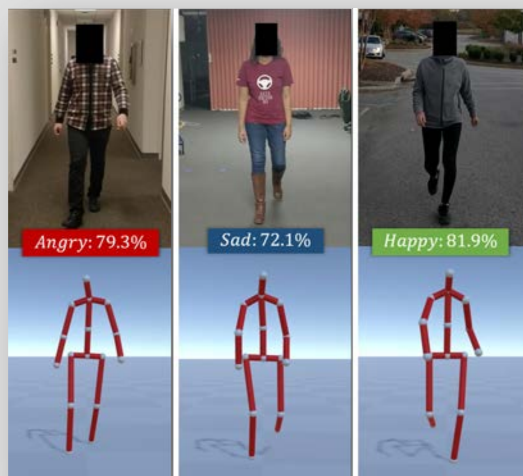
Prediction of Human Motion & Traffic Agents in Dense Environments

Dinesh Manocha

University of Maryland at College Park

dm@cs.umd.edu



Collaborators

- Aniket Bera (UNC/UMD)
- Uttaran Bhattacharya (UMD)
- Rohit Chandra (UMD)
- Ernest Cheung (UMD)
- Kurt Gray (UNC)
- Sujeong Kim (UNC/SRI)
- Yuexin Ma (UNC/Baidu/HKU)
- Tanmay Randhavane (UNC)

Real-World Pedestrian/Crowd Analysis

- Behavior Learning
- Culture characteristics
- Crowd prediction

Analyze Crowd Movements

Driverless Cars: Pedestrian Interaction

Source: **Oxbotica** at Oxford University

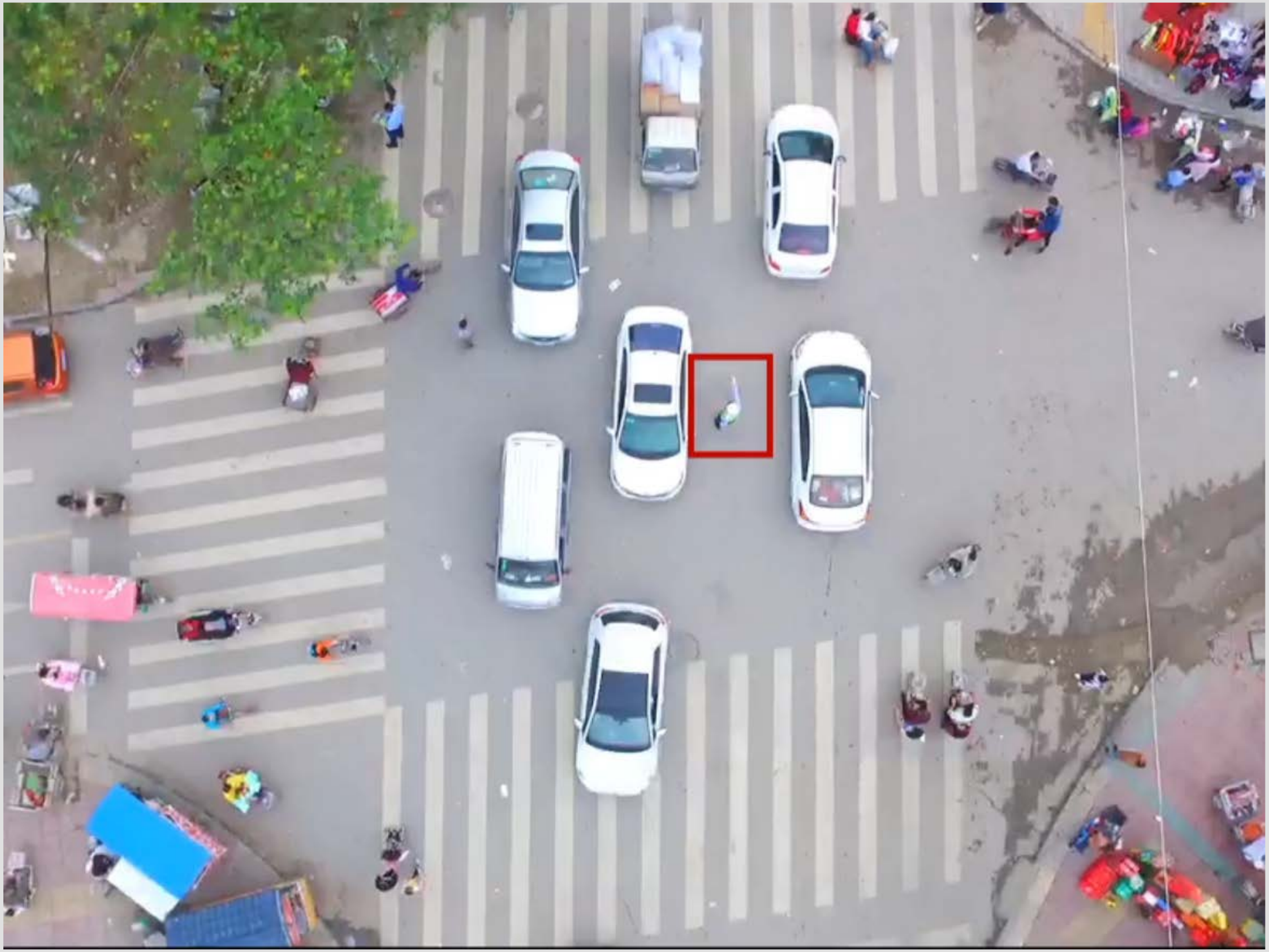
Current AD technology vs. Real-world Scenarios

- Many traffic situations are still too challenging for autonomous vehicles

Current Autonomous Driving

Urban Traffic Condition: China

Challenging Traffic Conditions: China



Current technologies and datasets for dense traffic are limited

More Challenging Conditions: India

No respect for rules; cultural norms, driver & pedestrian behaviors

Organization

- Pedestrian and Crowd Motions
- Heterogeneous multi-agent simulation
- Tracking urban traffic & Prediction
- Driver behavior modeling

Organization

- Pedestrian and Crowd Motions
- Heterogeneous multi-agent simulation
- Tracking urban traffic & Prediction
- Driver behavior modeling

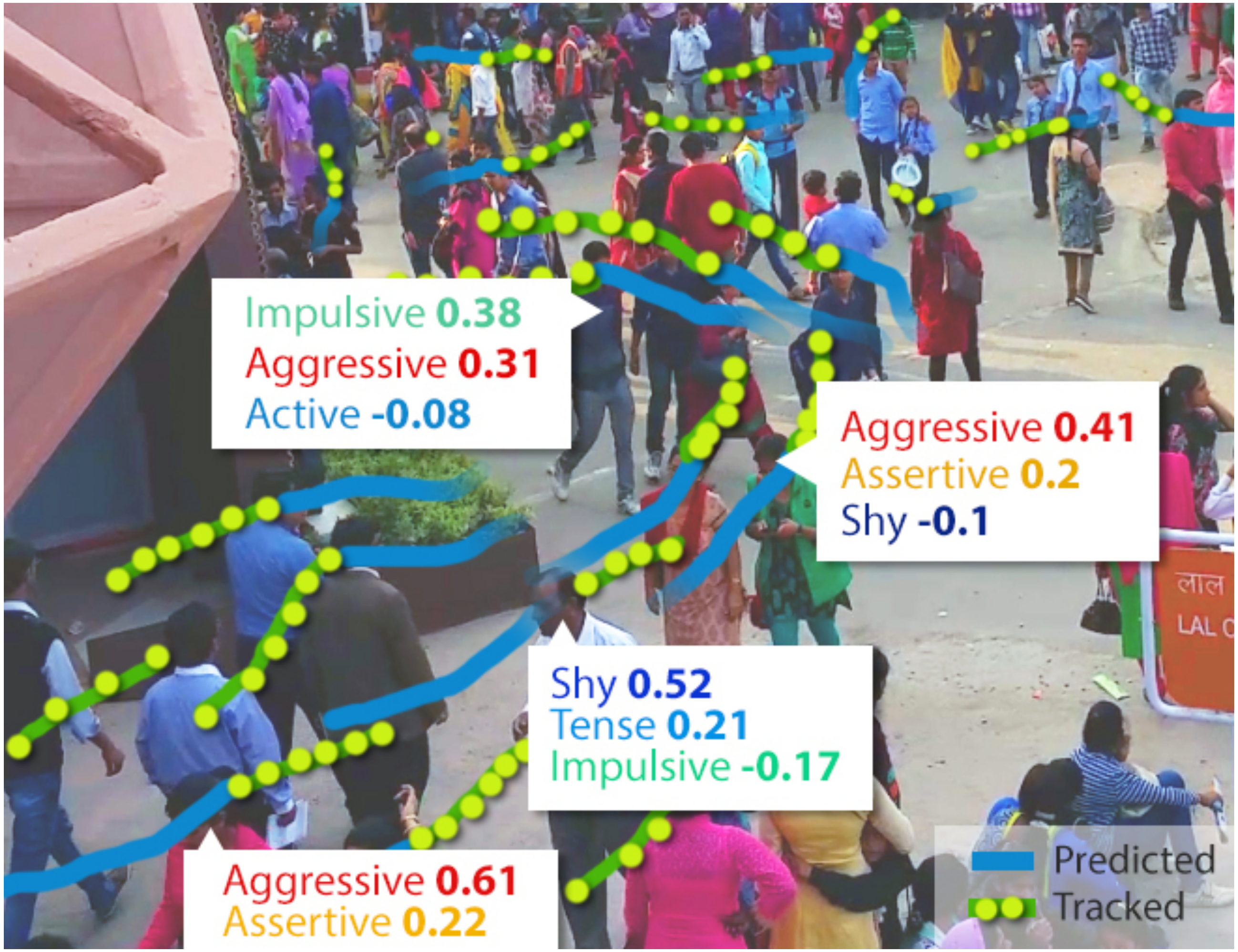
Pedestrian and Crowd Motion: Tracking & Prediction

- New motion models based on RVO (reciprocal velocity obstacles)
- Combine motion model with behavior models
- Real time tracking: deep learning + motion models
- Learning Pedestrian Dynamics using Bayesian Inferences
- Handling Dense Crowds

Realtime Pedestrian Tracking in Dense Crowds

[Chandra et al. 2019]

REALTIME TRACKING AND PERSONALITY MODELING



[Bera et al. 2017]

Pedestrian/Crowd Movement Prediction

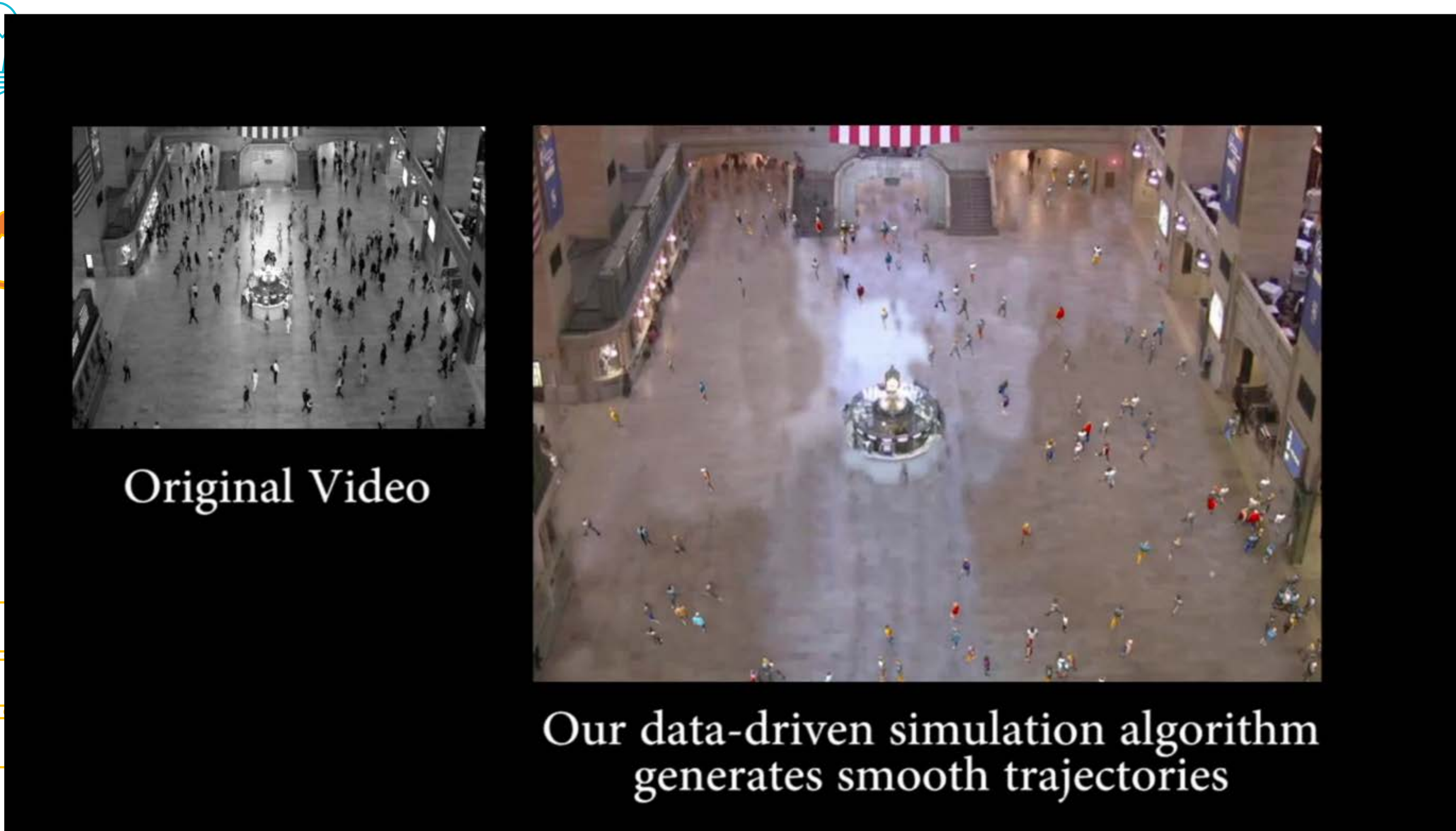
[Bera et al. 2016, 2017]

CROWD BEHAVIOR MOVEMENT PREDICTION

2017 PRESIDENTIAL INAUGURATION CROWDS

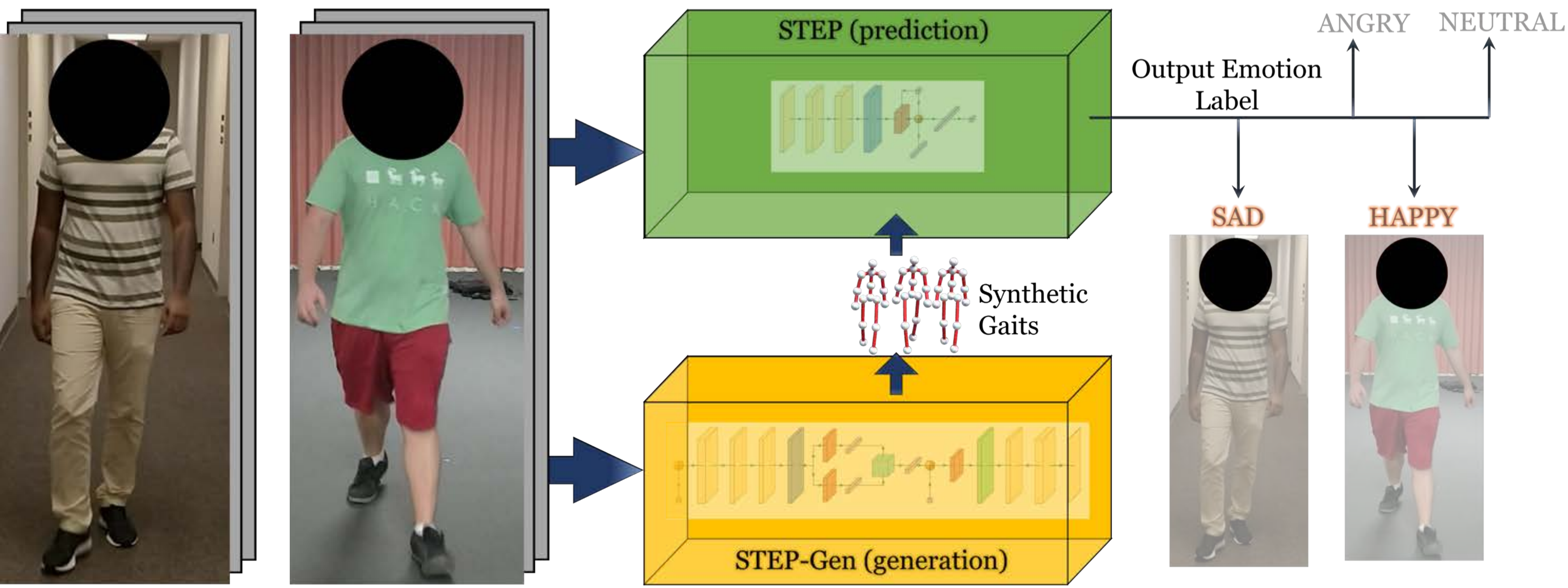
[Bera et al. 2017]

Data Driven Crowd Simulation & Prediction



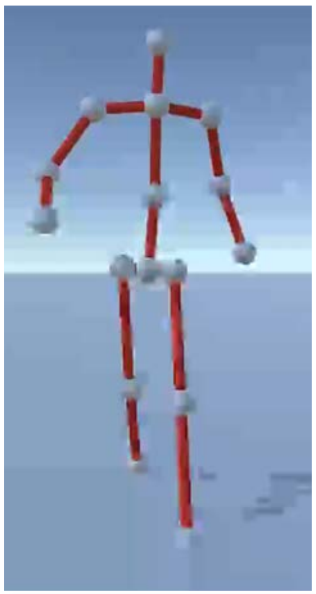
[Kim et al. 2017

STEP: Spatial Temporal Graph Convolutional Networks for Emotion Perception from Gaits

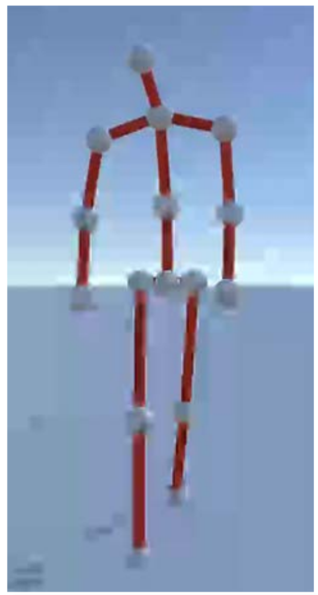


Input Walking Videos

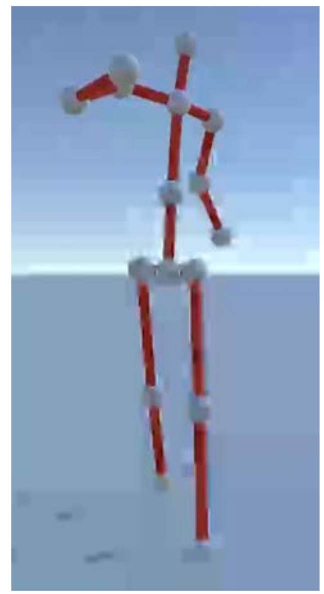
Can you Classify Emotions from Movements?



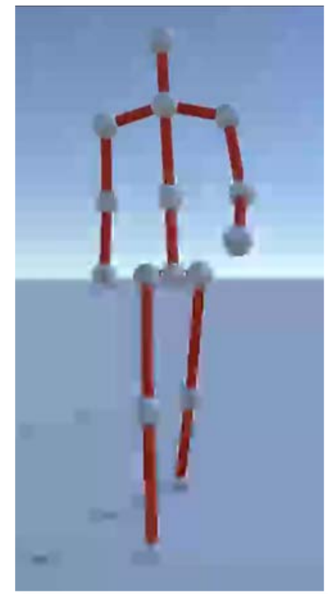
$Arousal = 0.68$
$Valence = -0.05$



$Valence = -0.67$
$Arousal = -0.25$



$Valence = 0.53$
$Arousal = -0.11$



$Valence = 0.10$
$Arousal = -0.07$

Which is angry, sad, happy, neutral?

Emotion Classification from Pedestrian Motion

[Randhavane et al. 2019]

$Arousal = 0.68$
$Valence = -0.05$

$Valence = -0.67$
$Arousal = -0.25$

$Valence = 0.53$
$Arousal = -0.11$

$Valence = 0.10$
$Arousal = -0.07$

Which is angry, sad, happy, neutral?
Forthcoming E-Walk Dataset

Organization

- Pedestrian and Crowd Motions
- Heterogeneous multi-agent simulation
- Tracking urban traffic & Prediction
- Driver behavior modeling

Dense traffic scenarios: Heterogeneous Agents

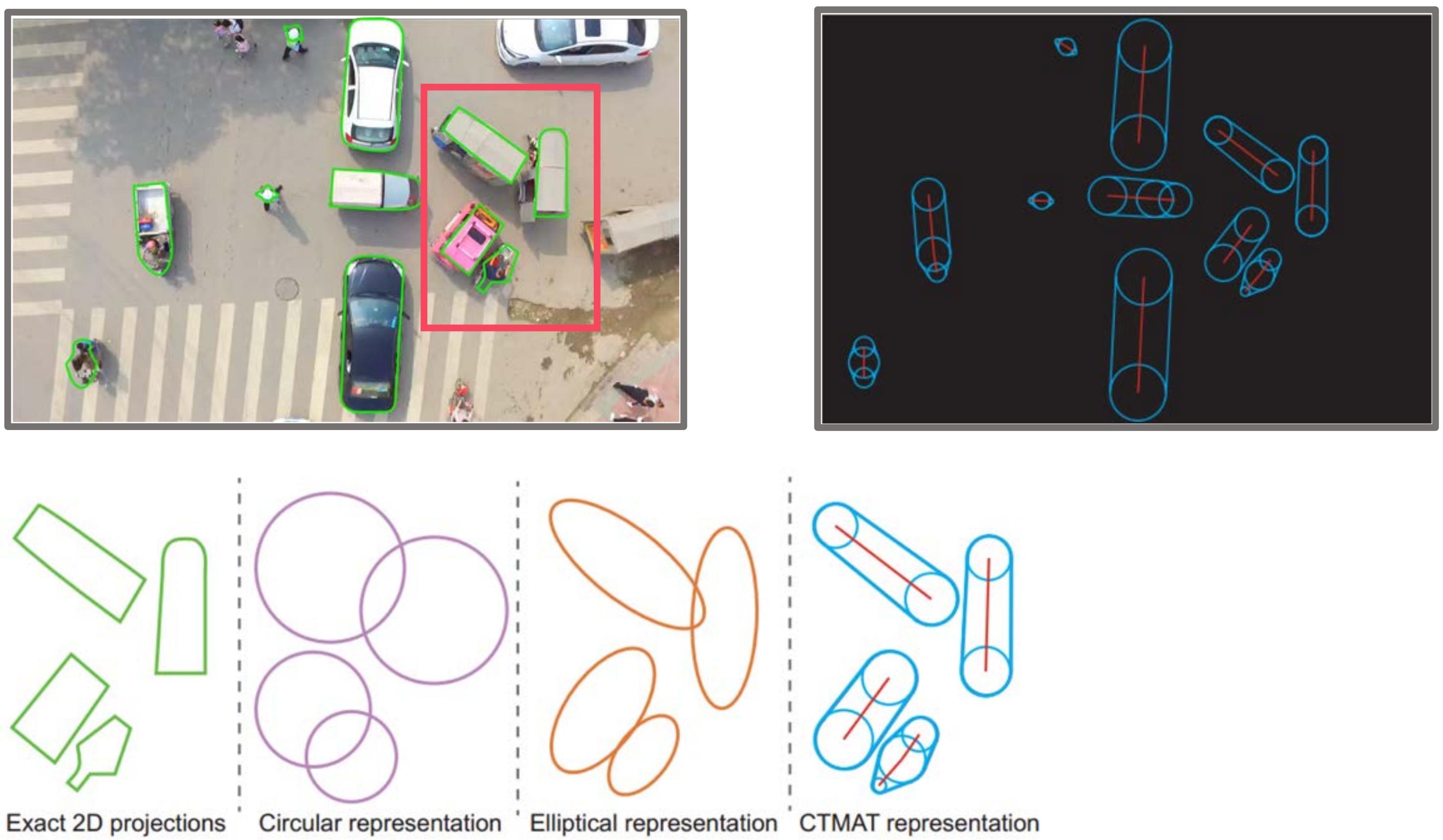
Vehicles (big and small), pedestrians, bicycles, tricycles, etc

Heterogeneous Multi-Agent Navigation

Agents:

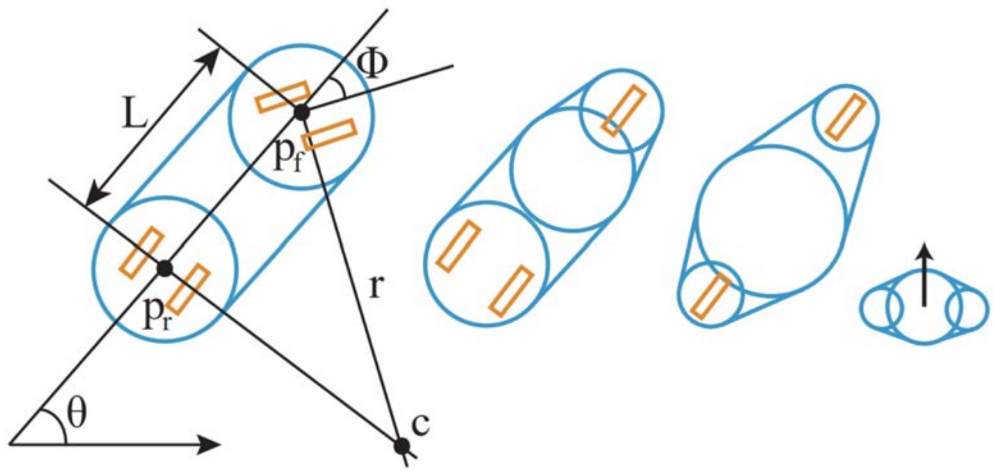
- Varying shapes
- Varying dynamics
- Different behaviors
- Operating in tight spaces

Heterogeneous Multi-Agent Representation



Ma et al. "Efficient reciprocal collision avoidance between heterogeneous agents using CTMAT." , AAMAS 2018

Kinematic Models: Different Agents

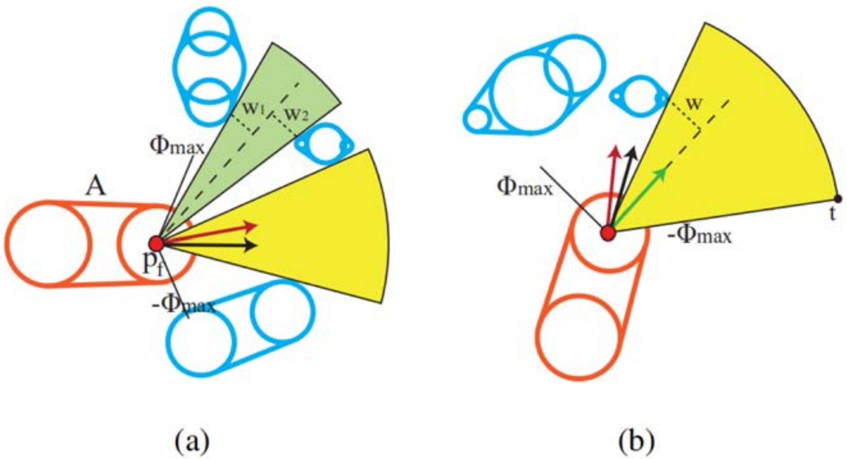


Car tricycle bicycle pedestrian

$$\dot{\vec{p}} = (v\cos(\theta), v\sin(\theta)), \qquad \dot{\theta} = \frac{\tan(\phi)}{L} v.$$

Simple Car Model [Laumond et al. 1998]

AutoRVO: Preferred Steering Computation

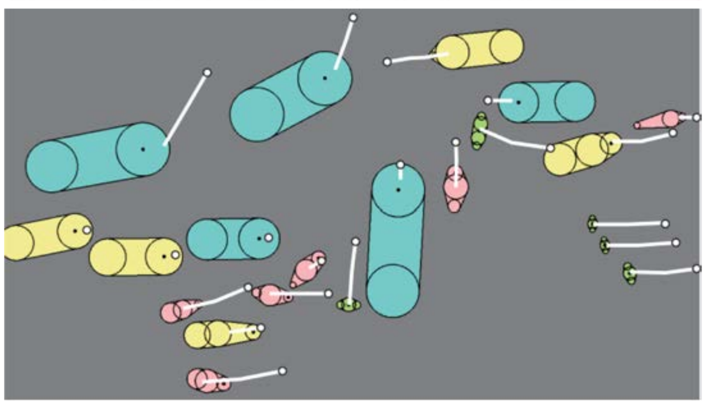


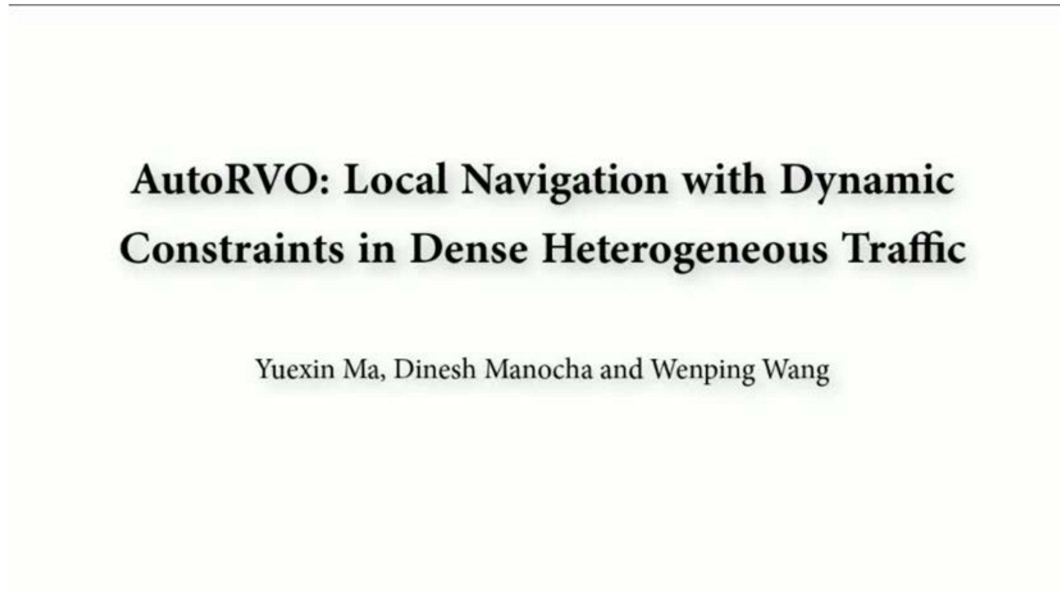
$$\phi^o = f(\vec{a}, \vec{d}^o)$$

$$\phi^o \in (-\phi_{max}, \phi_{max})$$

Search for free-space for collision-free local navigation

AutoRVO: Results





Comparisons: Multi-Agent Navigation

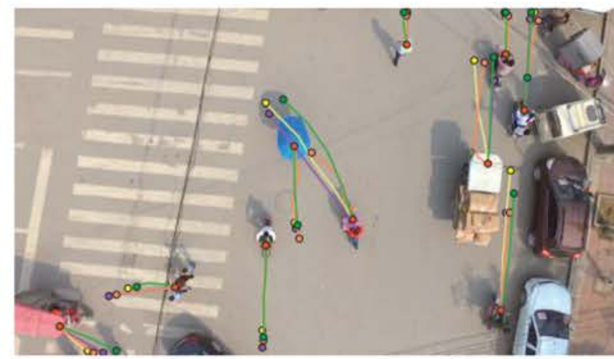
(a) traffic-1

(b) traffic-2

(c) traffic-3

(d) traffic-4

(e) traffic-5



(f) traffic-6

Figure 6: Comparison of real trajectories of 50 continuous frames and simulated trajectories. (a)-(c) are three different moments from one video. (d)-(f) are three different moments from three different videos. Green lines indicate the real trajectories extracted from videos captured using a drone. Trajectories generated by AutoRVO, ORCA with CTMAT representation, and ORCA with disk representation are drawn in yellow, purple, and orange respectively. We observe higher accuracy with AutoRVO. Red points represent beginning reference positions.

Organization

- Pedestrian and Crowd Motions
- Heterogeneous multi-agent simulation
- Tracking urban traffic & Prediction
- Driver behavior modeling

Highlights

- High degree of heterogeneity
- Dense Traffic
- No traffic protocols in place

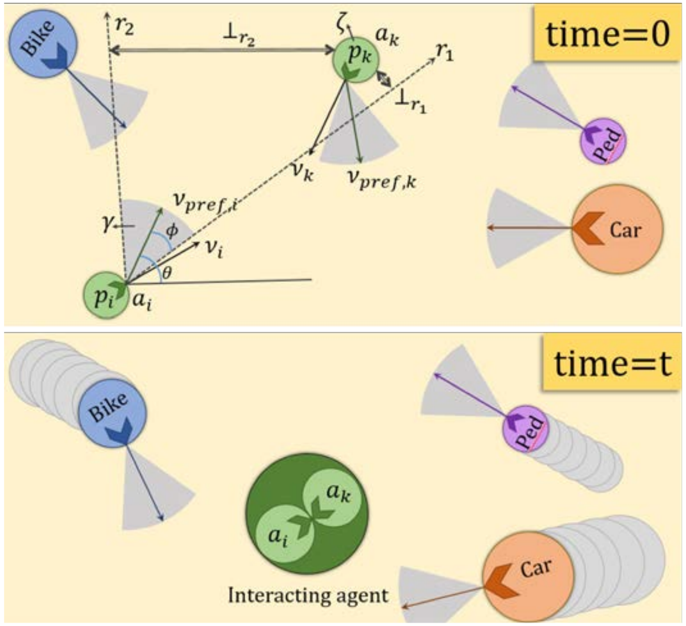
Approach
Combine model-based and learning-based methods

Consists of 4 stages

Stage 1 – Agent Detection Using Mask R-CNN

Use Mask R-CNN based agent detection to generate Segmented Boxes of each agent

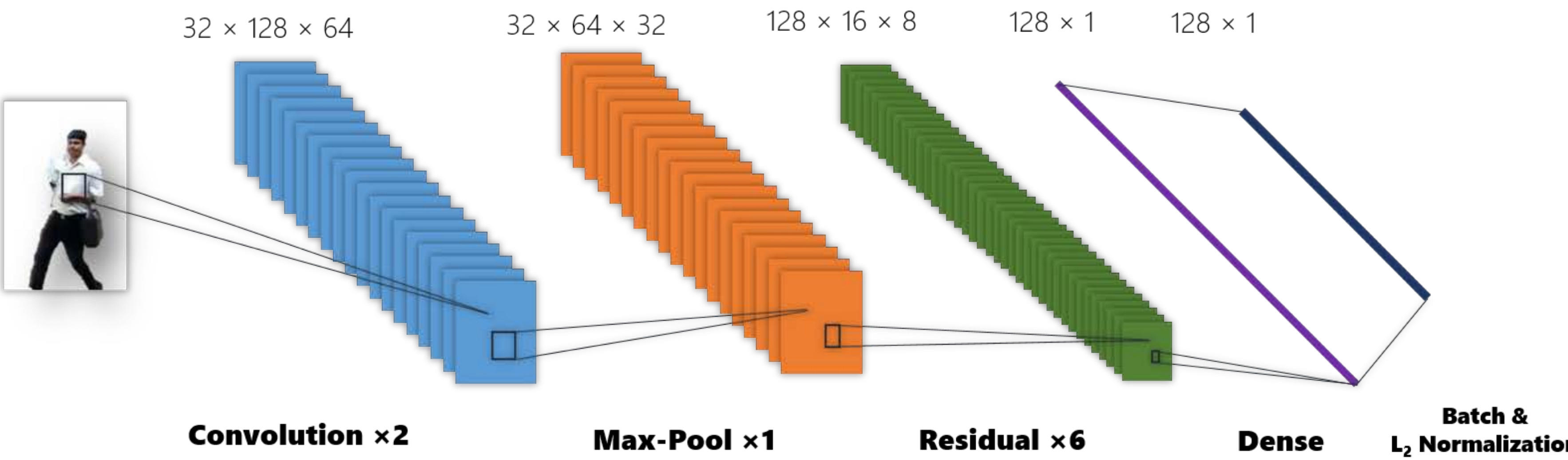
Stage 2 – Velocity Prediction Using HTMI



Use HTMI to model inter-agent interactions and collision avoidance

HTMI: Heterogeneous motion model

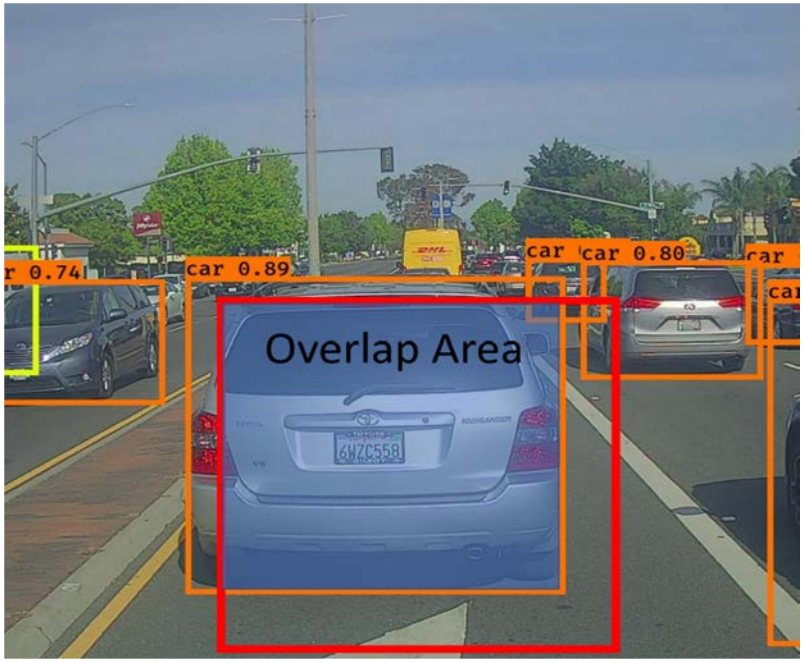
Stage 3 – Feature Extraction Using Segmented Boxes



Generate novel features called “**Deep TA-features**” from segmented boxes

Stage 4 – Feature Matching Using IOU Overlap

We use the cosine distance metric



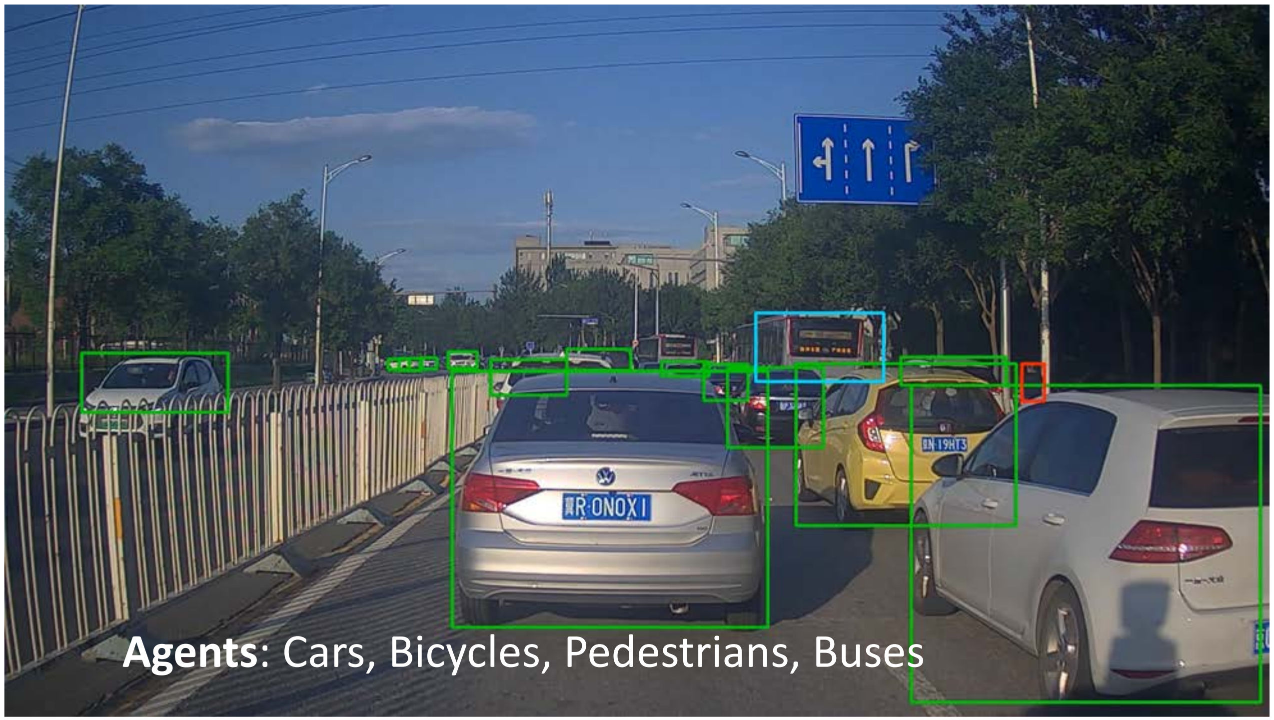
Results – Low Density Traffic

Car: Green

Pedestrians & Two-Wheelers: Red

Rickshaws: Purple

Results – Medium Density



Car: Green

Pedestrians & Two-Wheelers: Red

Buses: Cyan

Results – High Density

Car: Green

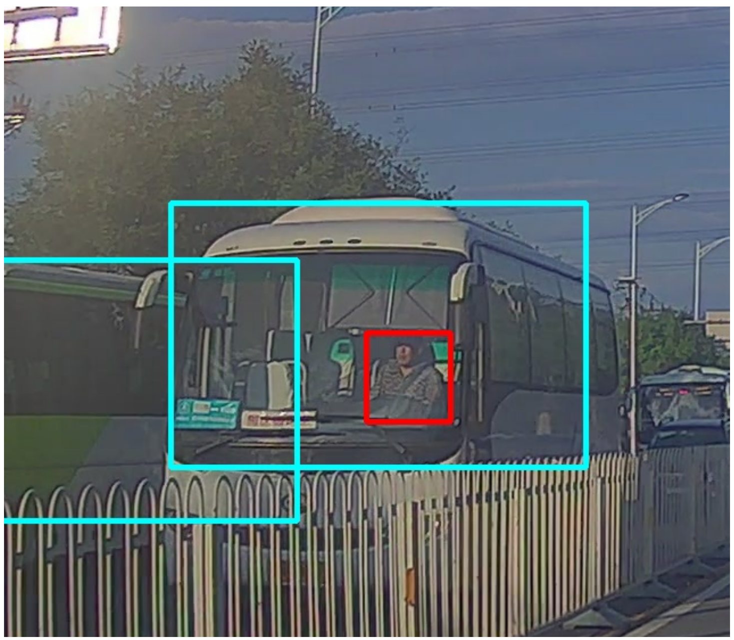
Rickshaws: Purple

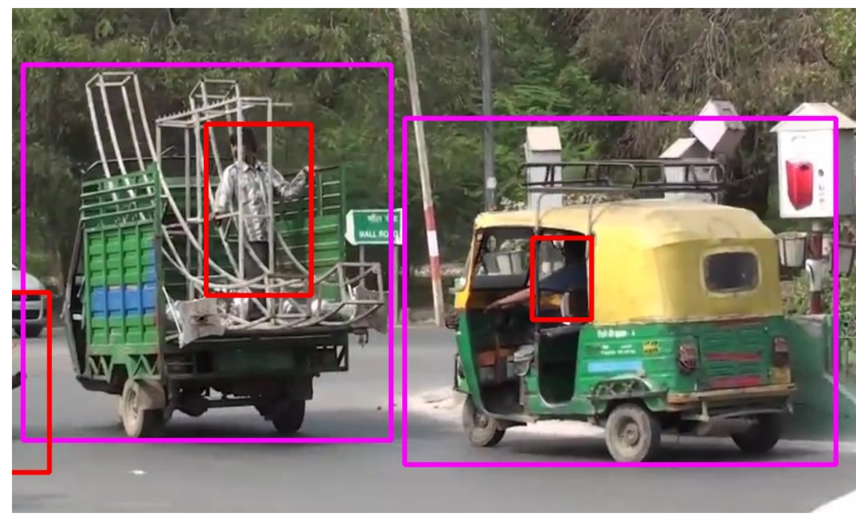
Pedestrians & Two-Wheelers: Red

Buses: Cyan

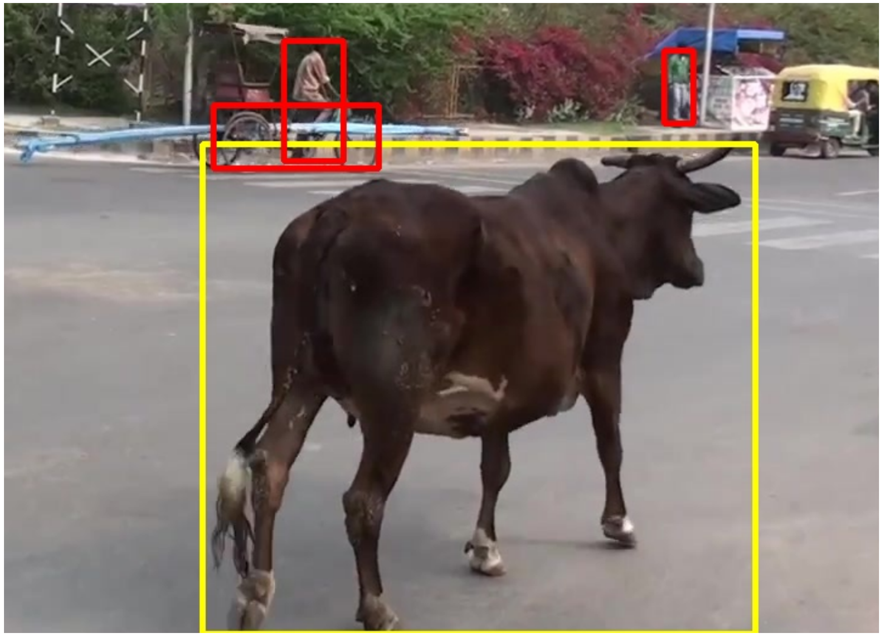
Animals: Yellow

Strengths – I
We can track drivers inside different road agents





Strengths – II
We can track atypical agents



Strengths – III
We can track agents in challenging conditions

- Night time with a jittery, moving camera with low resolution. There is heavy glare from oncoming traffic.

Dense Traffic Dataset

We introduce a novel dataset of 45 high resolution videos consisting of dense, heterogeneous traffic.

We have carefully annotated the dataset following a strict protocol.

The videos are categorized by camera motion, camera viewpoint, time of the day, and difficulty level.

Where to put your money in 2019 — it's not US stocks, according to Morgan Stanley (Emerging Economies)

https://www.cnbc.com/2018/11/26/stock-picks-morgan-stanley-upgrades-emerging-markets-downgrades-us.html

Where to put your money in 2019 — it's not US stocks, according to Morgan Stanley (Emerging Economies)

Traffic Prediction

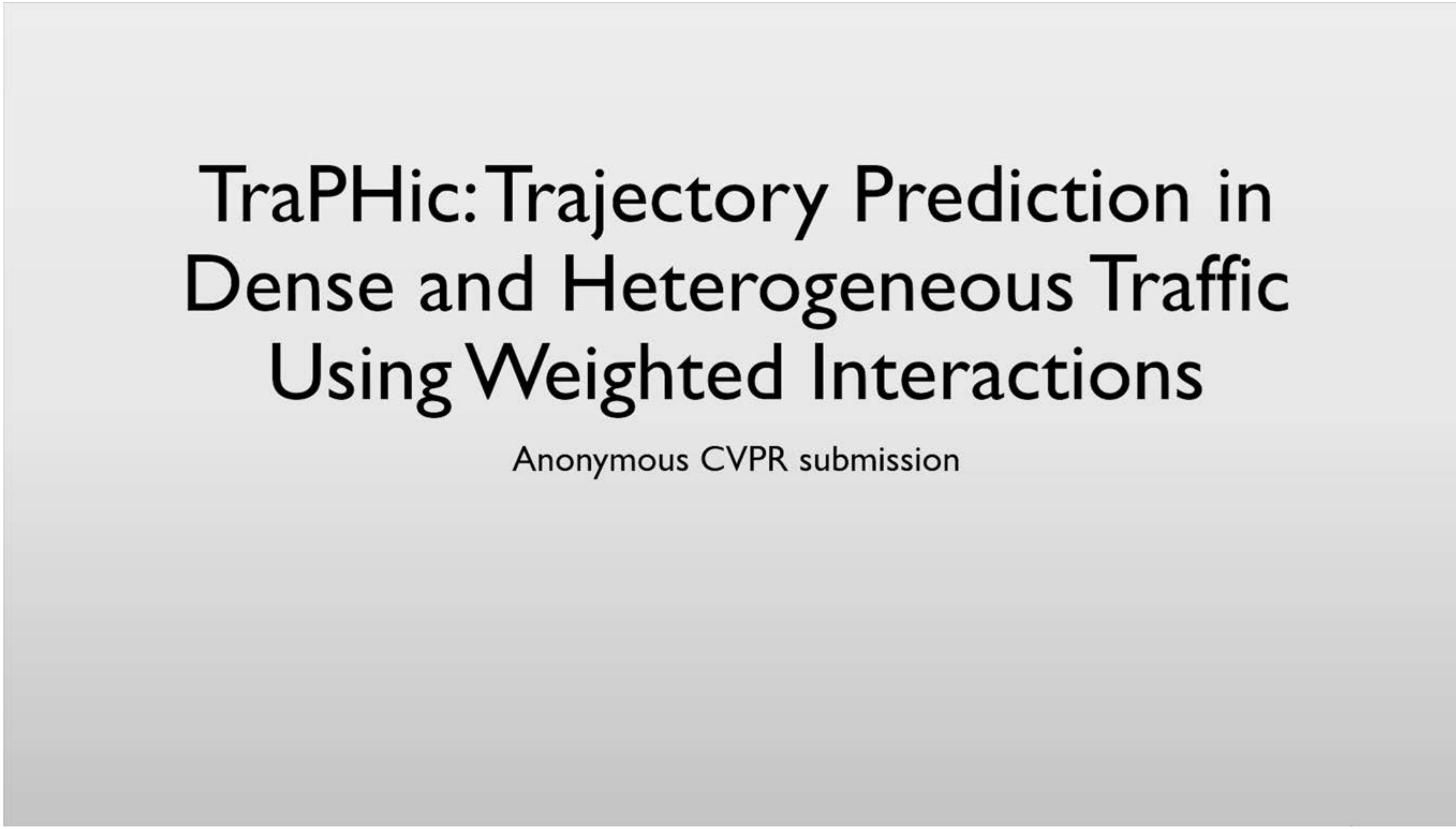
Dense

- Many agents (>3000) per Km of road length.

Heterogeneous

- Different types of road agents present simultaneously, e.g., pedestrians, two-wheelers, three-wheelers, cars, buses, trucks etc.

Traffic Prediction



Combining multi-agent navigation, deep learning & dynamics

Organization

- Pedestrian and Crowd Motions
- Heterogeneous multi-agent simulation
- Tracking urban traffic & Prediction
- Driver behavior modeling

Modeling Driver Behaviors

- Most traffic accident happens are caused by dangerous reckless drivers
- “Aggressiveness” and “Reckless” are subjective metrics
- Need navigation algorithms that can extract driving behaviors from sensors/trajectories and perform safe navigation (Behavior-based Navigation)

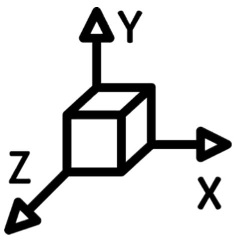
Aggressive Driving Behaviors

If you are driving, which driver will you pay attention to?

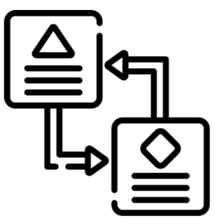
Identifying Driving behavior allows autonomous systems to:

- Pay extra "attention"
- Avoid getting close to them
- Re-run perception algorithms at higher resolution for those area

Main contributions



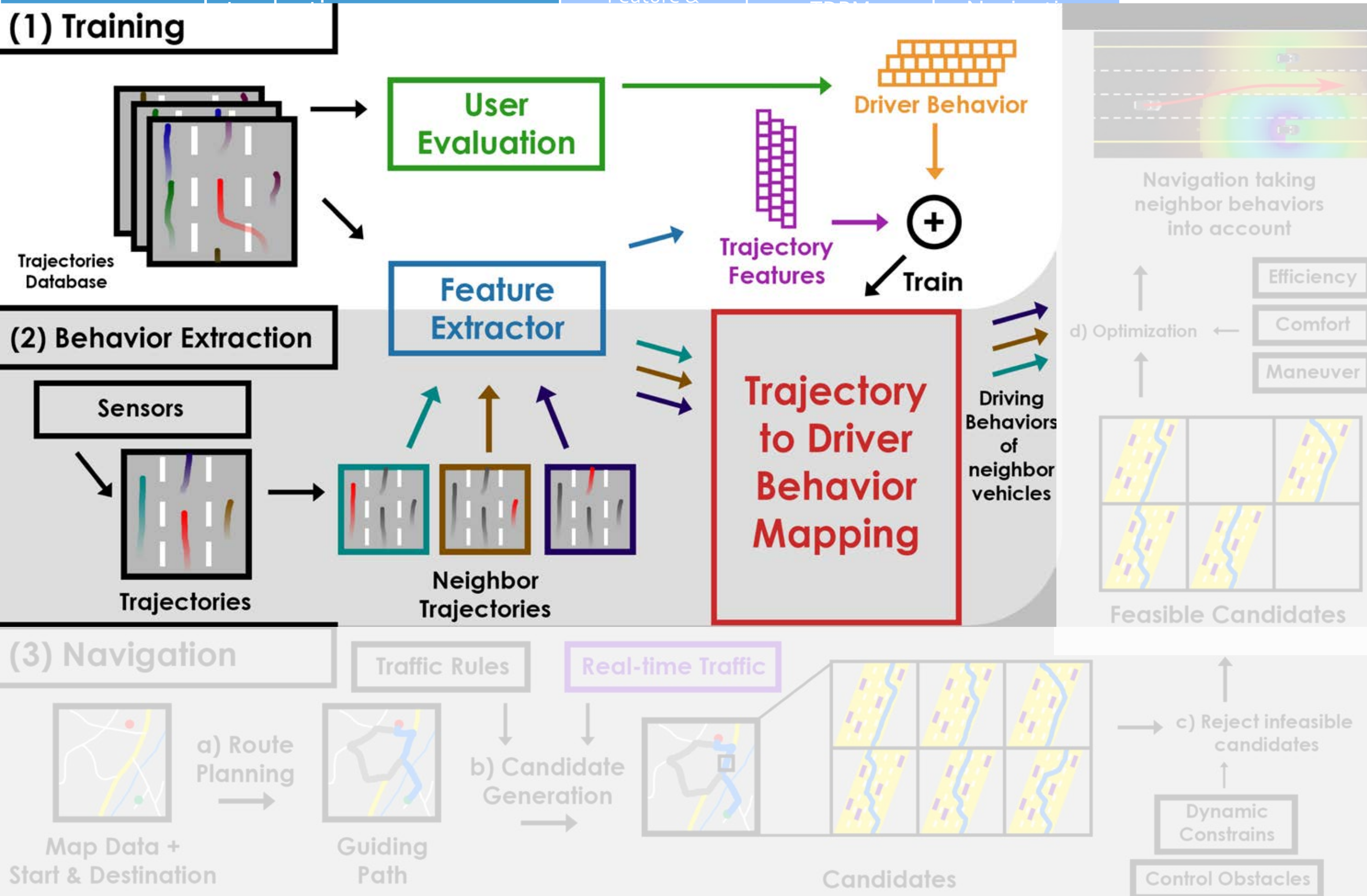
Feature extraction from trajectories in real-time



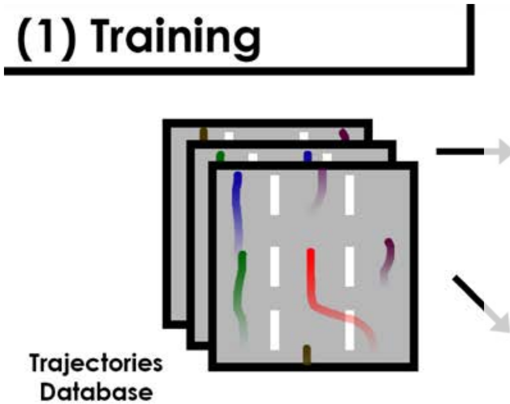
Trajectory to driver behavior mapping

Improved real-time navigation; Integrated with Autonovi-Sim

[Cheung et al. 2018, CVPR; Cheung et al. 2018 IROS]


(1) Training
Trajectories Database
User Evaluation
Driver Behavior
Trajectory Features
Train
Feature Extractor
(2) Behavior Extraction
Sensors
Trajectories
Neighbor Trajectories
Trajectory to Driver Behavior Mapping
Driving Behaviors of neighbor vehicles
Navigation taking neighbor behaviors into account
Efficiency
Comfort
Maneuver
d) Optimization
Feasible Candidates
(3) Navigation
Traffic Rules
Real-time Traffic
a) Route Planning
b) Candidate Generation
c) Reject infeasible candidates
Map Data + Start & Destination
Guiding Path
Candidates
Dynamic Constrains
Control Obstacles

Trajectory Database



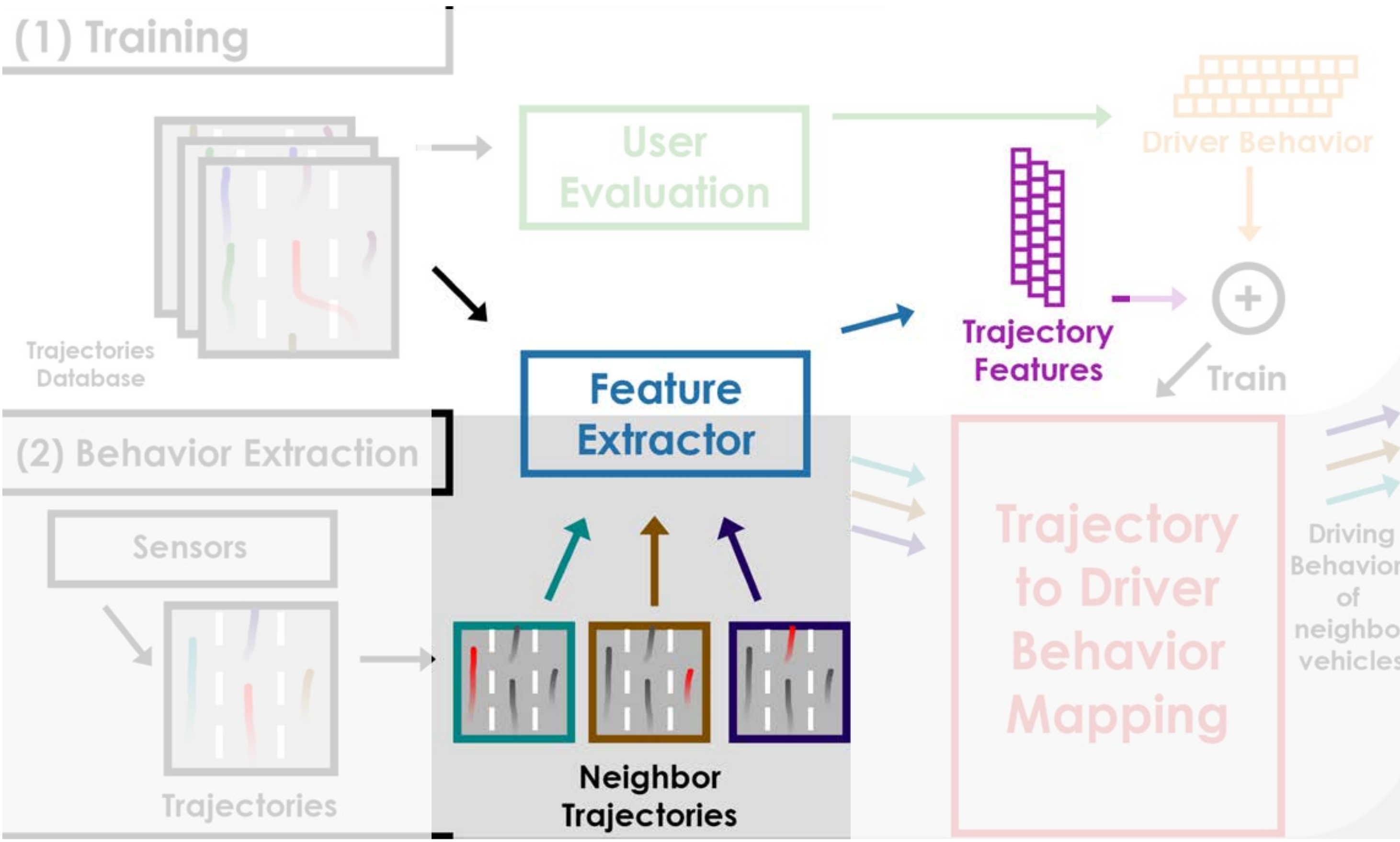
Interstate highway-80 dataset

- 45 minutes of trajectory
- 1650 feet section of highway
- captured with 7 cameras
- tracked automatically with manual verification

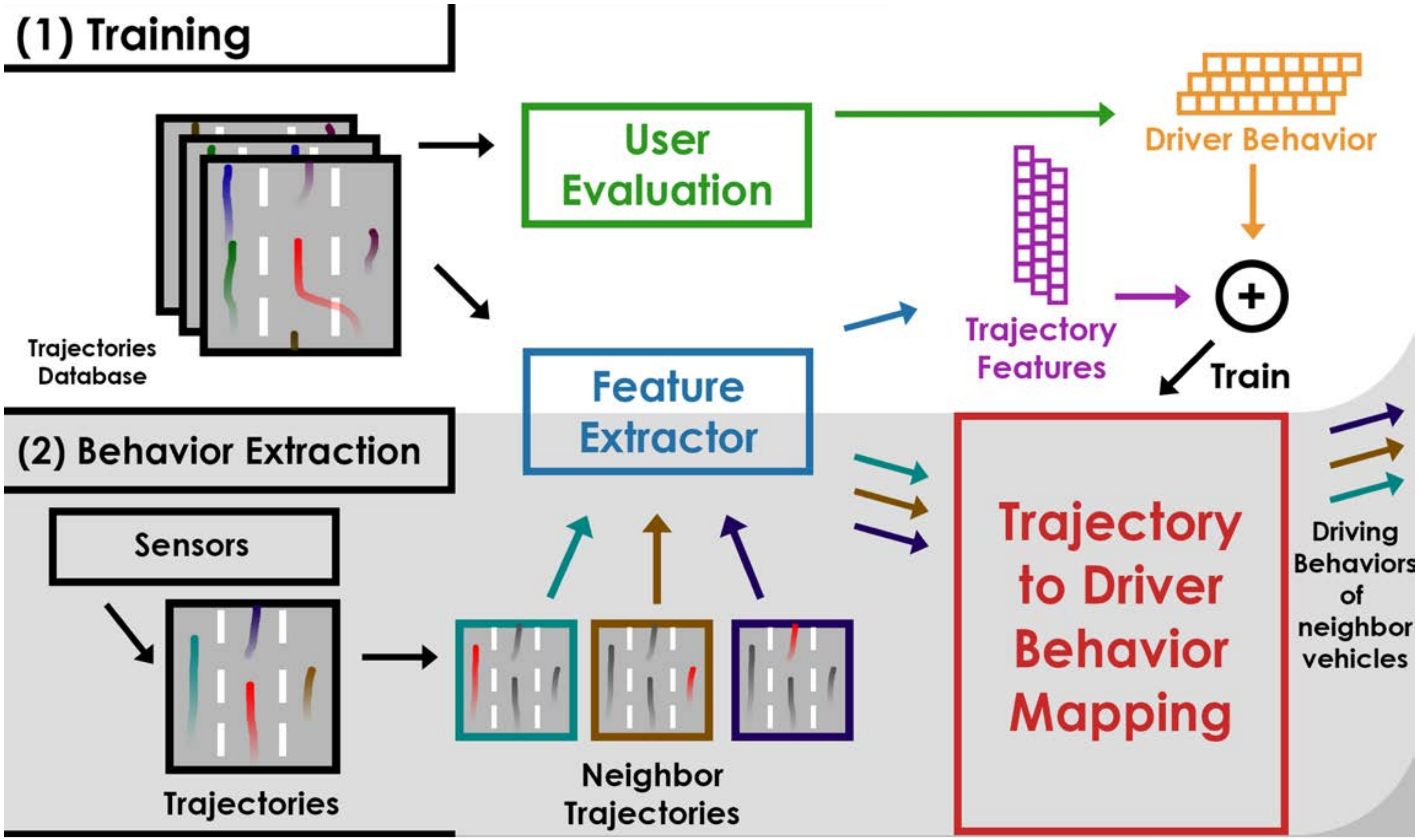
User Evaluation

Symbol	Description	Symbol	Level of Attention when
b_0	Aggressive	b_6	following the target
b_1	Reckless	b_7	preceding the target
b_2	Threatening	b_8	driving next to the target
b_3	Careful	b_9	far from the target
b_4	Cautious		
b_5	Timid		

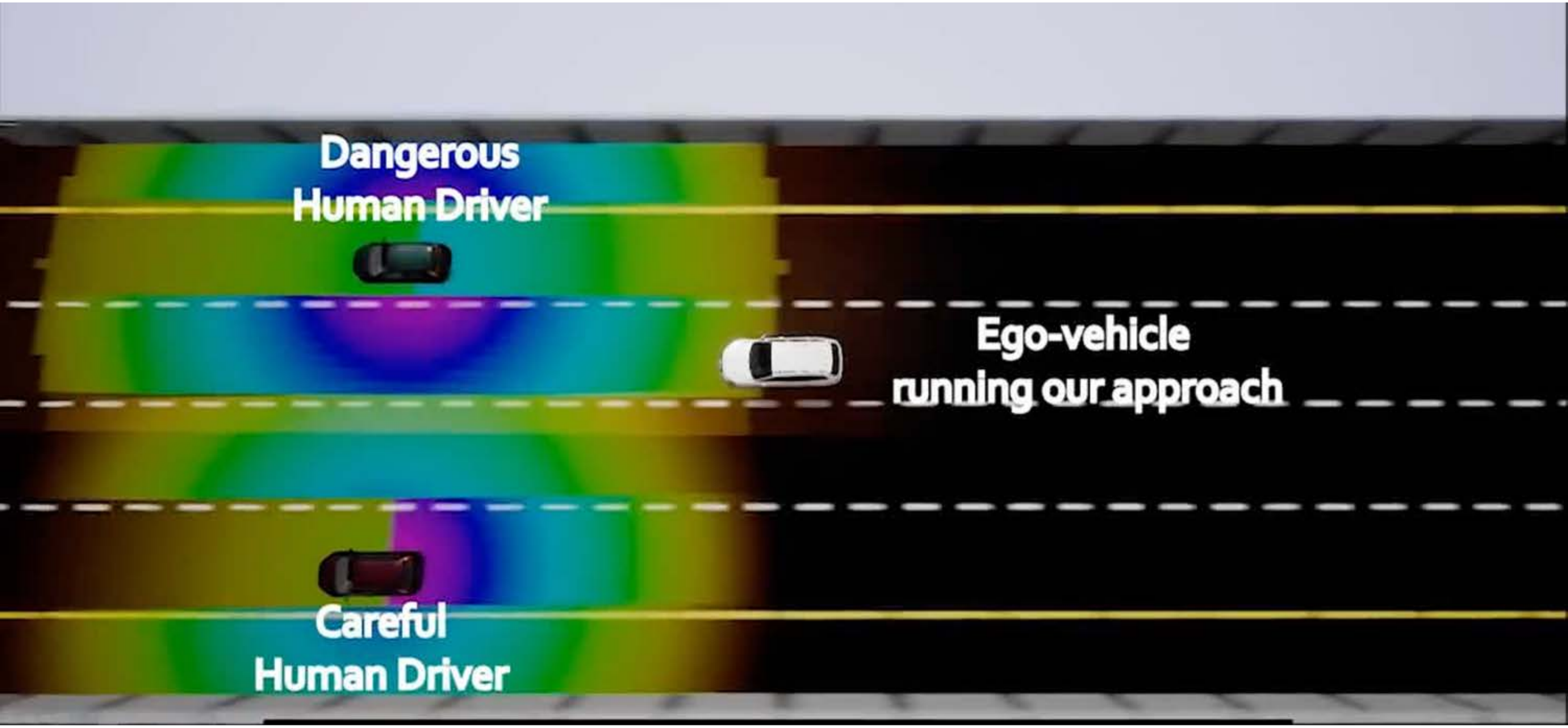
Feature Extraction



Trajectory to Driver Behavior Mapping

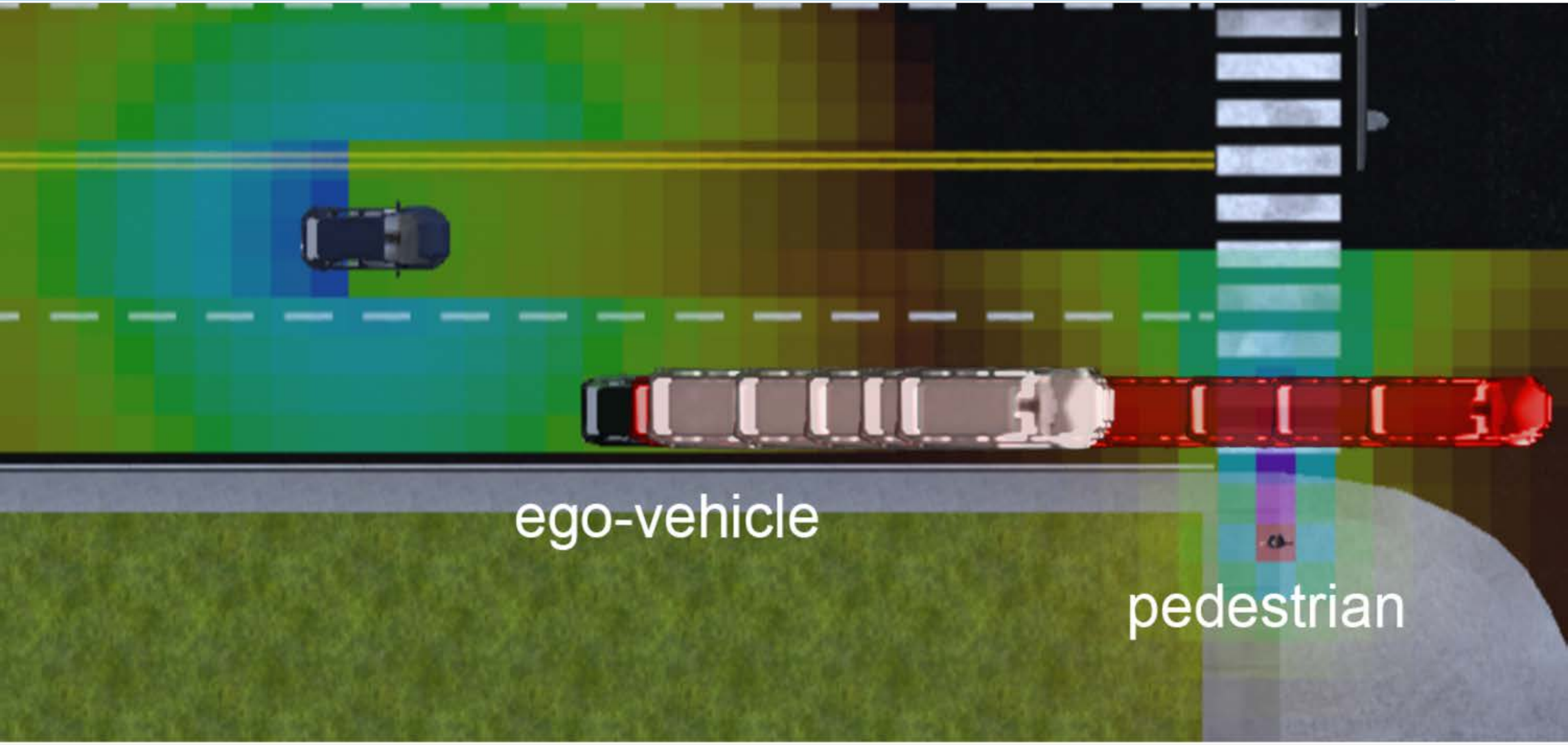


Navigation improvements



Navigation improvements

Navigation improvements



Conclusions

- Behavior modeling using feature extraction
- Applied to highway traffic data
- Safe and improved navigation

[Cheung et al. 2018, CVPR; Cheung et al. 2018 IROS]

Crowd and Traffic Motion

- New algorithms for tracking pedestrians & traffic agents
- Handle dense scenarios
- Use models from social psychology for behavior modeling
- Combine model-based and learning-based methods
- Applications to crowd scene analysis and autonomous driving

Acknowledgements

- Army Research Office
- Baidu
- DARPA
- Intel
- National Science Foundation

Questions: dm@cs.umd.edu