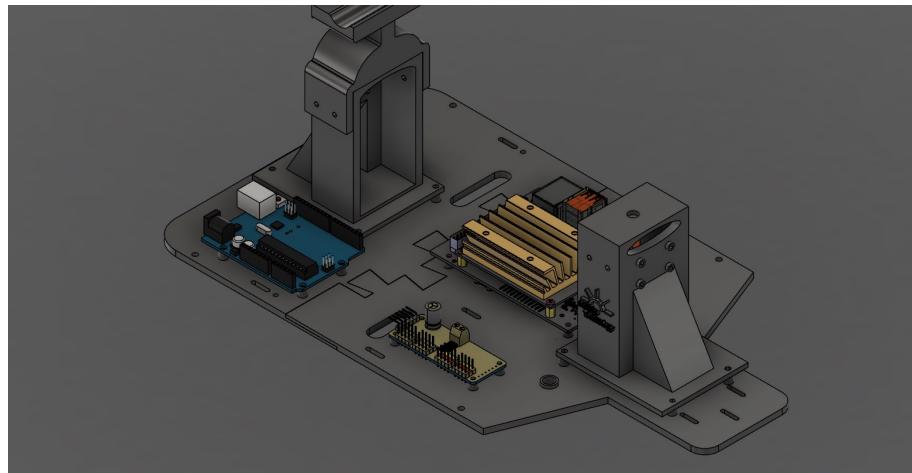


# A Deep Neural Network-Based Autonomous System Emulation Toolkits for Self-Driving Vehicle

Dantong Yu

Chief Scientist, Sunrise Technology,  
Associate Professor

July 2023





Cars



Trucks



Carts



Drones

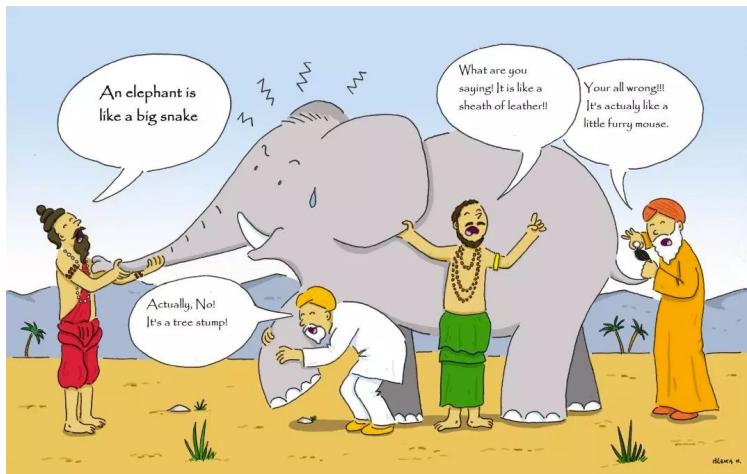
# Outlines

- Self-Driving Introduction and Motivation
- Self-Driving in real vehicles and RC-car
- Self-Driving Perceptron (Brain)
  - Behavior Clone Based Algorithm
- Our Openbot-based Architecture
- Topic 1: Build Reliable Self-driving Systems
- Topic 2: Obstacle Avoidance
- Self-Driving Challenges
- The drawback of Supervised Learning Algorithm
- Look into the Future and Conclusions



# What is Self-Driving

- ❖ What is the Self-Driving?
- ❖ An autonomous vehicle or driverless vehicle can operate itself and perform necessary functions based on road conditions without any human intervention through ability to sense its surroundings.



## Public Perception of V in Semi-Autonomous



SAE level	Name	Narrative Definition	Execution of Steering and Acceleration/Deceleration	Monitoring of Driving Environment	Fallback Performance of Dynamic Driving Task	System Capability (Driving Modes)
<i>Human driver monitors the driving environment</i>						
<b>0</b>	No Automation	the full-time performance by the human driver of all aspects of the dynamic driving task, even when enhanced by warning or intervention systems	Human driver	Human driver	Human driver	n/a
<b>1</b>	Driver Assistance	the driving mode-specific execution by a driver assistance system of steering or acceleration/deceleration, monitoring information about the driving environment and with the expectation that the human driver perform all remaining aspects of the dynamic driving task	Human driver and system	Human driver	Human driver	Some driving modes
<b>2</b>	Partial Automation	the driving mode-specific execution by one or more driver assistance systems of both steering and acceleration/ deceleration, monitoring information about the driving environment and with the expectation that the human driver perform all remaining aspects of the dynamic driving task	System	Human driver	Human driver	Some driving modes
<i>Automated driving system ("system") monitors the driving environment</i>						
<b>3</b>	Conditional Automation	the driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task with the expectation that the human driver will respond appropriately to a request to intervene	System	System	Human driver	Some driving modes
<b>4</b>	High Automation	the driving mode-specific performance by an automated driving system of all aspects of the dynamic driving task, except those that does not respond appropriately to a request to intervene	System	System	System	Some driving modes
<b>5</b>	Full Automation	the full-time performance by an automated driving system of all aspects of the dynamic driving task under all roadway and environmental conditions that can be managed by a human driver	System	System	System	All driving modes



→ AI → Deep Learning

## SOCIETY OF AUTOMOTIVE ENGINEERS (SAE) AUTOMATION LEVELS



					
0	1	2	3	4	5
<b>No Automation</b> Zero autonomy; the driver performs all driving tasks.	<b>Driver Assistance</b> Vehicle is controlled by the driver, but some driving assist features may be included in the vehicle design.	<b>Partial Automation</b> Vehicle has combined automated functions, like acceleration and steering, but the driver must remain engaged with the driving task and monitor the environment at all times.	<b>Conditional Automation</b> Driver is a necessity, but is not required to monitor the environment. The driver must be ready to take control of the vehicle at all times with notice.	<b>High Automation</b> The vehicle is capable of performing all driving functions under certain conditions. The driver may have the option to control the vehicle.	<b>Full Automation</b> The vehicle is capable of performing all driving functions under all conditions. The driver may have the option to control the vehicle.

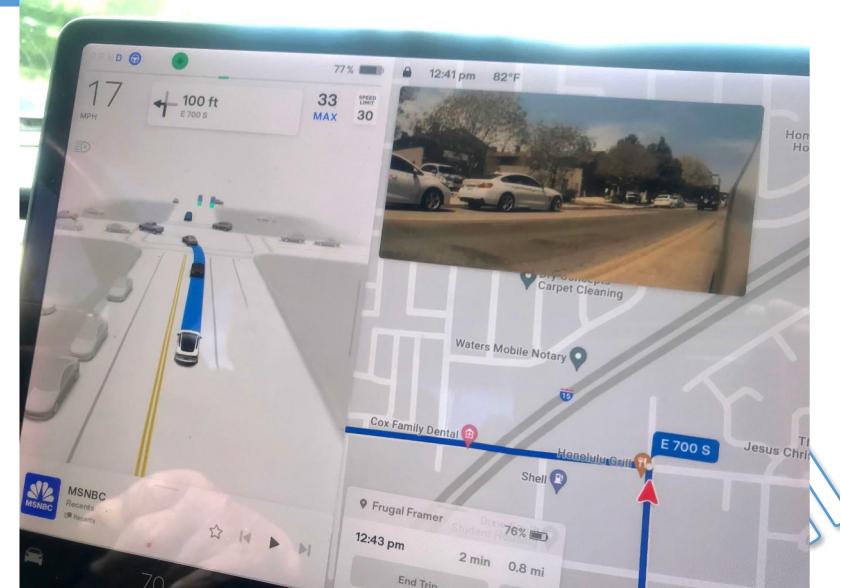
# Some Facts

- 1.35 million deaths worldwide due to vehicle crashes
- 94% of crashes involve human choice or errors.
- Three million Americans age 40 and older are not able to drive due to their vision/blindness
- 79% of seniors age 65 and older living in car-dependent communities
- 42 hours in traffic jams each year per person



# Machine Learning-based Self-Driving is a promising future of the workforce

- Autonomous Vehicle will Revolutionize our economy.
- The benefits of autonomous vehicles (AV) extend beyond safety – they could also greatly impact the labor market.
- AVs will create 300K US jobs in development, testing, maintenance, and deployment, as well as a need for technicians and mechanics with expertise in automation.
- <https://ts2.space/en/the-impact-of-autonomous-vehicles-on-employment-and-job-market/>



# Economy Fraction of Autonomous Driving

- ❖ How big the industry will be (Look at the celebrity company Tesla)
- ❖ It is one of the most exciting/critical applications of AI/Machine learning
- ❖ Solving it will bring changes to society, climate change, less dependence on fossil fuel, and bringing the sharing economy to a new level.

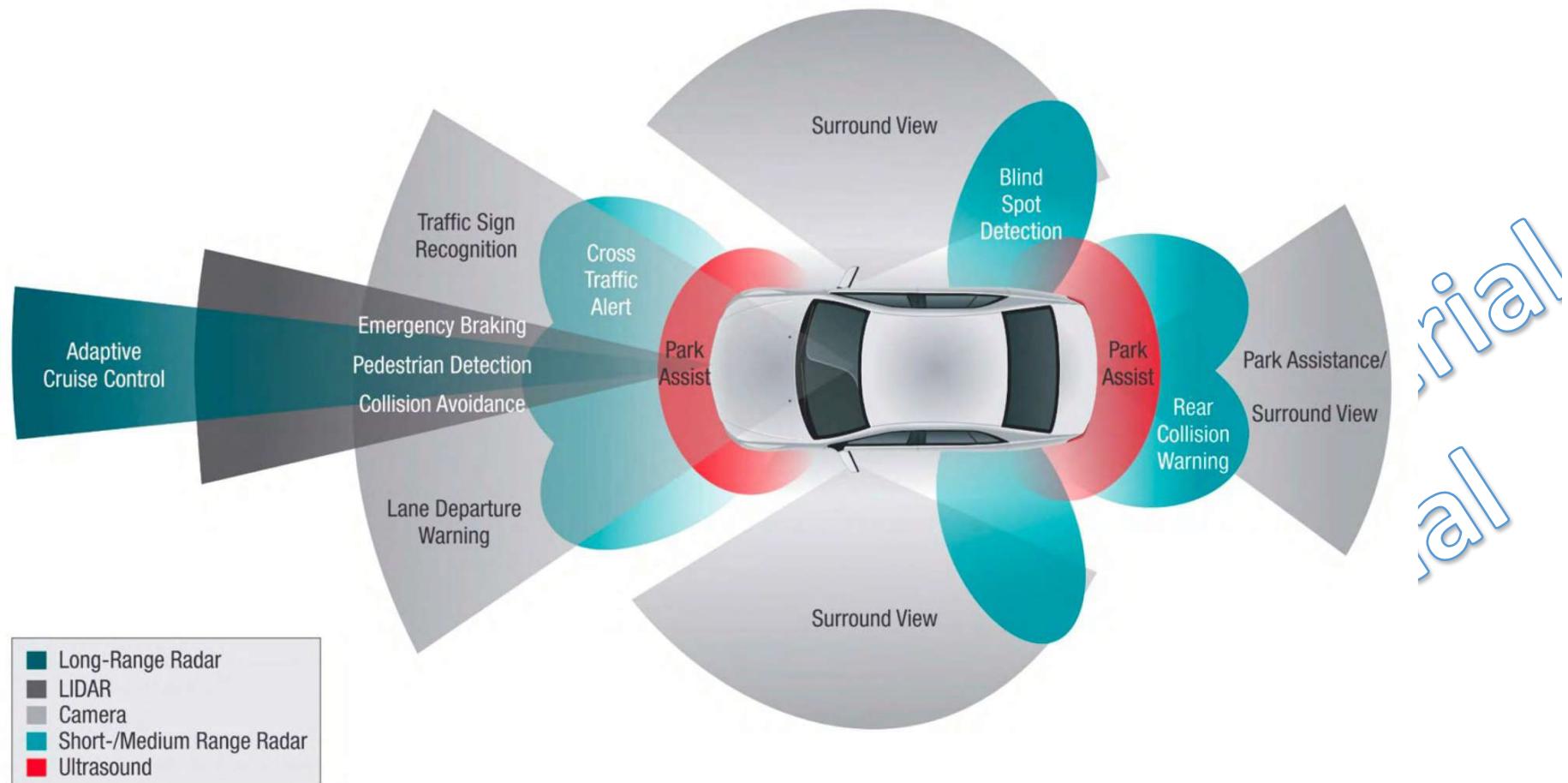


# Outlines

- Self-Driving Introduction and Motivation
- Self-Driving in real vehicles and RC-cars
- Self-Driving Perceptron (Brain)
  - Behavior Clone Based Algorithm
- Our Openbot-based Architecture
- Topic 1: Build Reliable Self-driving Systems
- Topic 2: Obstacle Avoidance
- Self-Driving Challenges
- The drawback of Supervised Learning Algorithm
- Look into the Future and Conclusions



# Tesla



# Similar Dynamics, different scales

**TRAXXAS XO-1**



**Tesla Model S**



# Traxxas 2WD Chassis



# F1/10 Used in College Self-Driving Course (\$5000)

**Chassis Design**

**F1/10 Platform Board**

**Hokuyo LiDAR**

**Zed Depth Camera**

**FLIR Flea Camera**

**Electronic Speed Controller**

**WiFi Telemetry**

**GPGPU Compute Platform NVIDIA Jetson TX1/TX2**

**System Integration**

**Software Architecture**

**ROS**

**Perception**    **Planning**    **Control**

**GPU Accelerated Libraries**

**Safe Autonomy**

**Secure Autonomy**

**Coordinated Autonomy**

**Efficient Autonomy**

**Lane Keeping Assist**

**AV Data Collection**

**F1/10 Simulator**

**Model-Predictive Control**

**Simultaneous Localization And Mapping (SLAM)**

**DNN Racing**

# We designed Two Systems for the Same Goal.

- We use two different hardware systems to implement autonomous driving.
- Enables flexibility in design and lowers the learning barriers

## Jetson

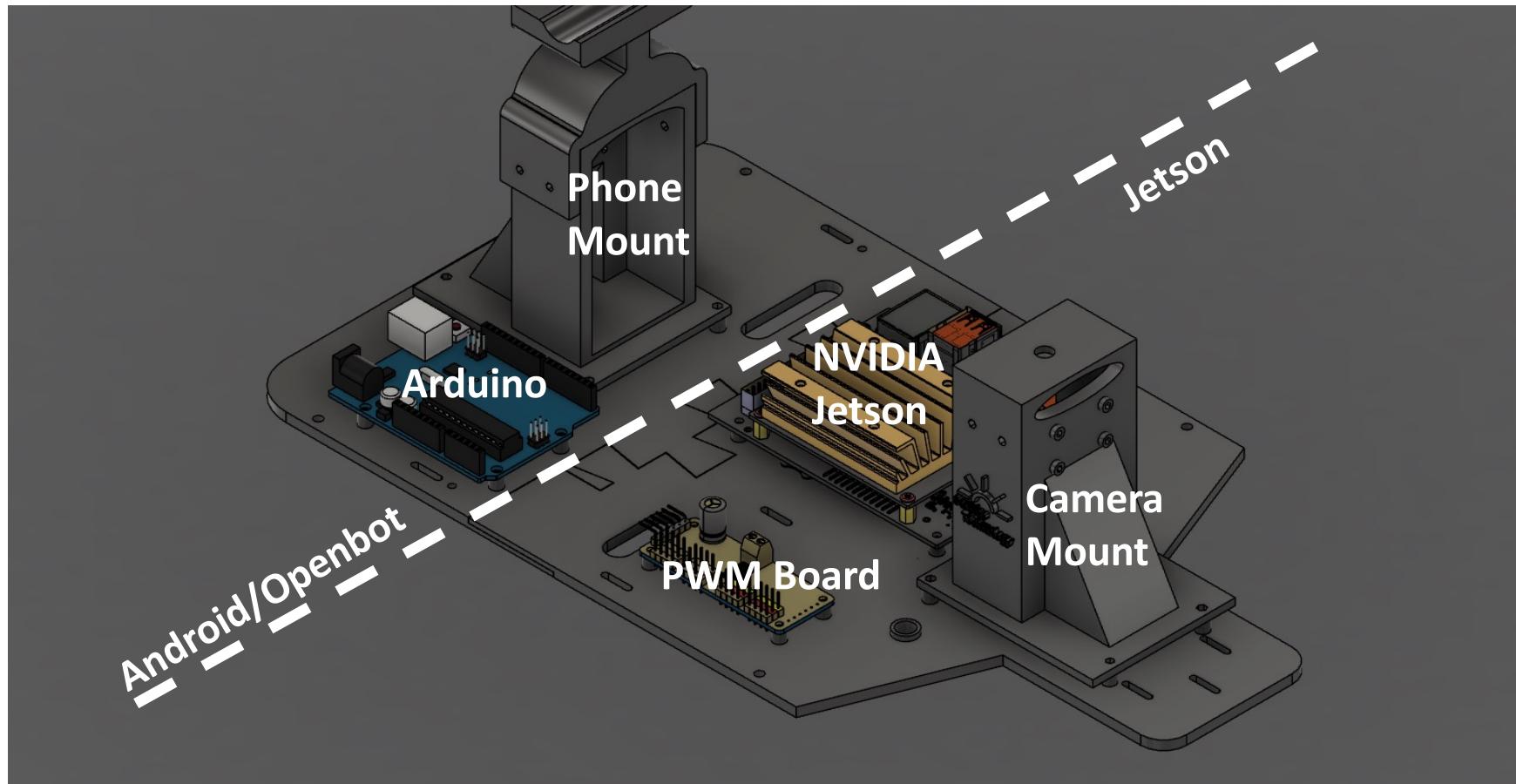
- Powerful embedded GPU
- ROS support for messaging
- Depth Camera

## Android

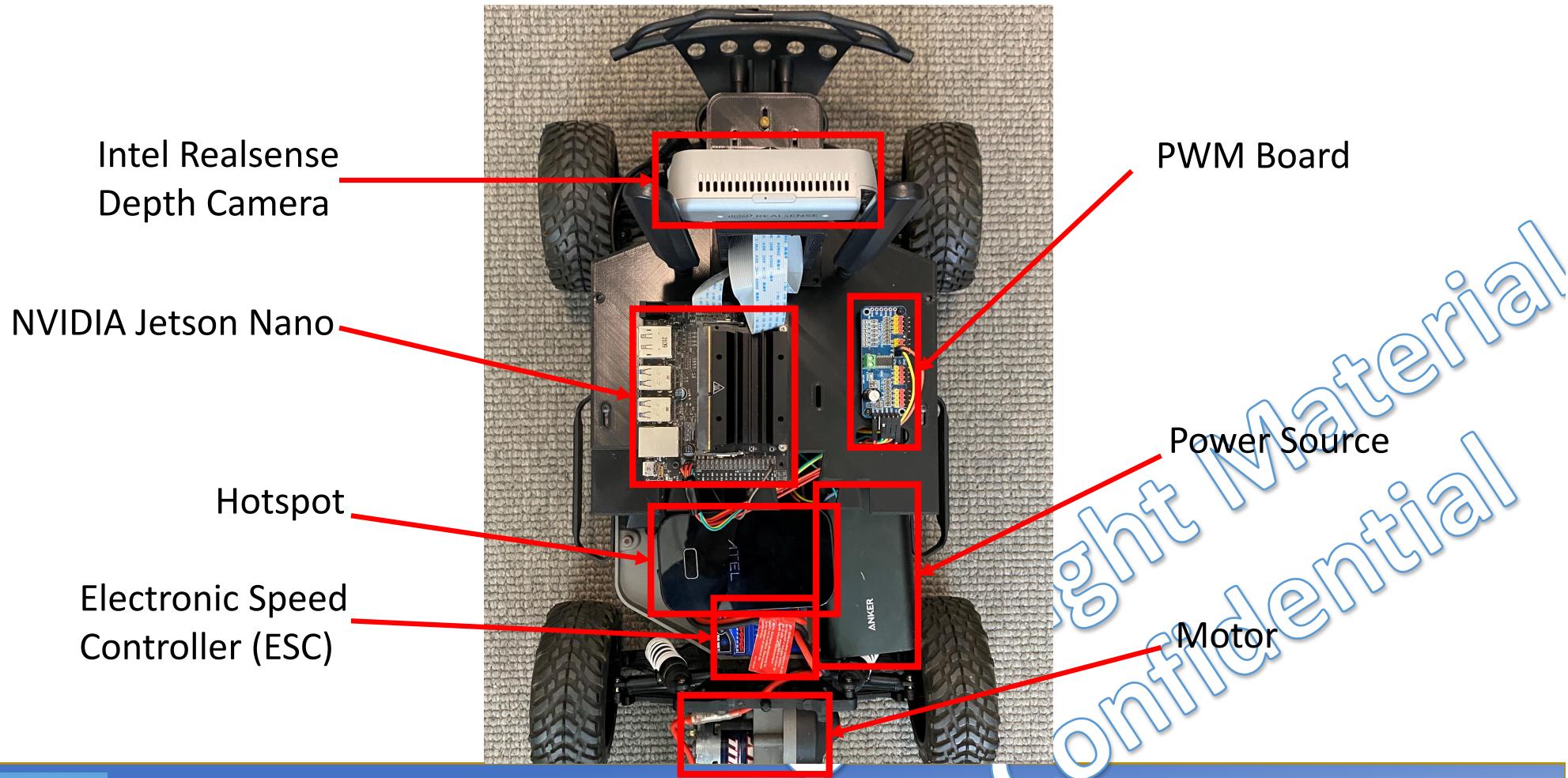
- Application based, very cheap
  - Google TPU
  - High camera resolution
- iPhone → Support LiDAR and 3-D scanning (Work-in-progress).



# Hardware support for two systems



# Jetson Hardware (Research)



# Openbot Based Self-Driving Car

## Sensor Integration



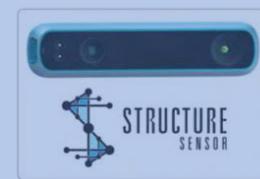
LiDAR



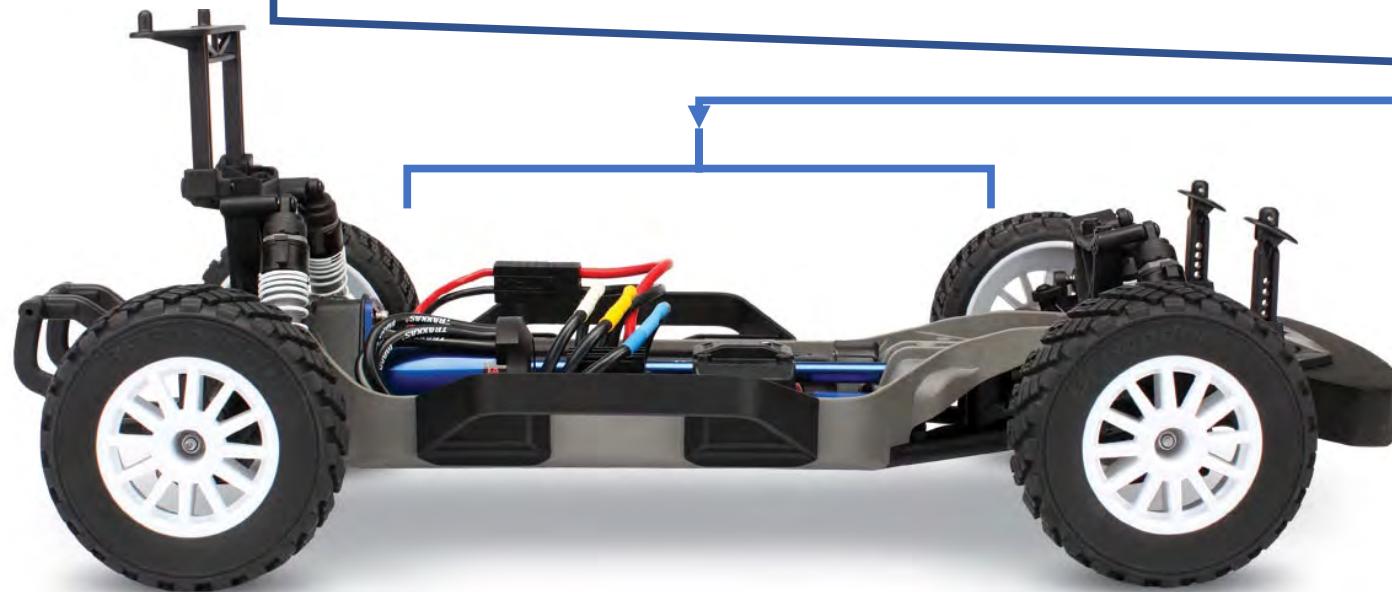
Camera



IMU



IR Depth Cameras

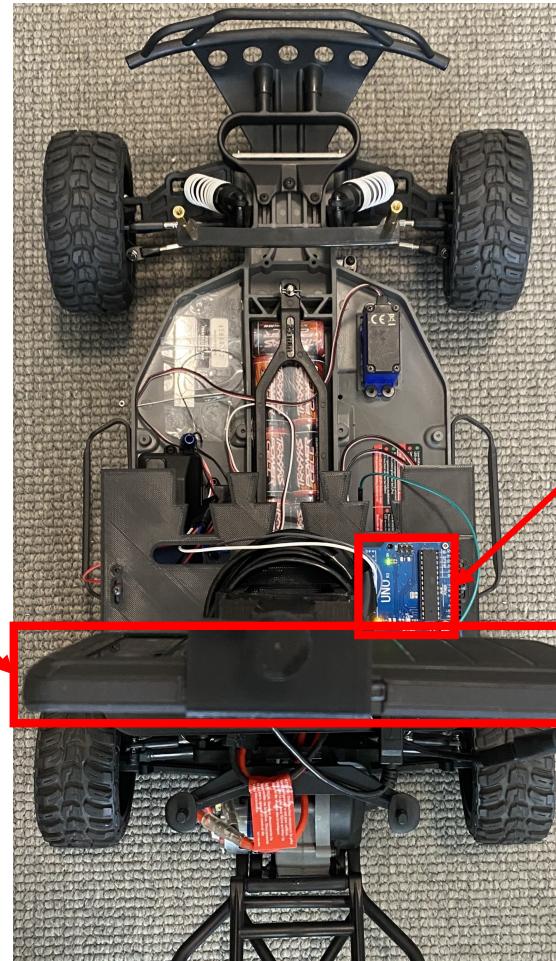


Self Driving

# Open Bot Hardware \$500

Samsung S23

Arduino UNO

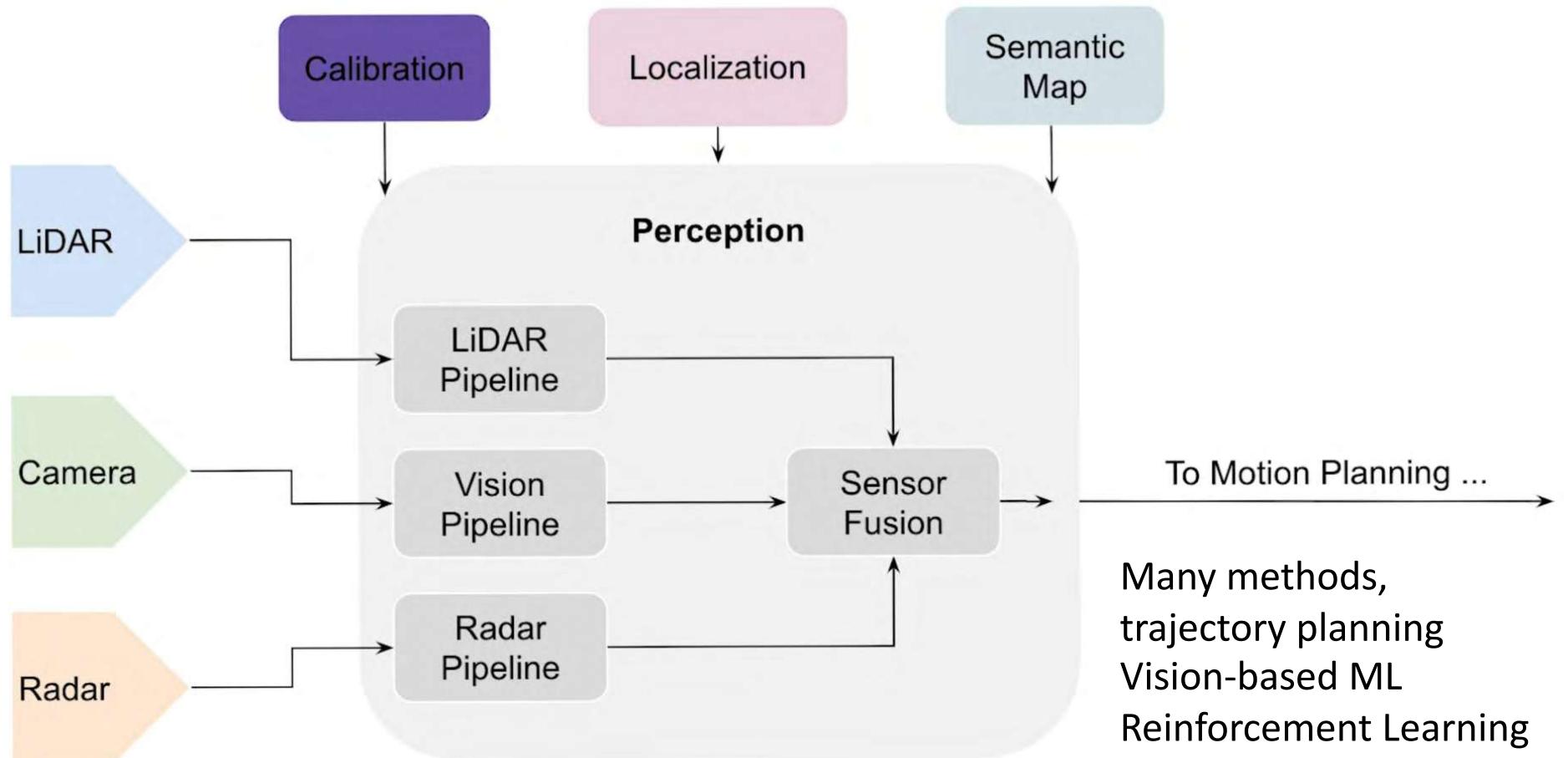


# Outlines

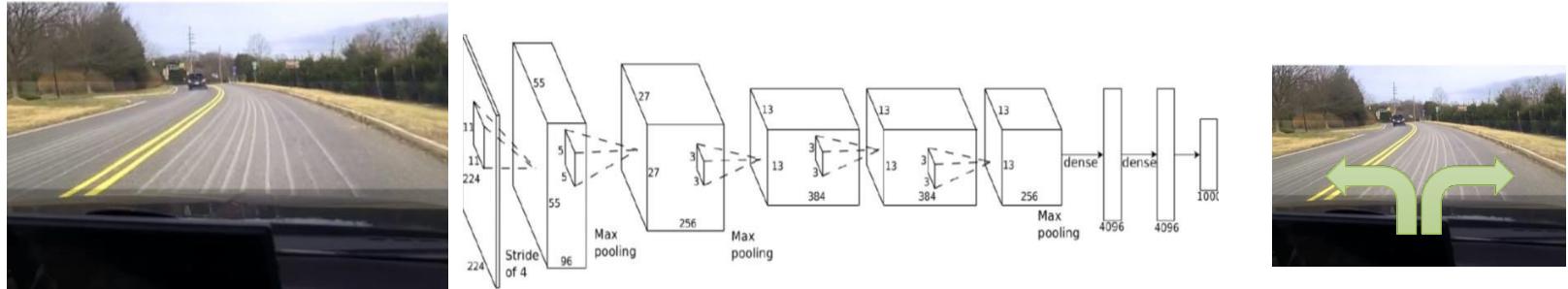
- Self-Driving Introduction and Motivation
- Self-Driving in real vehicles and RC-cars
- Self-Driving Perceptron (Brain)
  - Behavior Clone Based Algorithm
- Our Openbot-based Architecture
- Topic 1: Build Reliable Self-driving Systems
- Topic 2: Obstacle Avoidance
- Self-Driving Challenges
- The drawback of Supervised Learning Algorithm
- Look into the Future and Conclusions



# Perceptions in AV Stack



# Imitation Learning

 $\mathbf{o}_t$  $\pi_{\theta}(\mathbf{a}_t | \mathbf{o}_t)$  $\mathbf{a}_t$ 

Images: Bojarski et al. '16, NVIDIA

 $\mathbf{o}_t$  $\mathbf{a}_t$ 

supervised  
learning

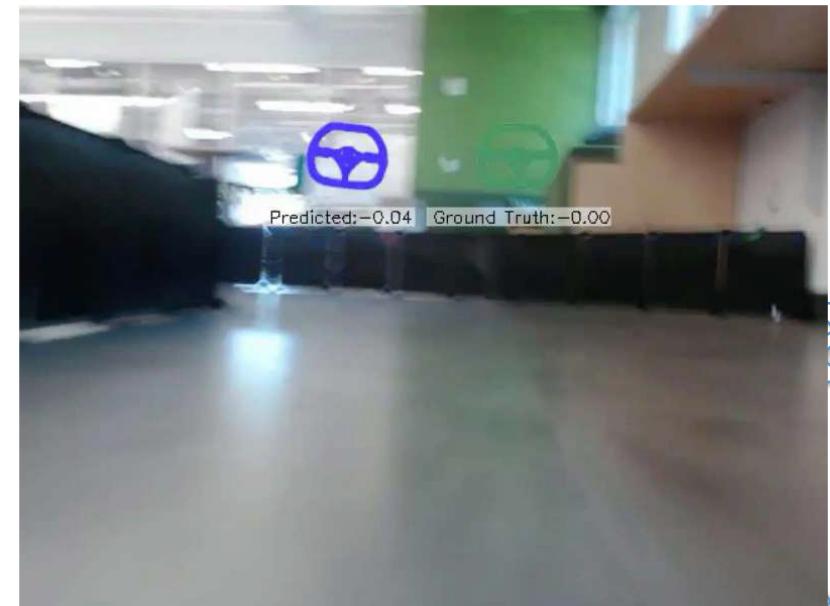
 $\pi_{\theta}(\mathbf{a}_t | \mathbf{o}_t)$

# End-to-End Driving

## Problem:

End-to-End driving with imitation learning and Behavior Clones.

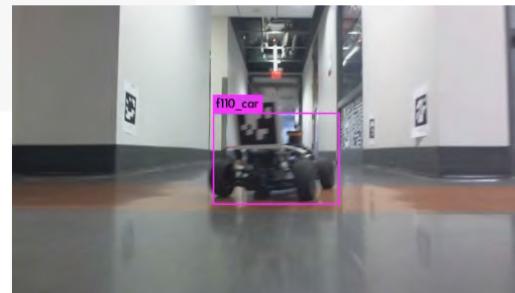
1. Train a network to map pixels to steering angle
2. Pixels to trajectory and throttles



# Examples



Wall Following



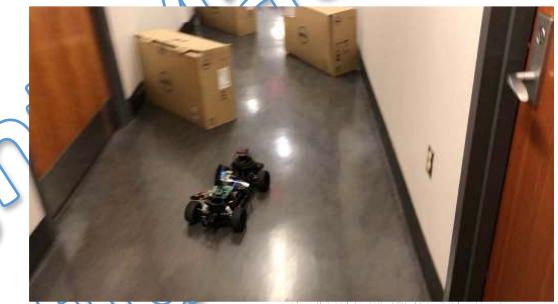
Object Detection and Tracking



Follow the track

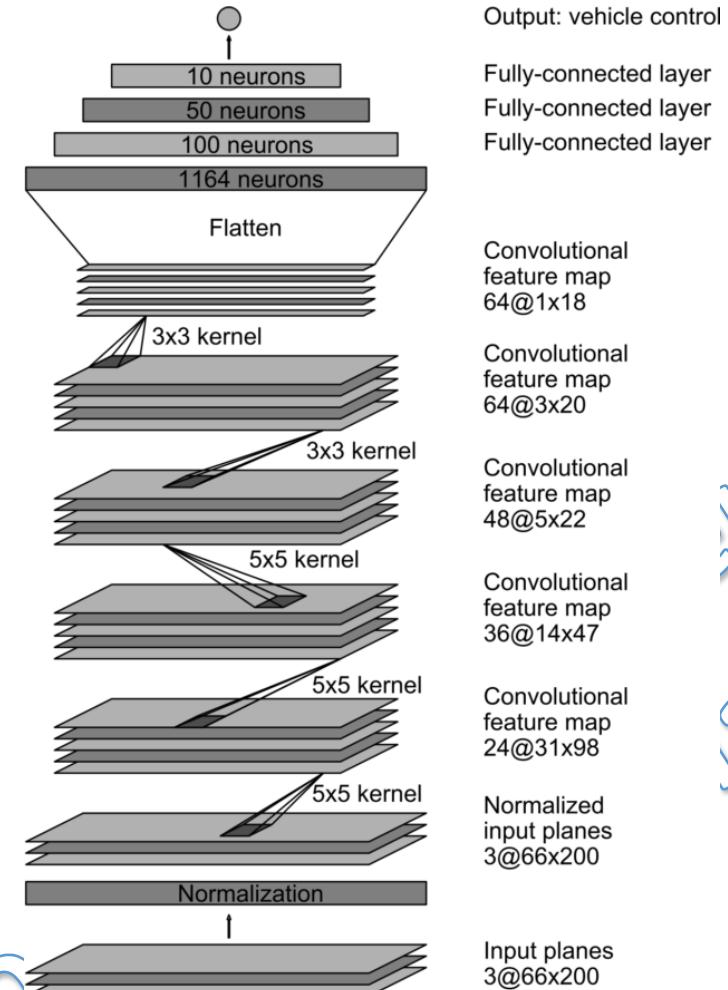
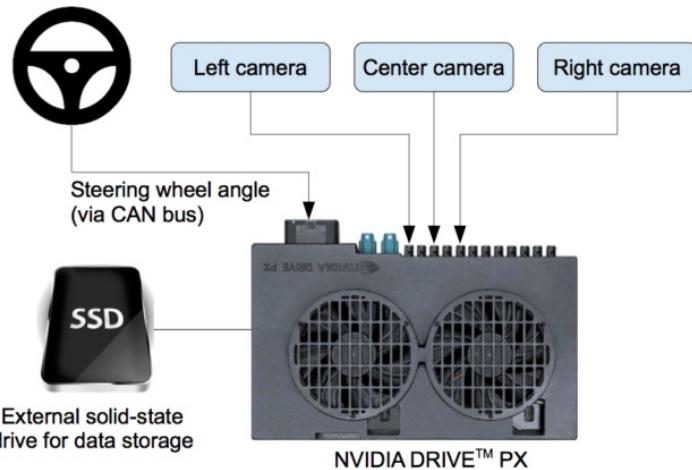


Your  
Project



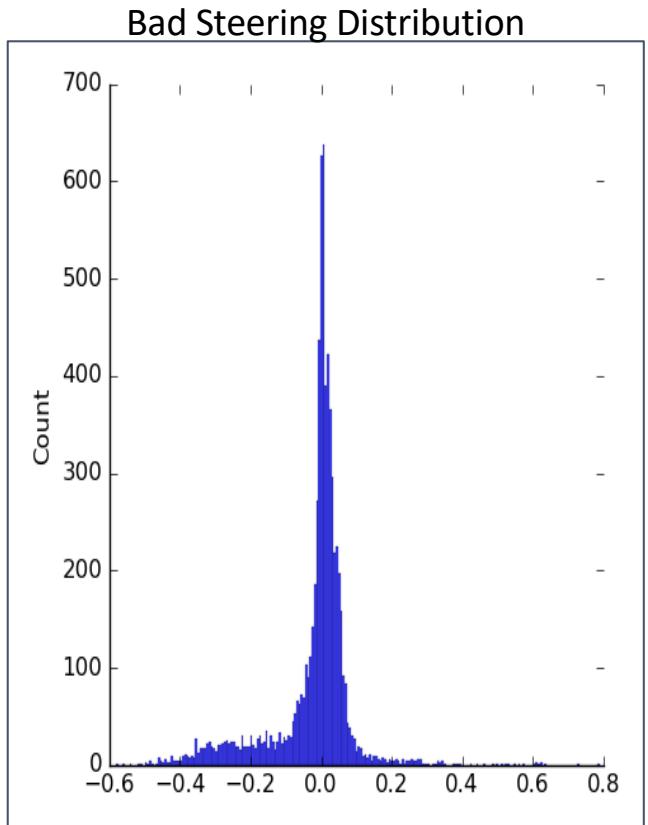
Obstacle Avoidance

# Behavior Cloning with Nvidia PilotNet

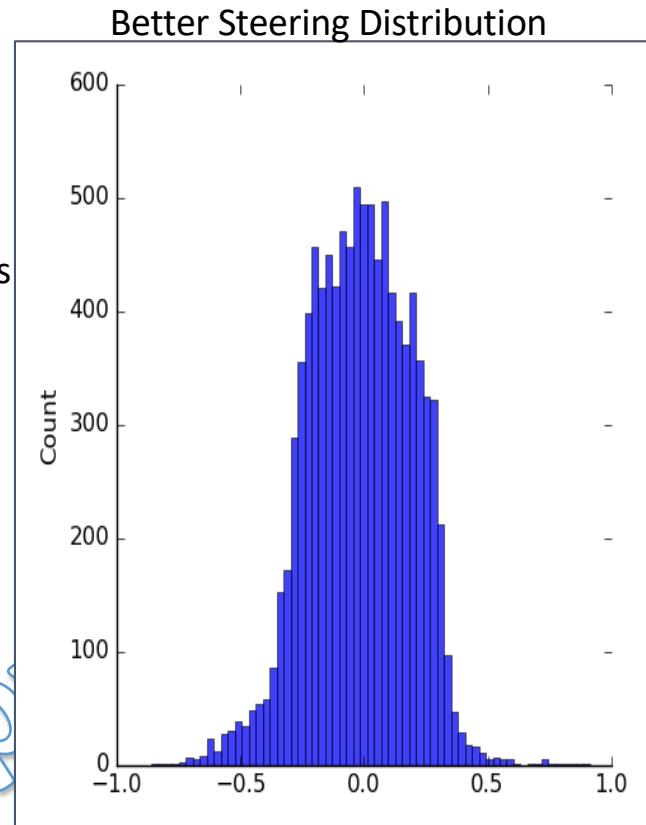


# Behavior Cloning : Bias

Understanding how the system collects data is only one part; good data collection involves being aware of bias.



Managing the bias of a dataset is crucial in getting good results.  
One way we can do this is by equally representing all events in the dataset, this can mean re-recording data several times.  
Considering most driving is done with 0 steering angle, it is good practice to record turns several times

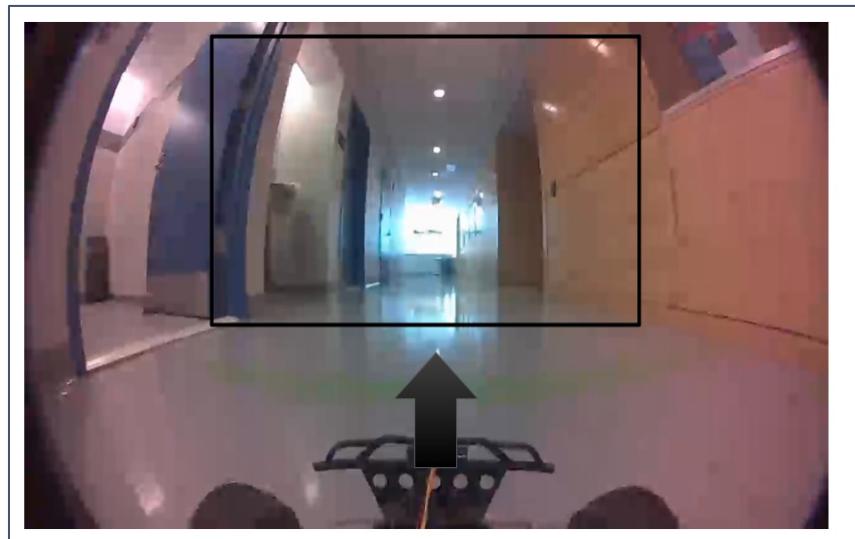


# Behavior Cloning : Data Augmentation

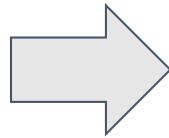
- Data augmentation is possibly the most critical part of behavior cloning.
- The bias of the data set can have a drastic effect on the performance of the model. If nothing is done for the bias, the trained model will not perform well, it will reach a situation that it has never seen before and fail.
- This leaves us with two choices, either collect more and more data on every possible situation or augment the current data to widen the distribution.
- Although clipping the image from  $1280 \times 720 \rightarrow 160 \times 90$  does help in terms of speed of computations, it also allows ourselves more room to generate data.

# Behavior Cloning : Data Augmentation

An example is shown below, the black box is the clipped image to be input to the model and the arrow is the adjusted steering angle.



Steering angle of 0.0



Steering angle of -0.5

## Behavior Cloning : Data Augmentation

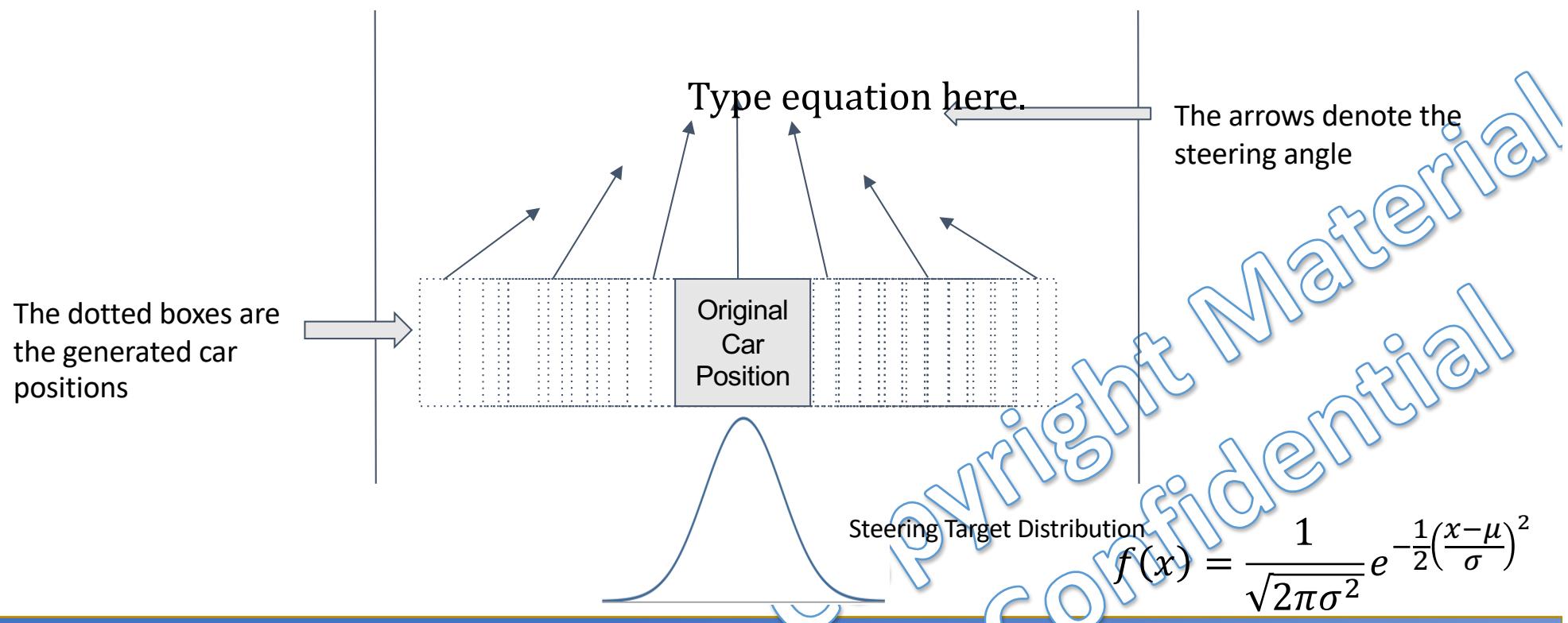
The first step in data augmentation that we take is to clip the image from 1280x720 to 160x90. Then by shifting the clipping to the left or right and adjusting the steering accordingly, we can teach the model how to respond without ever collecting that specific data point.

Using a normal distribution, we generate steering angles. By taking the difference between the current steering associated with the image and the sampled steering, we can shift the image by that amount to create a new data point.



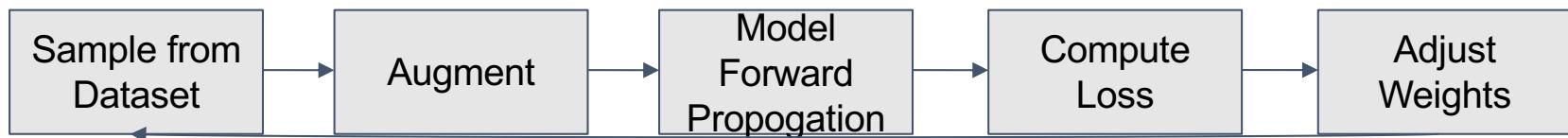
# Behavior Cloning : Data Augmentation

In this way we can simulate the left/right orientations of the car at every data point.



# Behavior Cloning : Training

Training the neural network is fairly straightforward; we randomly sample from the dataset, augment the data, then input it into the model.



The model produces steering and throttle predictions that we can test against the actual steering and throttle we recorded (or generated).

Using this, we compute the loss, and through backpropagation, we adjust the weights.

# Behavior Cloning : Loss

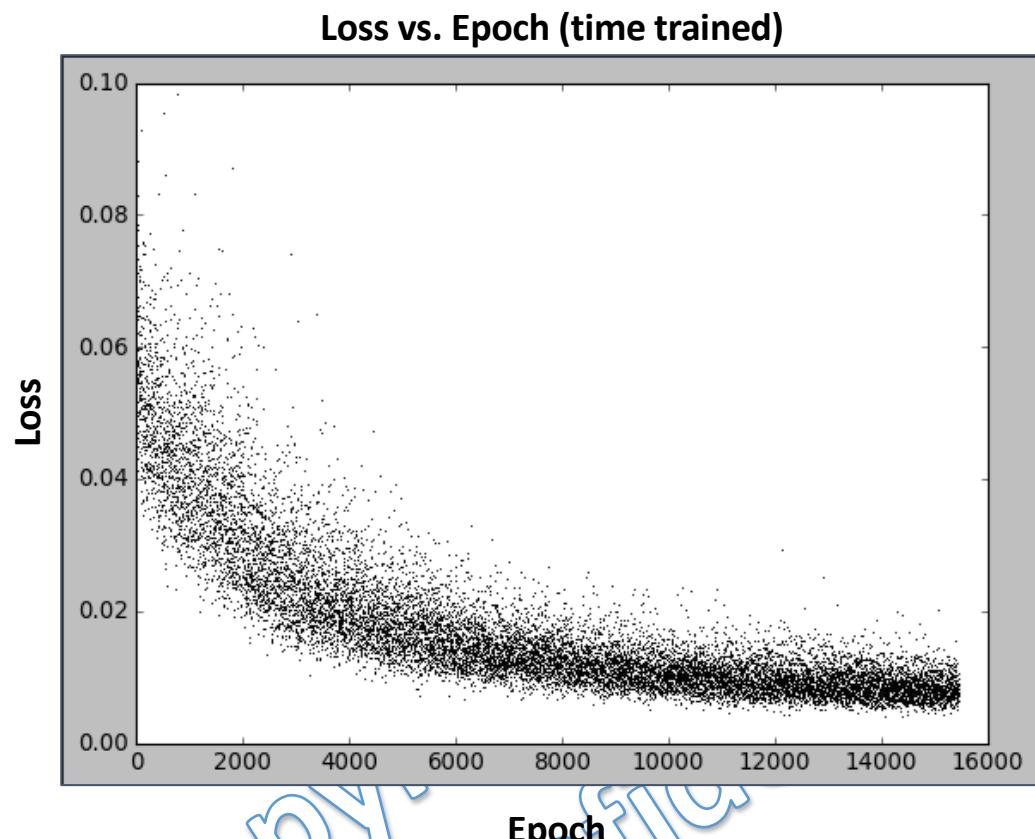
Loss is a metric by which we can test the validity of the model, there are many many functions that can be used. For our purposes we use MSE Loss (Mean Squared Error):

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2$$

This function simply takes the error at each point:

$$y_{\text{actual}} - y_{\text{predicted}}$$

Then squares and averages over all of them. This penalizes the model more for larger errors than for smaller ones. The graph shown is the loss as the model trains. As the loss tends towards 0 the model gets better and better at the task:

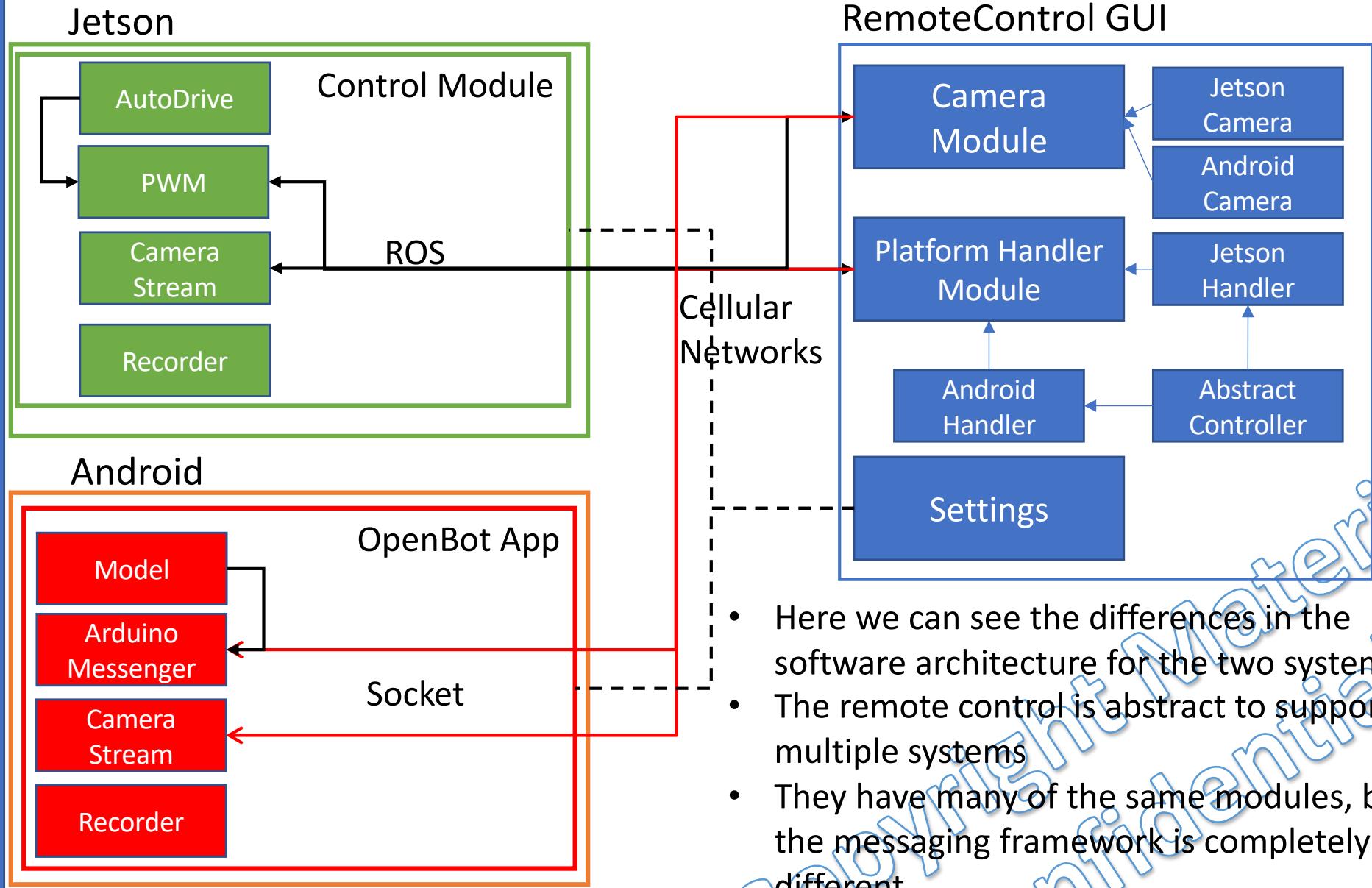


# Outlines

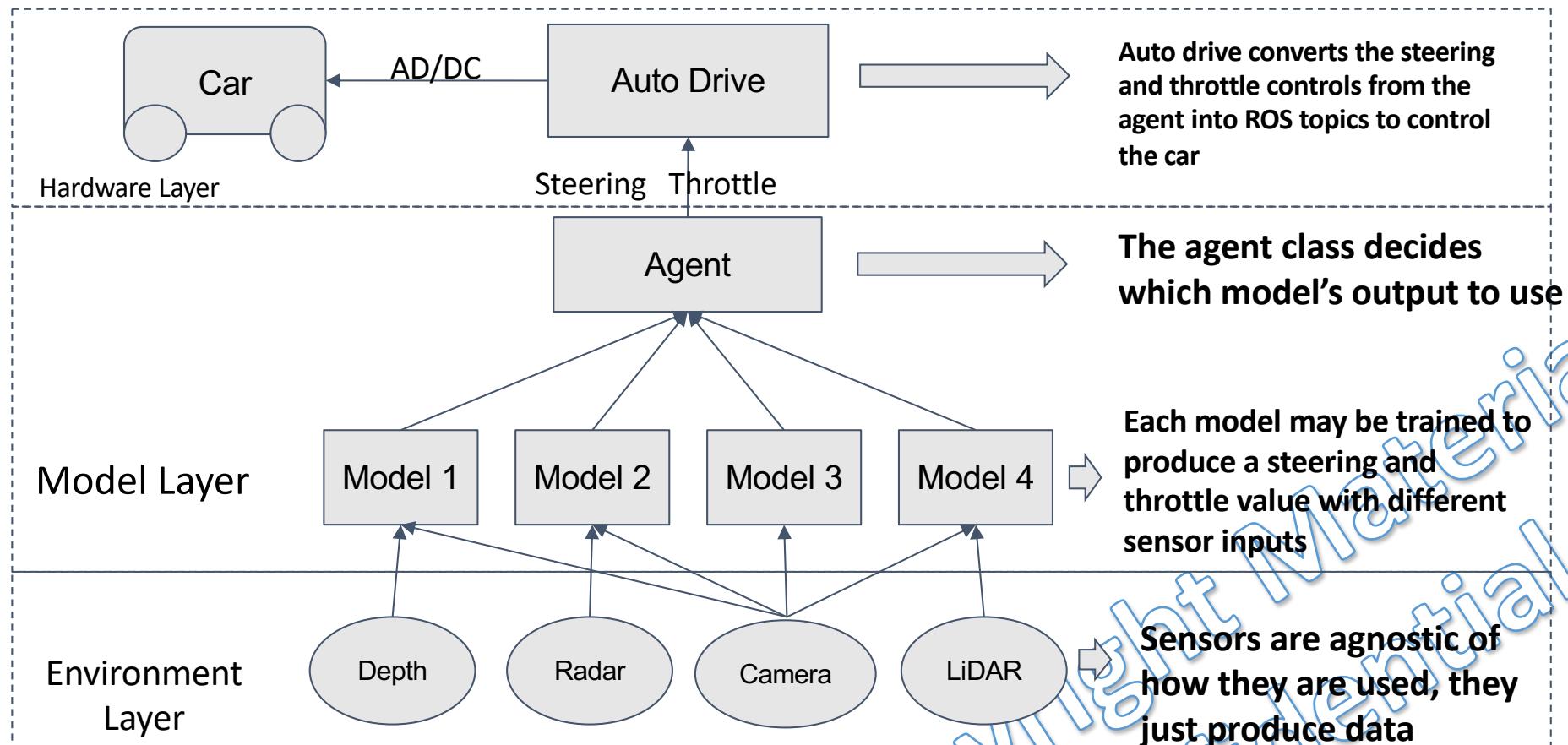
- Self-Driving Introduction and Motivation
- Self-Driving in real vehicles and RC-cars
- Self-Driving Perceptron (Brain)
  - Behavior Clone Based Algorithm
- Our Openbot-based Architecture
- Topic 1: Build Reliable Self-driving Systems
- Topic 2: Obstacle Avoidance
- Self-Driving Challenges
- The drawback of Supervised Learning Algorithm
- Look into the Future and Conclusions

# Demo of Openbot self-driving framework



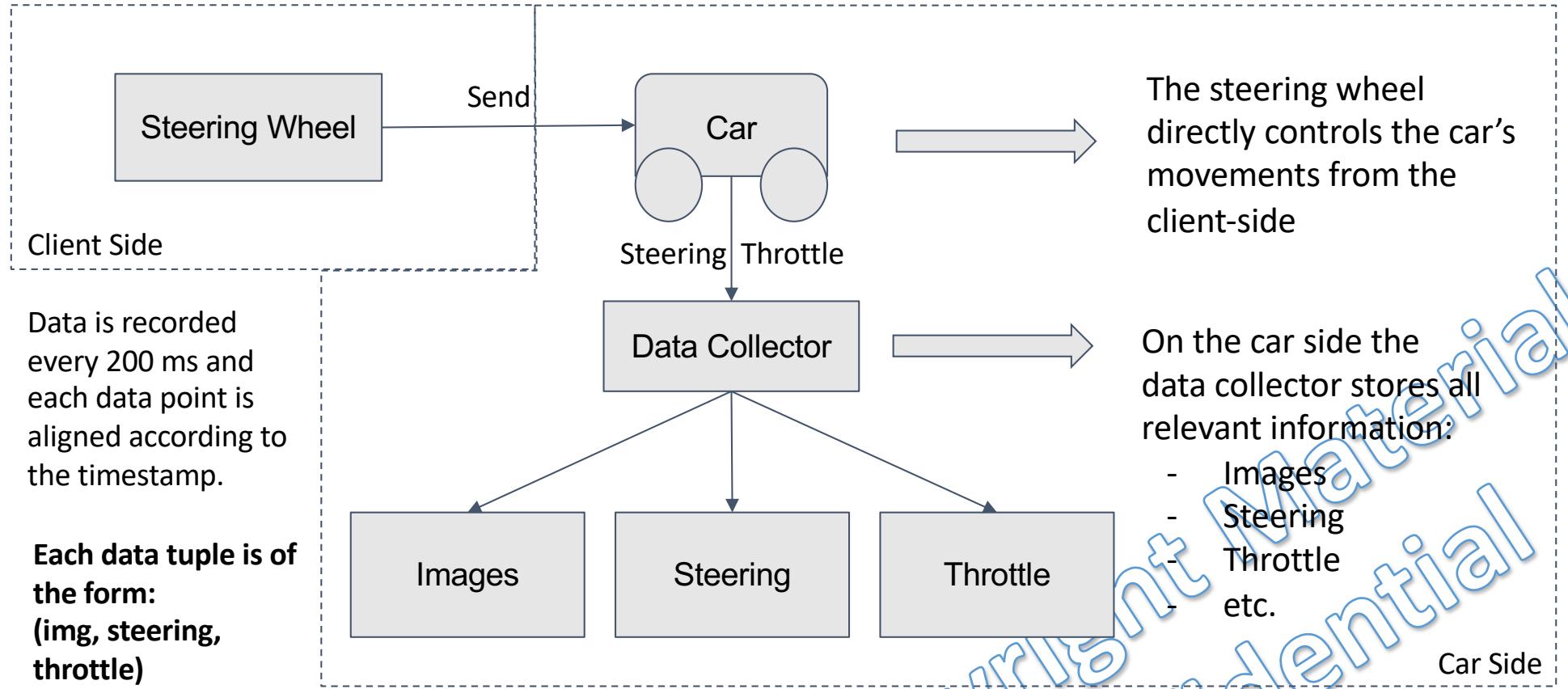


# Self-Driving Software Framework



\*\*Note: For the android system this is roughly the same pipeline, the only difference being the messaging framework (Socket vs. ROS). Also note that the android system does not have access to other sensors beside the camera

# Data Collections in Software Framework



# Outlines

- Self-Driving Introduction and Motivation
- Self-Driving in real vehicles and RC-cars
- Self-Driving Perceptron (Brain)
  - Behavior Clone Based Algorithm
- Our Openbot-based Architecture
- Topic 1: Build Reliable Self-driving Systems
- Topic 2: Obstacle Avoidance
- Self-Driving Challenges
- The drawback of Supervised Learning Algorithm
- Look into the Future and Conclusions

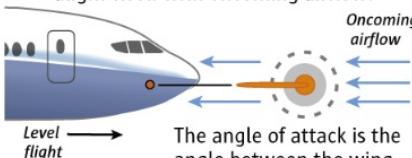


# Build Reliable Autonomous Vehicles



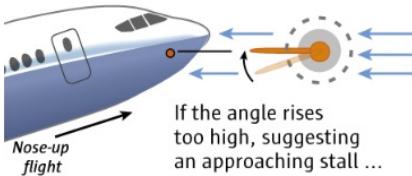
## How the MCAS (Maneuvering Characteristics Augmentation System) works on the 737 MAX

- 1.** The angle-of-attack sensor aligns itself with oncoming airflow.



The angle of attack is the angle between the wing and the airflow.

- 2.** Data from the sensor is sent to the flight computer.



If the angle rises too high, suggesting an approaching stall ...

... the MCAS activates.

- 3.** MCAS automatically swivels the horizontal tail to lift the plane's tail while moving the nose down.



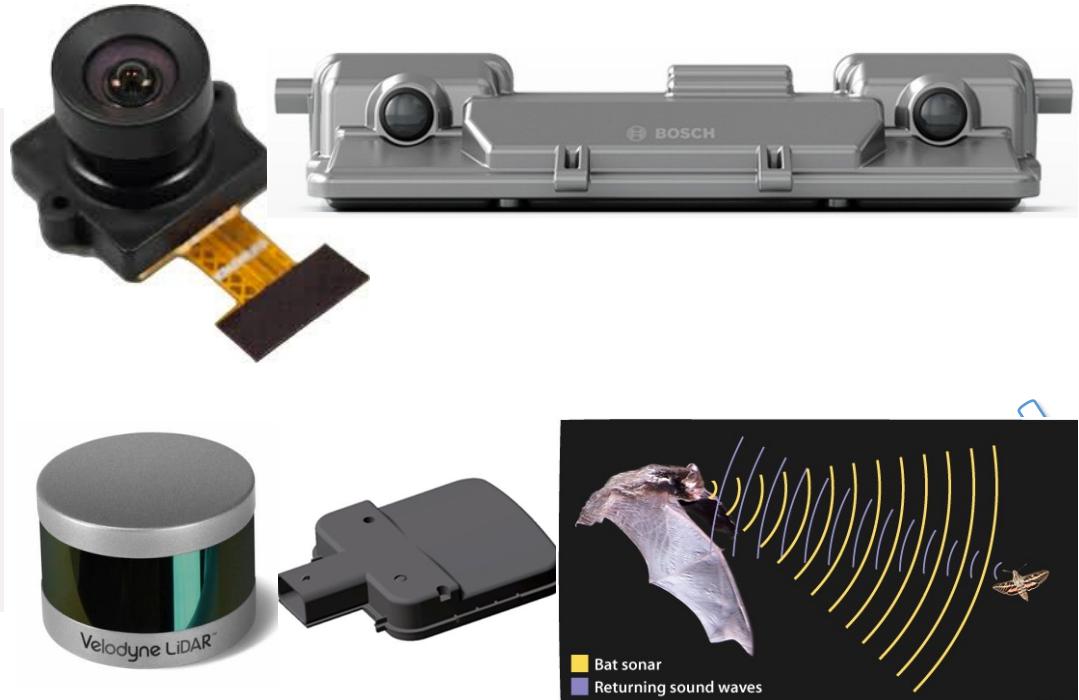
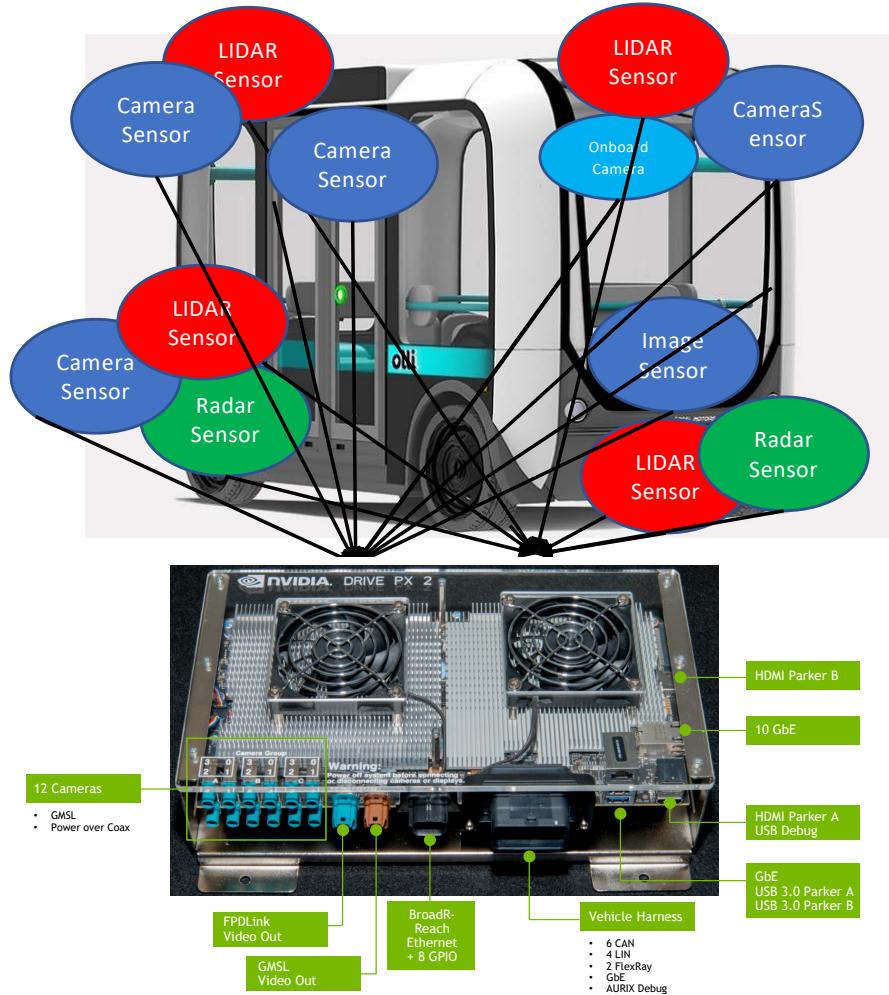
In the Lion Air crash, the angle-of-attack sensor fed false information to the flight computer.

Sources: Boeing, FAA, Indonesia National Transportation Safety Committee, Leeham.net, and The Air Current

Reporting by DOMINIC GATES,  
Graphic by MARK NOWLIN / THE SEATTLE TIMES

# Multi-Data Source Fusion to Improve Reliability

Olli Bus



**Ineffective during heavy rain or low hanging clouds:** LiDAR pulses may be affected by heavy rains or low hanging clouds because of the effects of refraction. No way to differentiate color



# Transfer Learning for Object Recognitions and Representations in Latent Space

Inputs  
(batch\_size: 416, 416, 32)

Conv 32x3x3 +  
Conv 64x3x3 s2  
(batch\_size: 208, 208, 64)

Residual Block 1x64  
(batch\_size: 208, 208, 64)  
Conv 128x3x3 s2  
(batch\_size: 104, 104, 128)

Residual Block 2x128  
(batch\_size: 104, 104, 128)  
Conv 256x3x3 s2  
(batch\_size: 52, 52, 256)

Residual Block 8x256  
(batch\_size: 52, 52, 512)  
Conv 512x3x3 s2  
(batch\_size: 26, 26, 512)

Residual Block 8x512  
(batch\_size: 26, 26, 512)  
Conv 1024x3x3 s2  
(batch\_size: 13, 13, 1024)

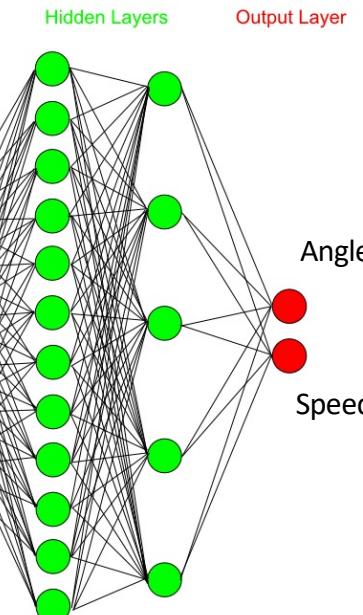
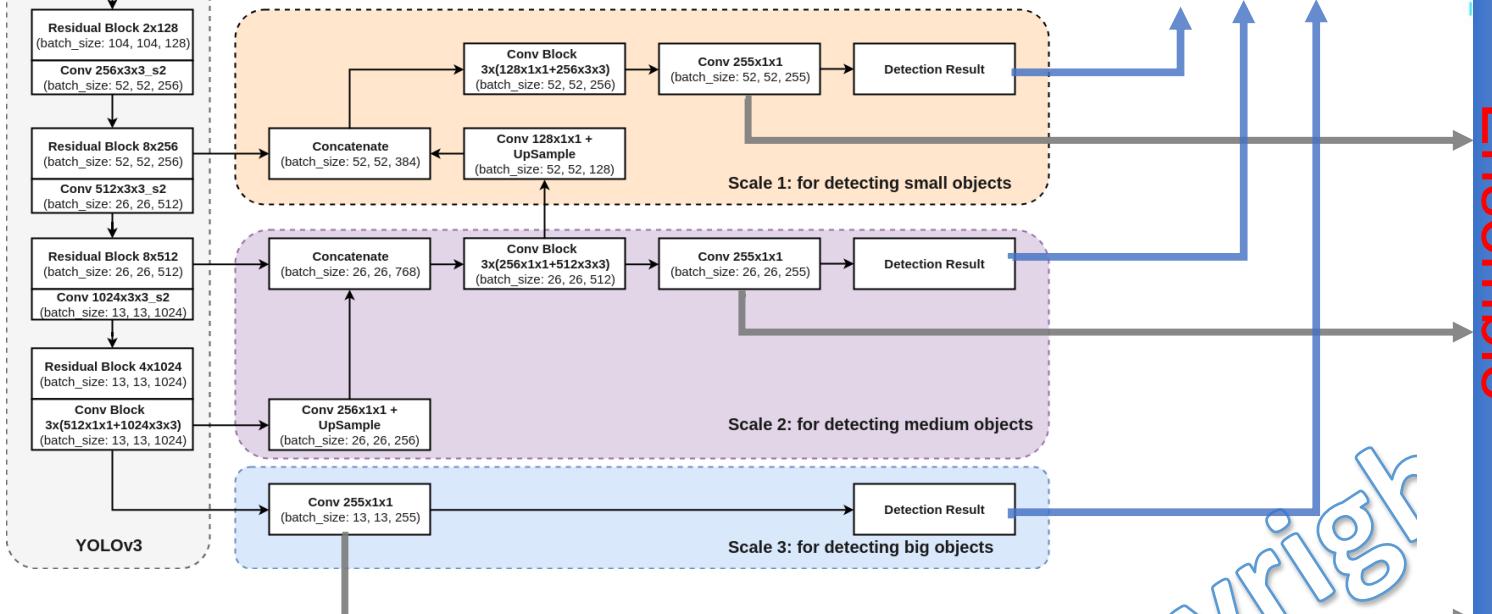
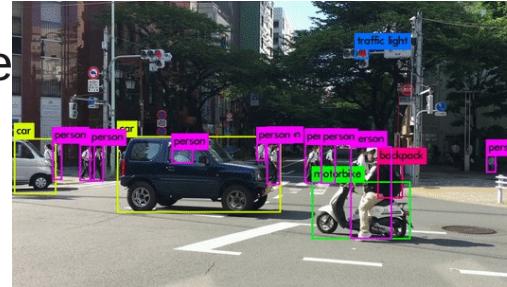
Residual Block 4x1024  
(batch\_size: 13, 13, 1024)  
Conv Block  
 $3 \times (512 \times 1 \times 1 + 1024 \times 3 \times 3)$   
(batch\_size: 13, 13, 1024)

## YOLOv3 Network Architecture

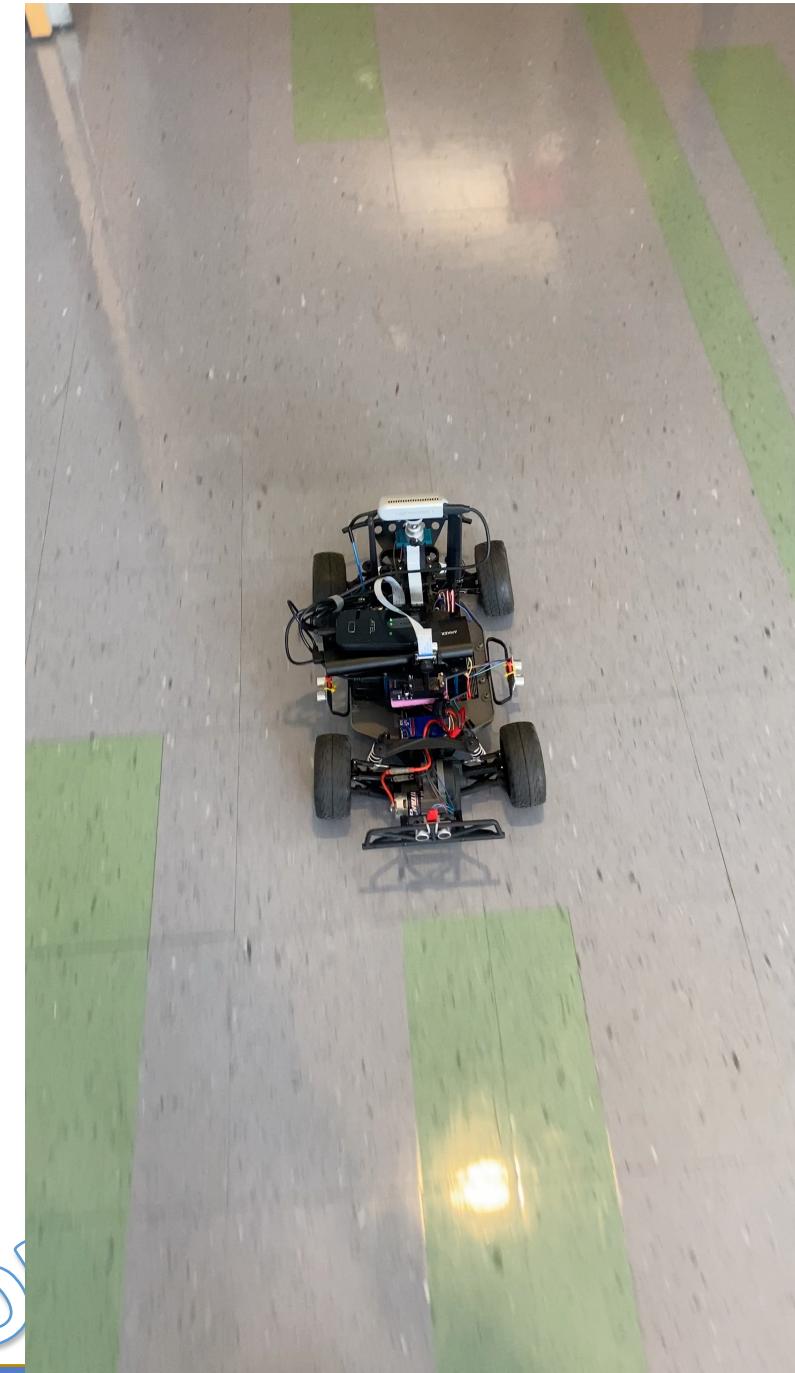
Conv: Convolutional layer   Concatenate: concatenate two inputs

\_s2: with stride of 2   batch\_size: the output size of this layer/block

Residual Block: repeated convolutional layers with ResNet structure

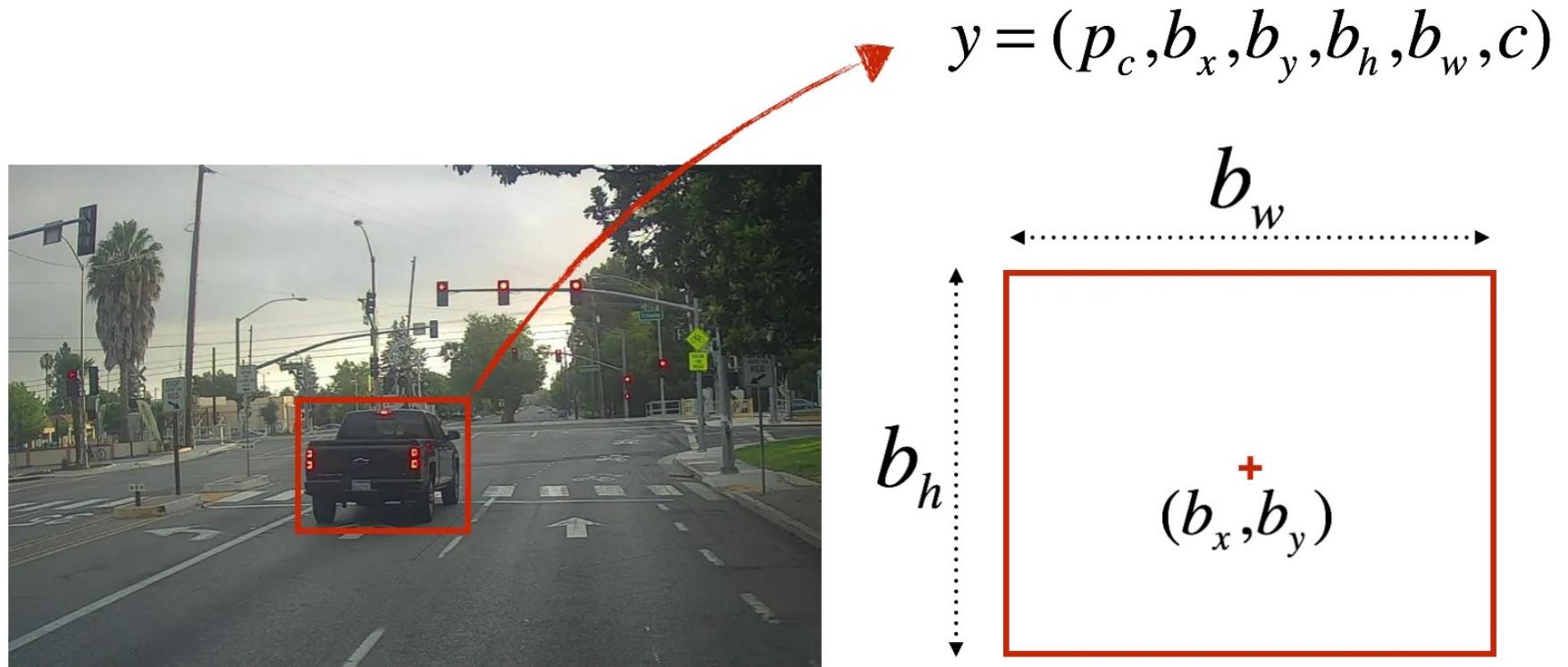


# Obstacle Avoidance



Sunrise Technology – Autonomous Driving

# Yolo (You only Look Once)



$p_c = 1$  : confidence of an object being present in the bounding box

$c = 3$  : class of the object being detected (here 3 for “car”)

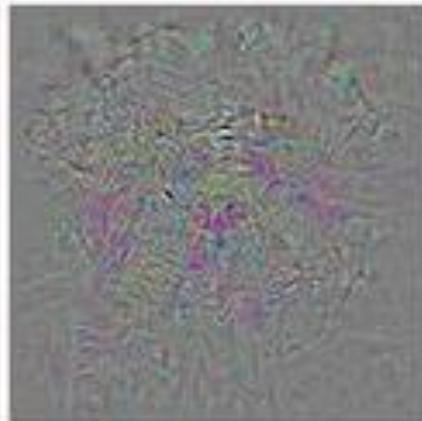
# Outlines

- Self-Driving Introduction and Motivation
- Self-Driving in real vehicles and RC-cars
- Self-Driving Perceptron (Brain)
  - Behavior Clone Based Algorithm
- Our Openbot-based Architecture
- Topic 1: Build Reliable Self-driving Systems
- Topic 2: Obstacle Avoidance
- **Self-Driving Challenges**
- Future Works and Conclusion

# Challenge: Self-Driving Security



Prediction: Dog



+ Distortion



Prediction: Ostrich



Prediction: Semi-Truck

+



+ Attacks



Prediction: Clouds

# Challenge: Will Behavior Clone Work?

Small Mismatch will accumulate

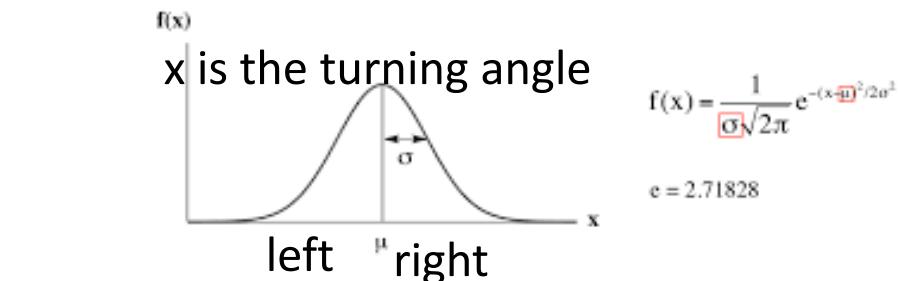


Katie Bradley

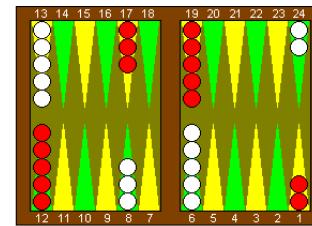
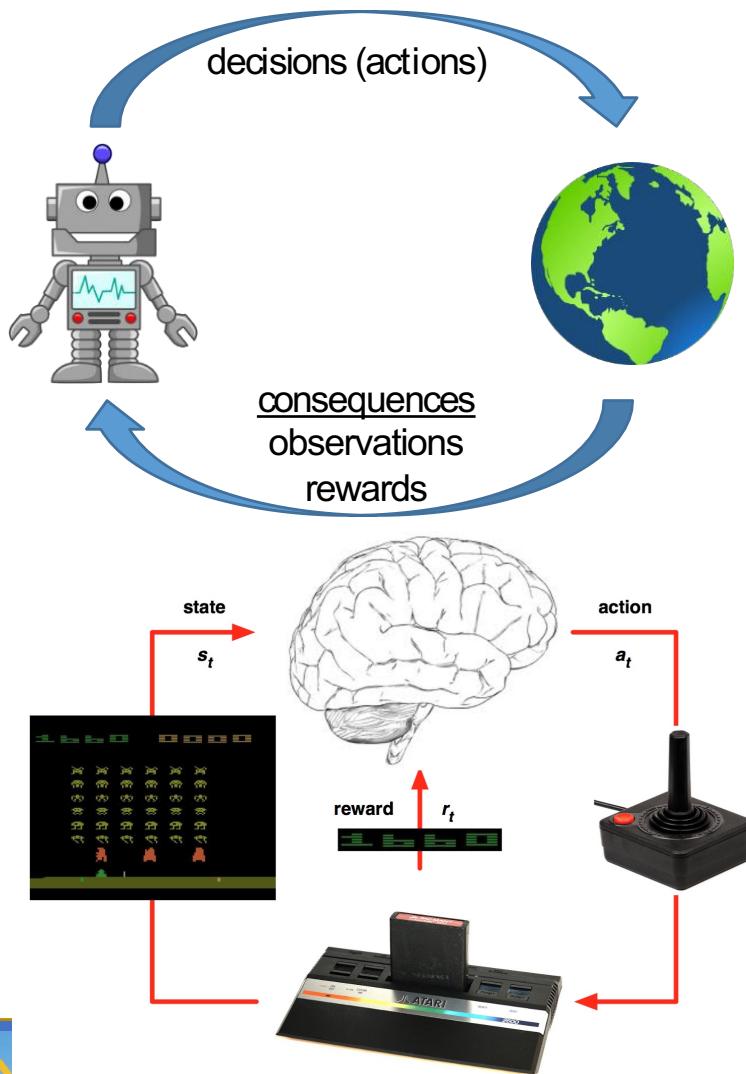
Corrective actions

Discrete Action is OK, Soft-max on Left or Right

What about Continuous Actions?



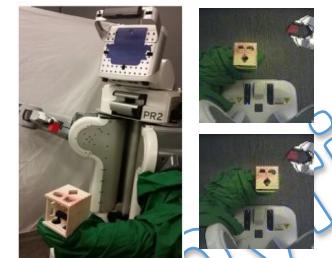
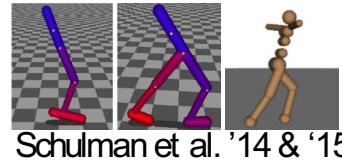
# Reinforcement learning provides a formalism for behavior



**Figure 2.** An illustration of the normal opening position in backgammon. TD-Gammon has sparked a near-universal conversion in the way experts play certain opening rolls. For example, with an opening roll of 4-1, most players have now switched from the traditional move of 13-9, 6-5, to TD-Gammon's preference, 13-9, 24-23. TD-Gammon's analysis is given in Table 2.



Mnih et al. '13



# Look into Future

- The question: When will full autonomous driving happen?
- What are the potential intermediate steps:
  - A) Algorithm Development for an Explainable AI system (phantom brake in Tesla)
  - B) Education and Outreach for public acceptance, Crowd-source, and Crowd development)
  - C) Infrastructure support: Star-Link and 5G/6G (99.999% wireless networks) to allow remote monitoring and control
  - D) Basic IT infrastructure development (fast, low-latency communications).

# Conclusions

- Self-Driving depends on an accurate learning model. Small errors accumulate and lead to substantial deviations from the track.
- No sensor works 100%; multiple sensors add reliability.
- Representation (embedding) is the key to uncovering the state information of observations and enhancing learning capability.
- Large training data is needed to mitigate the limitation of a supervised learning model.
- Performance and QoS are critical in self-driving: low power consumption, real-time decision making, and the flexibility to support transfer learning: incremental training, and independence from cloud and data center.
- Planning/controlling autonomous vehicles involves decision makings continuously in an episode to maximize future values rather than clone experts' behavior.

# Thank You!

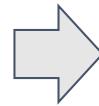
Copyright  
Confidential



# Behavior Cloning : Data Augmentation

Flipping: Flip the image horizontally and flip the steering angle to match

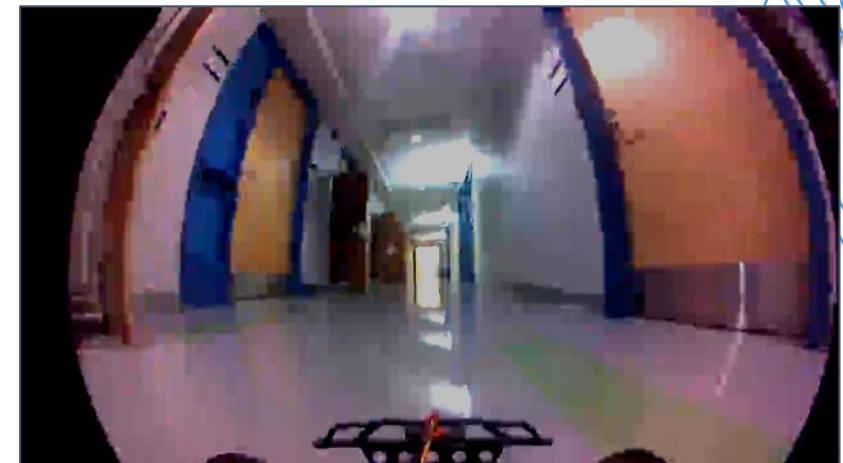
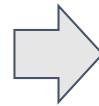
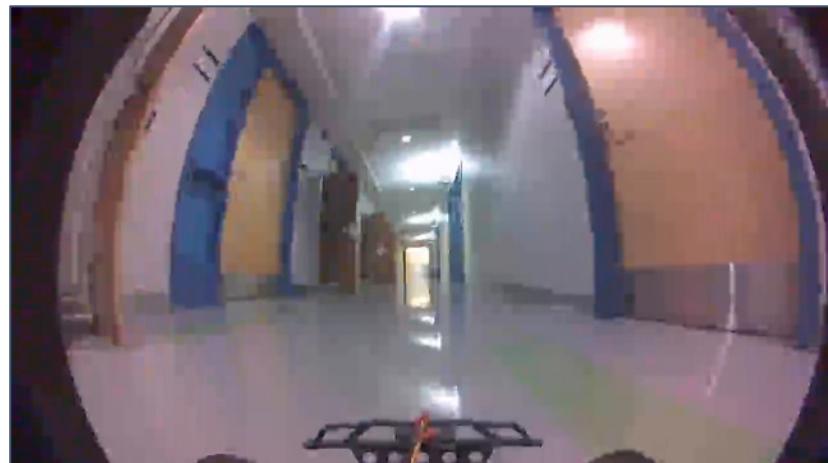
Equally trains left and right, further reduces bias



# Behavior Cloning : Data Augmentation

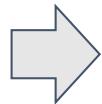
Lighting - Randomly adjust the contrast and brightness

Used to increase robustness of the model to lighting effects



# Behavior Cloning : Data Augmentation

Noise - add gaussian noise to the image



# Behavior Cloning : Data Augmentation

Grayscale - although not used, the option to convert the rgb image to grayscale remains in our training notebook. Furthermore, you can also equalize the histograms



This can be useful for reducing the computation cost associated with the channels of the image or for reducing the model's dependency on recognizing certain colors.