CS566
Fall 2025

Vision-Based Dynamic Objects Path Predictions for Safer Robot Navigation on Construction Sites.

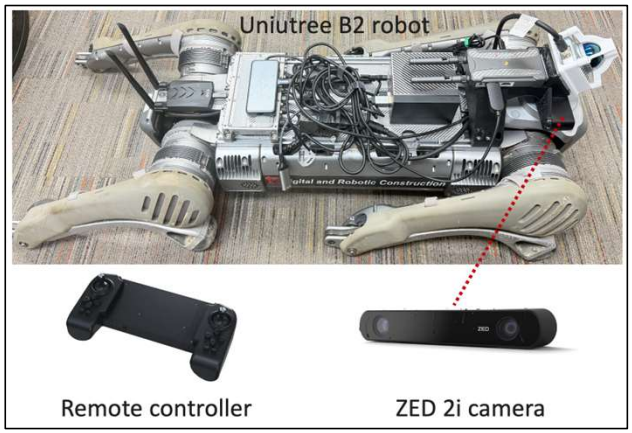
Liqun Xu, Pakorn Boonpetch

Project inspiration

UW-Madison Kellner Family Athletic Center.

Layout of the construction site, photographed on May 21, 2025

Pipeline – Data Collection

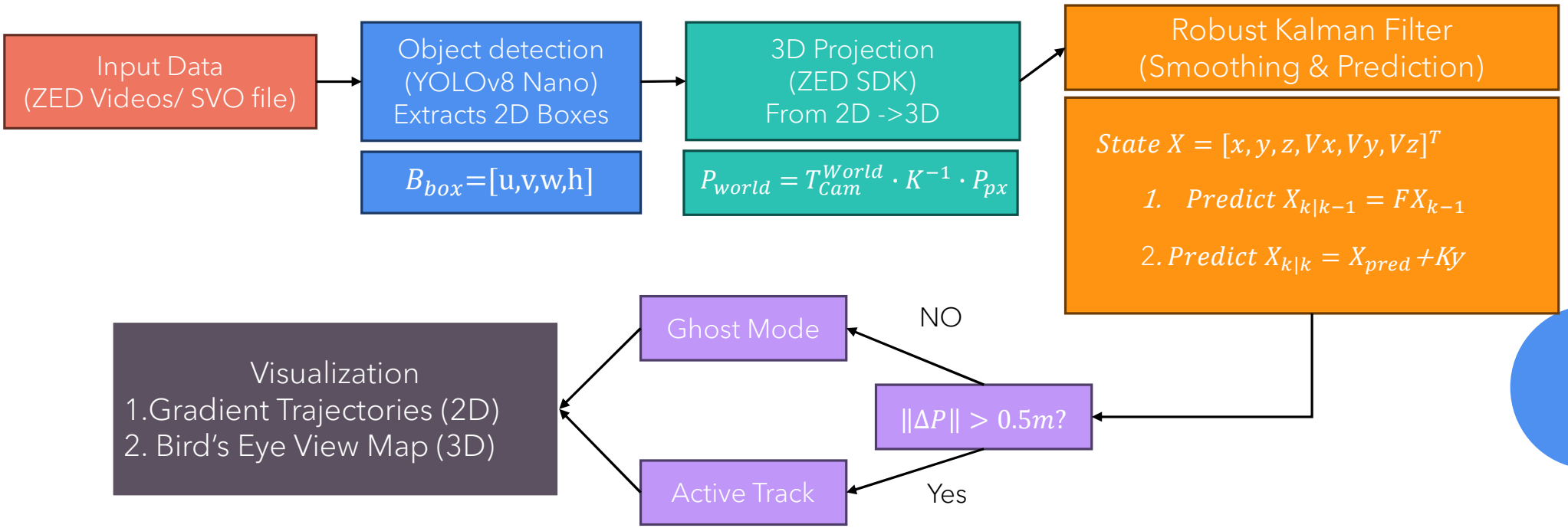


Unitree B2 used for field data collection

Field Data Collection

Example of the recorded Video

Pipeline – Our Approach

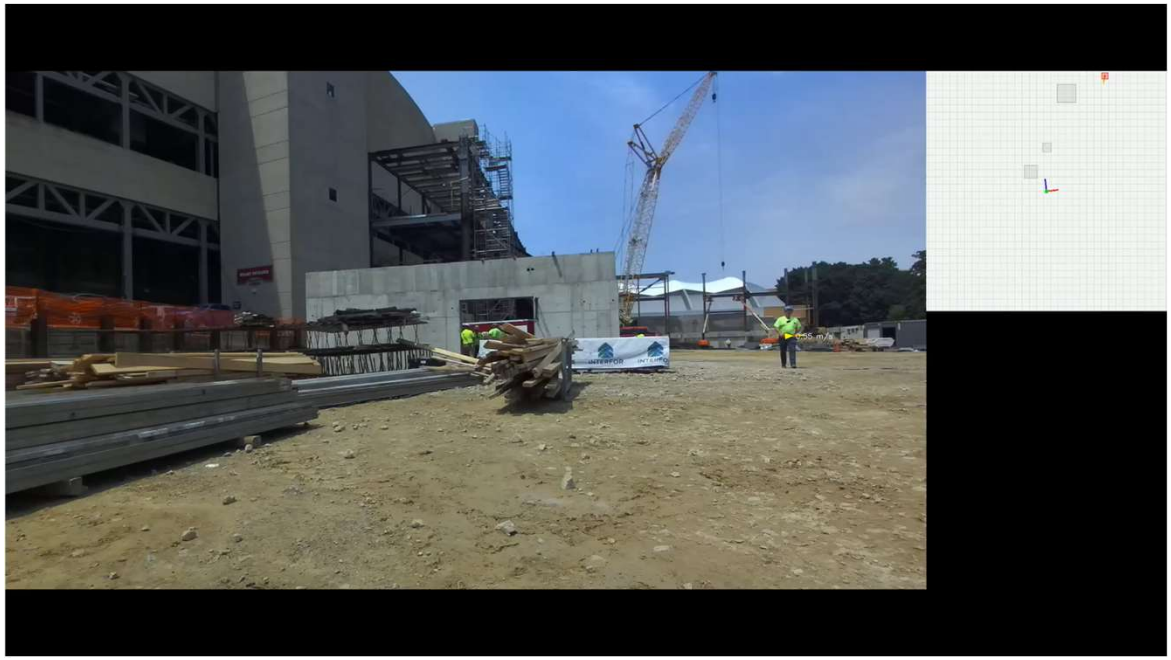


Demo-1 Worker

Bird's Eye View Map

Robot View

Demo-2 Worker Standing



Bird's Eye View Map

Robot View

Demo-3 Worker Stop & Go

Bird's Eye View Map

Robot View

Demo-4 Load and Truck

Bird's Eye View Map

Robot View

Discussion and Future Work

1. Comparison: Physics-Based vs. Data-Driven

Our Approach (Robust Kalman Filter):

- Pros: Real-time performance (low latency), highly interpretable ("Ghost Mode"), and works immediately without massive training datasets.
- Cons: Assumes linear velocity; struggles with sudden sharp turns.

Deep Learning Pipelines (e.g., LSTM / Transformers):

- Pros: Can model complex, non-linear paths (e.g., walking in a curve).
- Cons: "Black box" unpredictability, higher computational cost, and requires huge annotated datasets to avoid overfitting.

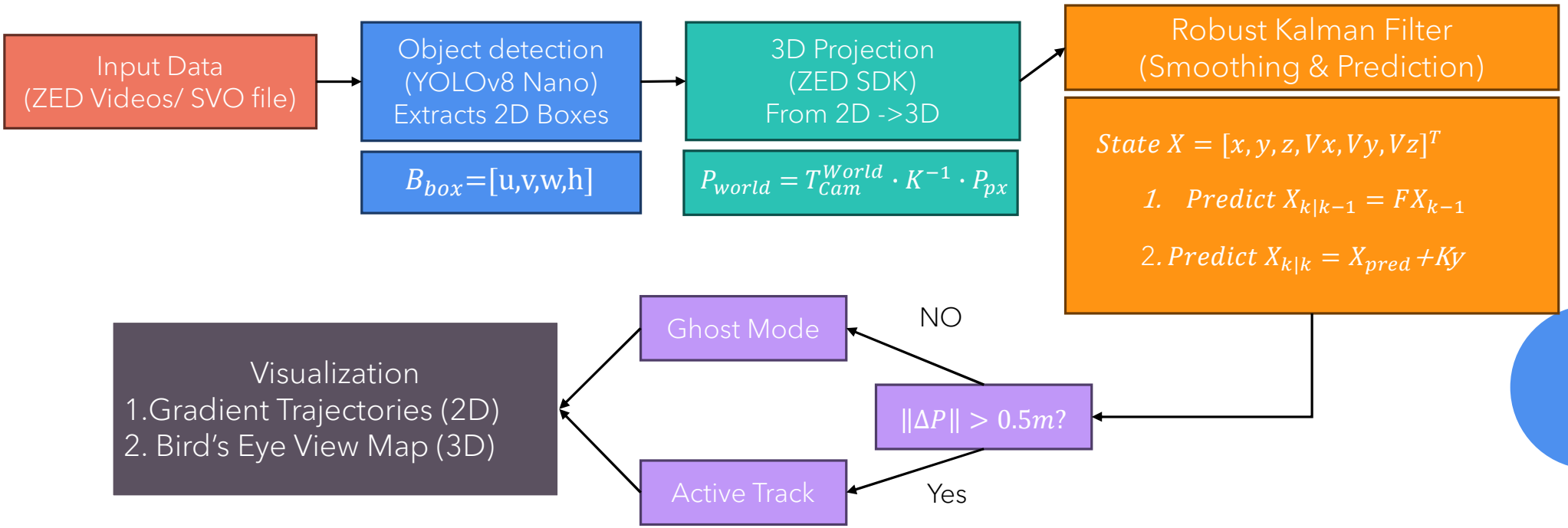
2. Future Improvements

Interaction-Aware Modeling: Implement Social-GAN to predict how workers avoid each other (currently, we track objects independently).

Sensor Fusion: Integrate LiDAR data with the ZED camera to solve the low-light/occlusion failures we observed in the field.

Long-Term Forecasting: Train an LSTM specifically for >5-second predictions where the linear Kalman assumption fails.

Pipeline – Our Approach




Thank you!
Q&A!