# AI and Computer Vision for Autonomous Vehicles

CS Katha Barta
**at NISER Bhubaneswar**

Dr. Ram Prasad Padhy

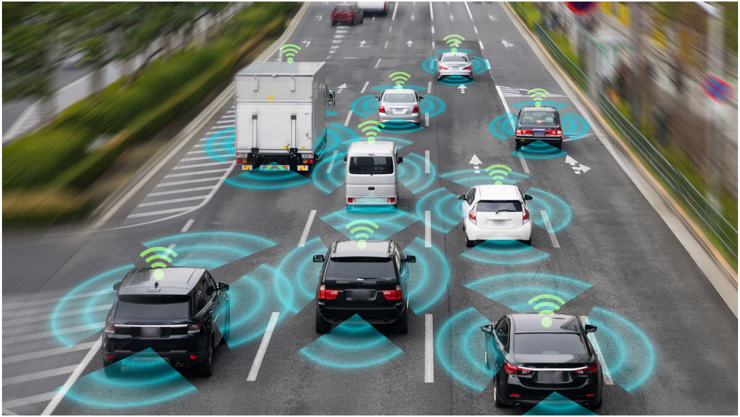
Computer Science & Engineering
School of Electrical and Computer Sciences
Indian Institute of Technology Bhubaneswar

August 28, 2024

# Table of contents

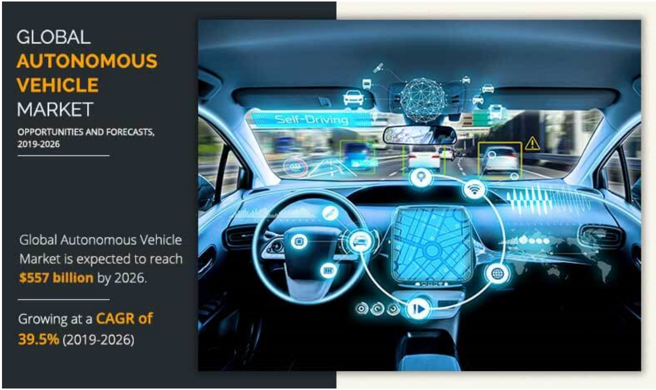
1. Introduction
   - Self-Driving Cars
   - AI-ML Revolution
   - AI-ML for Self-driving Cars
2. Situational Awareness in Self-driving Cars
   - Camera
   - LiDAR
   - RADAR
   - GPS and IMU
3. Multi Sensor Fusion
   - Camera+LiDAR Fusion
   - LiDAR+RADAR Fusion
4. Datasets
5. Demo
6. Summary

# Self Driving Cars



Source: https://www.walleniuswilhelmsen.com/insights/the-future-of-mobility-whats-the-road-ahead-for-self-driving-vehicles

# Self Driving Cars



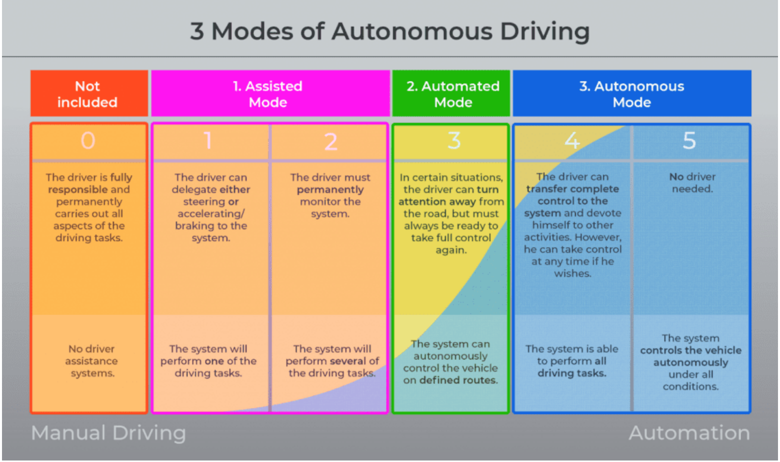

Source: https://www.alliedmarketresearch.com/autonomous-vehicle-market

- Also known as Autonomous Vehicles (AV), or Advanced Driver Assistance Systems (ADAS)

# Self Driving Cars Industry

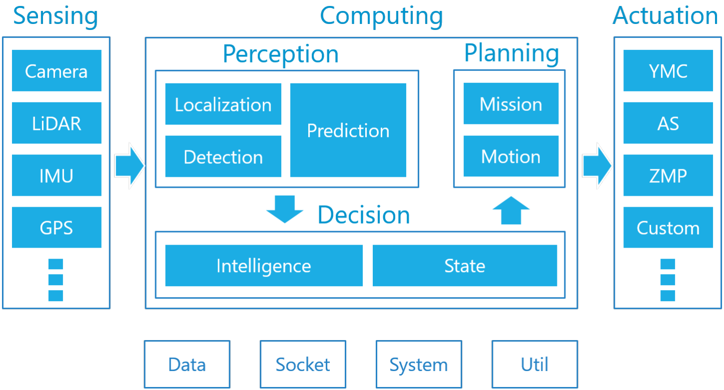
- **Key Market Players:** General Motors, Daimler AG, Ford Motor Company, Volkswagen Group, BMW AG, Renault-Nissan-Mitsubishi alliance, Volvo-Autoliv-Ericsson-Zenuity alliance, Groupe SA, AB Volvo, Toyota Motor Corporation, and Tesla Inc *etc.*
- **Auto suppliers:** Robert Bosch GMBH, Aptiv, Continental AG, and Denso Corporation *etc.*
- **Technology providers:** Waymo, NVDIA Corporation, Intel Corporation, Baidu, and Samsung *etc.*
- **Service provider:** Uber, Lyft and Didi Chuxing *etc.*

# Levels of Vehicle Autonomy




Source: https://www.blickfeld.com/blog/levels-of-autonomous-driving/

# ADAS Software Architecture

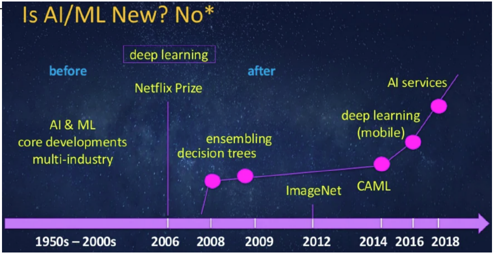


# AI-ML Revolution

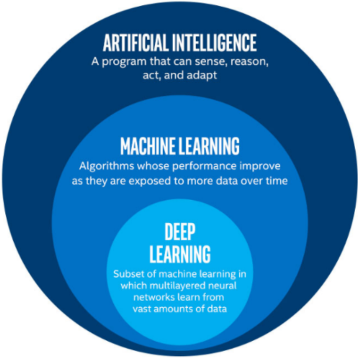
**McKinsey Global Institute**[a]

[a] https://mck.co/3mzGs5l

- AI is contributing to the transformation of the global economy
- 10 times faster, 300 times the scale, and 3000 times of 1st Industrial revolution.



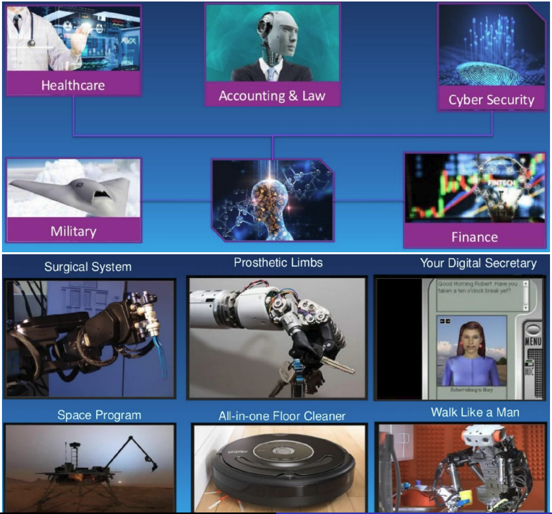
# AI-ML Revolution



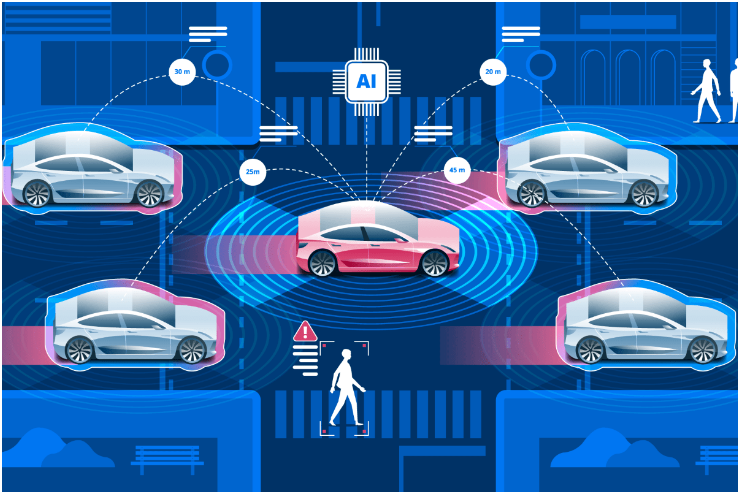

## AI vs ML vs DL

- AI is considered as a broad field of research that includes ML
- AI can manage data of a more generic and abstract nature than ML
- AI enables the transfer of common solutions to different types of data without the need for complete retraining
- ML includes deep learning
- DL is based on artificial neural networks

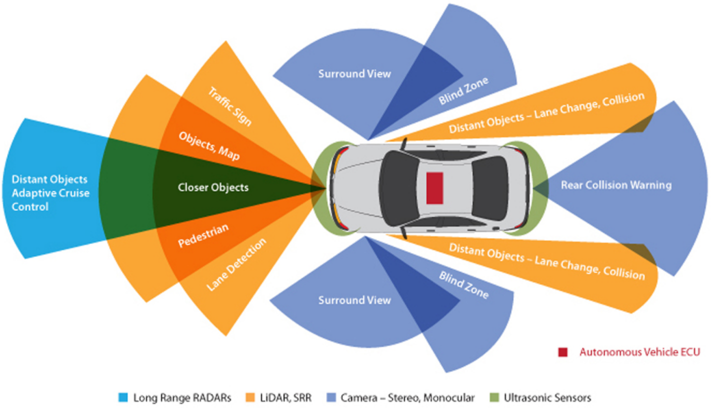
# AI-ML Revolution: Applications



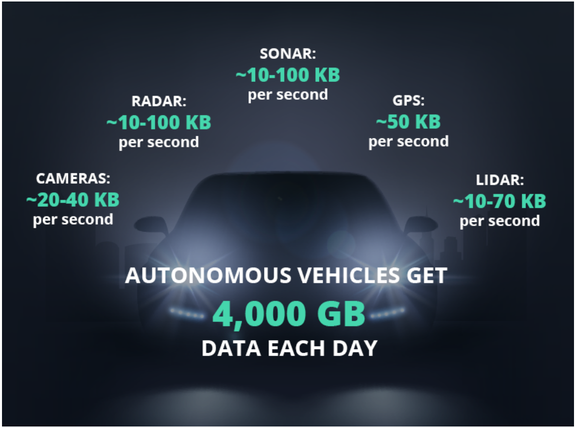
# AI-ML Revolution: Self Driving Cars



# AI-ML Revolution: Self Driving Cars



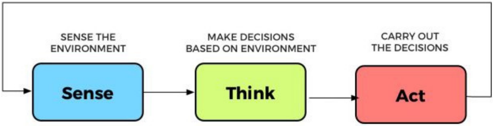
# AI-ML for Self-driving Cars




Source:
https://medium.com/analytics-vidhya/perception-in-self-driving-cars-7424e20b77c7

# Situational Awareness in Self-driving Cars

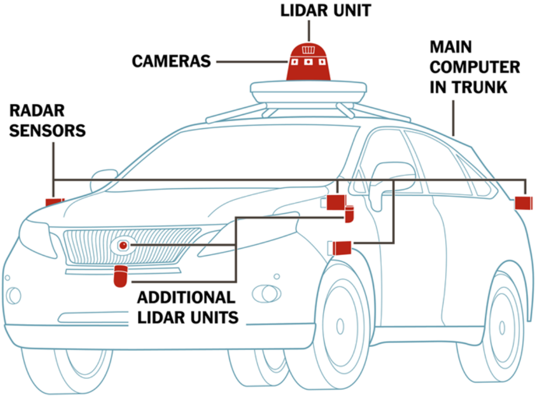
- Commonly known as **Perception**
- How the car senses the environment around it
- Where is its location w.r.t others in the environment
- Sensors: Camera, Lidar, Radar, GPS, IMU *etc.*
- Huge amount of data
- Use the computer intelligence to evaluate data and make something meaning of it



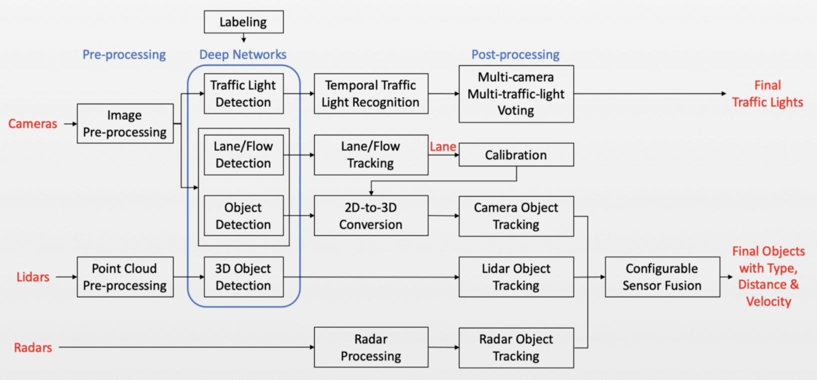
# Sensors

- **Camera**: 2D object detection, semantic segmentation, Traffic light detection, lane-following etc.
- **Lidar**: 3D object detection, range estimation
- **Radar**: Velocity and range estimates
- **GPS**: Puts the AV in global frame of reference
- **IMU**: Provides the acceleration and gyroscope measurements
- **Multi sensor Fusion**: Camera+Lidar, Lidar+Radar etc.

# Sensors


LIDAR UNIT
CAMERAS
MAIN COMPUTER IN TRUNK
RADAR SENSORS
ADDITIONAL LIDAR UNITS

# Situational Awareness Module



# Camera

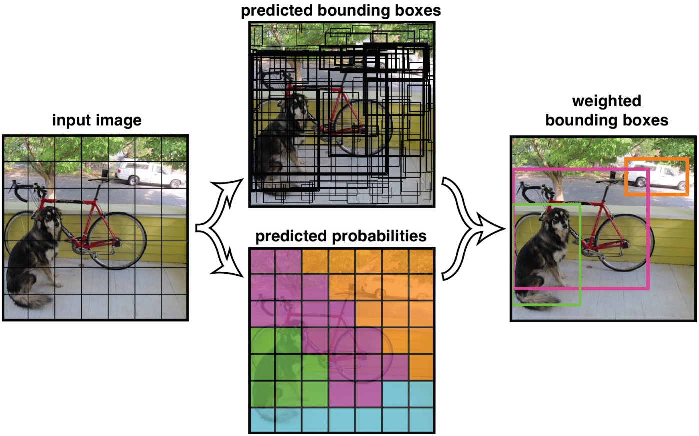
## 2D Object Detection

- Reads image data from the camera, and provides image-based object detection
- Popular Algorithms: R-CNN, SSD, and Yolo *etc.*
- Mostly Neural Network based
- Multiple classes of detection are supported, such as cars, trucks, bicycles, pedestrians, auto rickshaws, and many more

## 2D semantic Segmentation

- Pixel-wise classification is done for the whole image
- Each class is given a different color
- Mostly CNN based deep learning approaches are followed
- Semantic segmentation helps in free space detection, and thereby makes the safe trajectory estimation much easier for the AV

# Camera

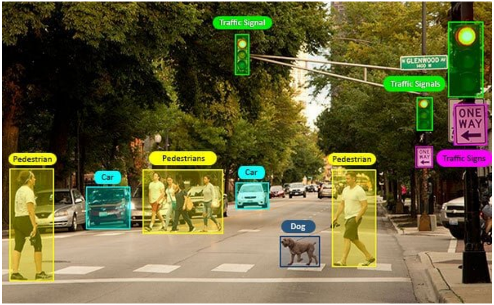

predicted bounding boxes
input image
weighted
bounding boxes
predicted probabilities

# Camera


(a) 2D Object detection

(b) 2D Semantic Segmentation

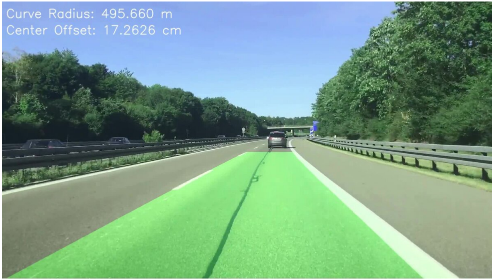
# Camera




**Traffic Light Detection**

- Detects the traffic light at traffic posts
- Classifies whether it is Green, Red or Yellow

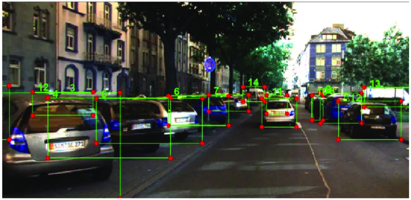
# Camera




## Lane/Flow Detection

- Detects the derivable area
- Follows the lane
- Combined with 2D object detection and semantic segmentation

# Camera



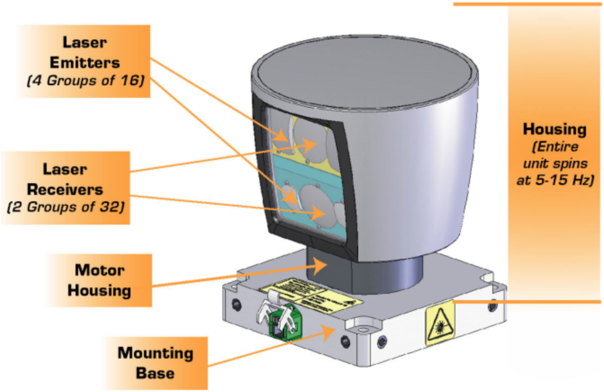
## 2D Object Tracking

- Takes the input from the 2D object Detection predictions
- Uses the algorithms, like Kalman Filtering
- Assigns a unique ID to each of the objects
- Helps in the measurement of kinematics parameters and trajectory prediction of the surrounding vehicles

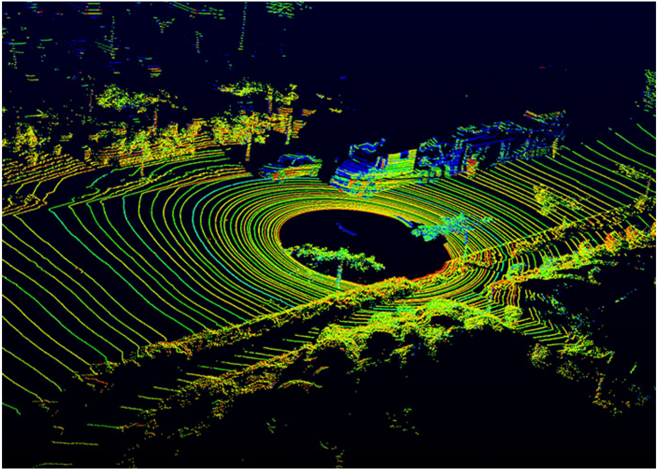
# LiDAR

- Light Detection And Ranging
- Sends out very short light pulses at different angles across the field of view and receives the photons reflected back from an object
- It measures the time difference and determines the distance to the object
- It generates the point cloud with x, y, z coordinates in 3D

# LiDAR


Laser
Emitters
(4 Groups of 16)
Laser
Receivers
(2 Groups of 32)
Motor
Housing
Mounting
Base
Housing
(Entire
unit spins
at 5-15 Hz)

# LiDAR

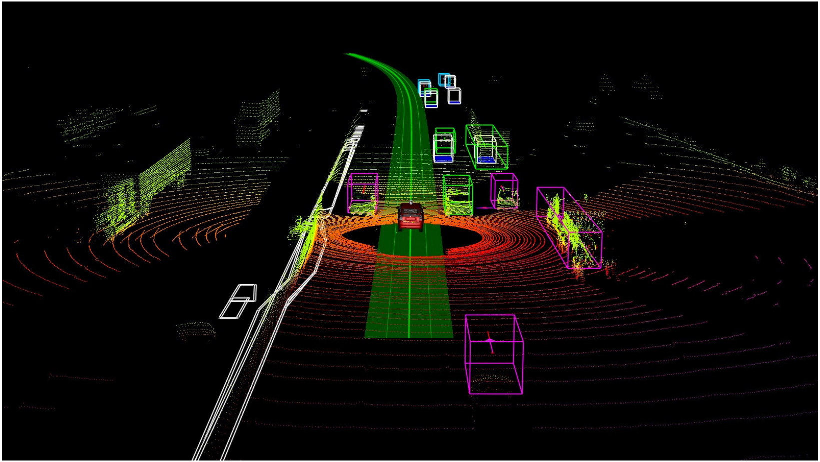


# LiDAR

## 3D Object Detection

- The 3D point cloud data from the Lidar sensor is passed through CNNs to detect 3D objects around the vehicle
- It detects the class of the objects (bus, truck, pedestrian etc.)
- It also estimates the position and dimension in 3D coordinate system along with the heading
- Requires heavy processing as compared to camera based object detection

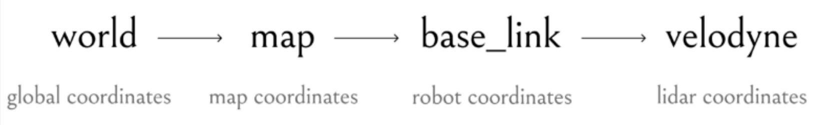
# LiDAR



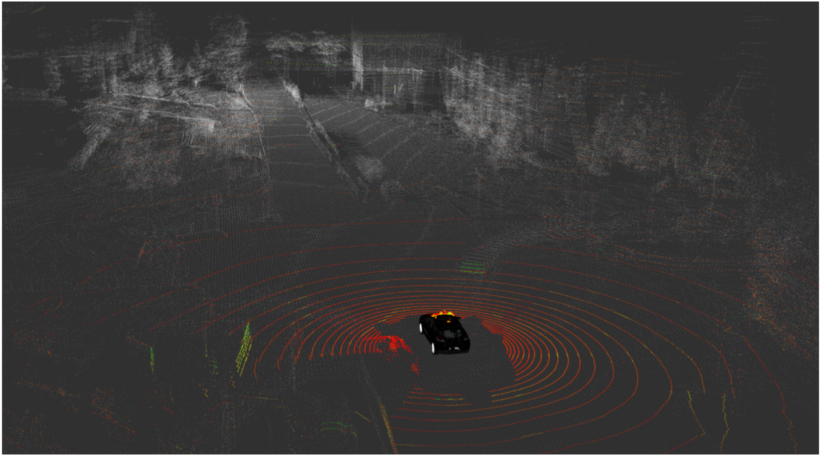
# LiDAR

## 3D Localization

- Creating an effective 3D map of the environment
- Memorizing the landmarks
- The vehicle knows its relative position w.r.t others in the environment



# LiDAR



# RADAR

- Radar is very efficient with a low cost and small footprint
- Capable to determine the targets at long range with accurate velocity and spatial information
- Its sensitivity in dark and poor weather conditions also helps to cover the domains where LiDAR or camera may fail
- Transmit electromagnetic waves, and receive reflected wave from targets
- **Comparison with LiDAR:**
  - **Pros:** Operates in the bad weathers, low-cost, accurate velocity detection
  - **Cons:** Lower resolution

# GPS/GNSS and IMU

## GPS/GNSS

- **GPS**: Global Positioning System
- **GNSS**: Global Navigation Satellite System
- Puts the AV in global frame of reference
- Helps in effective global localization

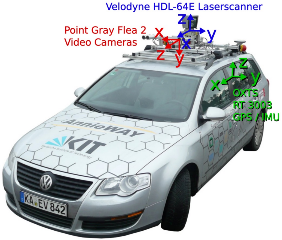
## IMU

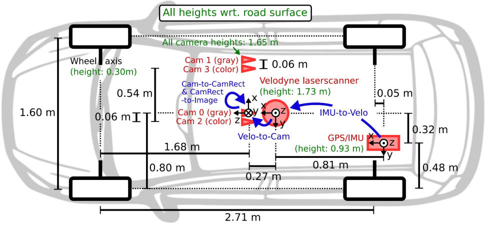
- Inertial Measurement Unit
- Provides the Acceleration and Gyroscope (Heading) data
- Can be combine with 3D object position and heading to measure the acceleration and gyroscope data of surrounding objects

# Multi Sensor Fusion

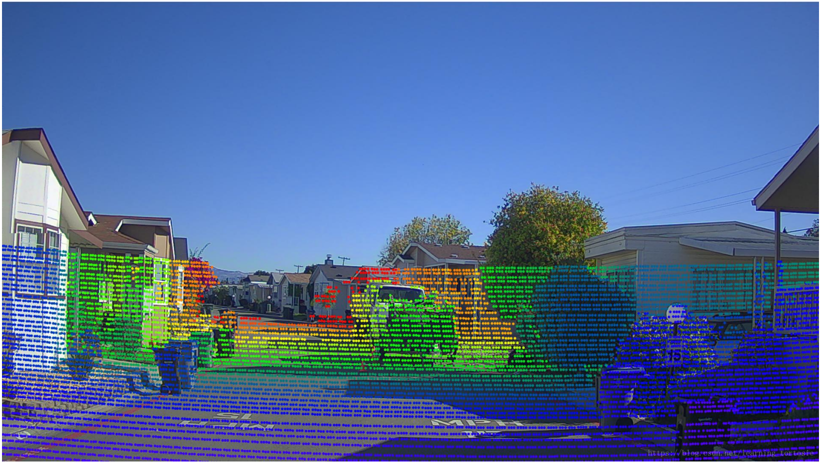
- Single-sensor approach is limited by the fact that each sensor has its own weakness in some situation.
- Combining data from multiple sources to increase the accuracy of prediction
- MSF can increase the accuracy of object detection as well as tracking
- Each sensor has its own frame of reference
- Needs to bring all to a common coordinate system
- External calibration is required

# Multi Sensor Fusion



# Camera+LiDAR Fusion

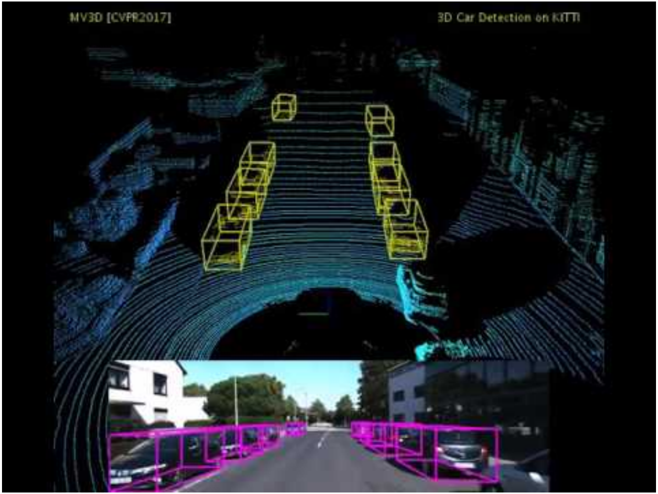


# Camera+LiDAR Fusion

## 3D Multi Object Tracking (MOT)

- MOT specifies tracking multiple objects simultaneously
- Generates the trajectory of different vehicles around the AV
- It requires the position sequence of objects for a period of time.
- Extended Kalman Filter or Unscented Kalman Filter are used to rectify the position of objects
- Unique ID is assigned to each of the identified objects in the scene.

# Camera+LiDAR Fusion

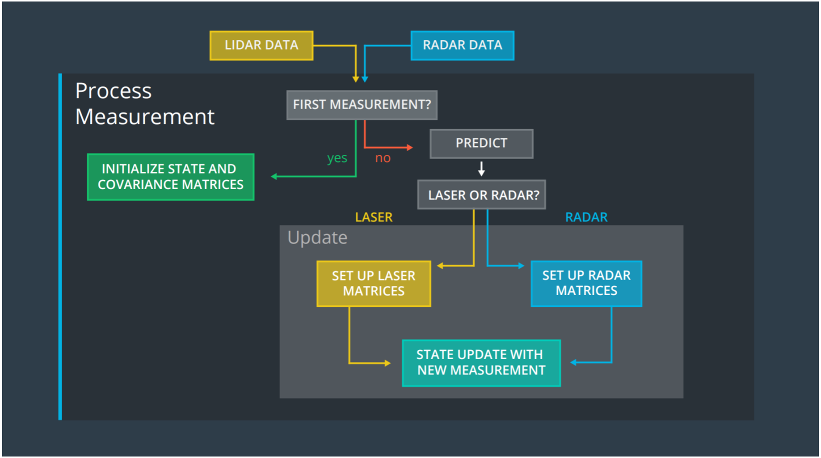


# LiDAR+RADAR Fusion

## Kalman Filter

- Receives initial measurement of the obstacle's position relative to the ego vehicle
- Initialize the obstacle's position based on the first measurement
- Ego vehicle receives another measurement after a dt
- Estimates the obstacle's position after the dt using constant velocity model
- Predicted location and measured location are combined to give an updated location
- Kalman filter will put weight on them depending on the uncertainty of each value
- Loop over another measurement after dt and predict/update the location

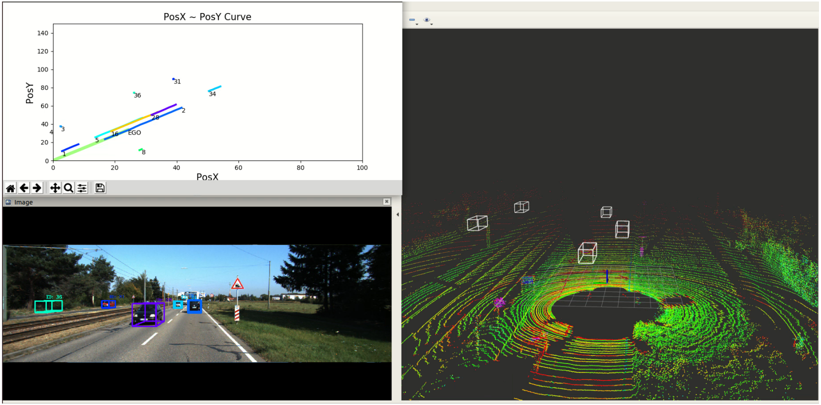
# LiDAR+RADAR Fusion



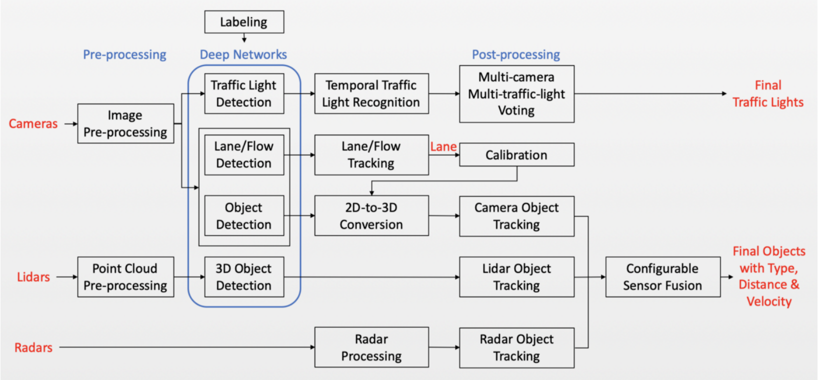
# Datasets

- KITTI, NuScenes, Waymo, ArgoVerse, A2D2, ApolloScape, DeepDrive, CityScapes, Comma2k19, Google-Landmarks, LeddarTech, Level 5 Open Data, Oxford Radar RobotCar Dataset, PandaSet, Udacity Self Driving Car Dataset *etc.*
- IDD Dataset: For Indian Scenarios
- **Challenges**: 2D Object detection/tracking, 3D Object detection/tracking, Lane following, Lidar and Image segmentation, Traffic Light Detection, Localization etc.

# Demo



# Summary


Labeling
Pre-processing
Deep Networks
Post-processing
Cameras
Image Pre-processing
Traffic Light Detection
Temporal Traffic Light Recognition
Multi-camera Multi-traffic-light Voting
Final Traffic Lights
Lane/Flow Detection
Lane/Flow Tracking
Lane
Calibration
Object Detection
2D-to-3D Conversion
Camera Object Tracking
Lidars
Point Cloud Pre-processing
3D Object Detection
Lidar Object Tracking
Configurable Sensor Fusion
Final Objects with Type, Distance & Velocity
Radars
Radar Processing
Radar Object Tracking

# Acknowledgment

- Some of the data, slides, code and images have been adopted from different internet sources. I am very thankful to the authors of these sources. The due credits are acknowledged.
- The slides are prepared for educational purpose only.


Email: ramprasad@iitbbs.ac.in / ramprasad.nitr@gmail.com
Profile link: link

**Internship, project researcher and Ph.D Positions are open throughout the year.**