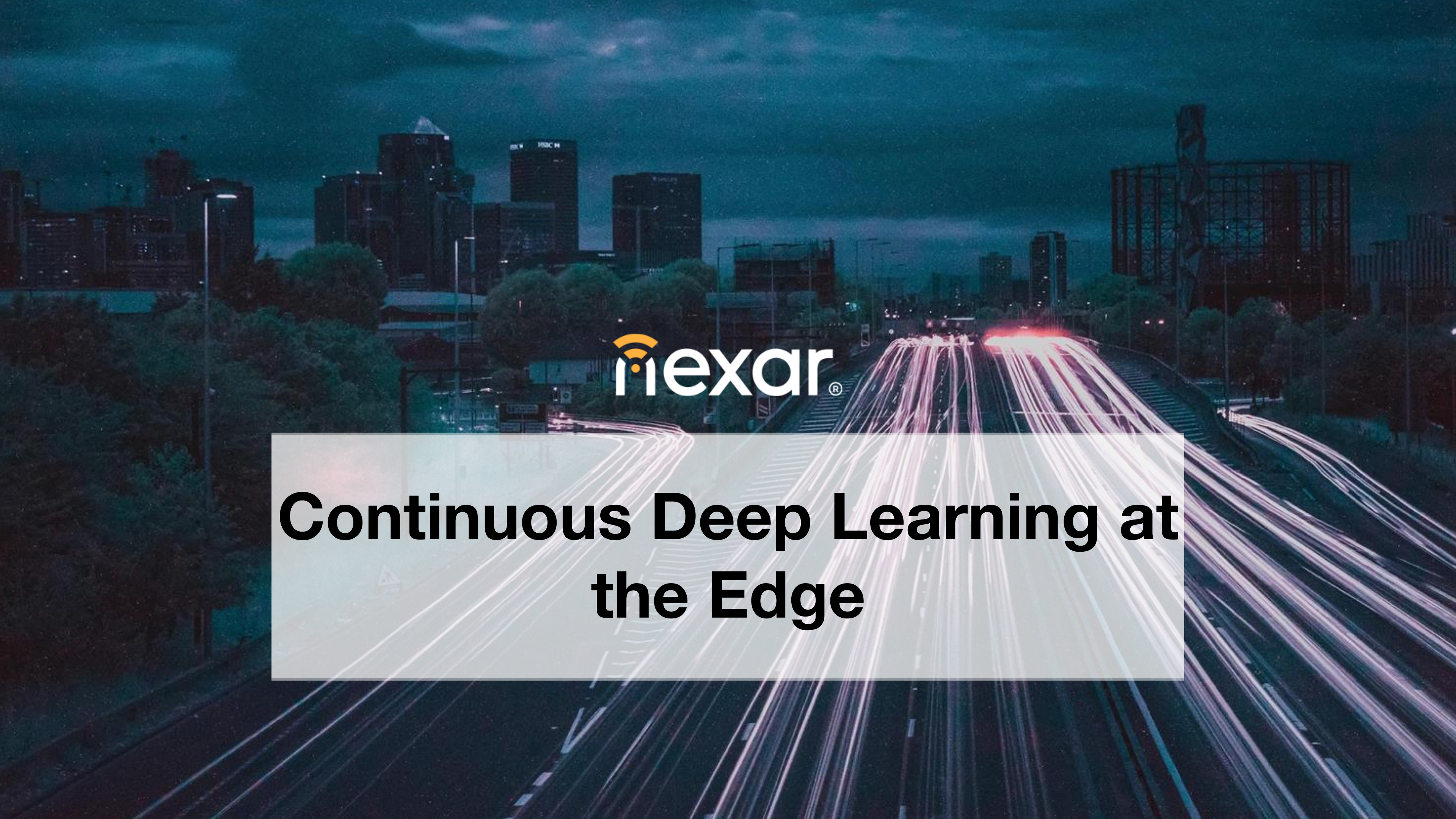


nexar®

Continuous Deep Learning at the Edge

Millions of eyes on the road

See, Understand and Communicate everything

- Millions of Nexar-connected cameras embedded in vehicles, bikes, and streets
- Drivers get 10-second warnings about dangerous road situations
- Autonomous Vehicles are plugged into a live stream of data about everything on the road
- Cities know exactly what's going on in every corner of the city

Nexar in 2018

A snapshot of where we've got to

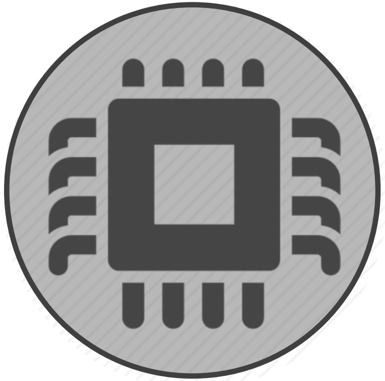
20m

Rare driving events

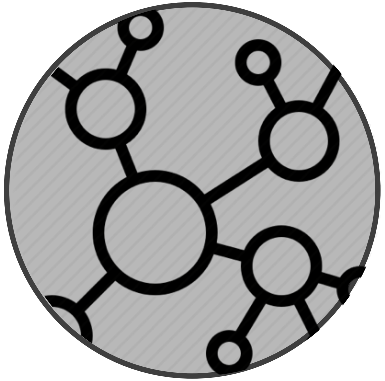
300m km

Road seen

Challenges Indexing the Road in Real-time



Limited compute at the edge



Limited networking at the edge

Limited exposure to corner cases

1. Addressing compute limitations in devices at the edge

Mass-Market AI in the Vehicle

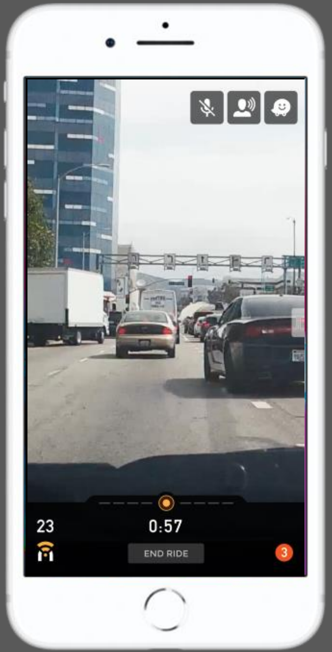
Commodity cameras paired with Nexar's software and cloud services using AI, deep learning and big data technology.

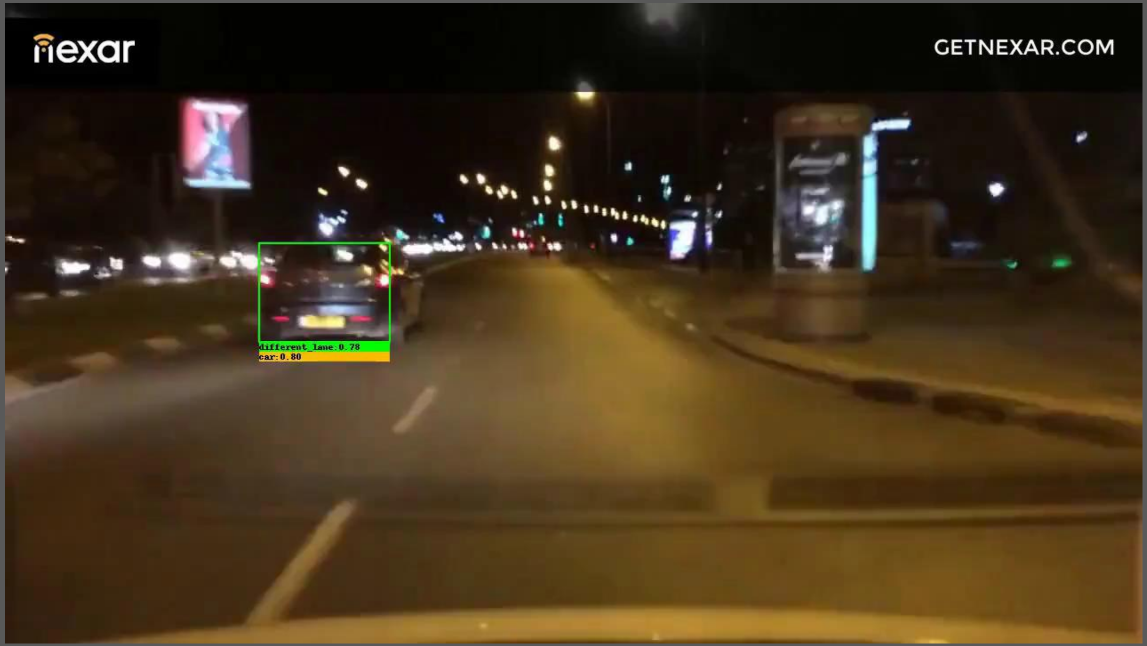
Starting at $50: **10x cheaper than alternative solutions.**

Turning simple dashcams into smart, **networked eyes**

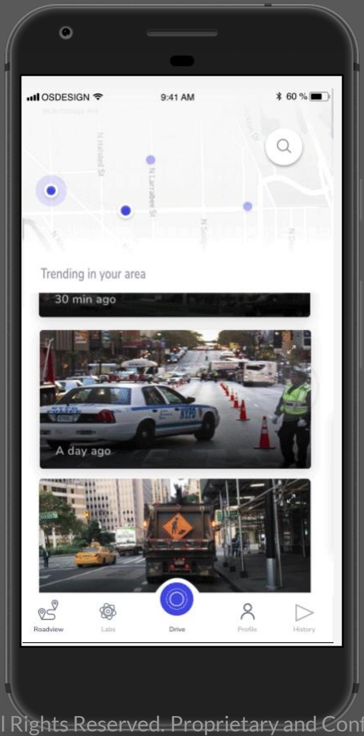
Networks inspired by SqueezeDet, YOLO, MobileNet, ShuffletNet and SkipNet.

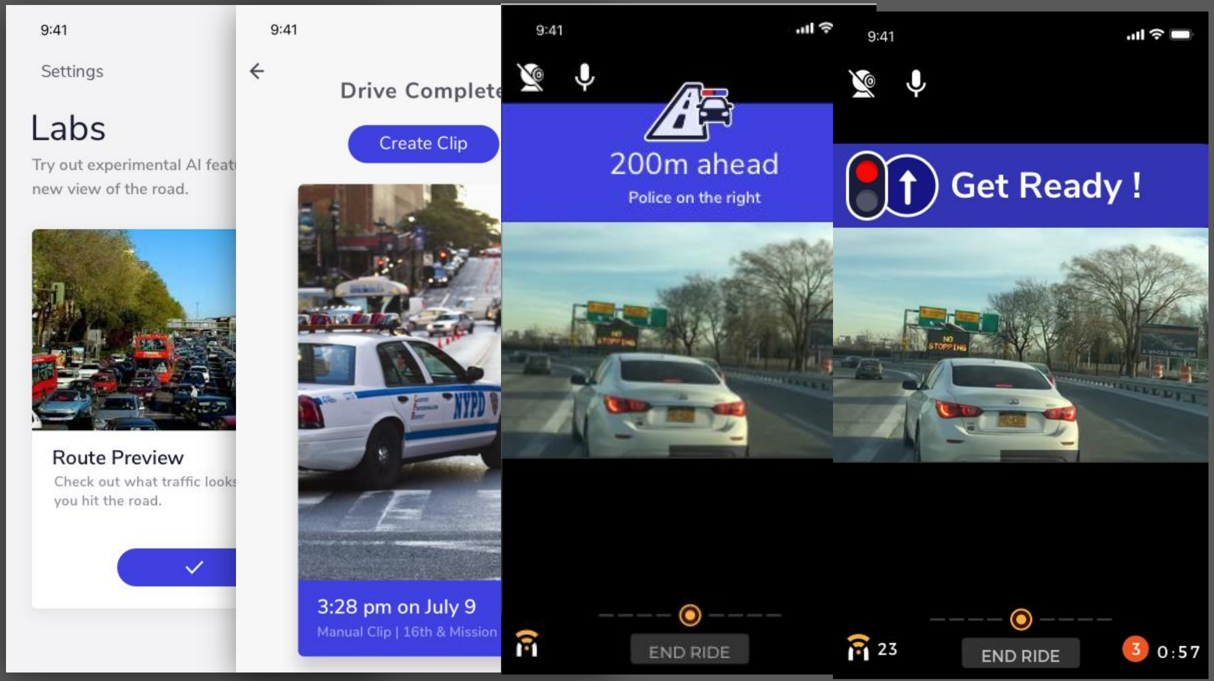
AI in the phone: Baby Steps



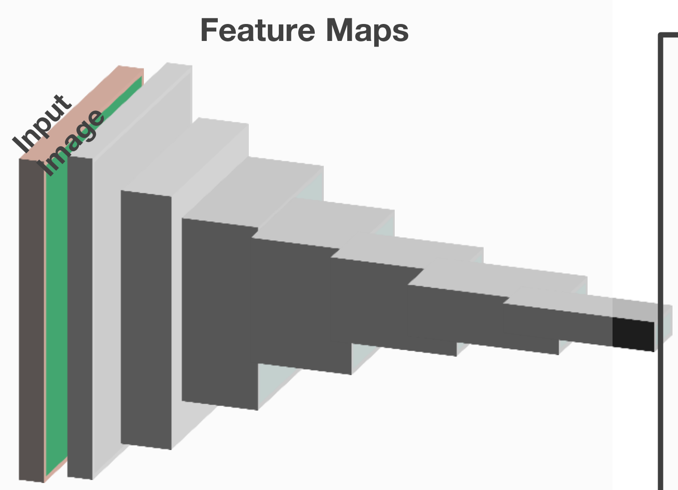


Increasing concurrent AI features





Mobile Unified Multi-tasking Network

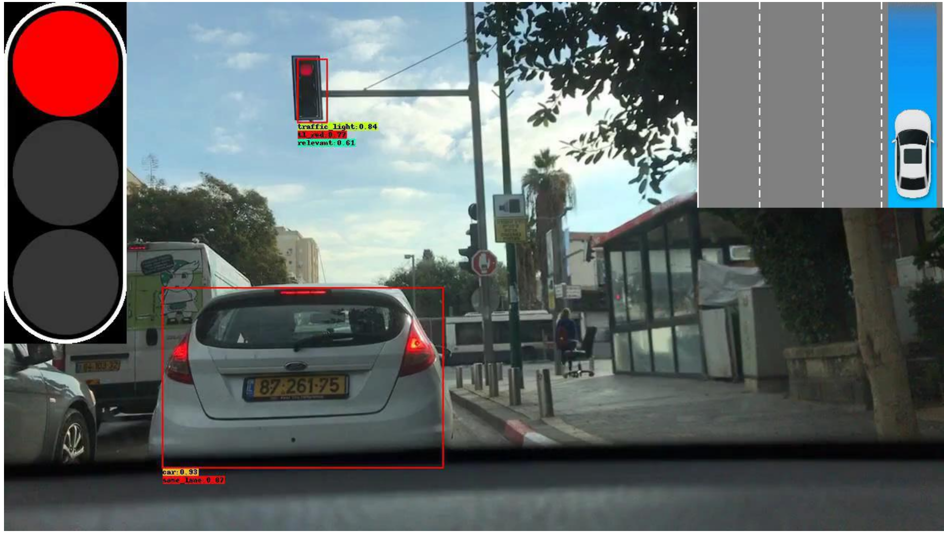


Model Size: 5.7MB | #Flops: 1.78G
Latency (mobile): 50ms

Single-Shot Road Understanding

- Detection
- Classification
- Relative Lane

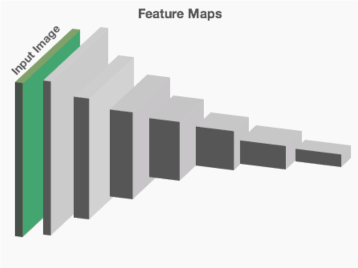
Network Output



Neural Net Performance as a Background Task

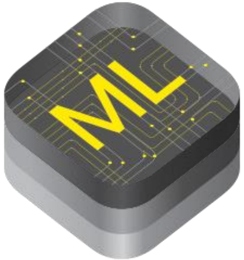
- App running on the background only gets access to the CPU.
- App must be highly efficient and respect the fair scheduling for foreground apps.

Efficient Architecture



Small Input Images (Patches)

Leverage CoreML



Background Results (CPU)

- ❑ Runtime: **3.19ms**
- ❑ CPU Usage: **0.53%**
- ❑ Classification Accuracy: **90%**

(iPhone7)

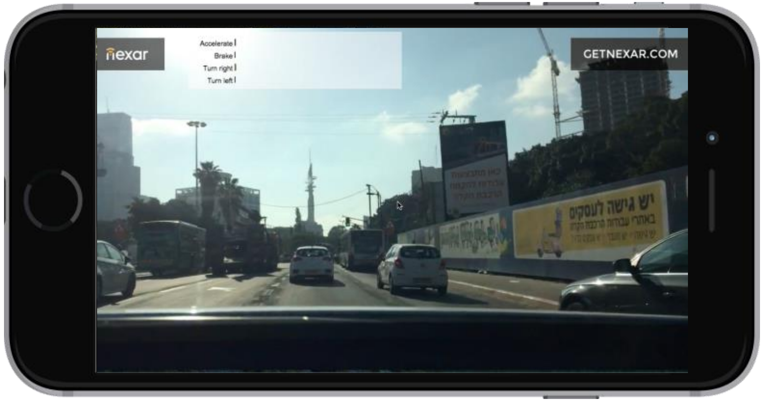
2. Addressing networking limitations in devices at the edge

+20m corner case videos

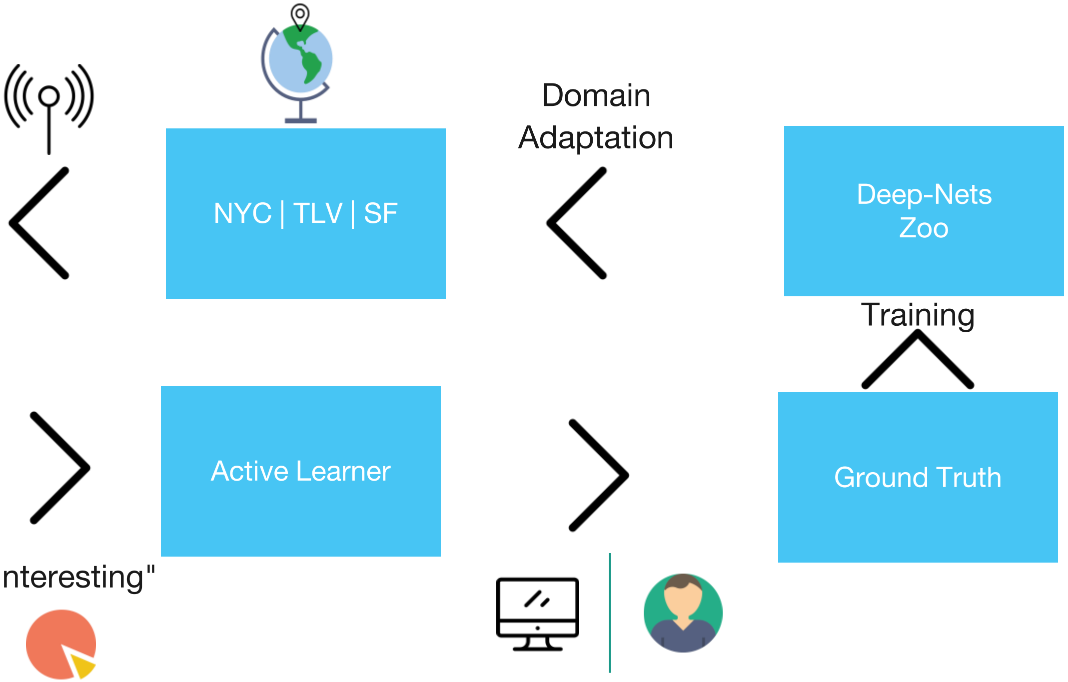
+100m annotated images

+80 countries

A Continuous Learning Pipeline



AI = Vision + IMU



Learning to detect "interestingness"

Start of road lane closure

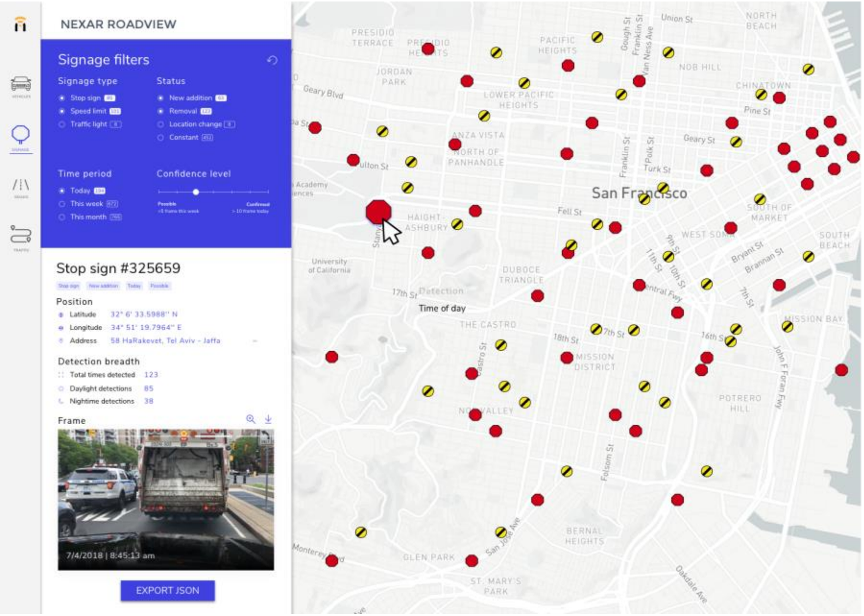
End of road lane closure

99% to 99.999999%: the power of corner cases

A Fresh View of the World

Leverage Nexar’s network **coverage** and **freshness** to:

- Understand Nexar drivers position, motion, path history and path self-prediction.
- Real time road radar, other cars, pedestrians, street signs, cyclists, street lights, police, etc.
- Proactively predict & prevent crashes. New Safety and Routing features above and beyond what’s available on the market.

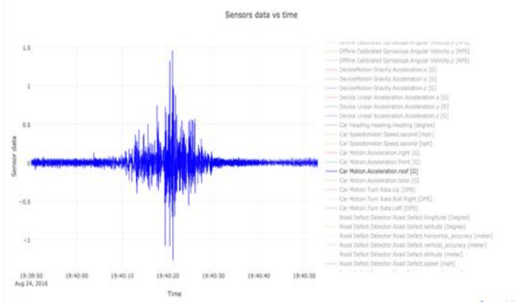


3. Addressing exposure bias in rule-based driving policies

Or … *once you have a LOT of data, what you can do differently with it*

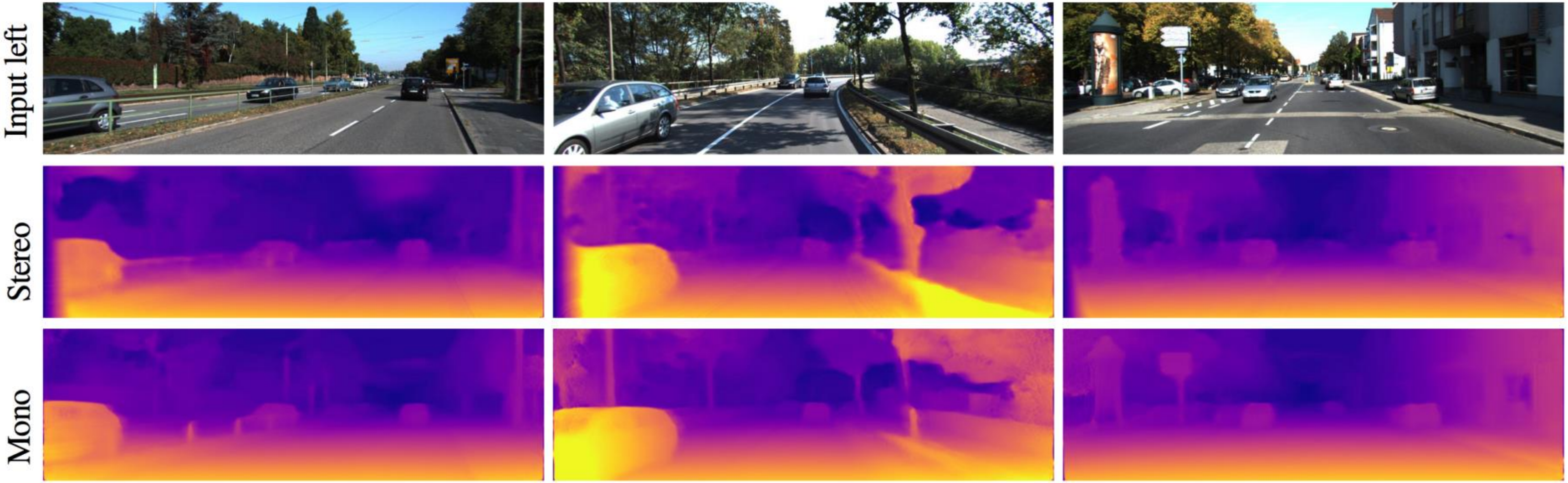
Cross-Modal Common Representation Learning

Using an additional modality (Inertial Measurement Unit – IMU) to learn visual road features: potholes, speedbumps, manholes.



Supervision using additional modalities

Pixel depth Estimation on monocular cameras

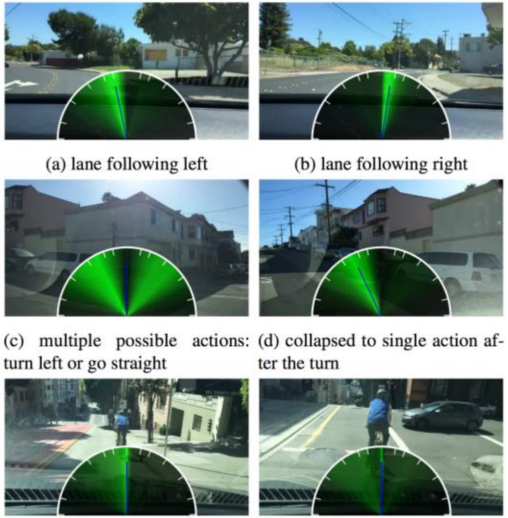


* Unsupervised Monocular Depth Estimation with Left-Right Consistency, Godard C. et al, CVPR 2017

End-to-end driving from large scale datasets

(a) go at yellow light (b) stop at red light

(c) stop & go equal weight at medium distance (d) stop when too close to vehicle ahead



(a) lane following left (b) lane following right

(c) multiple possible actions: turn left or go straight (d) collapsed to single action after the turn

(e) single sided prediction due to side walk (f) right turn becomes available at intersection

* End-to-end Learning of Driving Models from Large-scale Video Datasets, Huazhe X. et al, CVPR 2018

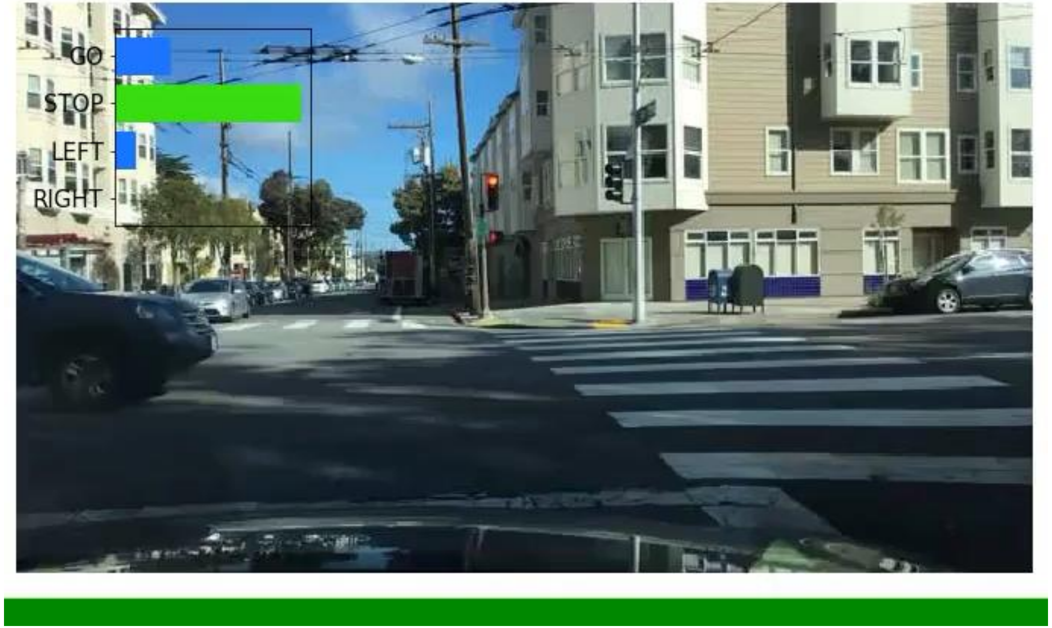
End-to-end learning of driving models from large scale video datasets

Holistic driving policy without requiring an intermediate semantic abstraction of the world and hierarchical selection of an optimal control policy.

* End-to-end Learning of Driving Models from Large-scale Video Datasets, Huazhe X. et al, CVPR 2018

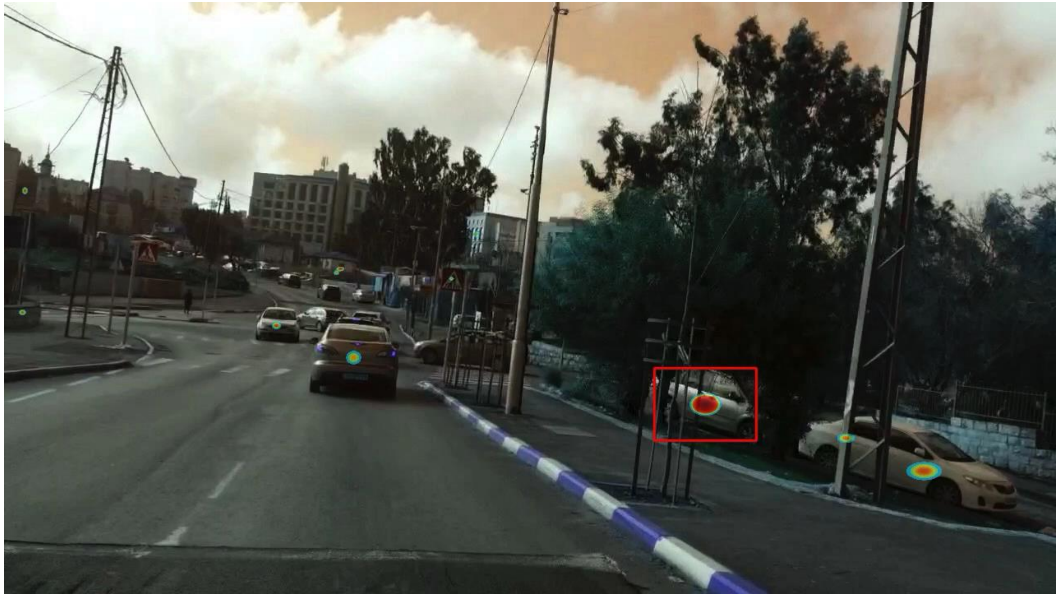
Learning holistic driving policies for mobile devices from privileged information

Using privileged information from road object detections at training time



Attention models for collision avoidance with end-to-end driving policies

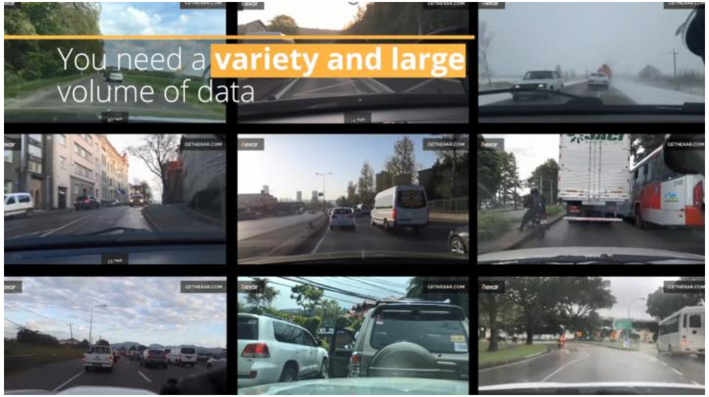
Using scene graphs with temporal encoding to detect collisions.



AV Driving in the Real World

Walking the talk: Nexar network open datasets

NEXET

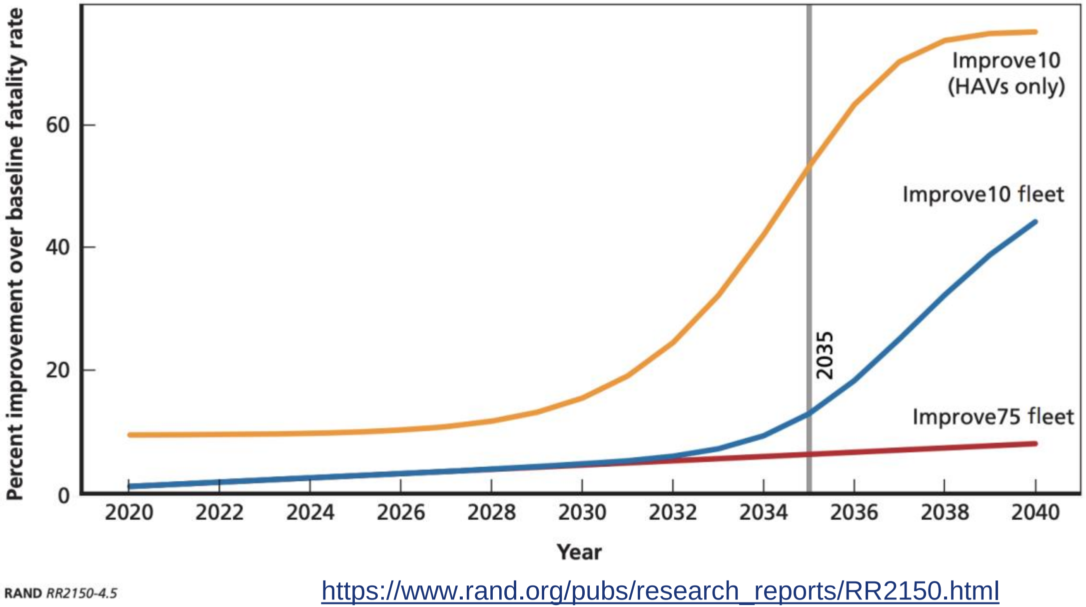


BDD100k

Multiple cities, multiple countries, multiple weathers, multiple times of day, multiple scene types, images, video sequences, image tags, road object bounding boxes, drivable area, lane markings, full-frame instance segmentation

The Enemy of the Good: The Need for Real-World Driving

- Waiting for AVs to be many times better than human drivers would be costly, 10s to 100s of thousands of lives over time.
- And just proving an AV drives just like a human would take 8.8 billion miles of testing.



https://www.rand.org/pubs/research_reports/RR2150.html

Thank You

Bruno Fernandez-Ruiz
b@getnexar.com
http://nxr.cm
@olympum