THE UNIVERSITY OF CHICAGO

Argonne NATIONAL LABORATORY

Intelligent Monocular Forward-Vision System (IMFS)

Agenda

Argonne
NATIONAL
LABORATORY

1. Client and Business Statement
2. Our Product - IMFS
3. Model in Action!
4. Data is Everything!
5. Future Development & Client Statement

Argonne National Laboratory, one of the U.S. Department of Energy's 17 national labs, is engaged in cutting-edge research in energy, transportation, and national security. This project is done in collaboration with Connected and Automated Vehicle (CAV) research.

Kevin Stutenberg leads the project, with **Michael Pamminger** and **Yihe Chen** providing technical research and implementation.

Advanced Computing & AI

Energy Innovation

Cross-Disciplinary Scientific Research

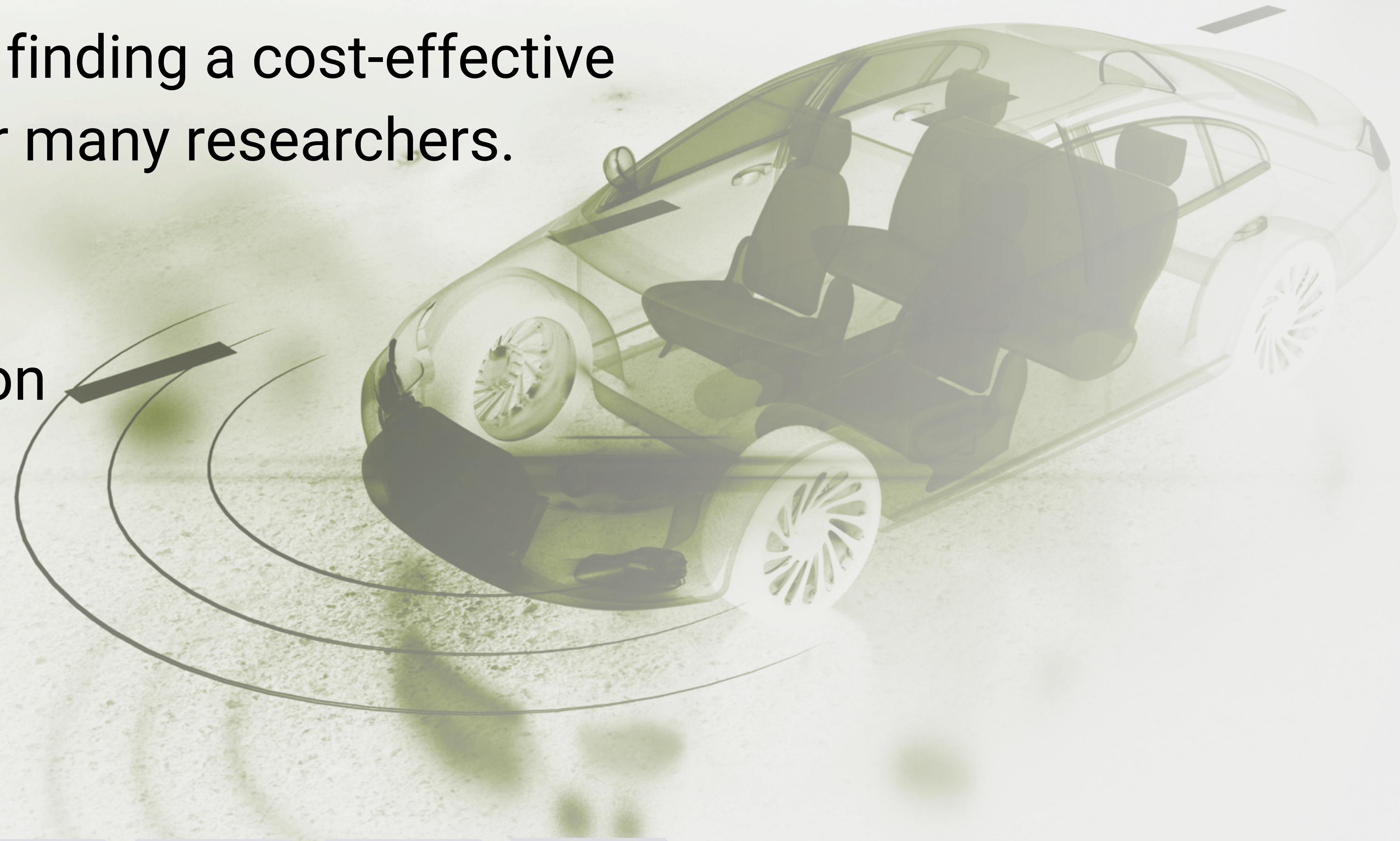
Business Problem & Opportunity

The Problem

Determining lead vehicle distances is crucial for analyzing driver behavior. However, the high costs associated with LiDAR and RADAR technology pose a barrier to research and finding a cost-effective alternative remains a significant hurdle for many researchers.

The Opportunity

Provide an accessible and scalable solution that delivers reliable real-time distance estimation using monocular video from common hardware like dash-cams.



Project Goals

Argonne NATIONAL LABORATORY

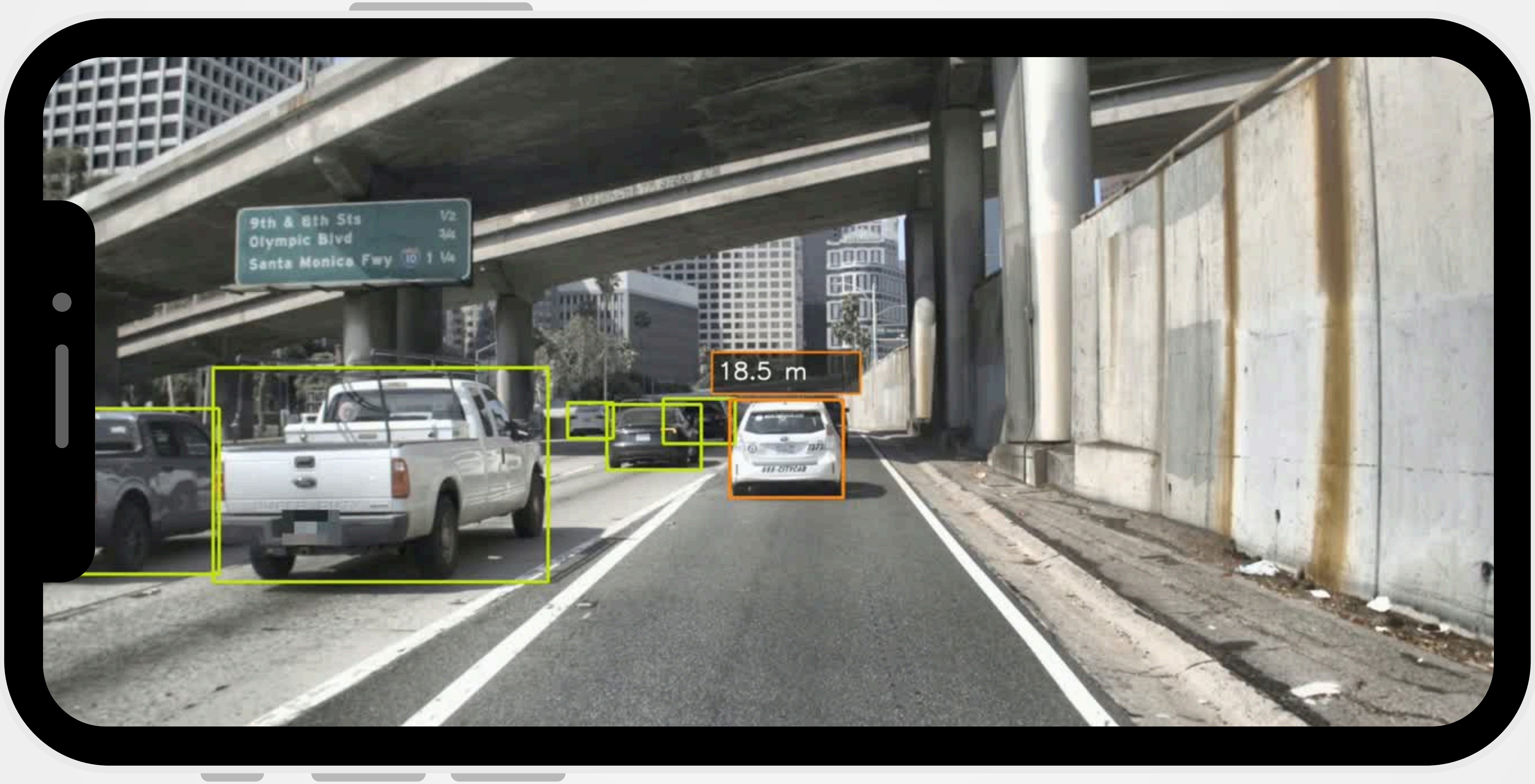
Core Project Goals

- Estimate distance to lead vehicles to understand driver behavior.
- Design and benchmark a lightweight, real-time AI model.

Target metrics

- Latency: 100 ms
- Model Size: Less than 100 MB
- Absolute Error: Approximately 2.00 m

What We Achieved



Client | Business Problem | Our Product | Model In Action | Data Assessment | Limitations | Future Scope

Model in Action: Preprocessing

Raw Image

Raw monocular dash-cam frame

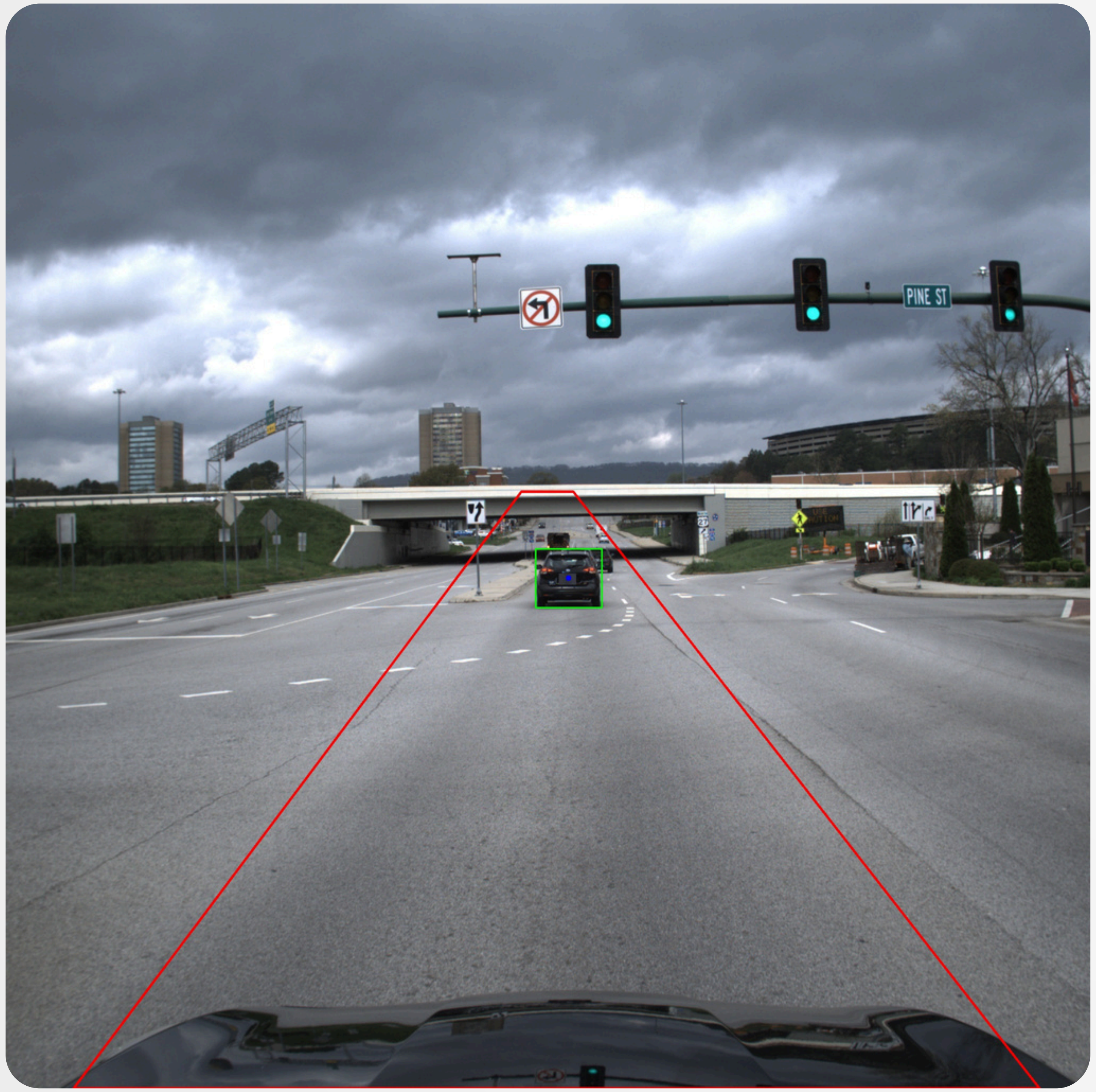
Undistorted Image

Lens distortion removed & perspective straightened for accurate geometry.

Fully Processed Image

Brightness/contrast normalized

Model in Action: Lead Vehicle Detection



Region Of Interest Polygon

- Restricts detection to the drivable area, emphasizing the lead vehicle zone.

YOLO - Real Time Object Detection Algorithm

- Detects vehicles in real time and provides accurate bounding boxes for the lead vehicle.

Center Focus

- Gives priority to vehicles near the vertical center to isolate the true lead vehicle.

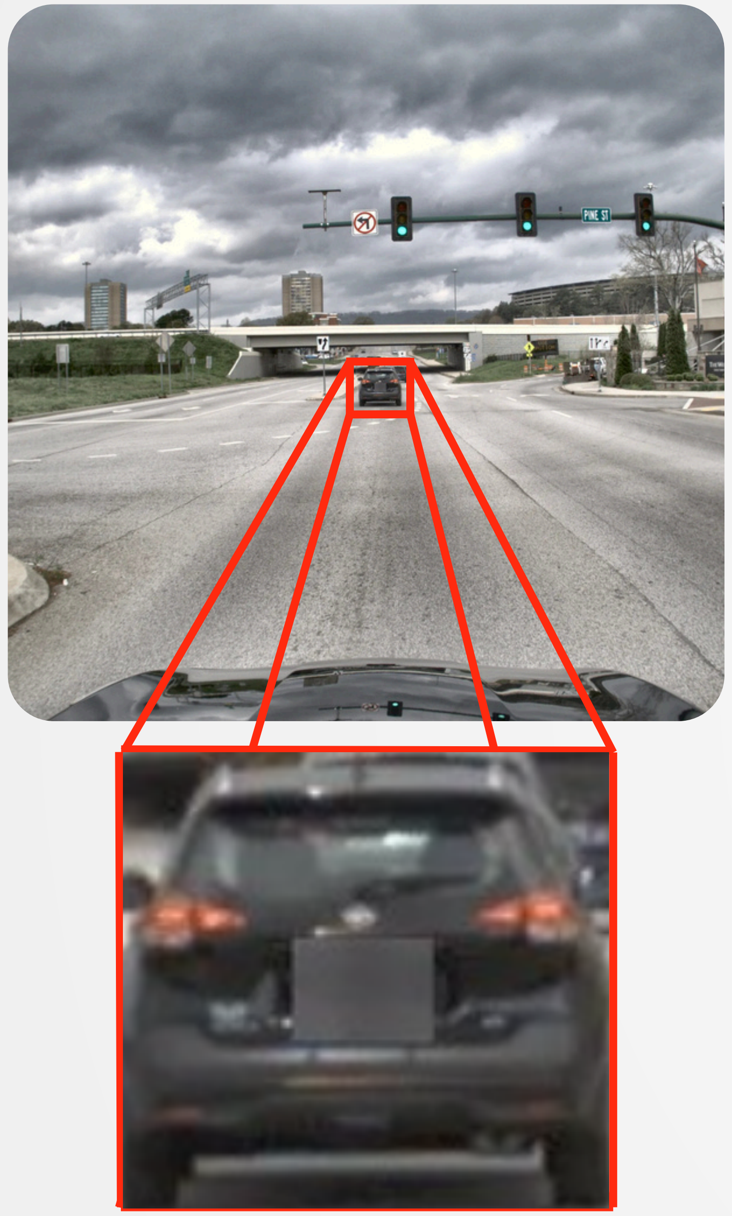
Bottom-Crop Heuristic

- Removes the bottom 20% of the frame to avoid hood reflections, occlusions, and irrelevant foreground clutter.

Model in Action: Model Selection

Model	Architecture	Remarks	Metrics
DepthPro + CNN	DepthPro Depth Map (Depth Values) + CNN (Car Patch) + Geometric Features → FC layers	Research baseline - trained on 300k+ industry-standard images; computationally expensive	MAE: 1.32 m, latency: 215 ms, throughput: 4.6 FPS
Hybrid 3-Branch (Our choice)	EfficientNetB4 (Full Image) +EfficientNetB3 (Car Patch) + Geometric Features → Fusion Head	Production-ready - balanced accuracy and speed; deployable	MAE: 1.41 m, latency: 104 ms, throughput: 9.7 FPS (~51.6% faster than baseline)
Student Model (Distilled)	MobileNetV3-Small (full image) + Custom 4-Layer CNN (Car Patch) + Geometric Features → Fusion Head	Edge-optimized - lightweight, real-time performance	MAE: 1.68 m, latency: 32 ms, throughput: 31+ FPS (real-time)

Model in Action: Distance Estimation



- Bounding box geometry (width, height, and aspect ratio)
- Position (vertical and horizontal) of lead vehicle

Full Image

Car Patch

Geometric Features

Final Ouput

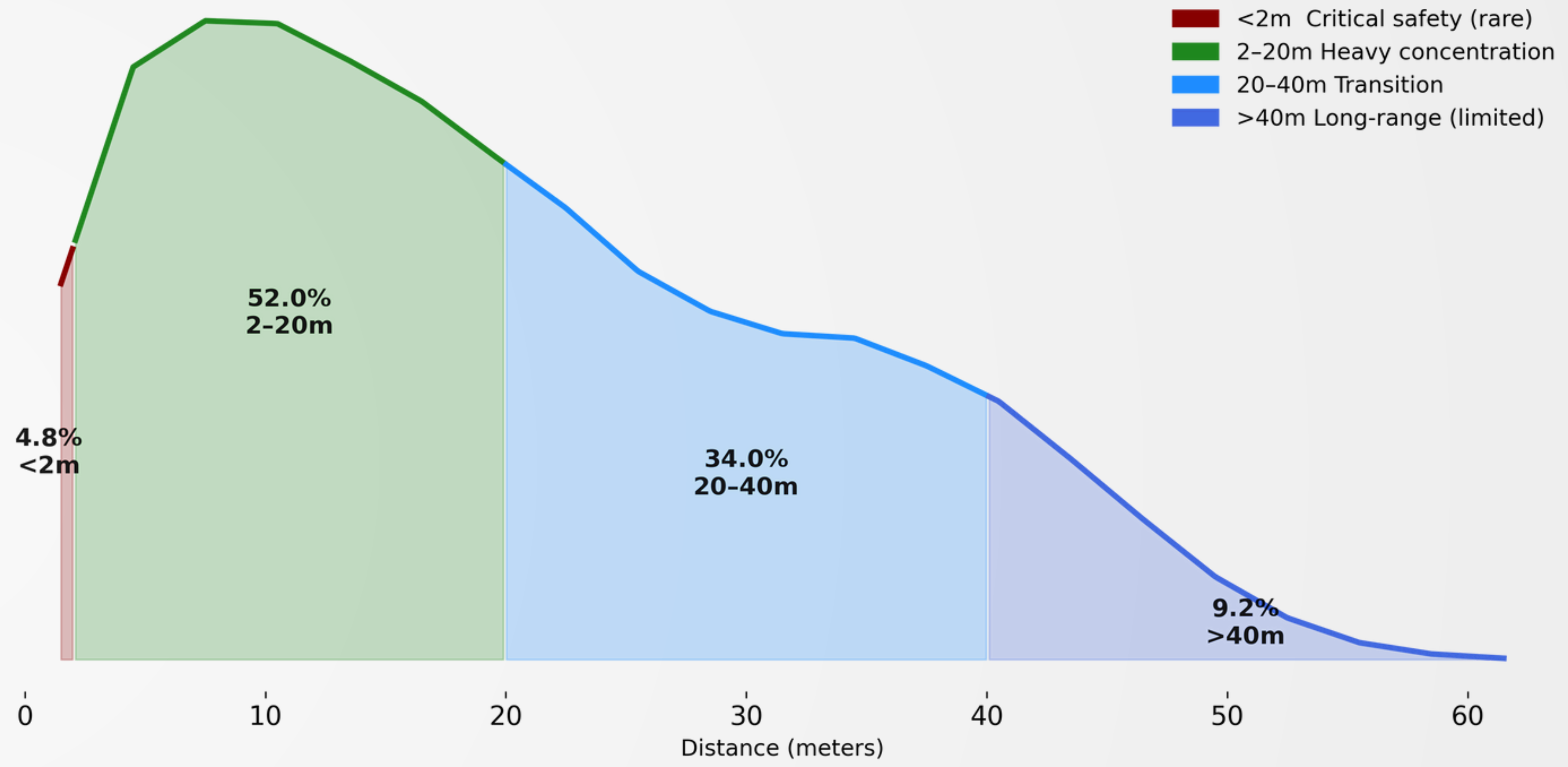
Client | Business Problem | Our Product | Model In Action | Data Assesment | Limitations | Future Scope

Understanding Our Data

Raw Dashcam Image

- Argonne National Lab dataset - PII filtered, proprietary
- 19,000+ images across LA, Nashville, Chattanooga
- Ground truth distances encoded in filenames

Dataset Distribution



- Heavy concentration in 0-30m range - reflects typical city driving patterns
- Critical safety range underrepresented - only 4.8% of samples under 2m
- Long-range samples (>40m) are limited

Improving Our Data

Method 1: Simulated data

CARLA simulator used to generate balanced distance samples
Synthetic-to-real domain gap too large - model failed to generalize from simulated to real dashcam images

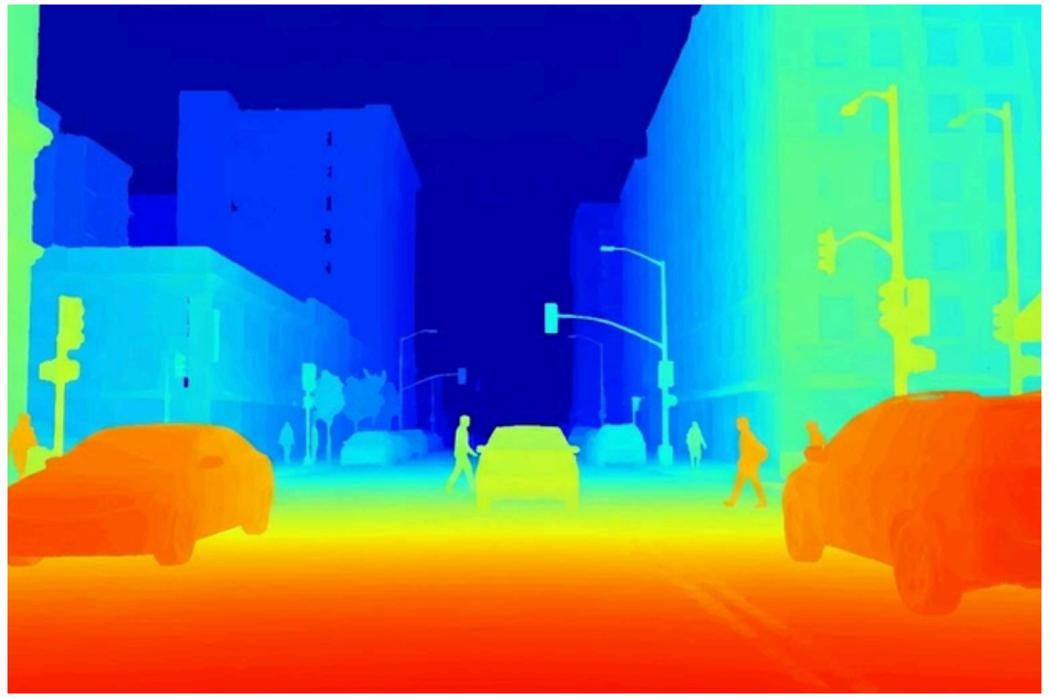
Method 2: Augmentation

Weather, lighting, and blur augmentations applied
Better Generalization compared to Synthetic CARLA Data

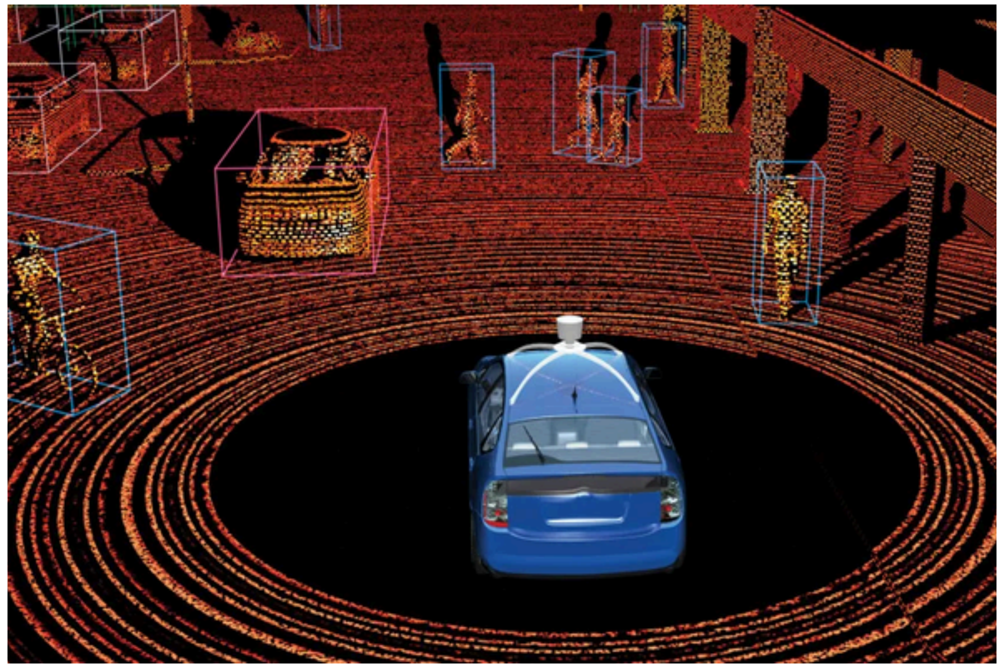
Monocular Vision Ambiguity

A single camera lacks true depth perception, making distance an estimation rather than a precise measurement (unlike Stereo Vision, RGBD camera, LiDAR)

CAMERA OUTPUT



RGB-D OUTPUT



LiDAR OUTPUT

Edge Cases

Sharp Turns

Extreme Light & Weather conditions

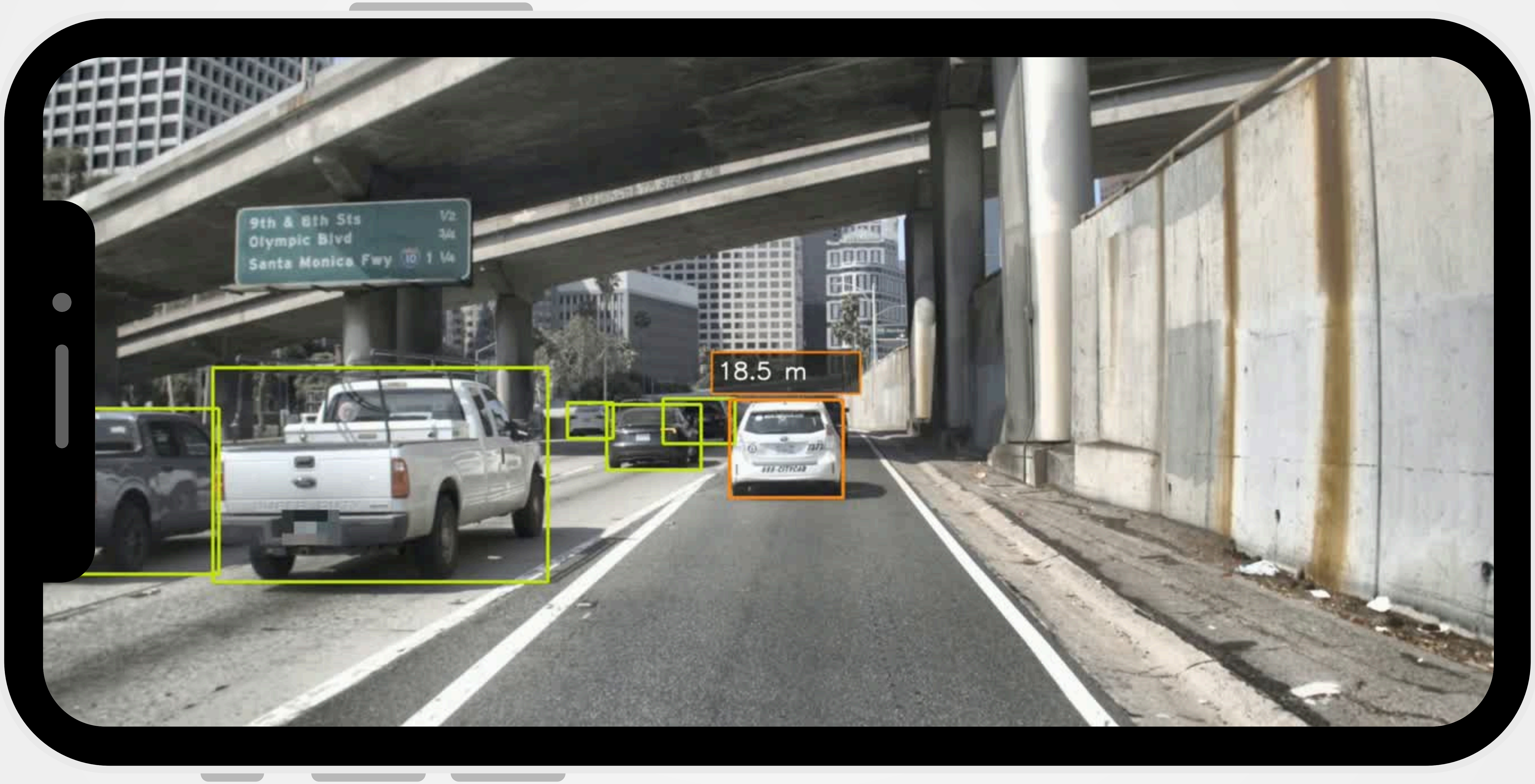
Future Improvements

Multi-modal Input (Vision + Vehicle Control Signals)
Fuse camera data with vehicle dynamics for robust perception.

Multi-frame Inference
Leverage sequential frames to improve distance estimation accuracy.

Uncertainty Quantification
Add an uncertainty head to provide confidence intervals for each distance prediction.

Intelligent Monocular Forward-Vision System



Client Statement

We are very thankful to the team for adding this capability to our research toolkit. By developing an algorithm that **replaces expensive sensors** like Radar and LiDAR with standard cameras, they have created a widely deployable solution that **drastically reduces hardware costs** and **accelerates our time-to-deployment**.

THE UNIVERSITY OF CHICAGO

Argonne NATIONAL LABORATORY

Thank you!