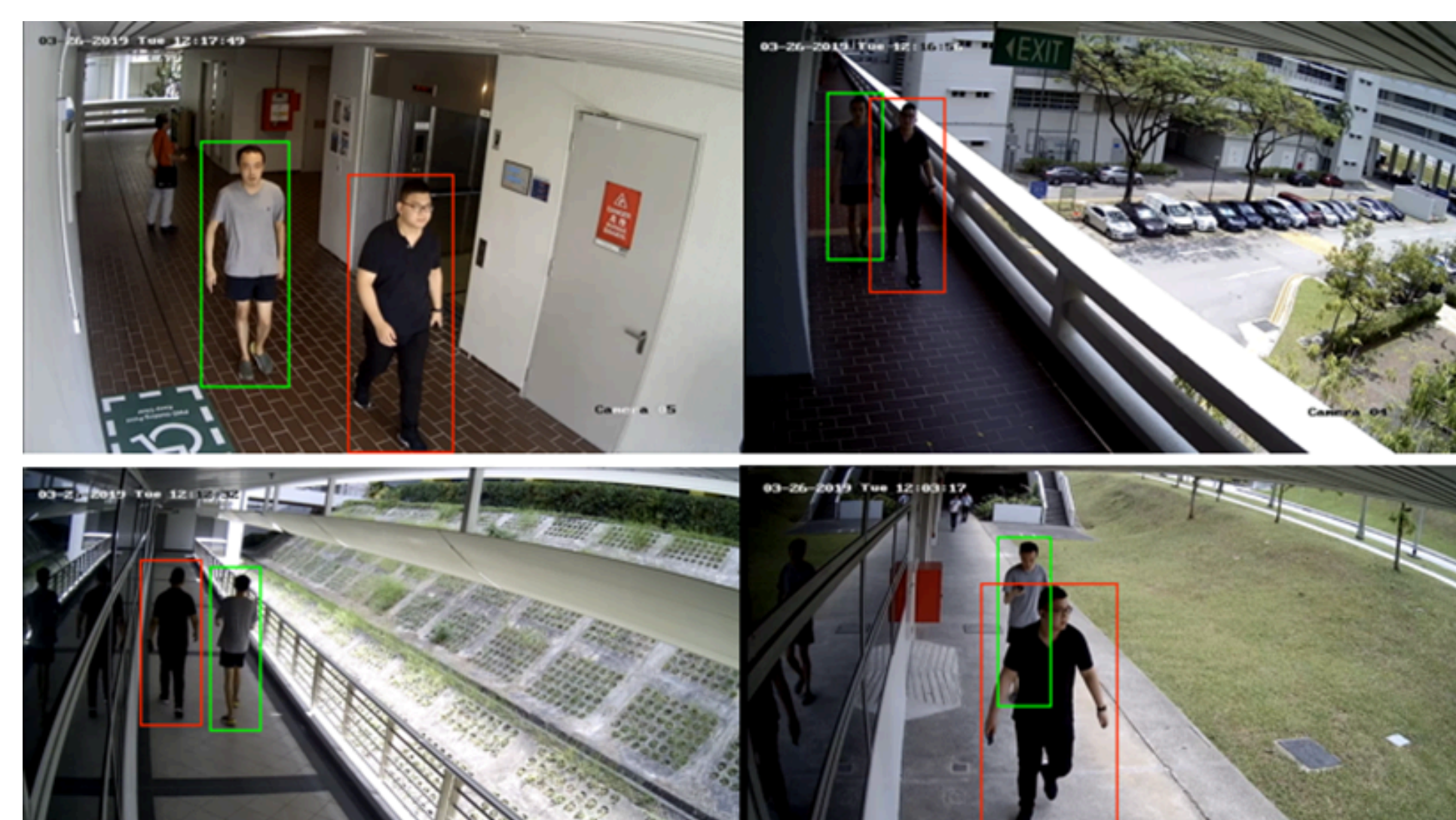


IMPROVEMENTS IN PERSON IDENTIFICATION IN VIDEO UNDER CONSTRAINTS

project code: 25-1-R-20
Capstone Project Phase B – 61999
Jul 2025
Asad Saffouri
Alexandra Simonishvili
Supervisor: Mr. Ilya Zeldner



Background

This project focuses on improving person identification in video streams by combining (YOLO), (FaceNet), and (FlowNet). The system is designed to support real-time tracking of individuals across frames, even when faces are partially hidden, or missing.

Main Requirements

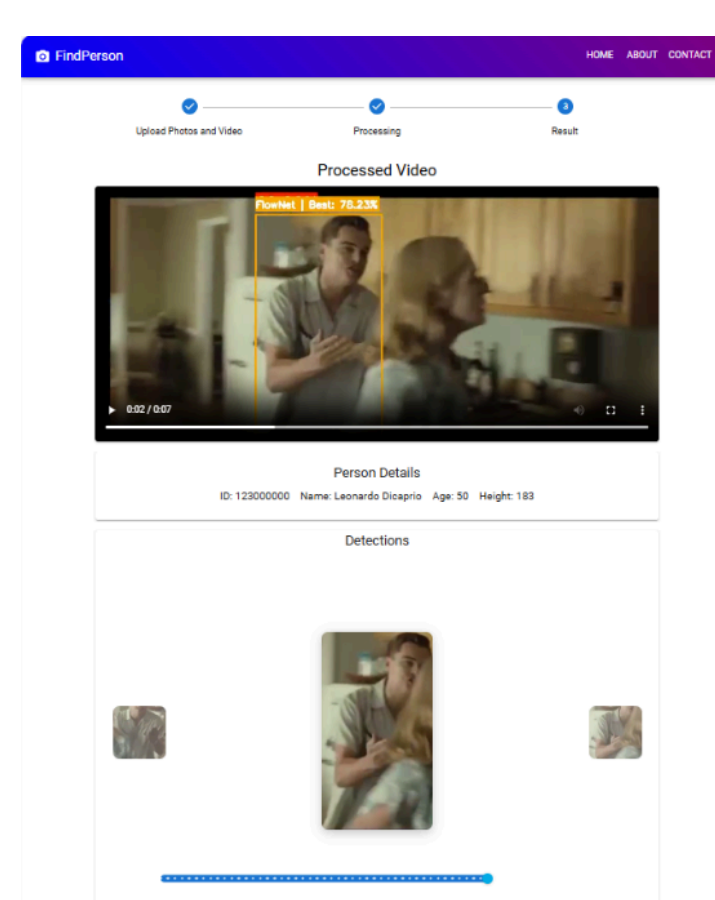
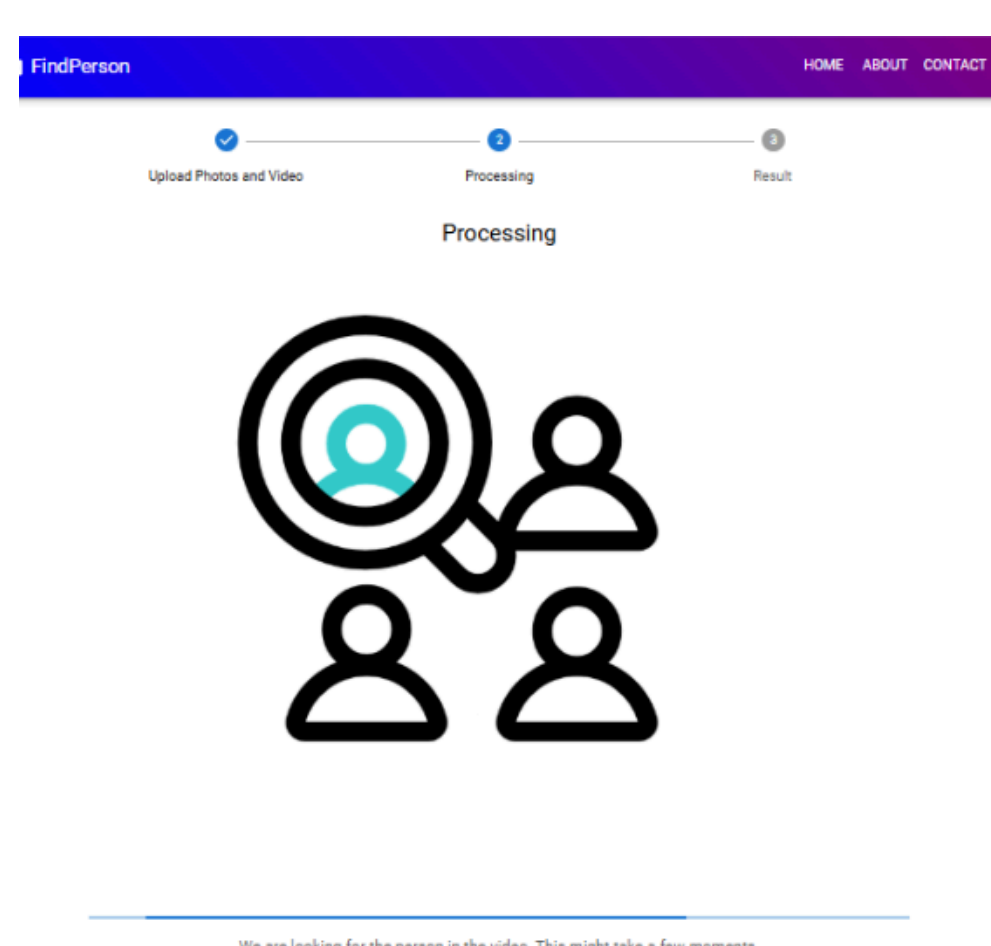
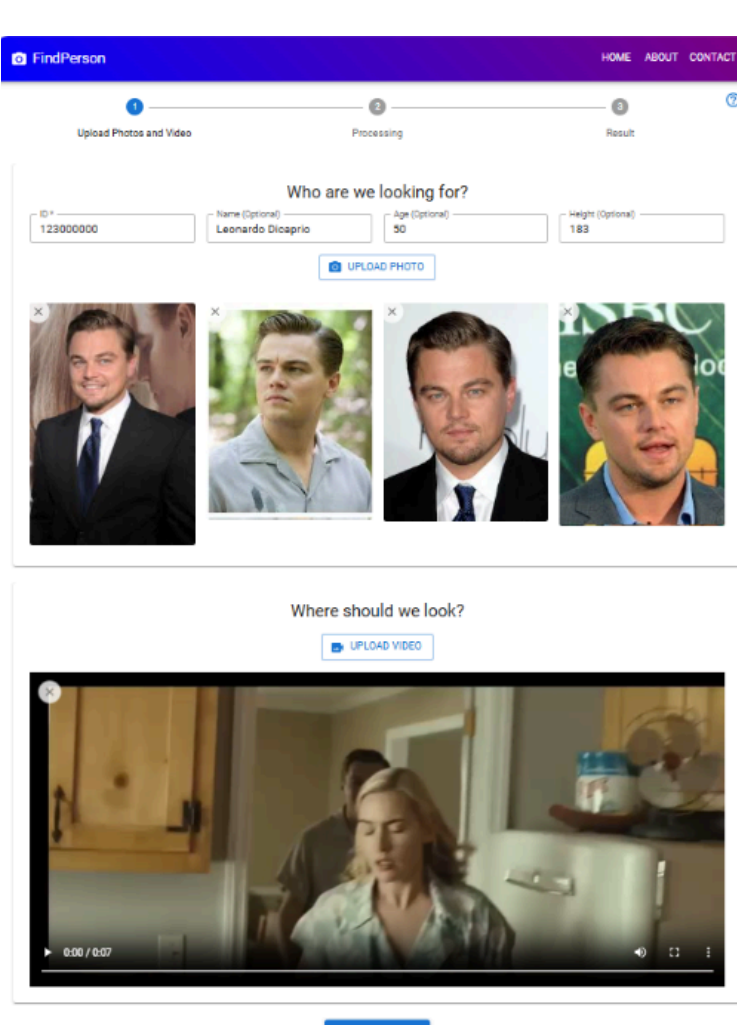
- Upload videos and images through a user-friendly interface.
- Detect people in video frames using YOLO.
- Extract and compare faces using FaceNet
- Display results with bounding boxes and similarity scores.
- FlowNet enables continuity in identity tracking when faces are turned or missing.

Our user interface

Loading Data Screen

Analyzing Screen

Result Screen



Challenges

- 1.Low Visibility:** Poor lighting or low-resolution videos often obscure faces, making recognition unreliable.
- 2.Heavy Processing Load:** Long or high-frame-rate videos strain system performance, causing delays or frame skips.
- 3.Complex Scenes:** Tracking one person is harder when multiple people or dynamic backgrounds are present.

Tests

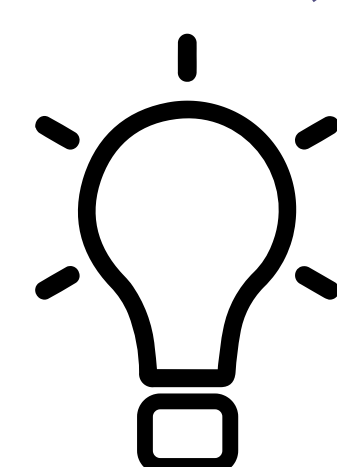
- 1. Purpose:** Detects people in video frames using object detection
Expected: Bounding boxes correctly appear around each person.
Result: Passed
- 2. Purpose:** Generates a unique embedding for each detected face.
Expected: Embedding vectors created successfully.
Result: Passed
- 3. Purpose:** Tracks person movement between frames, even with occlusion.
Expected: Motion vectors generated and accurate.
Result: Passed
- 4. Purpose:** Ensures tracking resumes after person becomes visible again.
Expected: Person is re-identified after temporary disappearance.
Result: Passed

conclusion

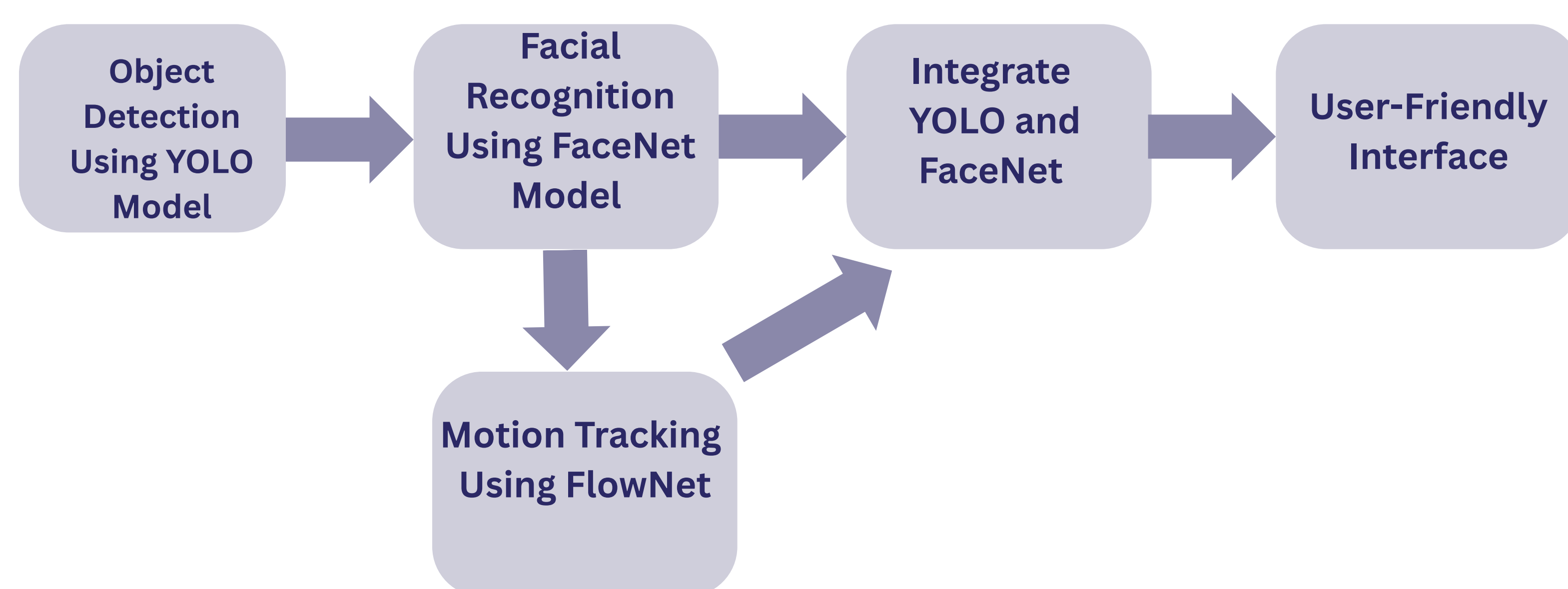
Adding FlowNet improved identity tracking across frames, especially when faces were blurred, turned, or missing. The system became stronger in videos with motion and occlusion.

Problem

Person tracking systems based on YOLO and FaceNet often fail when the face is not clearly visible due to motion, angle, or occlusion. In such cases, YOLO may detect the person, but identity tracking breaks, leading to inconsistent recognition across frames.



Our solution



ARCHITECTURE



Technologies and Tools

React (Frontend)

Builds the user interface for uploading photos, video, and displaying detection results.

FastAPI (Backend):

Handles video processing requests, coordinates between modules (YOLO, FaceNet, FlowNet) and returns annotated results

MongoDB (Database):

Stores facial embeddings and updated tracking data generated by motion analysis

OpenCV (Video Processing):

Manages frame extraction, annotations, and video output.

Results of Our Enhanced Tracking Solution

