Autonomy and AI

Dinesh Manocha

University of Maryland at College Park

dm@cs.umd.edu

https://www.cs.umd.edu/people/dmanocha

AI: Recent Developments

- Driven by advances in machine learning: (AI 2.0)
  - Deep learning; Big data; Commodity hardware (e.g GPUs, AI processors)
- Strong impact in image recognition, speech understanding, automated translation
- Rapid expansion in other fields
  - Natural sciences, social sciences, engineering, medicine
- Strong investments in industry
  - Google, Facebook, Microsoft, Apple, Samsung, Intel, Adobe
- International expansion (Canada, Europe, China)
- Strong educational demand: undergraduate and graduate

AI at University of Maryland

- Integral part of “Computing at UMD” since 1960s
- Strong research groups in computer vision, natural language processing, planning an game theory
- Recent growth in machine learning, robotics, data science
- Strong interest all over the campus
  - STEM fields (natural sciences, engineering)
  - Non-STEM fields (social sciences, humanities, business)

AI at University of Maryland

- Computer vision
- Neuro and swarm computing
- Robotics
- Machine learning
- Natural language processing

AI at University of Maryland

- Computer vision
- Neuro and swarm computing
- Robotics
- Machine learning
- Natural language processing

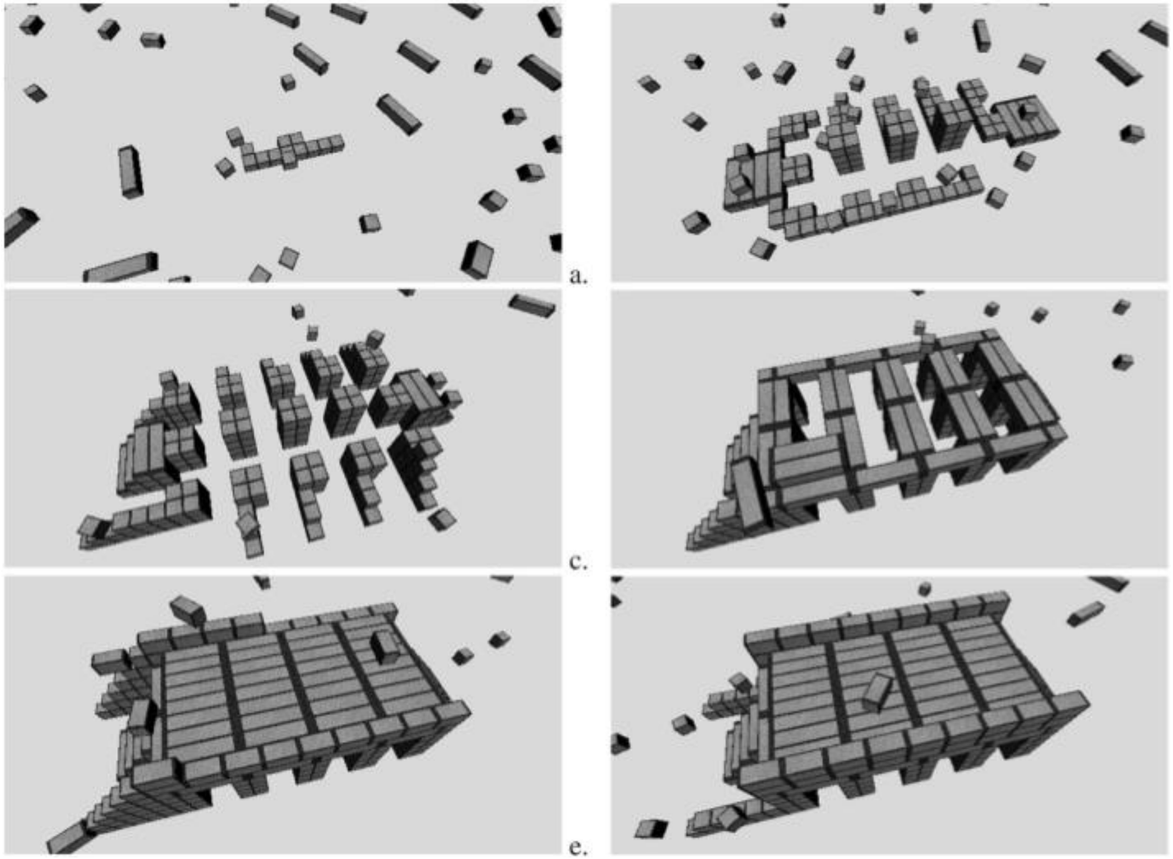
One of the Oldest and Largest Vision Groups in the World

- Founded by Azriel Rosenfeld in 1965.
- By quantitative measures (csrankings.org) #3 group in US, 1998-2018.
- Five computer vision faculty (ECE AND CS)
- Plus several research faculty
- Other related faculty
- ~60 grad students

AI at University of Maryland

- Computer vision
- Neuro and swarm computing
- Robotics
- Machine learning
- Natural language processing

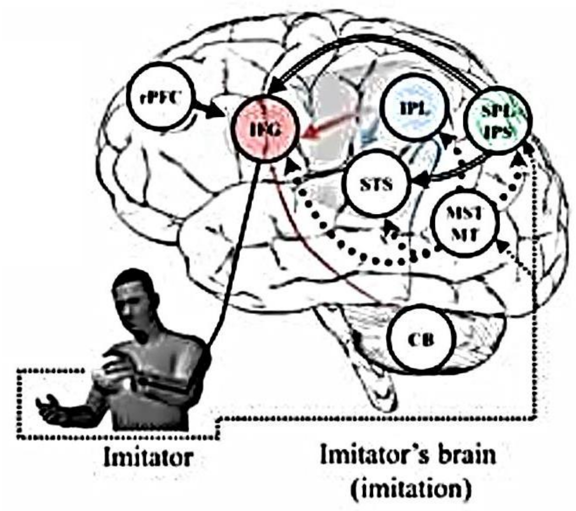
Swarm Intelligence



Example: self-assembly of components into complex structures (bridge here)

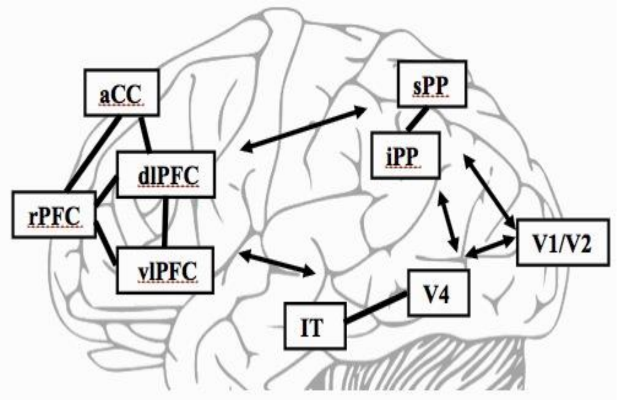
Grushin A, et al., *ACM Trans. Autonomous and Adaptive Systems*, 5, 2010.

Large-Scale Neurocognitive Architectures



imitation learning

Oh H, et al.. *Human Movement Science*, 2018, in press.



cognitive control

Sylvester J, et al.. *Neural Networks*, 79, 2016, 37-52.

AI at University of Maryland

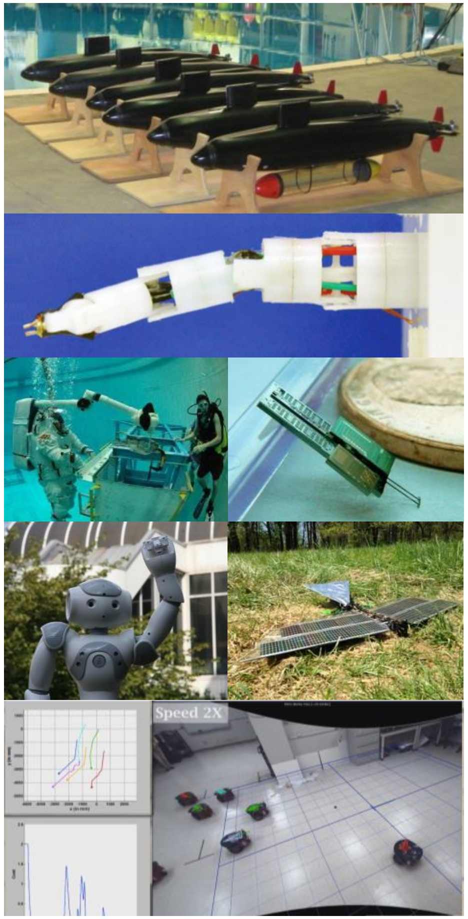
- Computer vision
- Neuro and swarm computing
- Robotics
- Machine learning

Maryland Robotics Center: Overview

- Housed in the Institute for Systems Research
- Consists of twenty-one labs
- Consists of 40 participating faculty members from eight academic departments
- Current activities cover most facets of robotics
- Educational programs including M. Eng in Robotics

Center Research Expertise

- Bio-Inspired Robotics
- Cognitive Robotics
- Cooperative, Collaborative, Networked Robotics
- Unmanned Vehicles
- Miniature Robots
- Medical Robotics
- Robotics in Extreme Environments
- Social Robotics



Center research projects supported by NSF, ARO, ARL, ONR, AFOSR, NIH, DARPA, NASA, and NIST.

New Robotics Labs in Iribe Building and Idea Factory

Natural Language Human-Robot Communication

[Park and Manocha 2019]

AI at University of Maryland

- Computer vision
- Neuro and swarm computing
- Robotics
- Machine learning

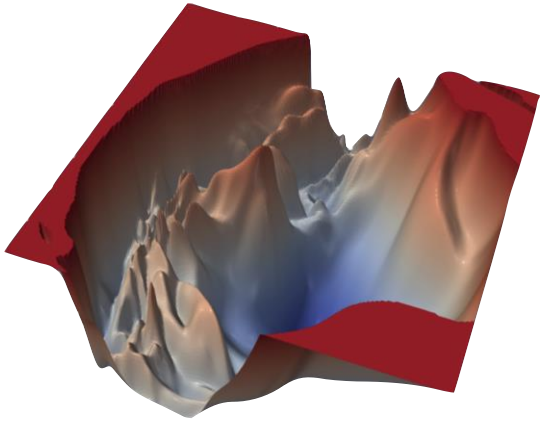
Machine Learning Center

https://cmns.umd.edu/news-events/features/4399

Research Areas

Distributed ML for big data in the cloud

Understanding Neural Nets



ML on portable low-power devices




- Optimization
- Distributed computing
- Computer vision
- Circuit design
- New MS program in ML

Adversarial Learning

“cat”

“traffic light”

Recent Research

- Autonomous Driving
- Robot Navigation
- Behavior and Emotion Classification

Recent Research

- Autonomous Driving
- Robot Navigation
- Behavior and Emotion Classification

Traffic and Transportation

- 1.2 billion vehicles on the roads today
  - 78.7M million new vehicles in 2018
    - China: 29M U.S.: 17.2M Europe 15.6M
- New features: electric cars, assisted driving…

New Delhi, India

Bangkok, Thailand

Olomuoc, Czech

Autonomous Driving

- Global Transportation Section: $4.8 Trillion Dollars

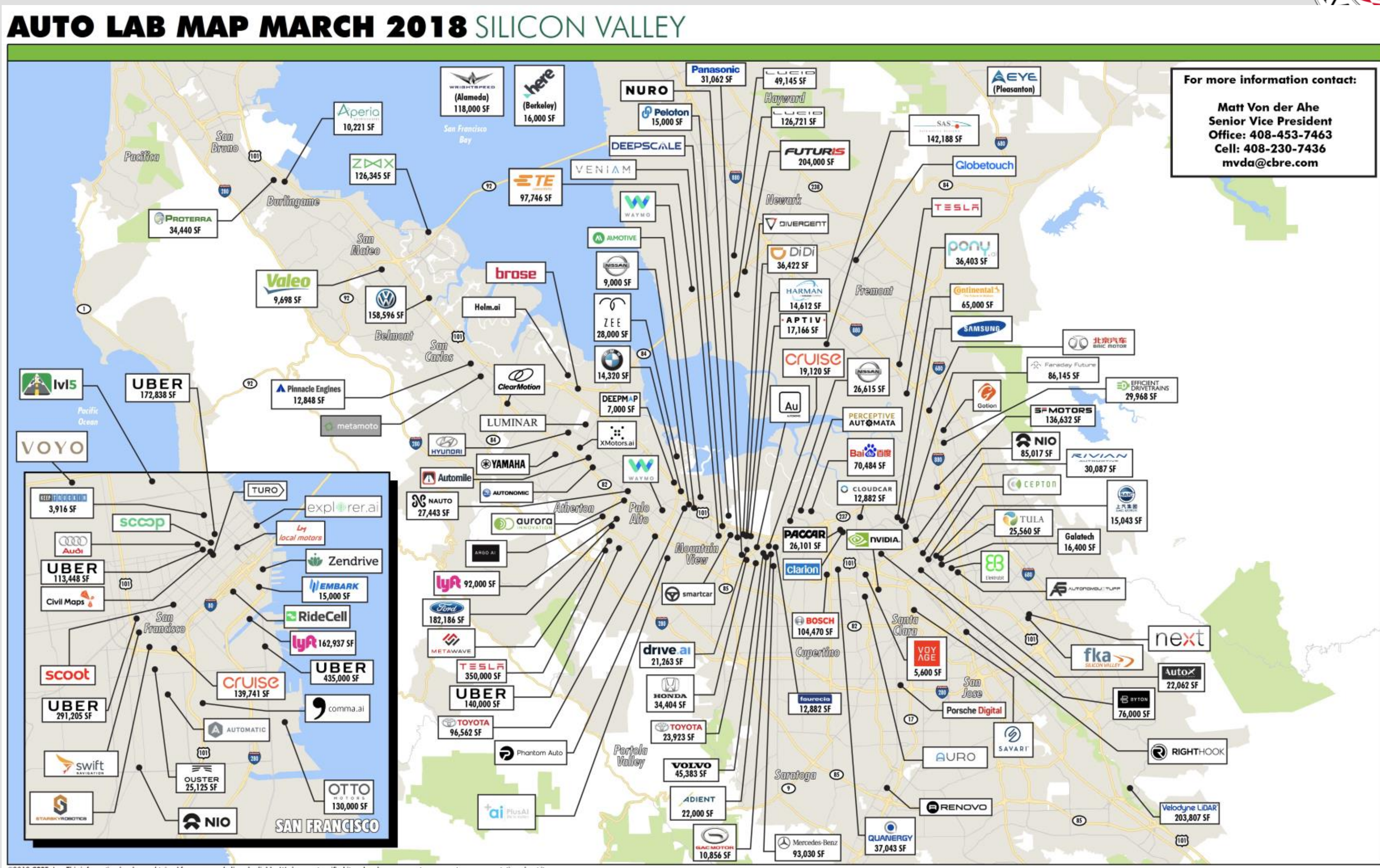
- Huge impact on safety and productivity
  - Reducing road deaths could considerably improve the GDP

- AI and Autonomy Technologies: Biggest advancement since Henry Ford's 1915 car

Henry Ford, 1915.

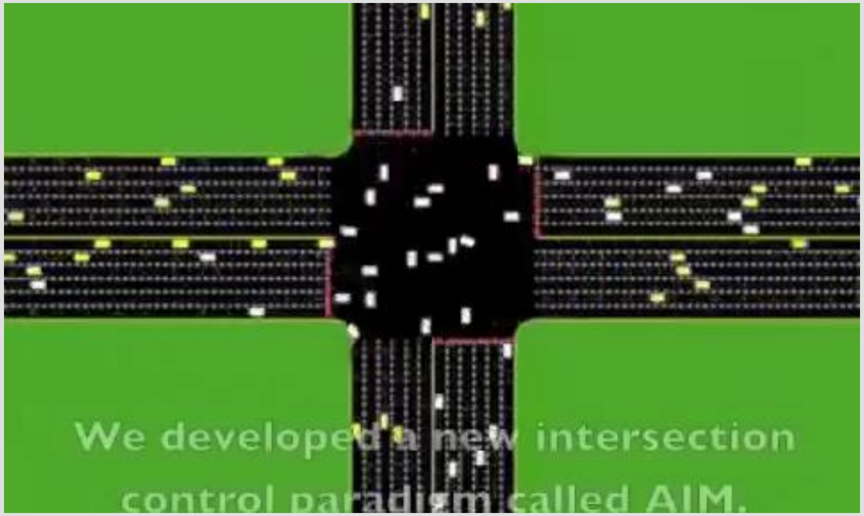
Fully Autonomous Car ???

AD: Commercial Developments



"**Autonomous Driving is the Mother of all AI Projects**", Tim Cook (Apple CEO), 2017

Autonomous Driving (AD): Ultimate Goal



Au et al. 2012

Kabbaj, TED 2016

Autonomous Driving

- First self-driving car (Ernst Dickmanns, 1987)

Autonomous Driving: Recent Predictions

- First self-driving car (Ernst Dickmanns, 1987)

- “First fully autonomous Tesla by 2018” (Musk, 2015)
- “Next generation Audi A8 capable of fully autonomous driving in 2017” (Moser, Audi, 2014)
- “first self driving cars on the market by 2019” (Volkswagen, 2016)
- “large number of self-driving cars on the road by 2019” (Ng, Baidu, 2016)

Autonomous Driving: Recent Predictions

- First self-driving car (Ernst Dickmanns, 1987)

- “First fully autonomous Tesla by 2018” (Musk, 2015)
- “Next generation Audi A8 capable of fully autonomous driving in 2017” (Moser, Audi, 2014)
- “first self driving cars on the market by 2019” (Volkswagen, 2016)
- “large number of self-driving cars on the road by 2019” (Ng, Baidu, 2016)

Is this **Irrational Exuberance?**

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

Tracking, Prediction, and Navigation

for Autonomous Driving in Dense or Urban Environments

PART-I

Tracking

PART-II

Prediction

PART-III

Navigation

PART-I
Tracking

- **Problem Statement:** Maintain the identity of a road-agent as they move in dense urban traffic.

Tracking

Easy (single agent, sparse)

Hard (Multiple agents, Dense, Occlusion)

Tracking by Detection

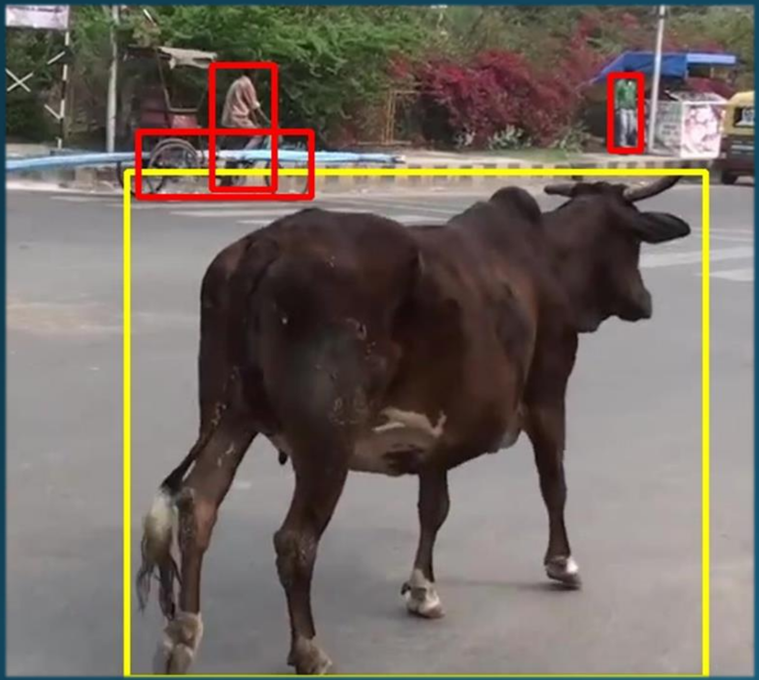
Tracking by Detection

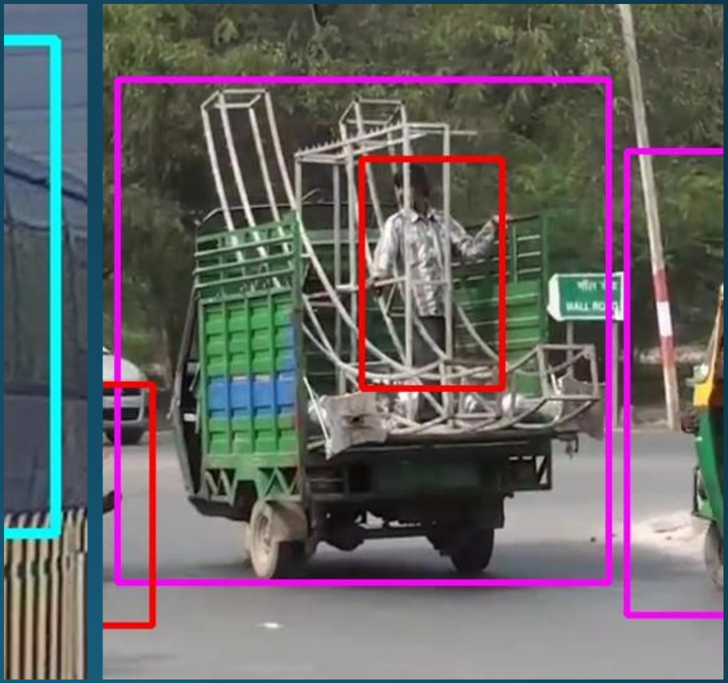
Object Detection and Segmentation achieves lowest false negatives among more than a 100 competitors on the latest benchmarks.

Results (**significant** improvement over prior work)

Tracking Challenging Cases

Unconventional Agents





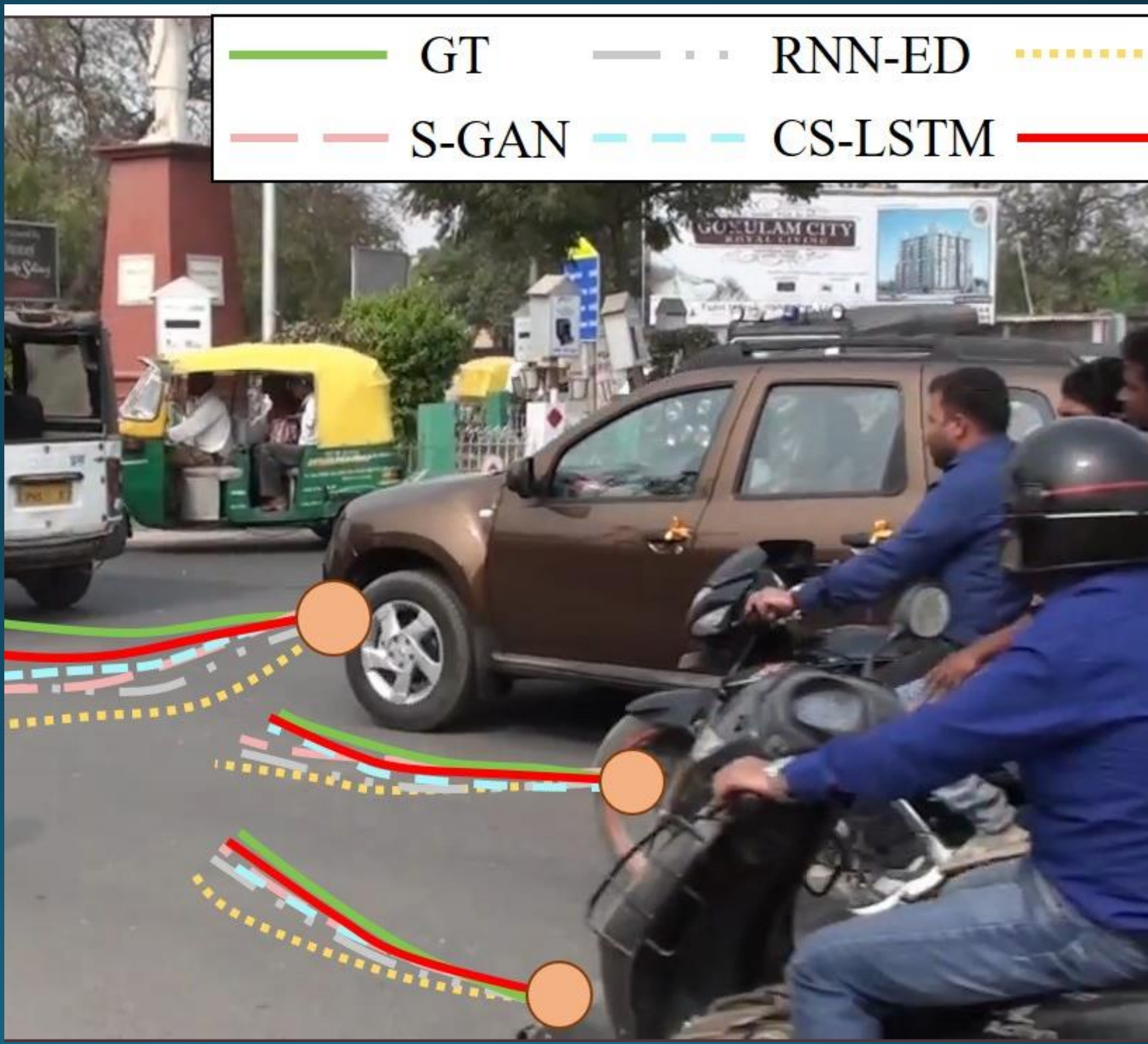
PART-II

Prediction

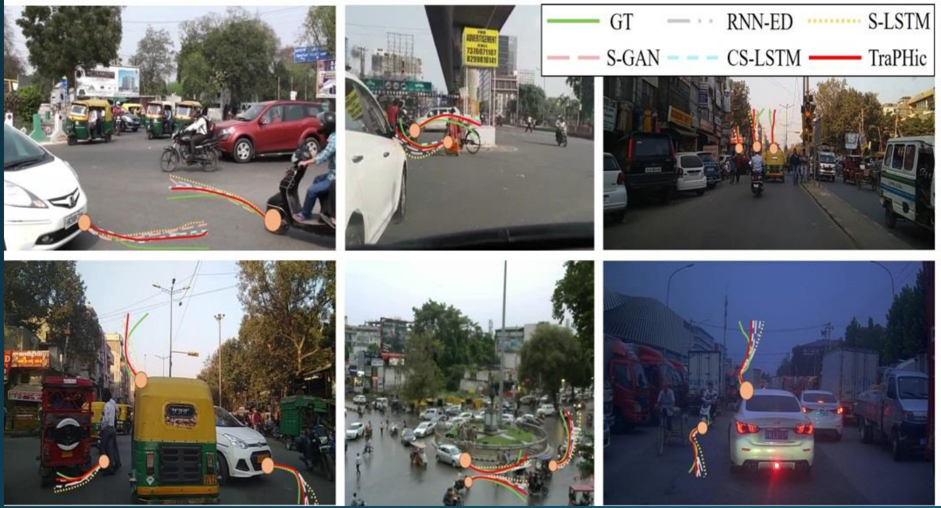
- **Problem Statement:** Given past 3-8 seconds of a road-agents trajectory, the task is to predict the trajectory for the next 5 seconds

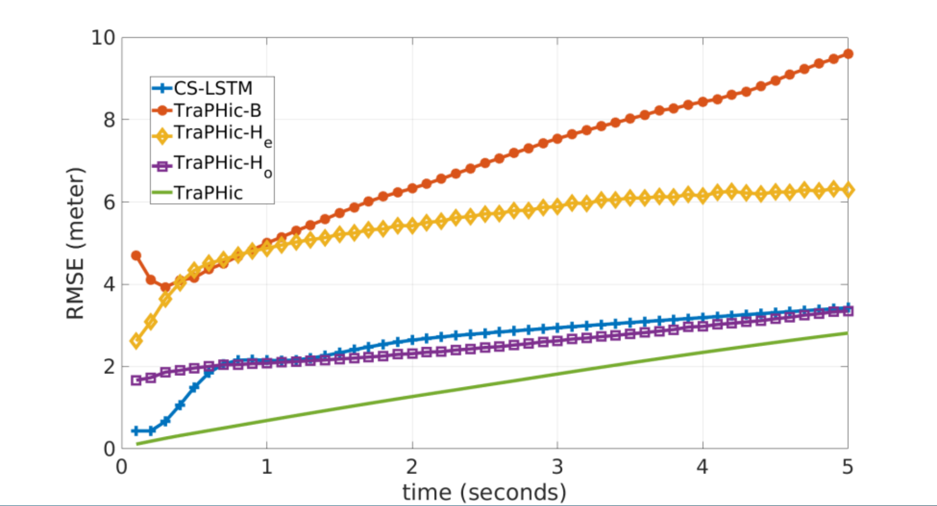
Trajectory Prediction

- Reduce the task to an end-to-end sequence prediction problem using neural networks.
    - Training Data: Trajectory history
    - Learned Output: Future trajectories
- In urban traffic, the difficulty of using learning algorithms increases with increasing density.



Results (**significant** improvement over prior work)





- **Problem Statement:** Safe movement of a road-agent from a starting location to a goal destination.

PART-III

Navigation

Autonomous Navigation in Dense Urban Traffic

- Dense
- Heterogeneous Agents
- Non-conformity to standard traffic rules such as:
    - Lane-Following
    - Traffic Signals
    - Speed Limits
    - Right of Way
    - Inter-Vehicular Distance

Combine ideas from robotics, geometry, physics and social sciences

Autonomous Navigation

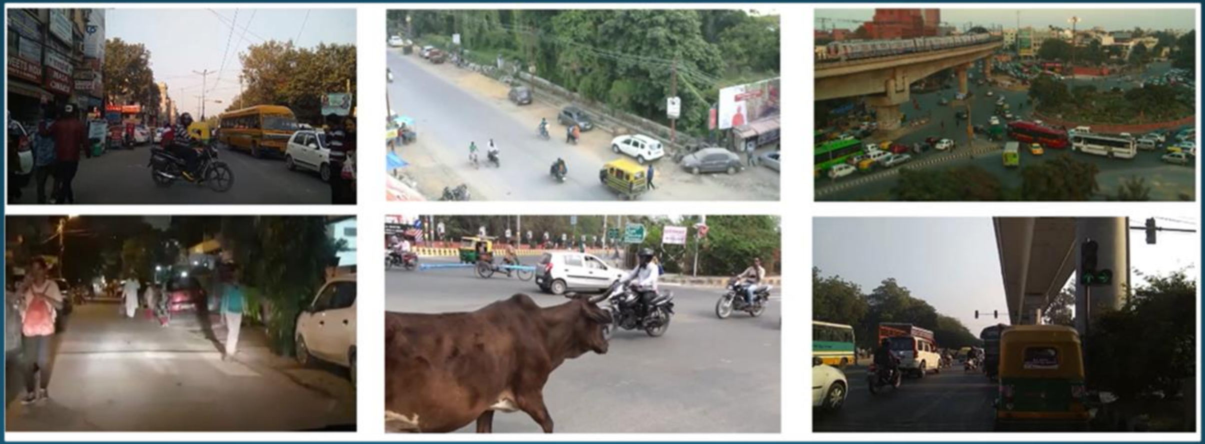
- To build a navigational model:
  - We leverage the behavioral effects of aggressive and conservative drivers.
  - Aggressive agents display maneuvers such as over speeding, tailgating, and overtaking. Conservative agents attempt to maintain a fixed trajectory.
- A driver behavior-driven navigational model results in theoretical guarantees for collision avoidance.

Our New Dataset

55 high resolution videos of dense, heterogeneous traffic.	Carefully annotated following a strict protocol.	Categorized by camera motion, camera viewpoint, time of the day, and difficulty level.

Link:
https://go.umd.edu/TRAF-Dataset

This dataset will be released as part of our CVPR’19 paper.



Ongoing Projects

- Develop an integrated pipeline: from perception to action
- Collaborate with industry on technology transfer: Intel and Baidu

Autonomous Driving: Challenges

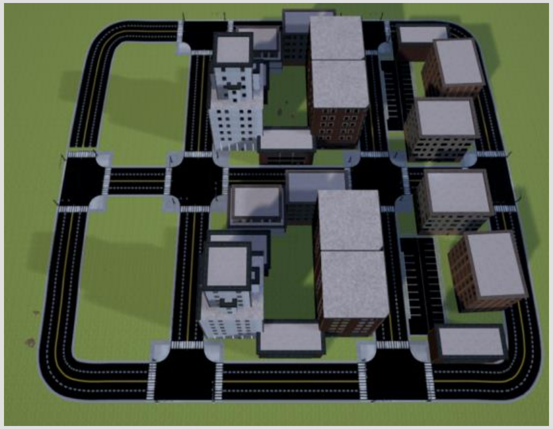
- Safety guarantees are critical
- Drivers, pedestrians, cyclists difficult to predict
- Road and environment conditions are dynamic
- Laws and norms differ by culture
- Huge number of scenarios

Autonomous Driving: Development and Evaluation

- Development and testing of autonomous driving algorithms
  - On-road experiments may be hazardous
  - Closed-course experiments may limit transfer
  - High costs in terms of time and money

- Solution: develop and test robust algorithms in simulation
  - Test novel driving strategies & sensor configurations
  - Reduces costs
  - Allows testing dangerous scenarios
  - Vary traffic and weather conditions

Parking lot mock-up



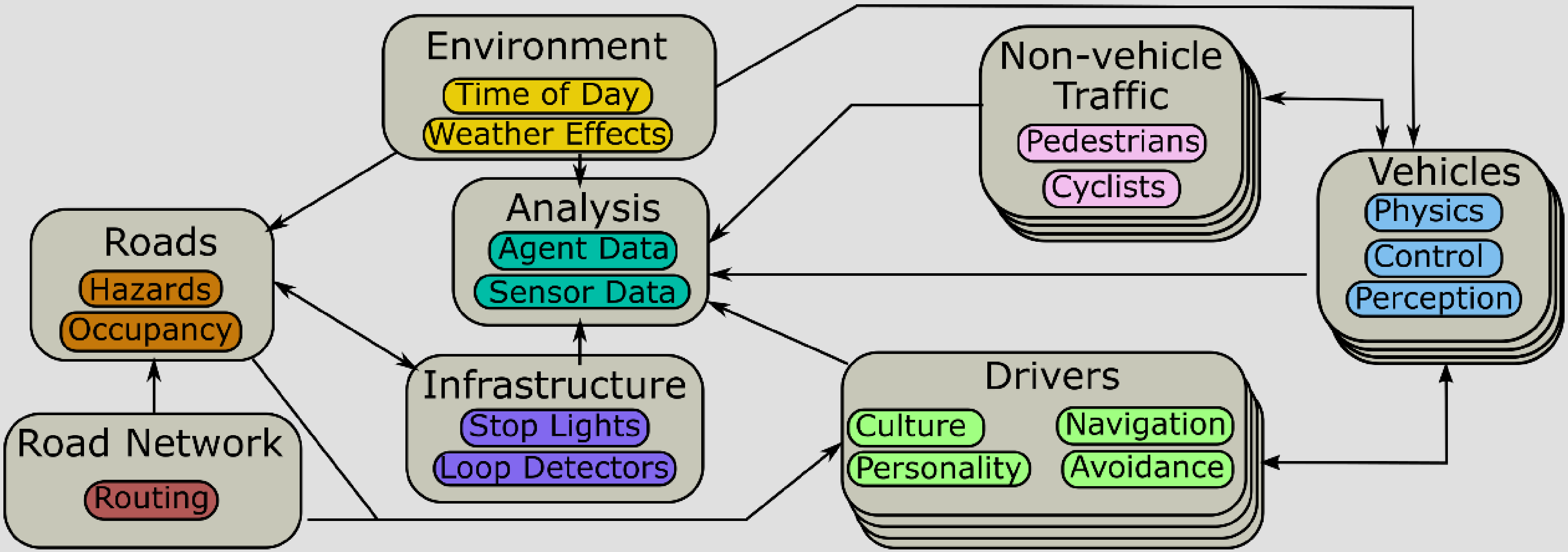
Simulated city

Our Autonomous Driving Simulator

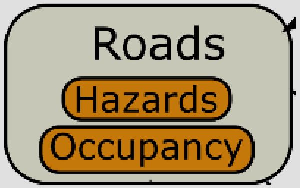
- AutonoVi-Sim : high fidelity simulation platform for testing autonomous driving algorithms
  - Varying vehicle types, traffic condition
  - Rapid Scenario Construction
  - Simulates cyclists and pedestrians
  - Modular Sensor configuration, fusion
  - Facilitates testing novel driving strategies

Autonovi-Sim

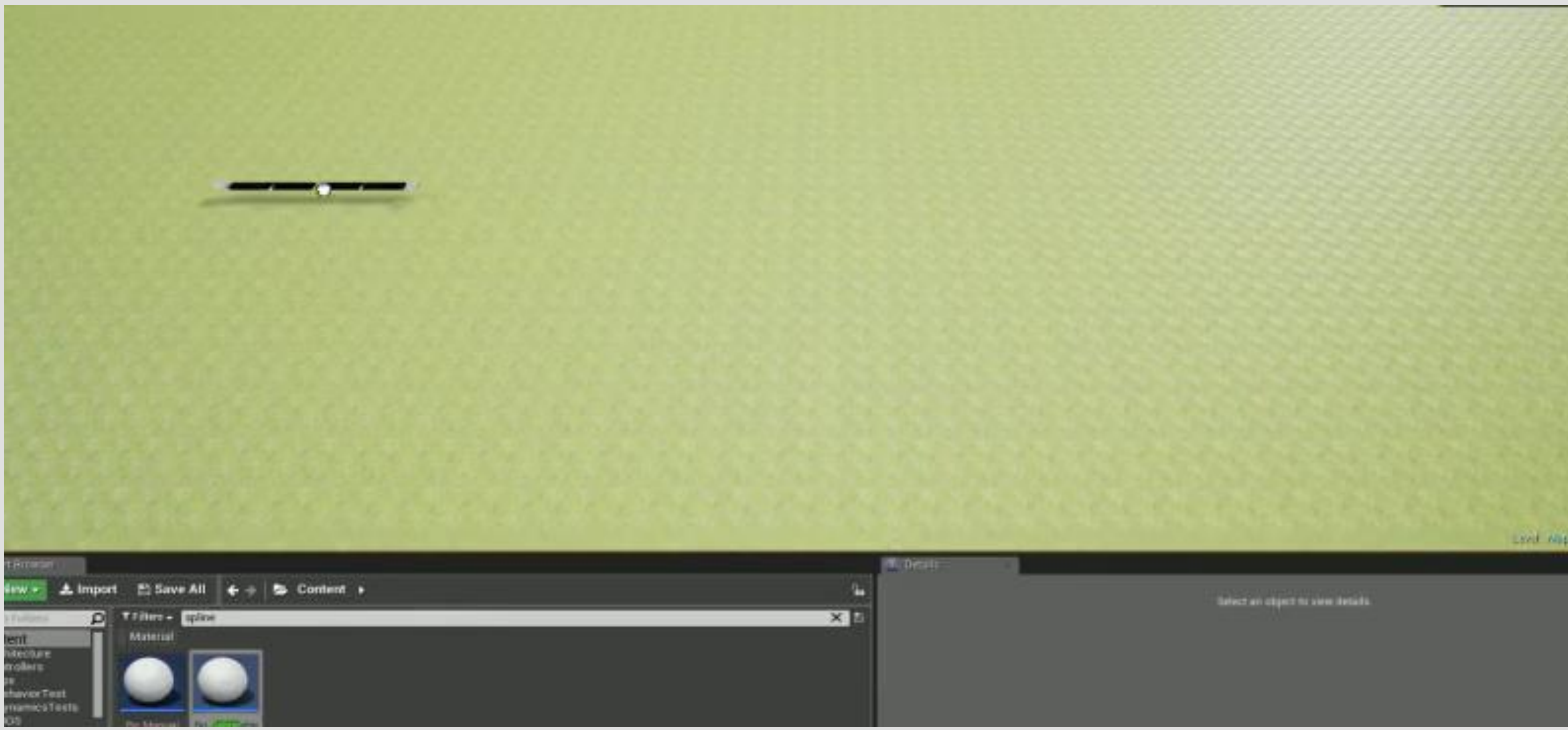
- Modular simulation framework for generating dynamic traffic conditions, weather, driver profiles, and road networks
- Facilitates novel driving strategy development
- On top of UnReal Engine



Autonovi-Sim: Roads & Road Network

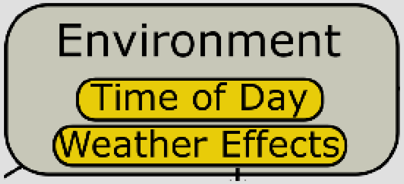


- Roads constructed by click and drag
- Road network constructed automatically



Road layouts

Autonovi-Sim: Environment

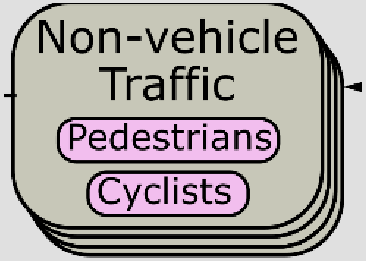


- Goal: Testing driving strategies & sensor configuration in adverse conditions
- Simulate changing environmental conditions
  - Rain, fog, time of day
  - Modelling associated physical changes

Fog reduces visibility

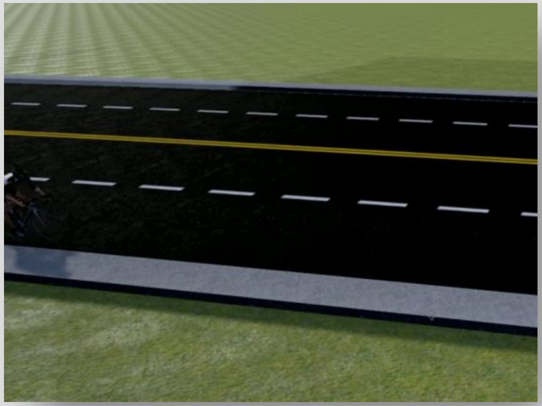
Heavy rain reduces traction

Autonovi-Sim: Non-vehicle Traffic



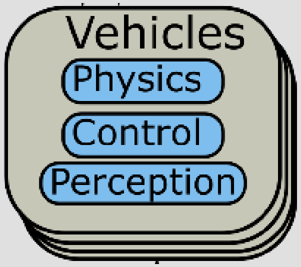
- Cyclists
  - operate on road network
  - Travel as vehicles, custom destinations and routing
- Pedestrians
  - Operate on roads or sidewalks
  - Programmable to follow or ignore traffic rules
  - Integrate prediction and personality parameters

Pedestrian Motion

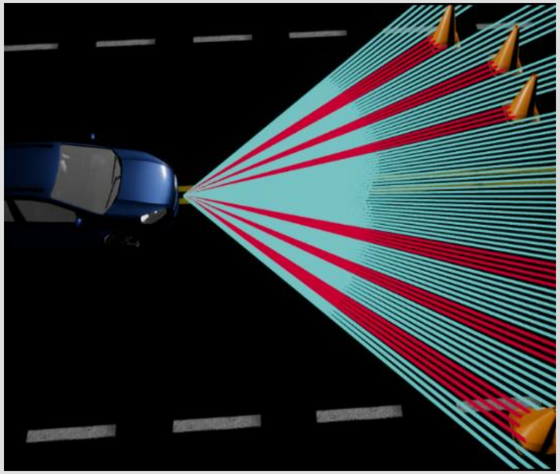


Cyclist Motion

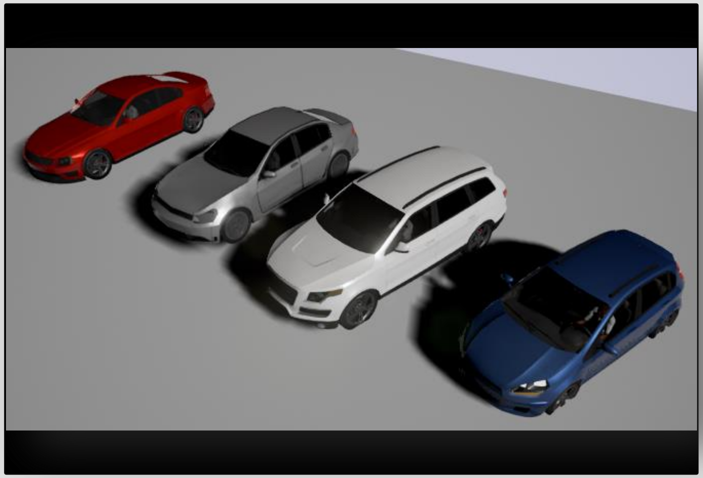
Autonovi-Sim: Vehicles



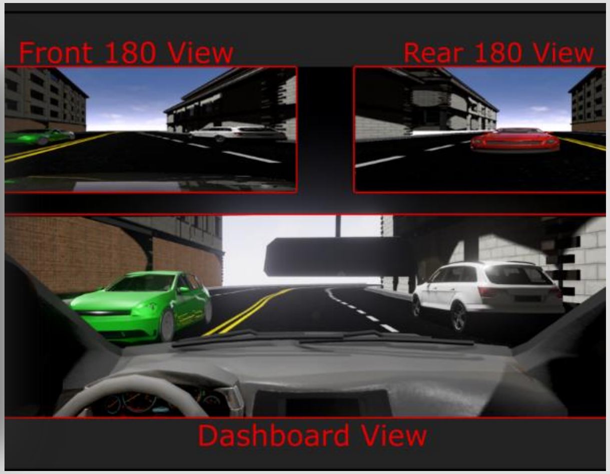
- Various vehicle profiles:
  - Size, shape, color
  - Speed / engine profile
  - Turning / braking
- Manage sensor information



Laser Range-finder



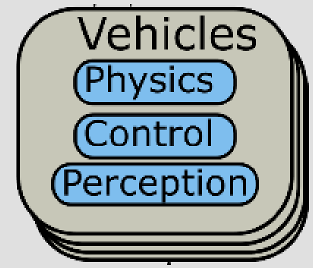
Multiple Vehicle Configurations



Multi-camera detector

Autonovi-Sim: Vehicles

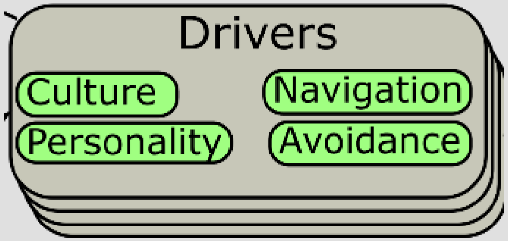
- Sensors placed interactively on vehicle
  - Configurable perception and detection algorithms



[Best et al. 2018, CVPR]

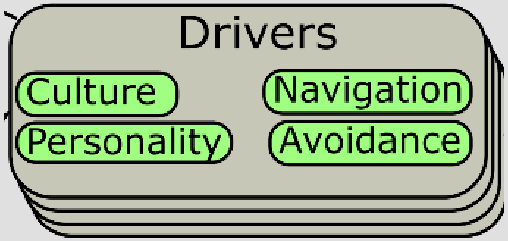
Autonovi-Sim: Drivers

- Control driving decisions
  - Fuse sensor information
  - Determine new controls (steering, throttle)
- Configurable parameters representing personality
  - Following distance, attention time, speeding, etc.
- Configure proportions of driver types
  - i.e. 50% aggressive, 50% cautious



Autonovi-Sim: Drivers

- 3 Drivers in AutonoVi-Sim
  - Manual
  - Basic Follower
  - AutonoVi



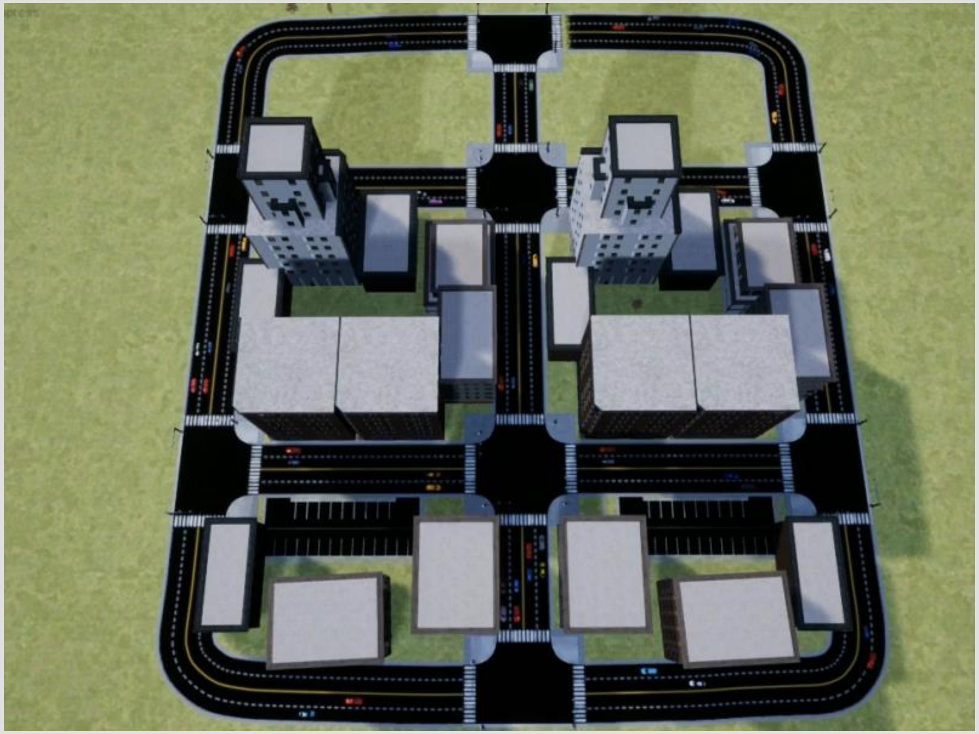
Manual Drive

Basic Follower

AutonoVi

Autonovi-Sim: Results

- Simulating large, dense road networks
- Generating data for analysis, vision classification, autonomous driving algorithms



50 vehicles navigating (3x)

AADS: Augmented Autonomous Driving Simulator

- Combine data-driven methods with model-based simulation
- Realistic rendering and behavior modeling
- Science Robotics (March 27, 2019)

Wei Li, Chengwei Pan, Rong Zhang, Jiaping Ren, Yuexin Ma, Jin Fang, Feilong Yan,
Qichuan Geng, Xinyu Huang, Huajun Gong, Weiwei Xu, Guoping Wang,
Dinesh Manocha, Ruigang Yang

Baidu Research; National Engineering Laboratory of Deep Learning Technology and Application, China; Peking University; Deepwise AI Lab;
Zhejiang University; University of Hong Kong; Beihang University; Nanjing University of Aeronautics and Astronautics; University of Maryland.

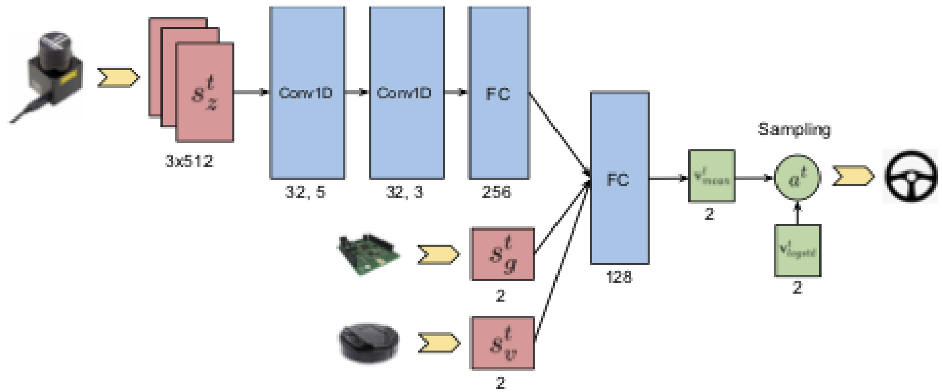
Recent Research

- Autonomous Driving
- Robot Navigation
- Behavior and Emotion Classification

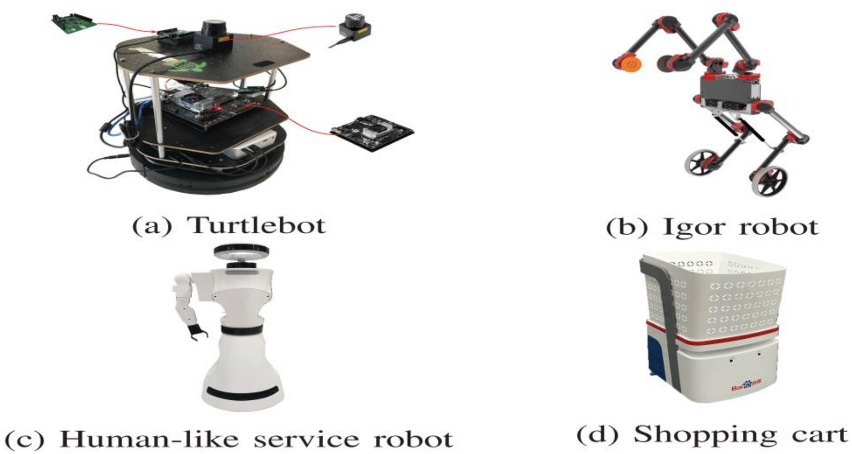
Robot Navigation in Dense Crowds

Dinesh Manocha

- New method for automatic mapless navigation
- Combination of reinforcement learning and collision avoidance methods
- Optimizes maples navigation policy with a robust policy gradient algorithm
- Applicable to different robots

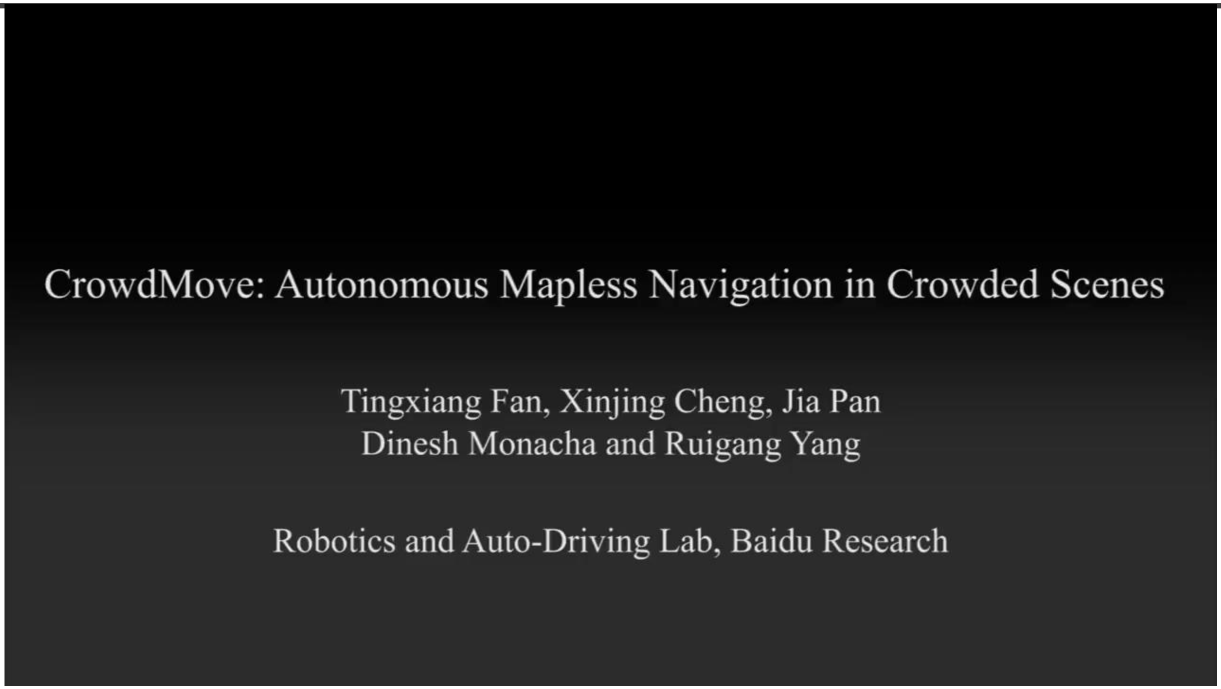


Our Collision Avoidance Neural Network



Different Robots used for Evaluation

Robot Navigation in Dense Crowds

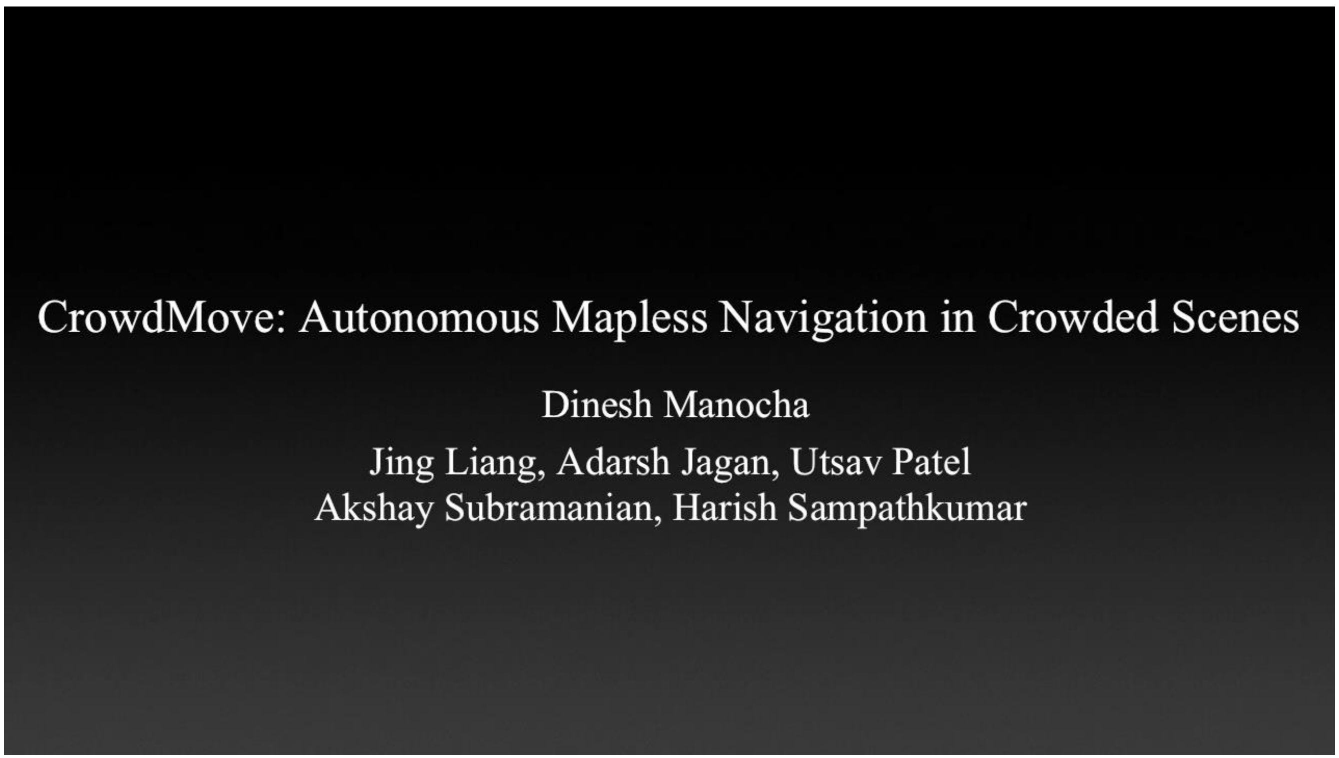


Automatic navigation in unstructured environments

Collision Avoidance in Dense Crowds

Can be used for delivery and transportation robots

Robot Navigation in Dense Crowds



Live UMD Demonstration

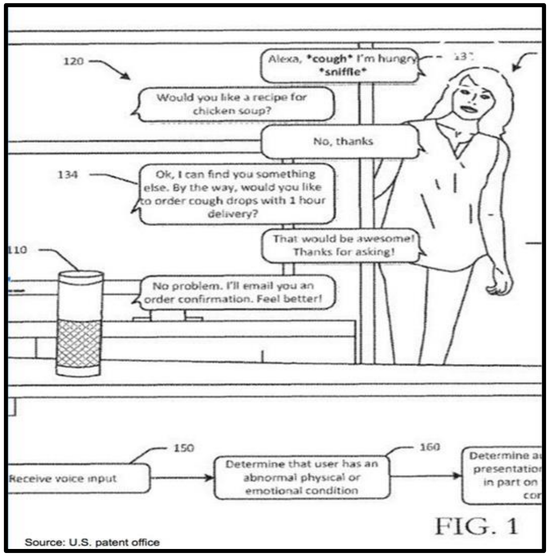
Recent Research

- Autonomous Driving
- Robot Navigation
- Behavior and Emotion Classification

How can an autonomous car navigate here?

- **Dense** and **heterogeneous** traffic flow
- Will the pedestrian stop? Will the agent overtake? Wil it over-speed?
- Agent looks at **emotional/behavioral state** of surrounding agents and then decides its next move

Emotion/Behavior Modeling is gaining attention. Let’s look at a few recent developments!

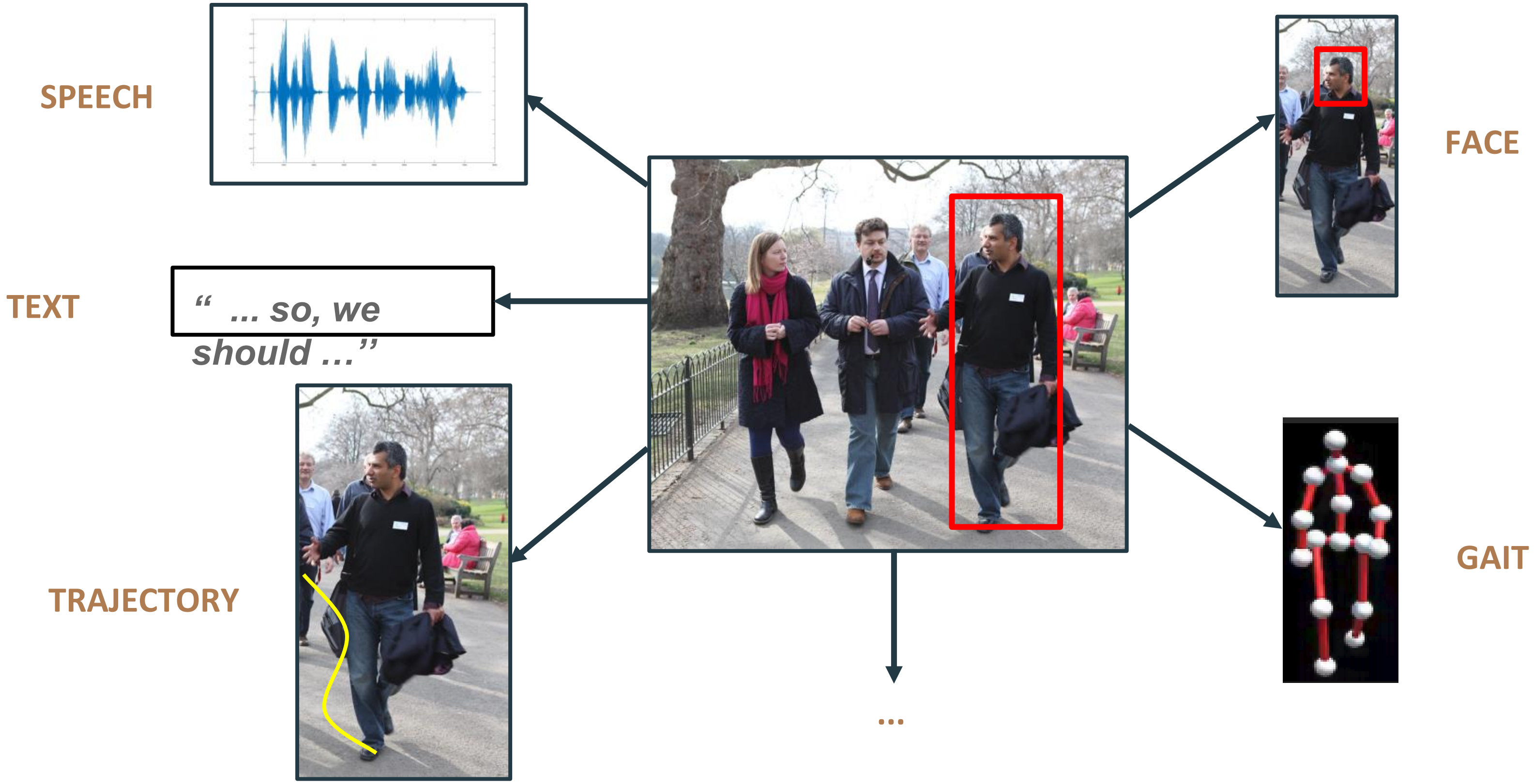


No Boarding Passes: Facial Emotion Recognition is clearing passengers for security in some airports in US and China.

Amazon is working on a wearable device to read human emotions simply from the sound of your voice.

The BioEssence wearable detects stress or pain and releases a scent to help the wearer adjust to the negative emotion.

Modalities/Cues



Learnings from Psychological Studies

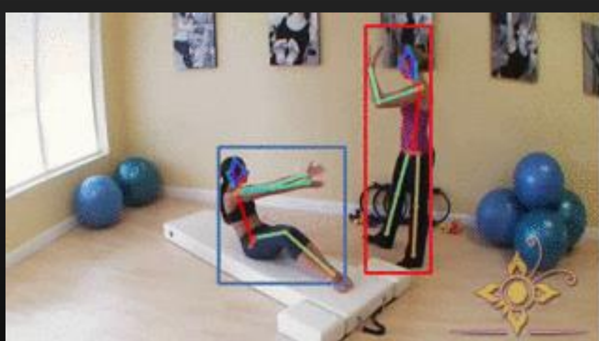
- Humans perceive emotions from **all possible available** bodily cues.
- Humans decide from **observing all these cues simultaneously** and not by observing each of the cues separately.

Other Practical Advantages

- Easy to **fake** certain modalities, while not so others.
- **Not easy** to obtain such **datasets** with a modality captured perfectly. Do away by treating others as proxy modalities.

Perceiving Emotions and Behavior from Pedestrian Gait Videos

What are Gaits?



Girdhar et al., "Detect And Track: Efficient Pose Estimation in Videos", ICCV 2017

- A *sequence of such poses* for a person is that person's *gait*
- Track people in videos and extract their poses in each frame

Why Gaits?

- Easy to observe
  - Can be seen from a long distance, even when face is not clear, and speech cannot be heard
- Hard to fake
  - Studies in psychology show people have less conscious control over their gaits

A Shopping Mall in Birmingham on Boxing Day, Courtesy: BMC HD Videos

Emotions and Behavior from Gaits are Extremely Useful for Navigation

Emotion: Happy, Sad, Angry, …

Behavior: Aggressive, Tense, Shy, …

Autonomous Navigation

- Predict pedestrian behaviors from their walking styles to determine when to stop and go.
- Becomes extremely challenging in dense traffic.

But Building a Dataset is Tough!

Dubai International Airport, Courtesy: Getty Images

Pedestrian crossing in downtown Seoul, Courtesy: Storyblocks Video

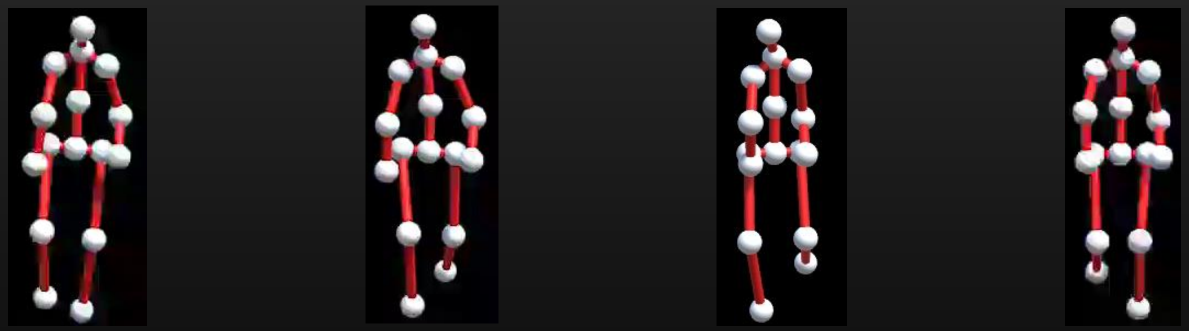
- High cost of annotation
- Video annotation requires tracking

- Subjectivity
- Mostly neutral labels in the wild

Start Simple: Our Emotion-Walk Dataset

342 single-person gaits with emotion labels

Collected in a controlled setting where participants were asked to walk towards the camera while acting out emotions.



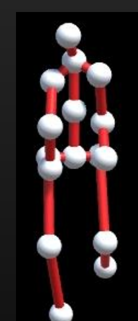
Start Simple: Our Emotion-Walk Dataset

342 single-person gaits with emotion labels

Collected in a controlled setting where participants were asked to walk towards the camera while acting out emotions.

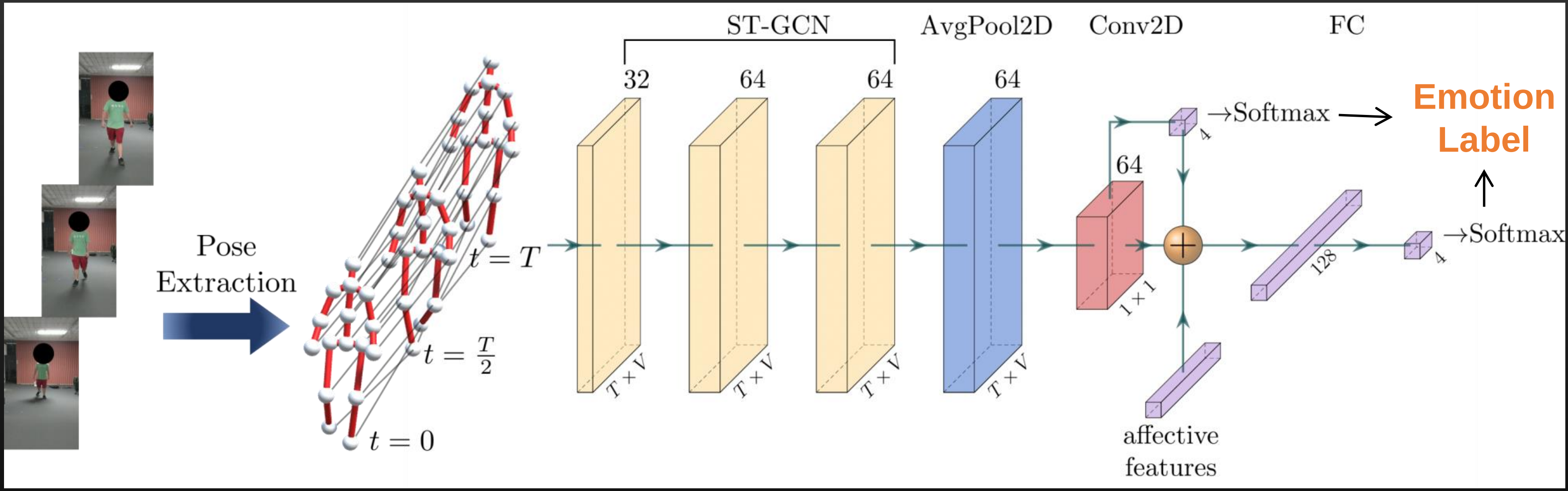

Angry


Happy


Sad


Neutral

Train a Classifier



Gait Videos

Temporal sequence of pose graphs

Graph Convolution Layers

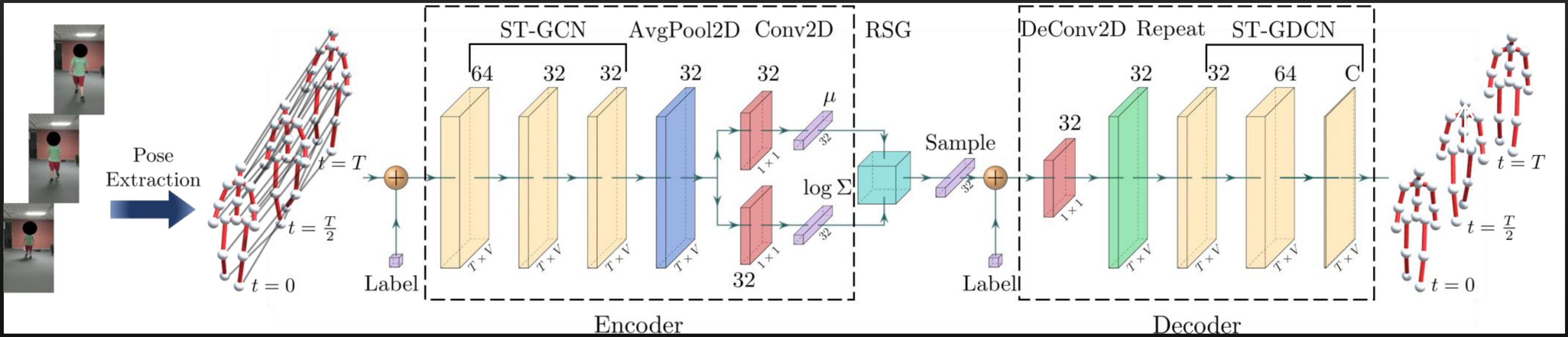
Performance

SVM	58.42%
Random Forest Classifier	66.22%
LSTM	75.10%
ST-GCN	76.32%
ST-GCN + Affective features	81.25%

Joint velocities, acceleration: walking speed,
Angles between joints: posture, and so on

More Data? Can Generate!

A generative network that can generate synthetic gaits annotated with emotion labels.



Gait Videos

(used for training)

Random Samples

(used for generation)

Synthetic Gaits

Conclusions

- UMd has major research efforts in autonomy and AI
- Industry collaboration: Capitol One, Adobe, Amazon HQ2 (local); Silicon Valley companies
- Research on autonomous driving, robot navigation and behavior classification
- New educational program in ML
- ARLIS: A new UARC on Intelligence and Security