中央研究院
資訊科技創新研究中心
Research Center for Information Technology Innovation

Video Analytics for AI City Smart Transportation

December 25, 2017

Ming-Ching Chang

College of Engineering and Applied Sciences (CEAS)
University at Albany
State University of New York (SUNY)

UNIVERSITY AT ALBANY
State University of New York

About me

Ming-Ching Chang 張明清

http://www.albany.edu/faculty/mchang2/

- SUNY Albany ECE (2016 – present)
- SUNY Albany Adjunct (2012 - 2016)
- GE Global Research (2008 - 2016)
- Brown Univ. PhD (2008)
- National Taiwan Univ. MS (1998)
- ITRI Research Assistant (1998)
- National Taiwan Univ. BS (1996)

Industrial Technology Research Institute
工業技術研究院

BROWN

UNIVERSITY AT ALBANY
State University of New York

Talk Overview

- SUNY Albany and New Engineering College
- AI City & Smart Transportation
- DETRAC Dataset & Benchmark
- 2017 NVidia AI City Challenge
- 2017 AVSS T4S Workshop & Challenge
- Video Analytics
- On-going Activities

Talk Abstract

- The prosperity of AI, Deep Learning and IoT are making our world smarter and impacting our life in every aspects. Among many changing fronts, smart transportation represents the core of smart city, that sits at an unique spot with strong technology readiness and boost. With millions of traffic and street video cameras around the world capturing data, there are far less automated analysis that aim to create value to support decision making for traffic optimization, safety, and management. In this talk we will present recent advancements in video analytics technologies for (1) the detection and tracking of pedestrians, vehicles, and motorists for traffic analysis, and (2) behavior and event recognition methods that can automate the decision making support. With the fast advancement of powerful GPU servers and GPU-enabled embedded platforms, automatic video data analytic technologies are mature for improving traffic control, reducing congestions, preventing accidents, supporting surveillance, and upgrading transportation infrastructure, to make our transit systems safer, smarter, and cheaper. We will also summarize results from two recent public contests --- the IEEE SmartWorld NVIDIA AI City Challenge and the IEEE AVSS Workshop on Traffic and Street Surveillance for Safety and Security Challenge, and share thoughts for future developments.

Bio

- Dr. Ming-Ching Chang is an assistant professor at the electrical computer engineering department at the College of Engineering and Applied Sciences (CEAS), University at Albany, State University of New York (SUNY). He was a lead computer scientist in the Visualization and Computer Vision laboratory at the GE Global Research Center in Niskayuna, NY from 2008 to 2016. His expertise includes video analytics, computer vision, image processing, and artificial intelligence. He was a research assistant in the Laboratory for Engineering Man/Machine Systems (LEMS) at Brown University, where he received his doctoral in Engineering in 2008. He was a research assistant in the Mechanical Industry Research Laboratories, Industrial Technology Research Institute (Taiwan). He received his M.S. in computer science and information engineering in 1998 and the B.S. in civil engineering in 1996, both from the National Taiwan University. He is the recipient of the IEEE Advanced Video and Signal-based Surveillance (AVSS) 2011 Best Paper Award - Runner-Up, the IEEE Workshop on the Applications of Computer Vision (WACV) 2012 Best Student Paper Award, the GE Belief - Stay Lean and Go Fast Management Award in 2015, and the IEEE Smart World NVIDIA AI City Challenge 2017 Honorary Mention Award. He has authored more than 50 peer-reviewed journal and conference publications.

University at Albany - SUNY

https://en.wikipedia.org/wiki/University_at_Albany,_SUNY

Who is famous at Albany, the Capital District of New York

NY state gov

Rensselaer Computer Vision IN RPI'S COMPUTER SCIENCE DEPARTMENT

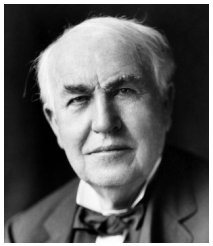
Rensselaer Polytech Inst.

Kitware

Anthony Hoogs

GE Energy

GE Research

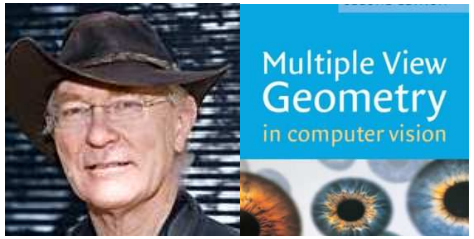


Thomas Edison

Joe Mundy

Bill Lorensen

Peter Tu



Richard Hartley

SUNY Office

U. Albany

SUNY Nanotech

Global Foundries

College of Engineering and Applied Sciences (CEAS)

- Started in 2015
  - was College of Computing and Information in 2005
  - Dept. of CS (25+ faculty), ECE (13+ faculty), Env. & Sus. Engin. (1 faculty). We keep growing.
- New home @ Schuyler building in downtown campus in 2019

Dean Kim Boyer

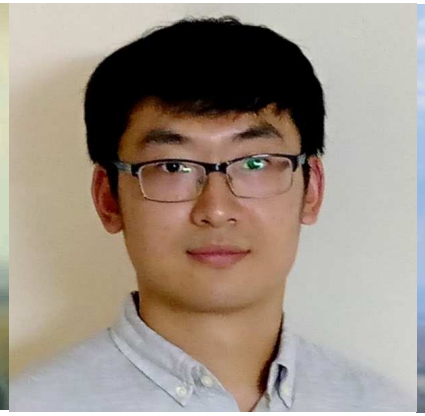
Computer Vision and Machine Learning (CVML) Lab

2 faculty, 4 affiliate faculty, 8 Ph.D. students, 8 alumni, ...

Siwei Lyu
Director

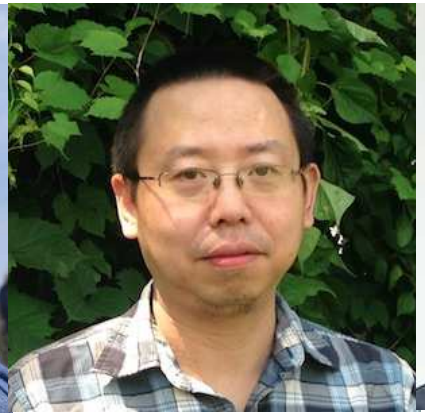
Ming-Ching Chang
Co-Director

Wenbo Li



Yuezun Li

Yi Wei



Nenghui Song

Lipeng Ke

A Smarter Future Full of Technologies

What will be the next rising star?

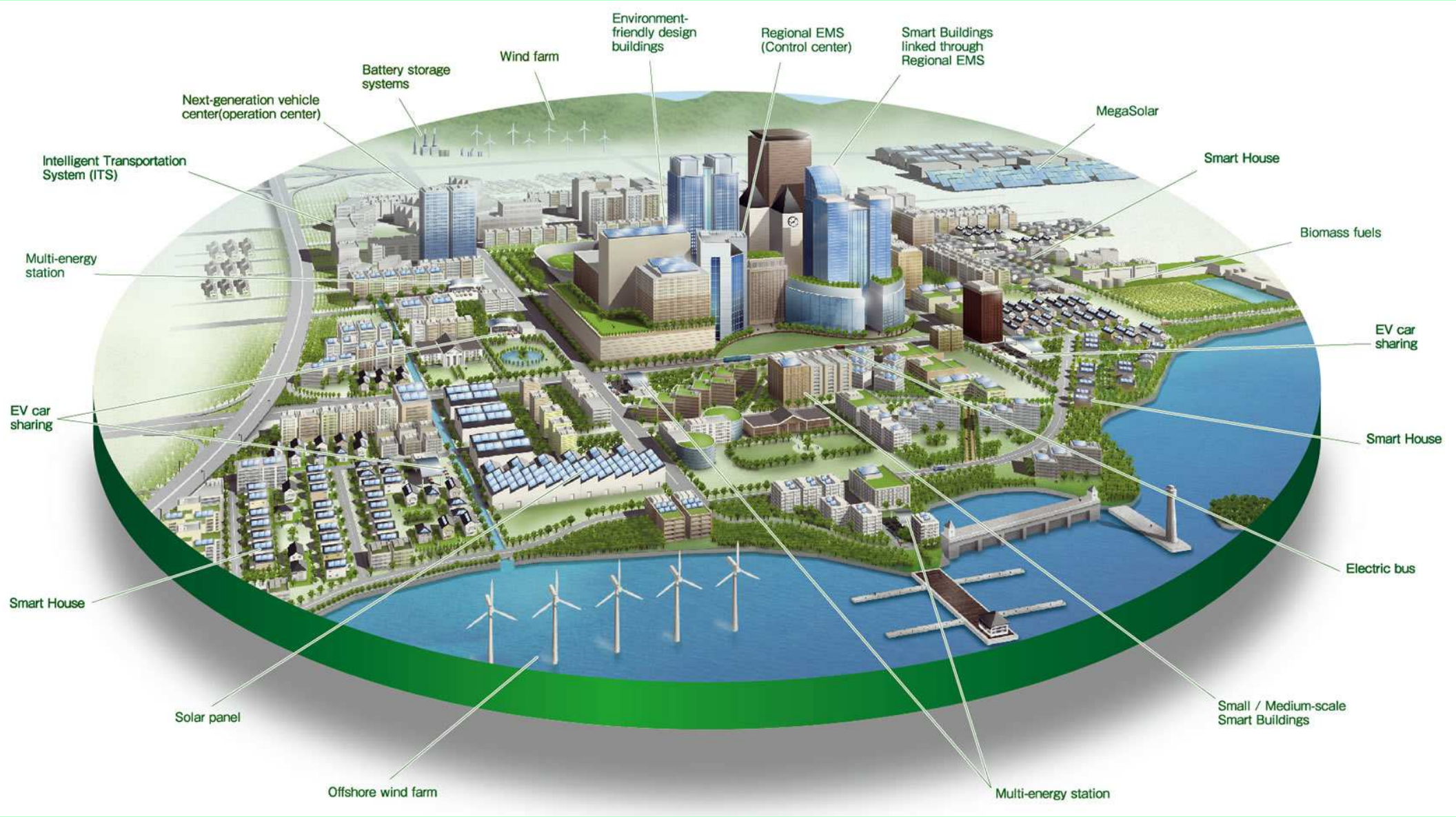


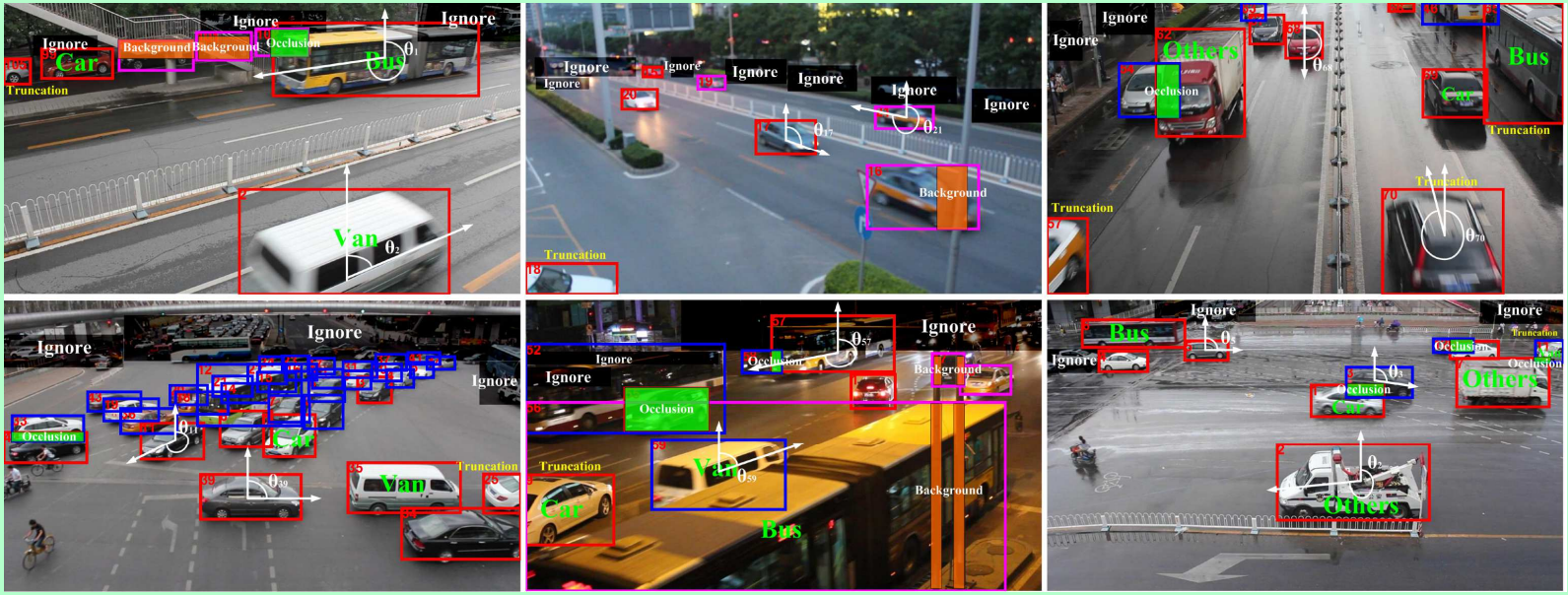
Image from http://www.districtoffuture.eu/

Motivation

- Smart Transportation is a core of Smart City/World
  - Make public transit systems safer, smarter, and cheaper using ubiquitous street cameras
  - Support strategic decisions for surveillance, safety, traffic control, parking, infrastructure investments, assisting self-driving cars

- Core technology: vehicle/person detection & tracking and analysis

UA-DETRAC Vehicle Dataset & Benchmark



UA-DETRAC Benchmark

University at **A**lbany **DE**tection and **TRAC**king dataset and benchmark

http://detrac-db.rit.albany.edu/

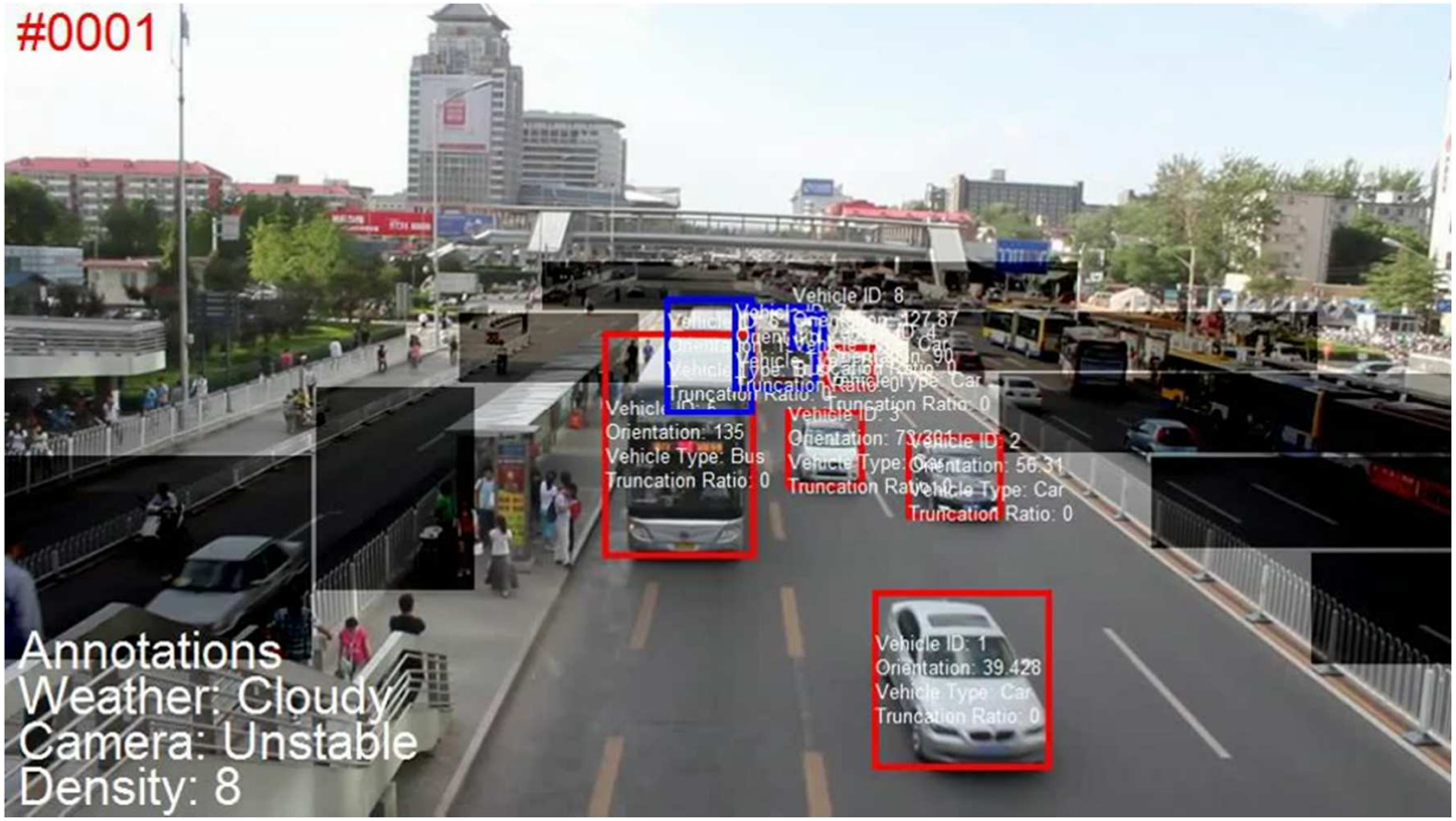
- Large scale vehicle detection and tracking evaluation framework
- Provide *dataset*, *open-src codes*, & *evaluation benchmark*

[L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M. Yang, and S. Lyu. "UA-DETRAC: A new benchmark and protocol for multi-object tracking." CoRR, abs/1511.04136, 2015.]

UA-DETRAC Dataset

University at **A**lbany **DE**tection and **TRAC**king dataset and benchmark

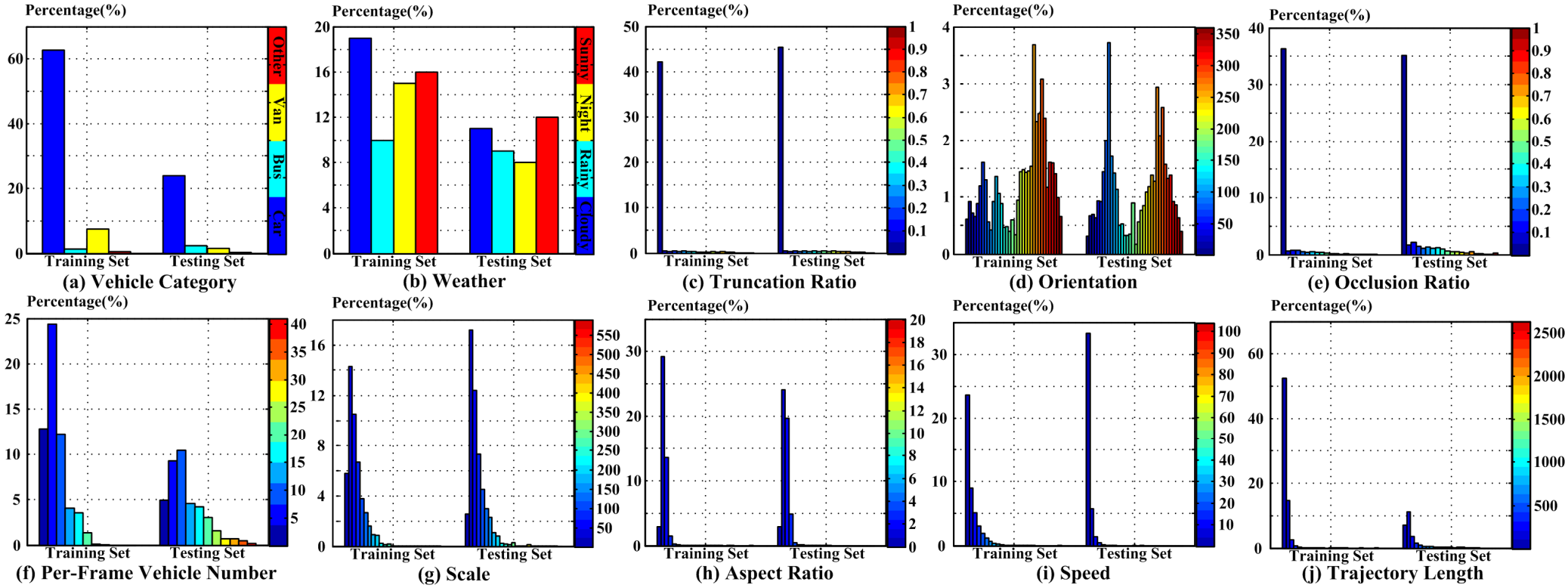
- 100 video sequences of traffic surveillance videos (~1 min. each)
- 140K frames of over 8K vehicles and 1.2M annotated bounding boxes



[L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M. Yang, and S. Lyu. "UA-DETRAC: A new benchmark and protocol for multi-object tracking." CoRR, abs/1511.04136, 2015.]

UA-DETRAC Dataset

Annotation attributes: weather, vehicle type, orientation, speed, occlusion ratio



(a) Vehicle Category (b) Weather (c) Truncation Ratio (d) Orientation (e) Occlusion Ratio
(f) Per-Frame Vehicle Number (g) Scale (h) Aspect Ratio (i) Speed (j) Trajectory Length

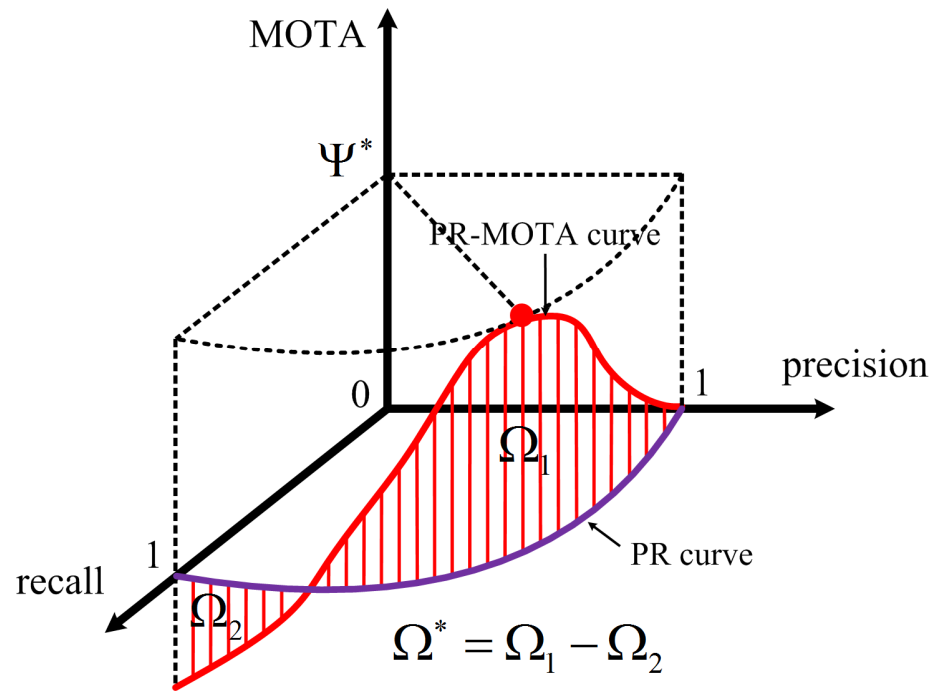
[L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M. Yang, and S. Lyu. "UA-DETRAC: A new benchmark and protocol for multi-object tracking." CoRR, abs/1511.04136, 2015.]

DETRAC Evaluation Metric

- Detection: average precision (AP)
- Tracking: DETRAC PR-MOTA

  - Precision Recall
  - Multi-Object Tracking Accuracy

$$MOTA = 1 - \frac{FN + FP + IdSw}{GT}$$



[L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M. Yang, and S. Lyu. "UA-DETRAC: A new benchmark and protocol for multi-object tracking." CoRR, abs/1511.04136, 2015.]

IEEE Smart World
2017 NVidia AI City Challenge

IEEE Smart World NVIDIA AI City Challenge

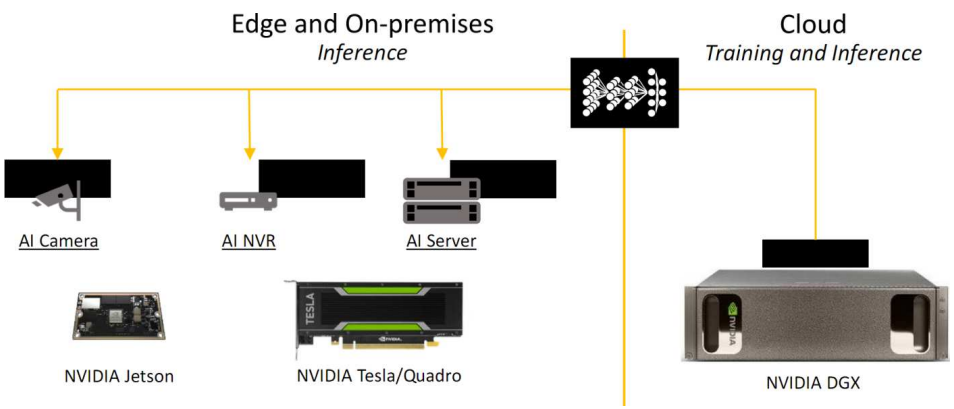
August 5, 2017, San Francisco, USA

http://smart-city-sjsu.net/AICityChallenge/index.html

2017 NVidia AI City Challenge

Videos captured at Silicon Valley and Virginia

NVidia platform support



worldwide participation by 29 teams.

- Track 1: vehicle detection
  - winner – UIUC
- Track 2: traffic applications
  - winner – U. Washington
  - honorable mention: UAlbany

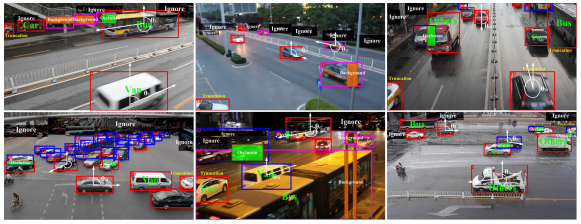
[M. Naphade, D. Anastasiu, A. Sharma, V. Jagrlamudi, H. Jeon, K. Liu, M.-C. Chang, S. Lyu, Z. Gao, "The NVIDIA AI City Challenge", IEEE Smart World Congress, San Jose, CA, USA, 2017]

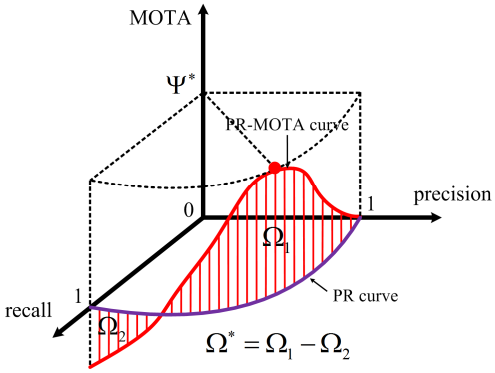
2017 AVSS Traffic and Street Surveillance for Safety and Security (T4S) Workshop

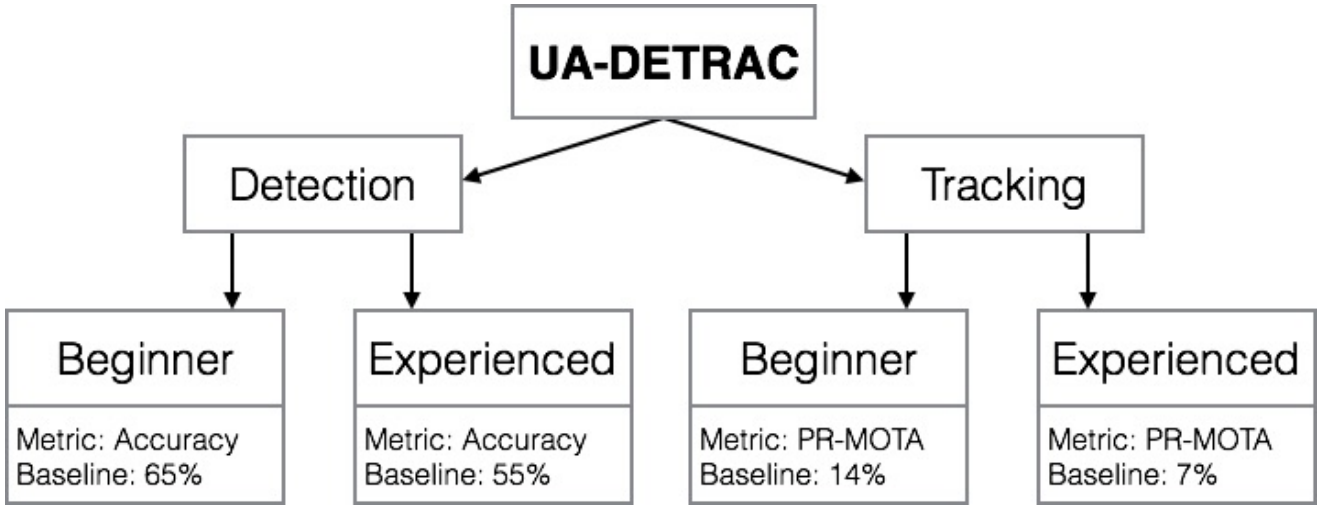
https://iwt4s.wordpress.com/

2017 AVSS T4S Challenge

- Dataset – UA-DETRAC
- Evaluation:
  - Detection: average precision (AP)
  - Tracking: DETRAC PR-MOTA
- Challenges:







[S. Lyu, M.-C. Chang, D. Du, L. Wen, H. Qi, Y. Li, Y. Wei, L. Ke, T. Hu, M. D. Coco, P. Carcagni, *et al.*, "UA-DETRAC 2017: Report of AVSS2017 & IWT4S Challenge on Advanced Traffic Monitoring", AVSS, 2017]

AVSS T4S – Detection Challenge

7 Submissions:

1. Geometric proposals for faster R-CNN (GPFRCNN): OSRAM GmbH, Munich, Germany — Experienced Winner
2. Evolving boxes for fast vehicle detection (EB): University of Washington, Seattle, USA
3. Lightweight SSD based on ResNet10 with dilations (SSDR): Intel, Nizhny Novgorod, Russia
4. R-CNN with Sub-Classes (RCNNSC): National Electronics and Computer Technology Center, Bangkok, Thailand — Beginner Winner
5. Faster R-CNN with ResNet101 (FRCNN-Res): University at Albany, Albany, USA
6. Region-based Deformable Fully Convolutional Network (DFCN): Iowa State University, Des Monies, Iowa, USA
7. CERTH Single Shot multibox Detector (SSD) for vehicle detection (CERTH-SSD): Information Technologies Institute (ITI), Thessaloniki, Greece**

AVSS T4S – Tracking Challenge

8 Submissions:

1. Higher-order graph and flow network based tracker (HGFT): Nanjing University of Posts and Telecommunications, Nanjing, China
2. Real-time multi-human tracking using a probability hypothesis density filter and multiple detectors (GMPHD): Technische Universität Berlin, Berlin, Germany*
3. Online distance based and offline appearance based tracker with correlated color dissimilarity matrix (CCM): University of Missouri, Columbia, USA
4. Intersection-over-union tracker (IOU): Technische Universität Berlin, Berlin, Germany* — Beginner Winner
5. Joint tracking with event grouping and temporal constraints (JTEGTC): Karlsruhe Institute of Technology, Karlsruhe, Germany — Experienced Winner
6. Multi-task deep learning for fast online multiple object Tracking (MTT): Institute of Automation, Chinese Academy of Sciences, Beijing, China
7. Gaussian Mixture Probability Hypothesis Density Filter extended by Kernelized Correlation Filters (GMPHD-KCF): Technische Universität Berlin, Berlin, Germany*
8. CERTH Kernelized Correlation Filters (KCF) tracking algorithm for vehicle tracking (CERTH-KCF)**

Take Home Messages

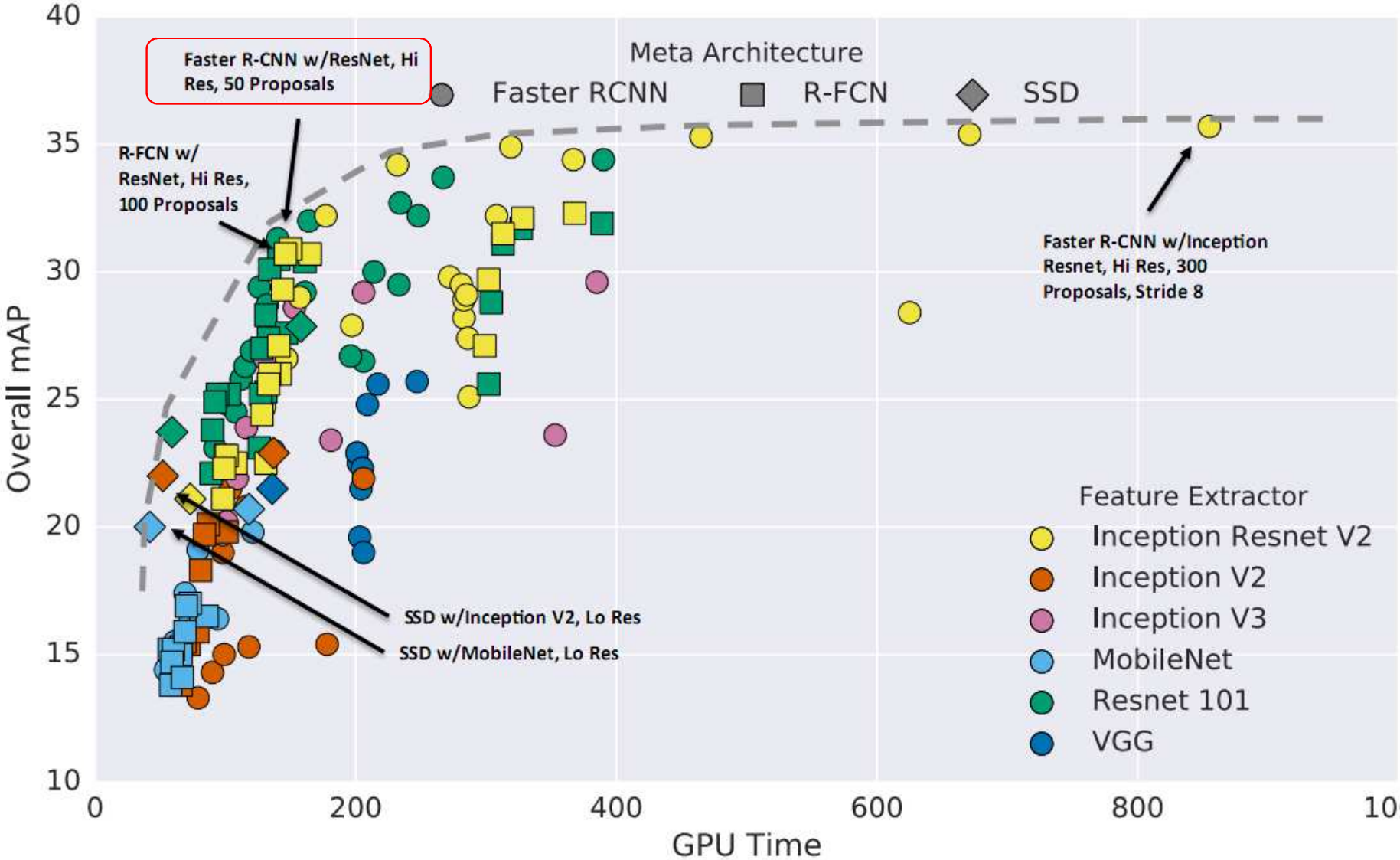
- Detection capabilities keep improving, fast
  - Faster R-CNN, SSD, YOLO9000, ...
- Tracking is relatively less attentive (but still important)
  - Is tracking a solved problem given good detections?
    - multi-view, re-acquisition, site-wide, city-scale?
- Analysis is hard, potential is large
  - many approaches / topics, rule-based? data-driven?
- Impact to real-world applications is huge
  - surveillance, security, smart transportation, parking, traffic control, self-driving cars, IoT, V2V, V2X, ...

Video Analytics

- **Detection** of Street Objects (Vehicles, Pedestrians)
- **Tracking** – Hypergraph Based Tracker
- **Mapping** – Camera Calibration & Projection
- **Analysis** – Bayesian or Learning Based

Detector: Speed/Accuracy Trade-offs

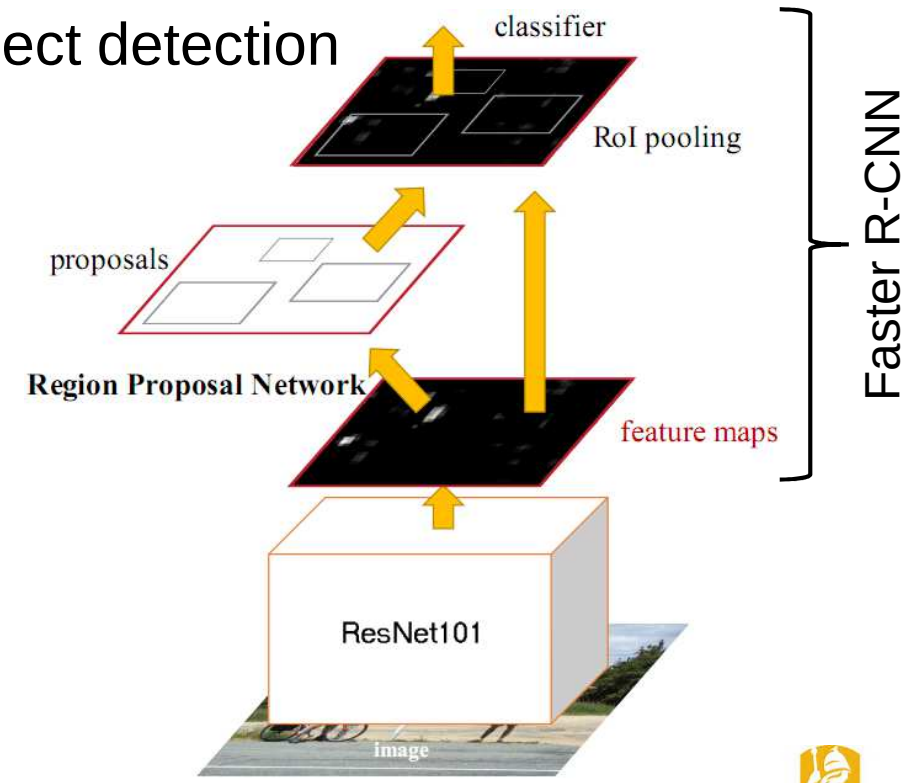
Figure 2 from "Speed/accuracy trade-offs from modern convolutional object detectors", Huang et al, CVPR 2017



Deep Object Detector

- Model architecture: **Faster R-CNN + ResNet101**
  - Inspired from Google Object Detection API *

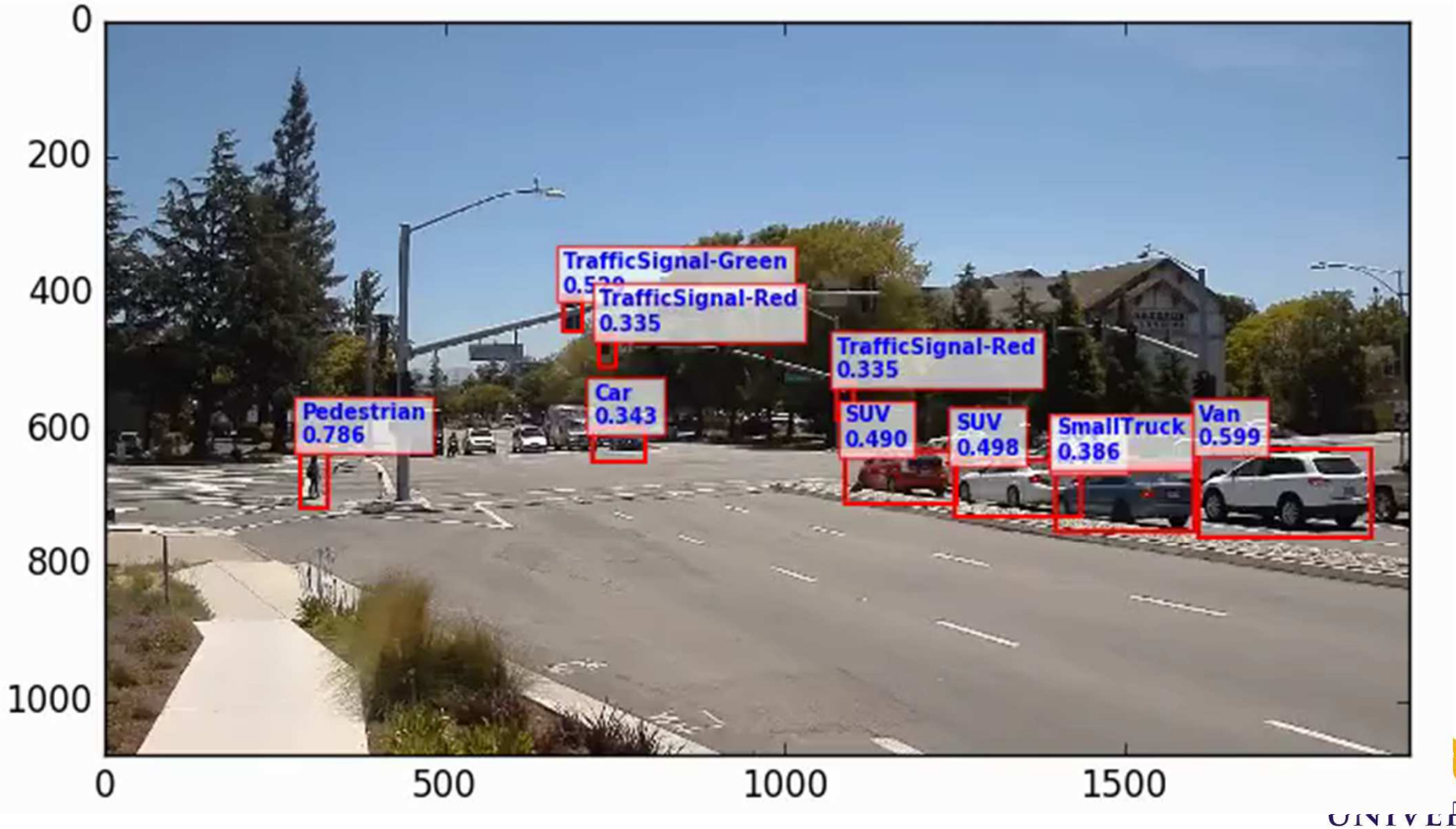
    "Speed/accuracy trade-offs from modern convolutional object detectors", Huang et al, CVPR 2017
  - Feature extraction using ResNet101
  - Faster R-CNN for vehicle/person/object detection
- Platform: TensorFlow (python)
  - Speed: 0.5 FPS on DGX-100



* Figure adopted from Ren et al. Faster R-CNN: Towards real-time object detection with region proposal networks, PAMI, 39 (6), 2017

2017 AICity Detection Test Results

- Detection and Classification Results (vehicles, pedestrians, traffic signals, *etc*…)



1800 frames

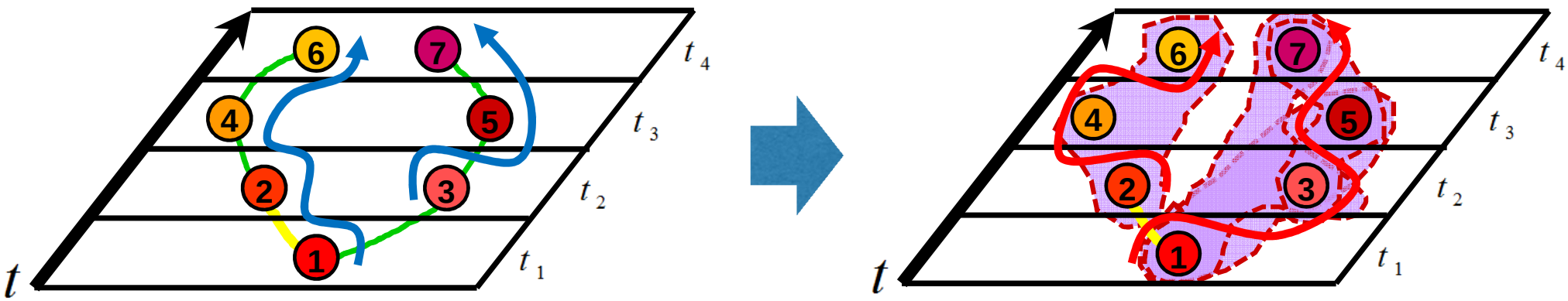
Hypergraph Tracker

Dense structure search on hypergraph to determine final tracks*

[L. Wen, Z. Lei, M.-C. Chang, H. Qi, and S. Lyu, "Multi-Camera Multi-Target Tracking with Space-Time-View Hypergraph", IJCV Special Issue, Sep., 2016]

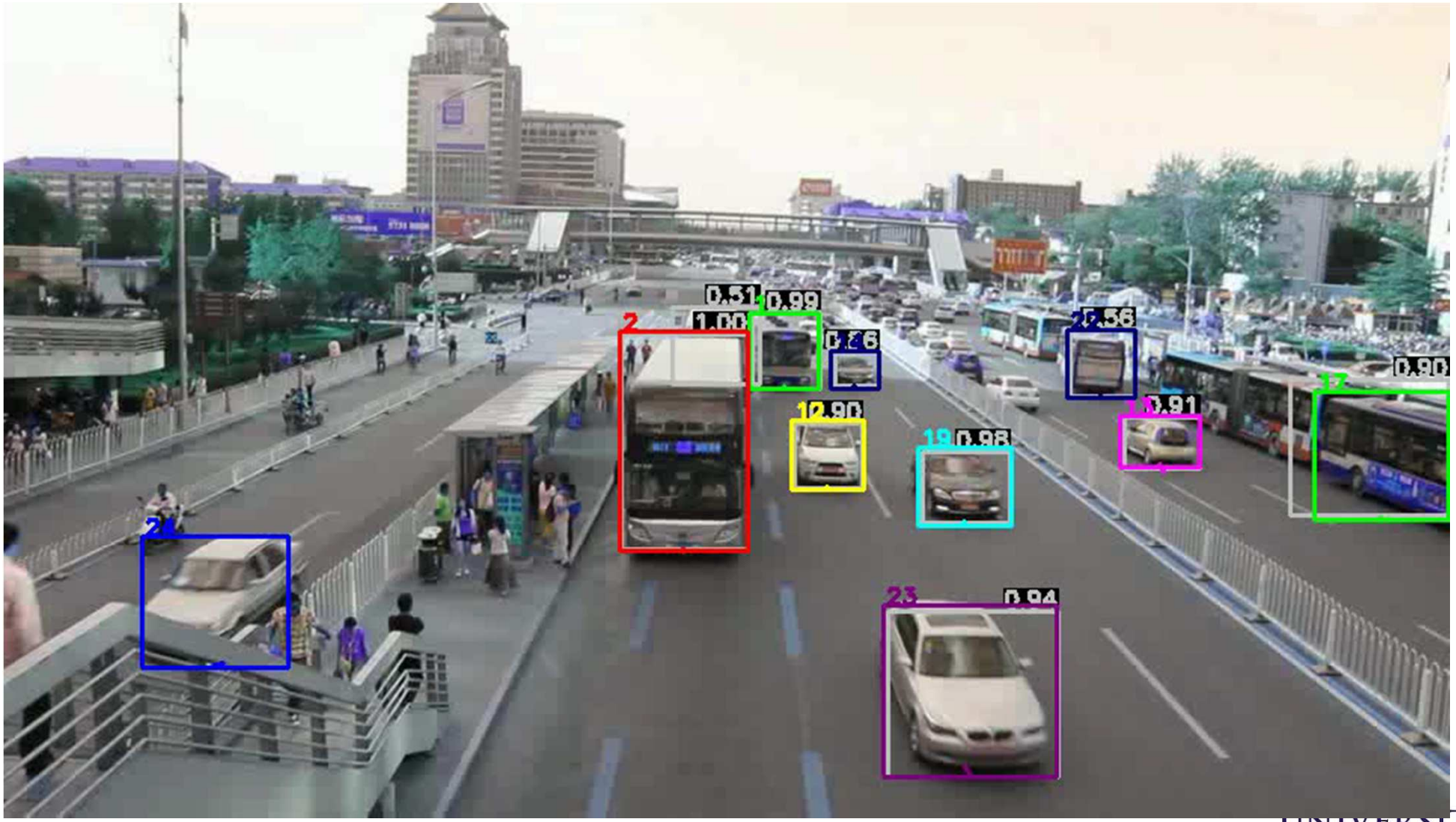
Graph (pairwise association) Hypergraph (high-order association)

in considering appearance, motion, track smoothness…



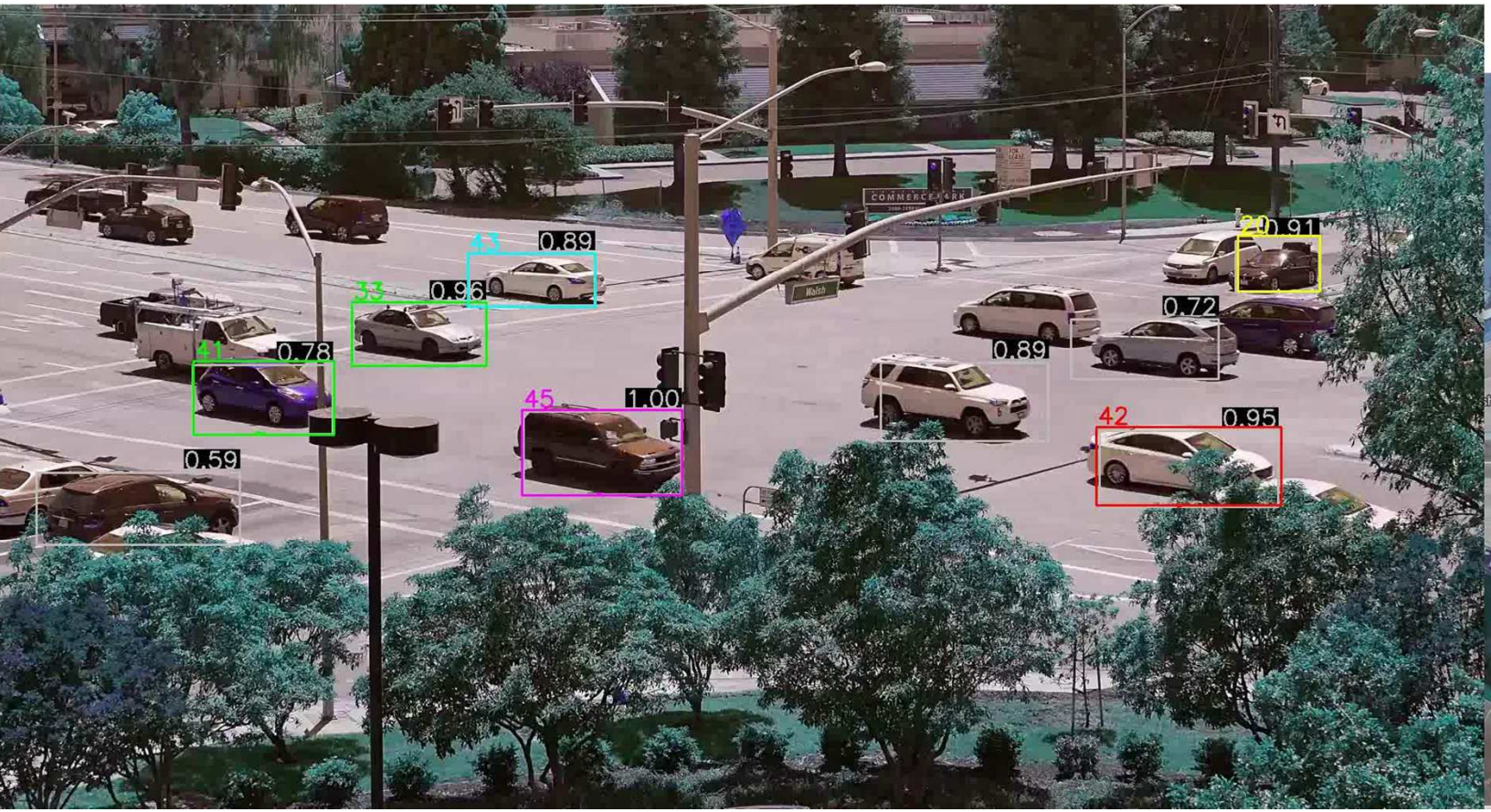
UA-DETRAC Test Results

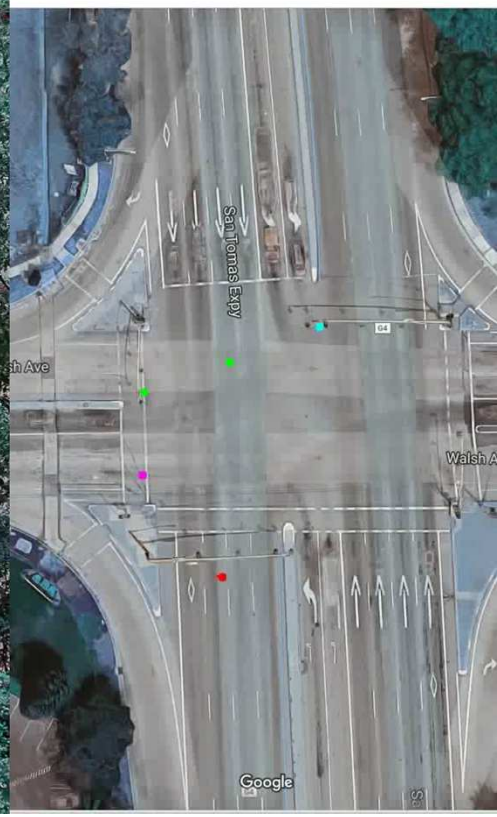
- Sequence 39511



2017 AICity Vehicle Tracking Results

- Walsh & Santomas 47





video length 30:00

Top-down view from Google Map

Site Calibration

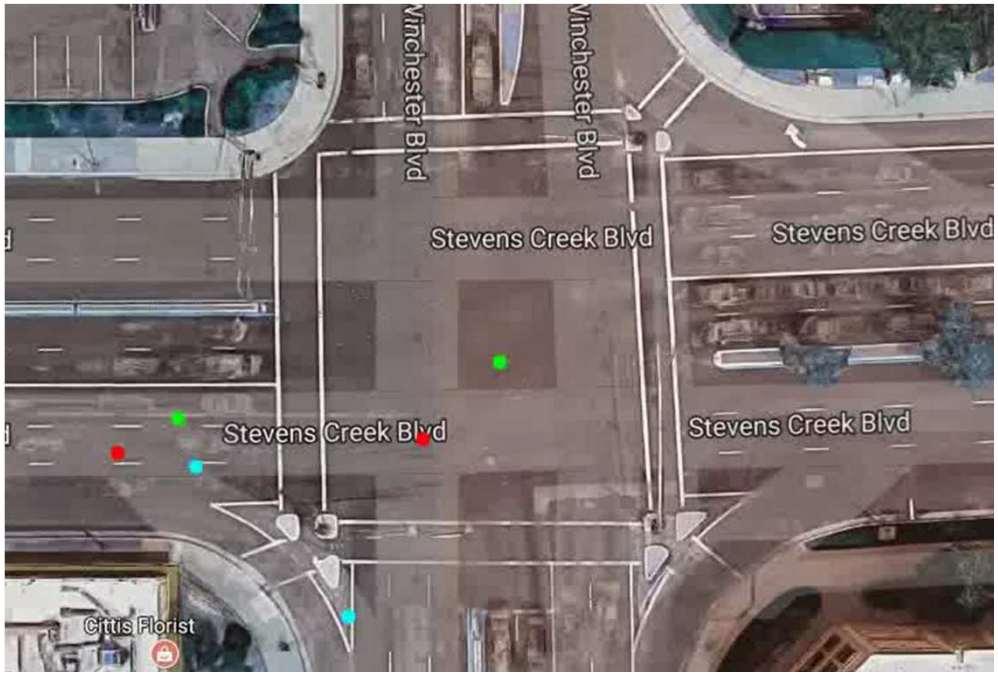
- Establish a mapping between pixels and the physical world
  - Manually estimate landmark dimensions
  - Calculate camera projection matrix
  - Project video view into a top-down ground-plane view

Traffic Analysis

- Site calibration to convert pixels into physical world coordinates
- Machine learning based on detection/tracking results

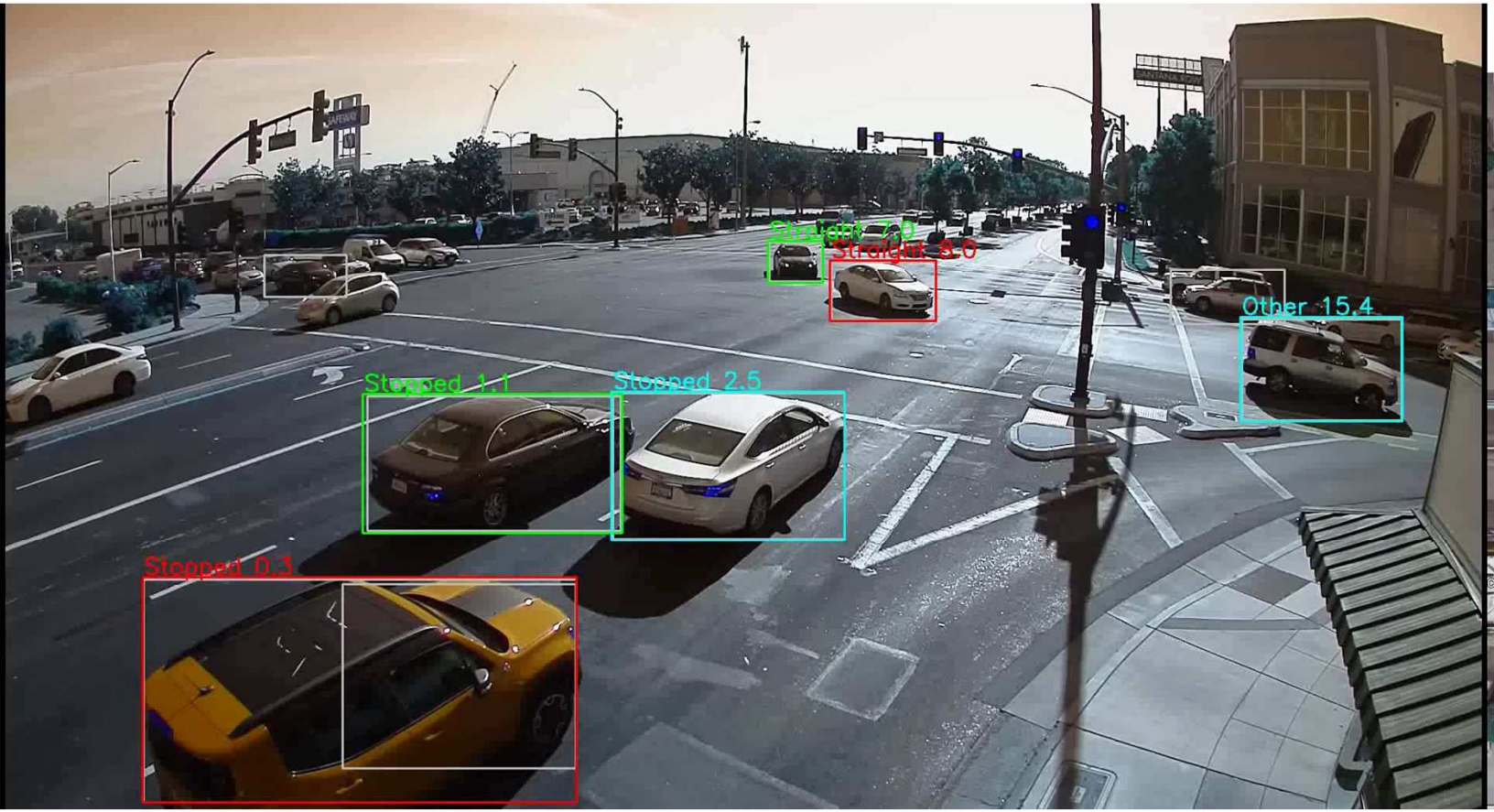
- For each vehicle track determine:
  - Incoming directions (1 of 4 at an intersection)
  - Traffic classification (going straight, left / right turns, U-turns, stopped, *etc*…)

Top-down view from Google Map 300 FPS



2017 AICity Traffic Analysis Results

Stevens & Winchester 1:
Visualizing the moving direction & velocity (MPH) for each vehicle.



video length 1:10:20

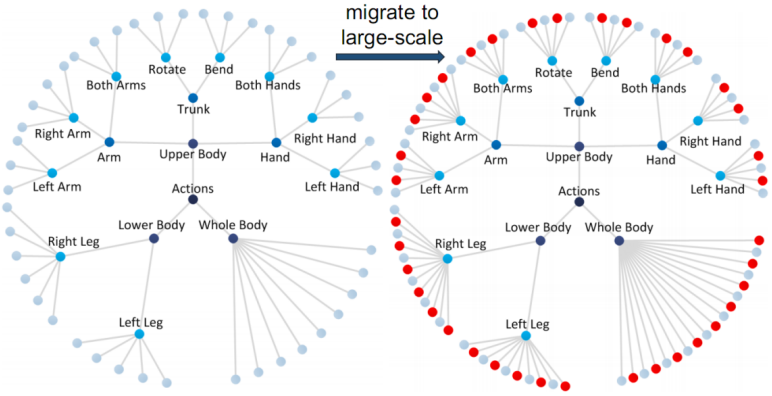
Speed unit: mile per hour

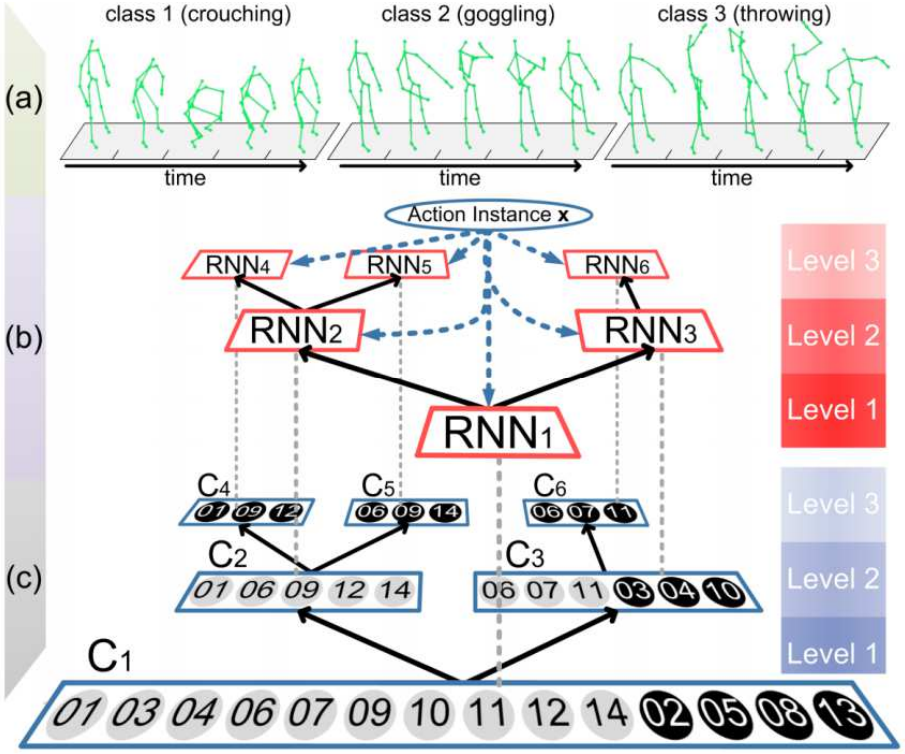
Top-down view from Google Map

Video Analysis Works

Human Action Recognition

[Wenbo Li, Longyin Wen, Ming-Ching Chang, Sernam Lim, and Siwei Lyu. "Adaptive RNN Tree for Large-Scale Human Action Recognition", ICCV'17]



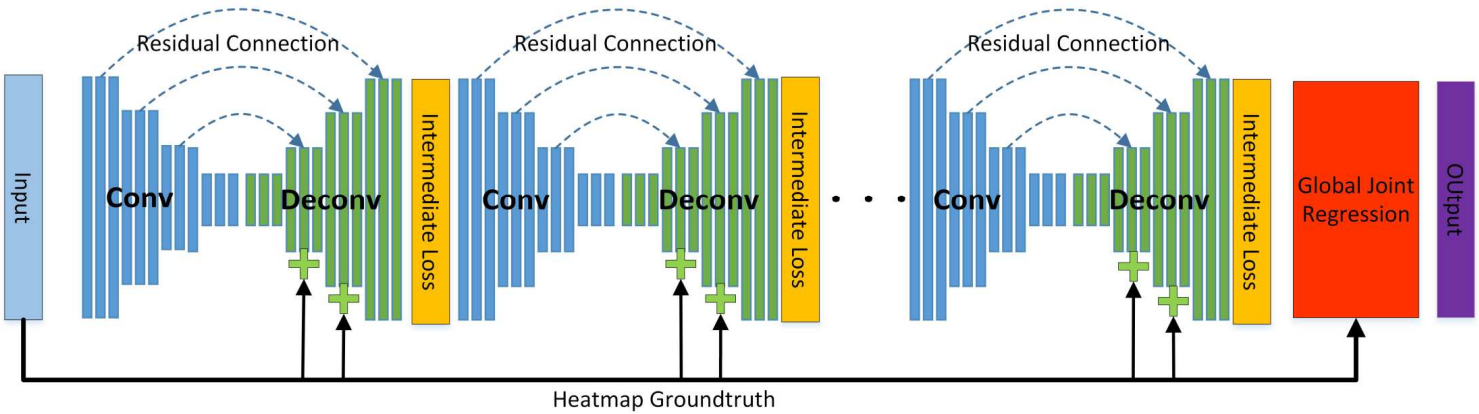


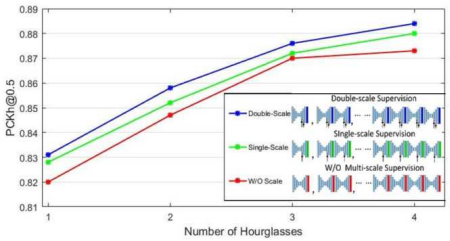
(a) Visualization of action instances from three action classes.
(b) A three-level RNN Tree (RNN-T) associated with the learned Action Category Hierarchy (ACH) in (c).
(c) Each circle represents an action class. Grey circles represent ambiguous classes, and black circles represent unambiguous ones. Action classes in the same box form one action category.

Human Pose Estimation

Hourglass modules with multi-scale supervision (for learned features) and global structural regression (to better handle scene clutters and multiple persons)

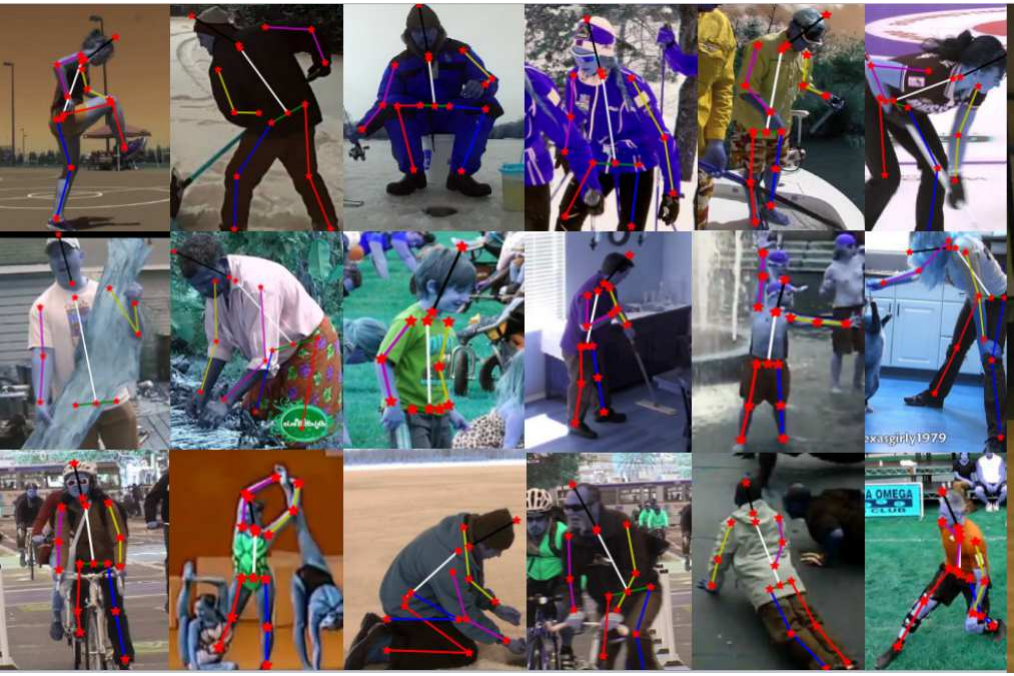
[under review]

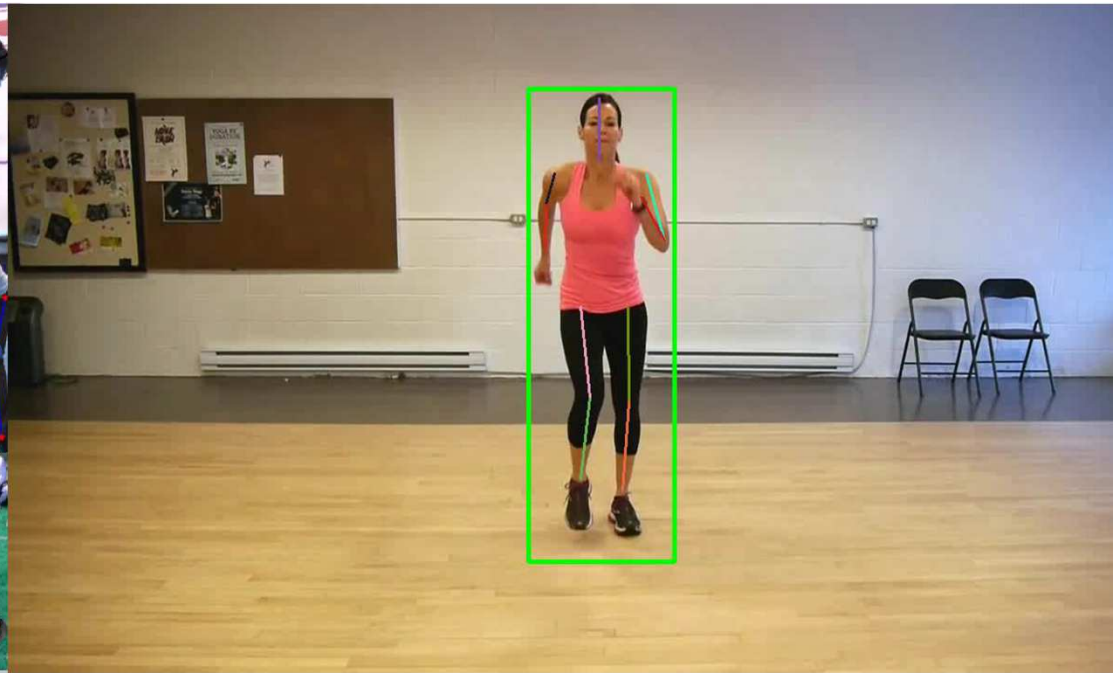




Example video: YouTube Hi-lo Aerobic Combo Set #1

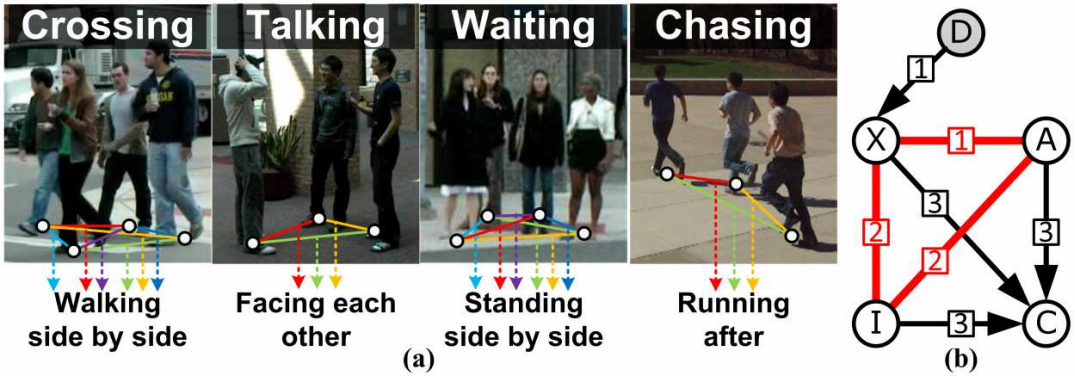
Per-frame person & pose detection, no tracking.





Group Behavior Recognition

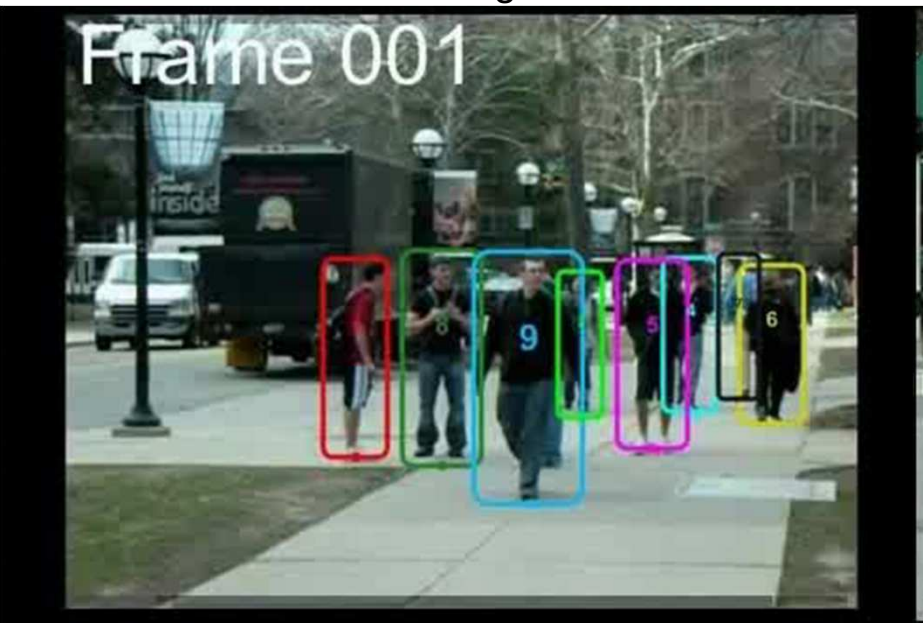
Who did what at where and when – Simultaneous tracking and group activity recognition



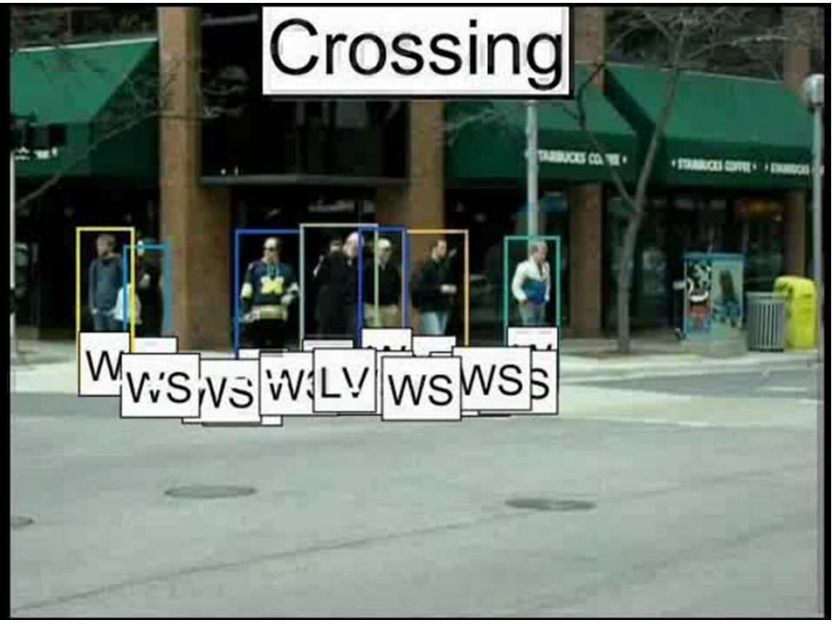
[ICCV'11]

[under review]

Tracking



Group Activity Recognition



S standing
W walking
R running
AP approaching
LV leaving
PB passing by
FE facing each other
WS walking side-by-side
SR standing in a row
SS standing side-by-side
NA no-interaction
WO walking in oppo dir
WR walking after
RS running side-by-side

Michigan Collective Activity Datasets

On-going Activities

CVPR 2018 Workshop

NVIDIA AI CITY CHALLENGE

CVPR 2018
Salt Lake City

Contact: nvidiaaicitychallenge@gmail.com

http://www.aicitychallenge.org/

http://cvpr2018.thecvf.com/program/workshops

Organization team: NVidia, U. Maryland, U. Albany, SJSU, U. Washington, U. Iowa

3 Challenge tasks:

- Traffic flow and volume estimation
- Abnormality detection (land violation, illegal U-turns, *etc.*)
- Multi-cam tracking and re-identification

AVSS 2019 Taiwan

- General Chair: Ming-Hsuan Yang, Yi-Ping Hung
- Core Committee: Ming-Ching Chang (PC), Y.-C. Frank Wang, JunWei Hsieh, Siwei Lyu, James Ferryman, Cosimo Distante, ...
- A back-to-back event with **ICIP 2019 Taipei**

Thank You