# Outline

First 45 min

- **第三次课程内容**
    - Mobileye 工作介绍
    - 车辆动力学

Second 45 min

- **自动驾驶数据集综述**
    - 感知类数据集
    - 预测、归控、地图类数据集

Third 45 min

- **数据算法闭环体系**
    - 车企闭环体系简介
- **榜单与生态**
    - Autonomous Grand Challenge by OpenDriveLab

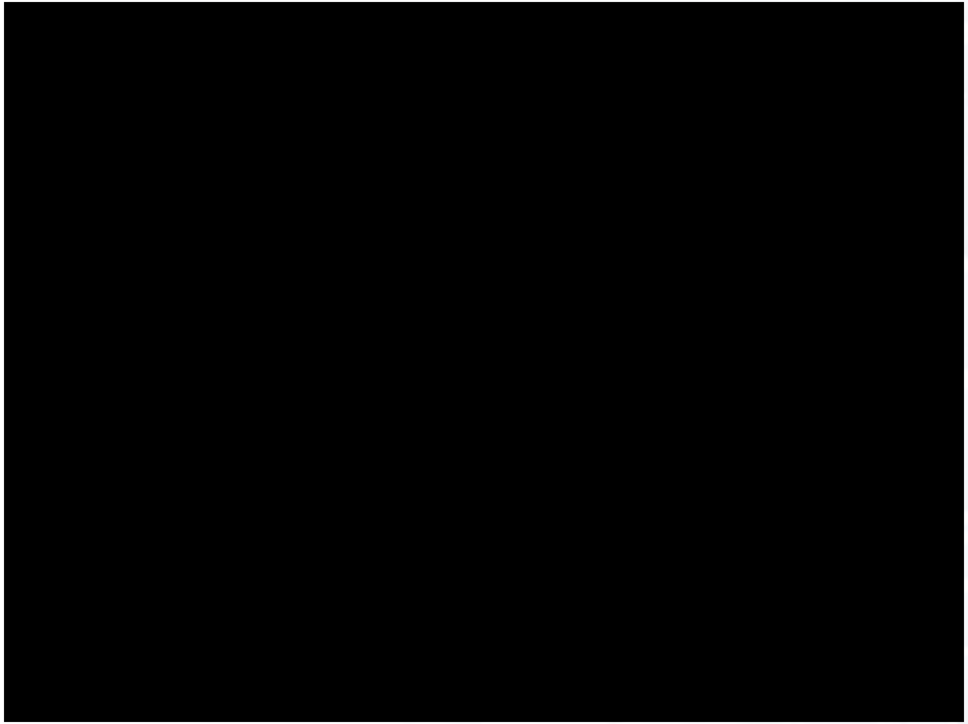
OpenDriveLab

# Mobileye

# What’s Mobileye

- Tier 1
- 卖摄像头+算法
- 主要做ADAS
- 性能稳定、配套健全
- 很不“自主原创、自主品牌”，ME前景堪忧

# What's Mobileye - From "吉利汽车" Perspective

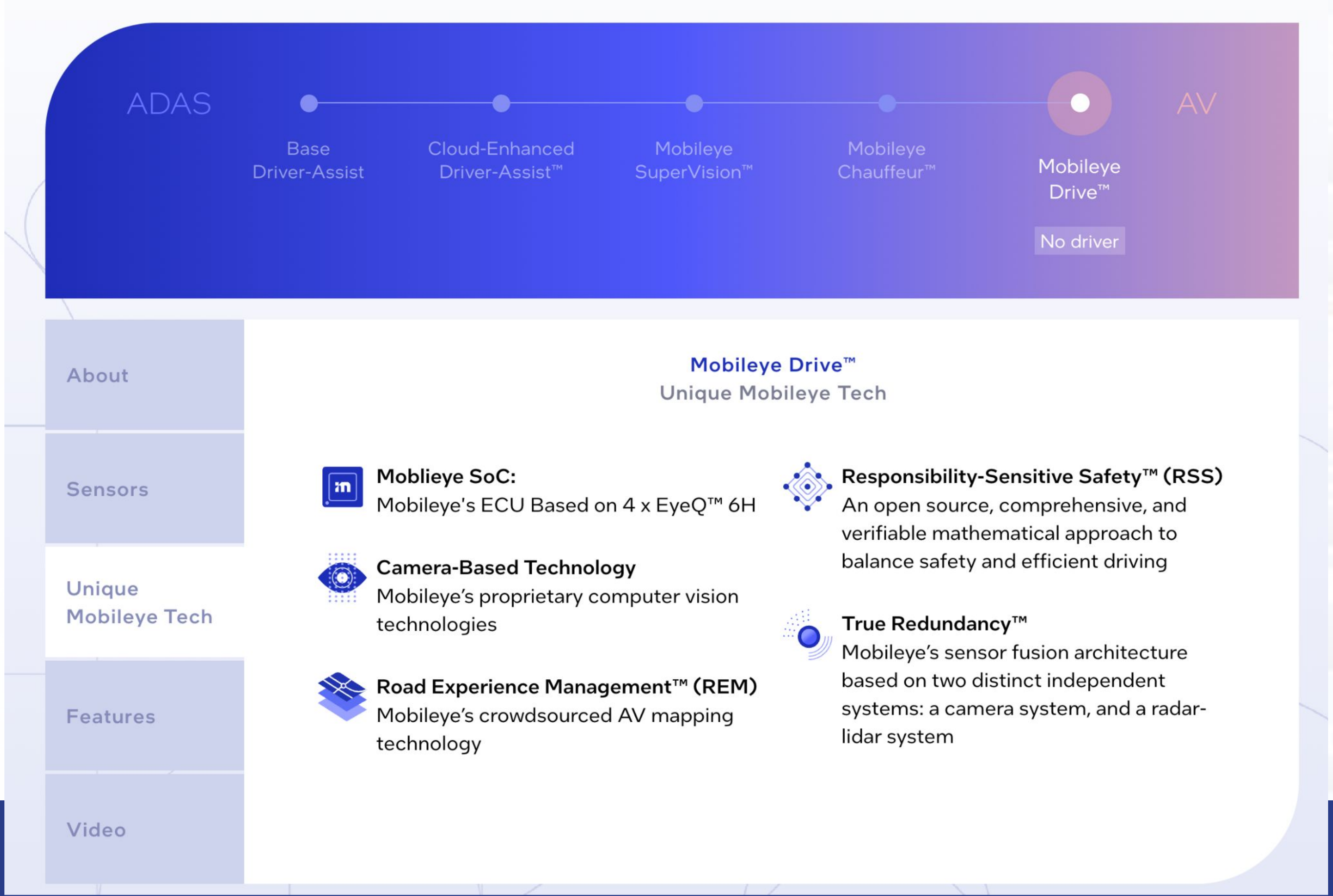


# Mobileye 技术路线

https://www.mobileye.com/ces-2024/

# Mobileye 技术路线

https://www.mobileye.com/ces-2024/

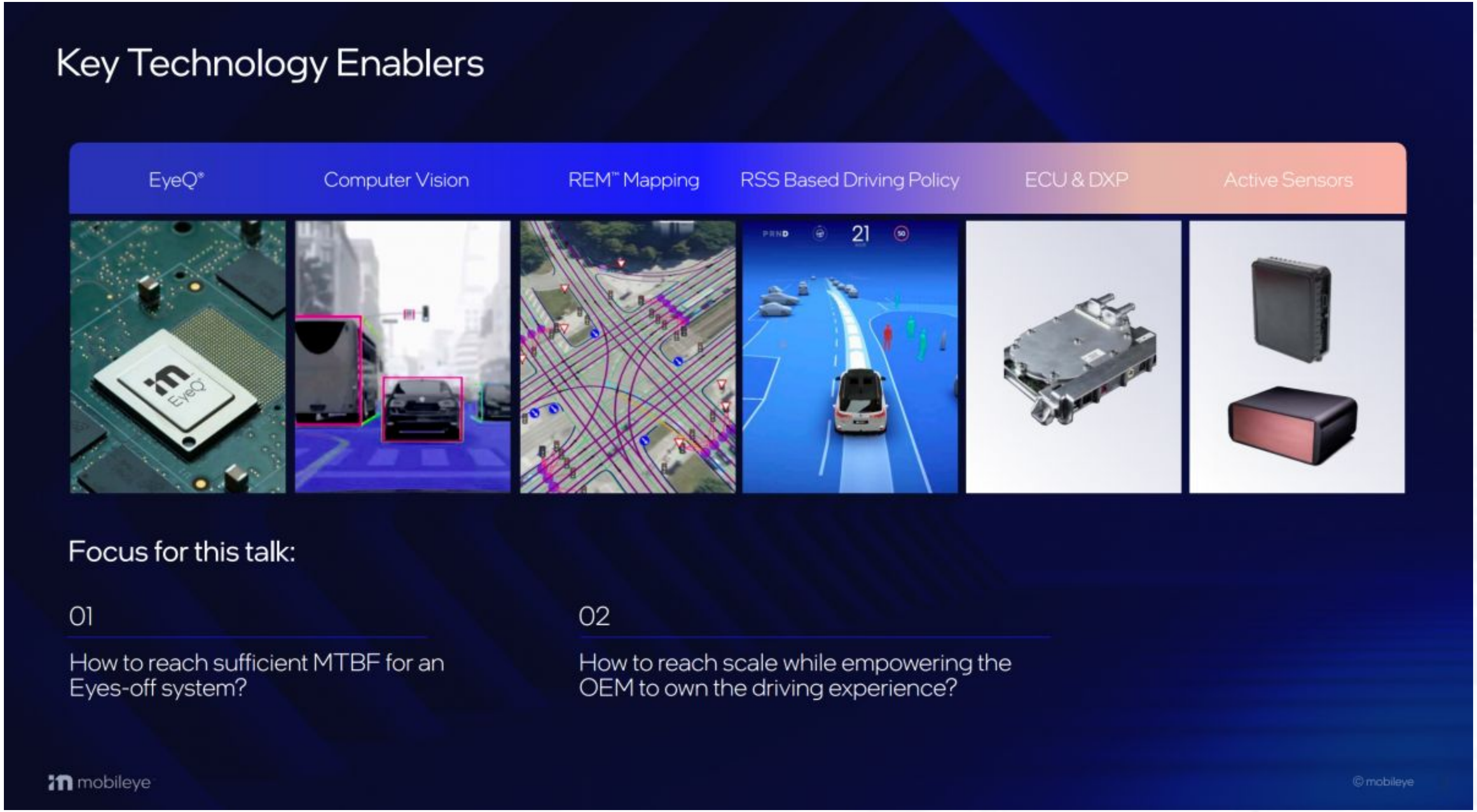


https://www.mobileye.com/ces-2024/

# Mobileye 技术路线 - Pipeline

MTBF：
Mean time between failures

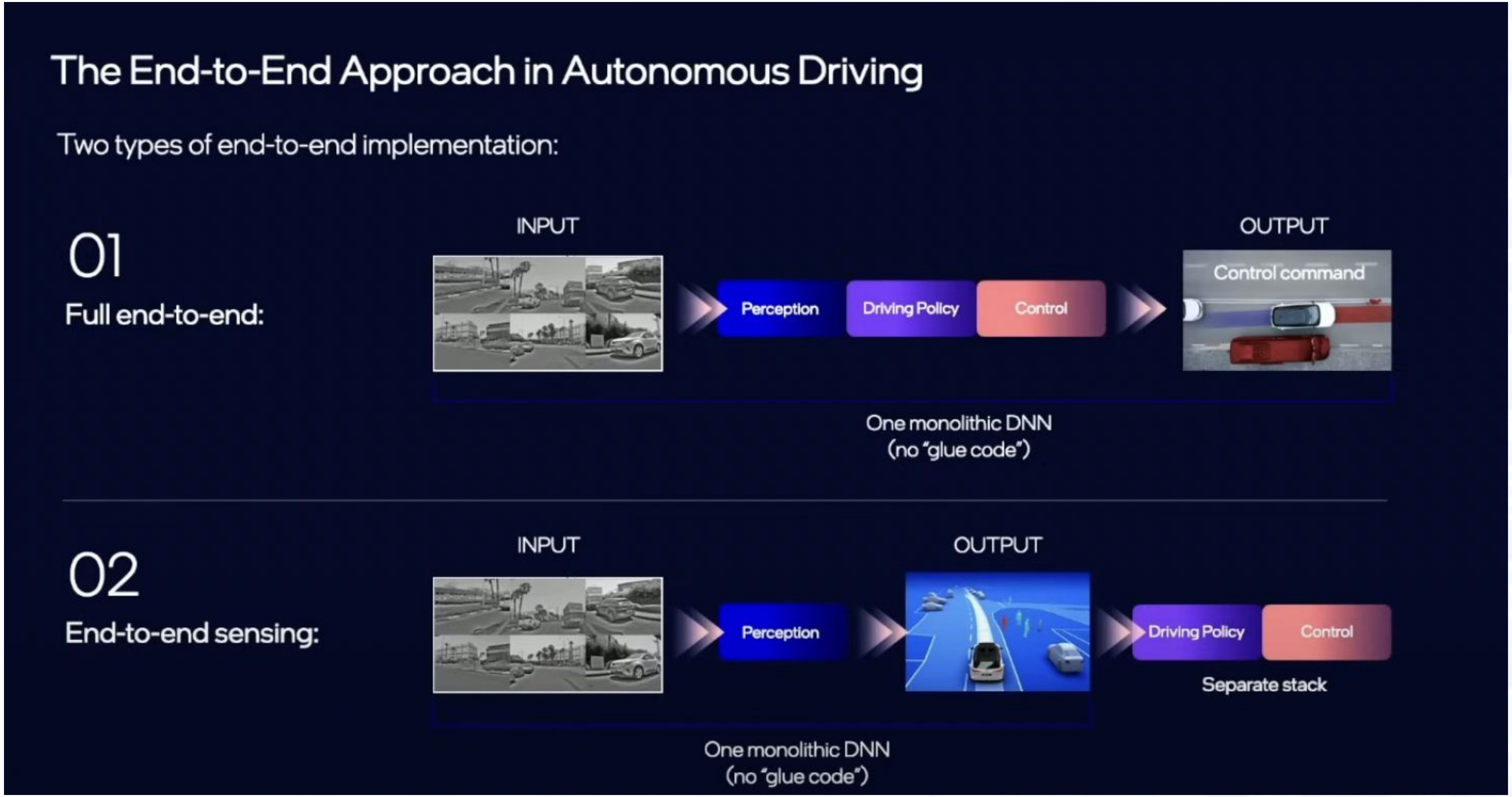
平均接管时间越大越好
（和测试环境有关）



https://www.mobileye.com/ces-2024/

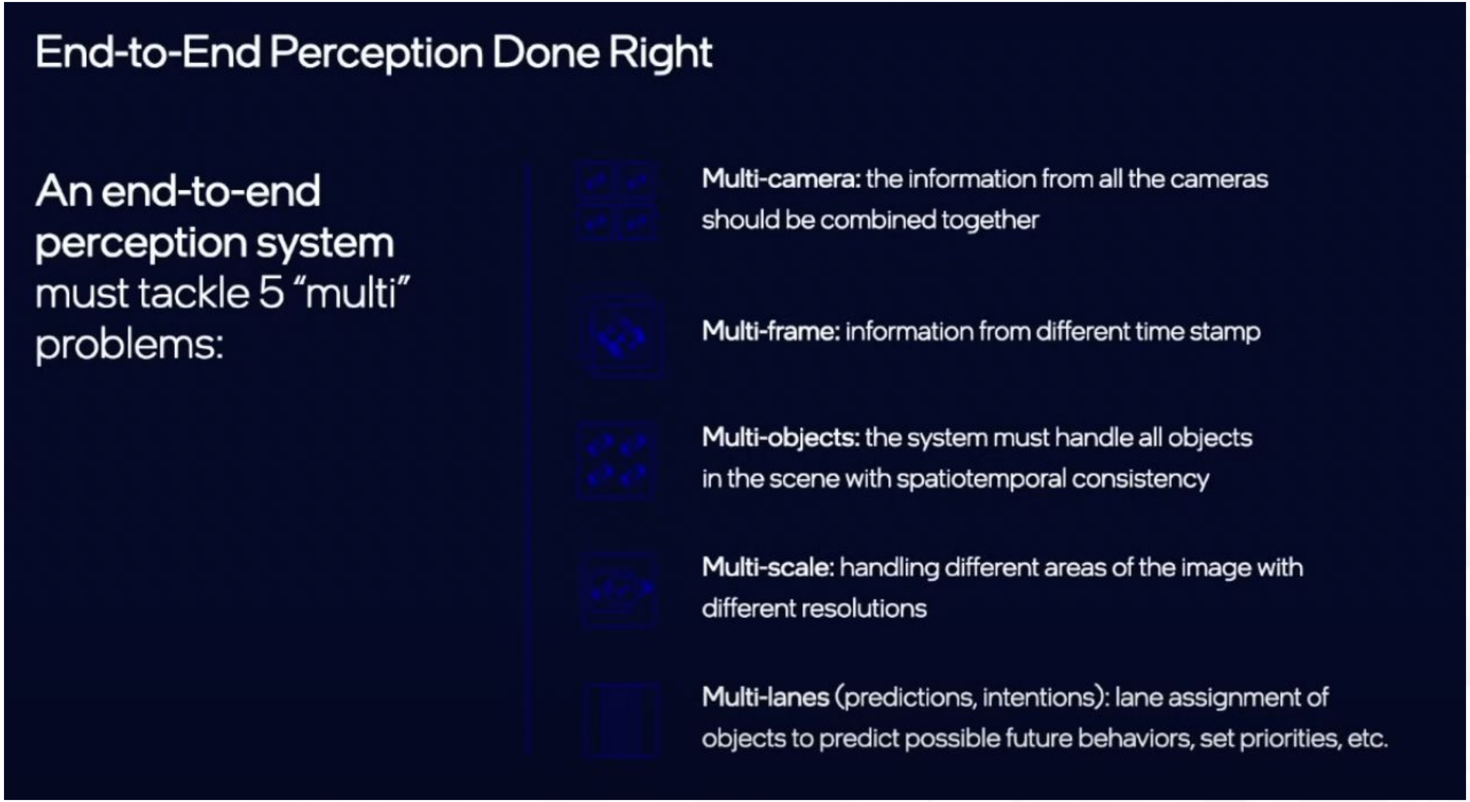
# Mobileye 技术路线 - 两种端到端

Monolithic:
完整的、
统一的
*integrated*

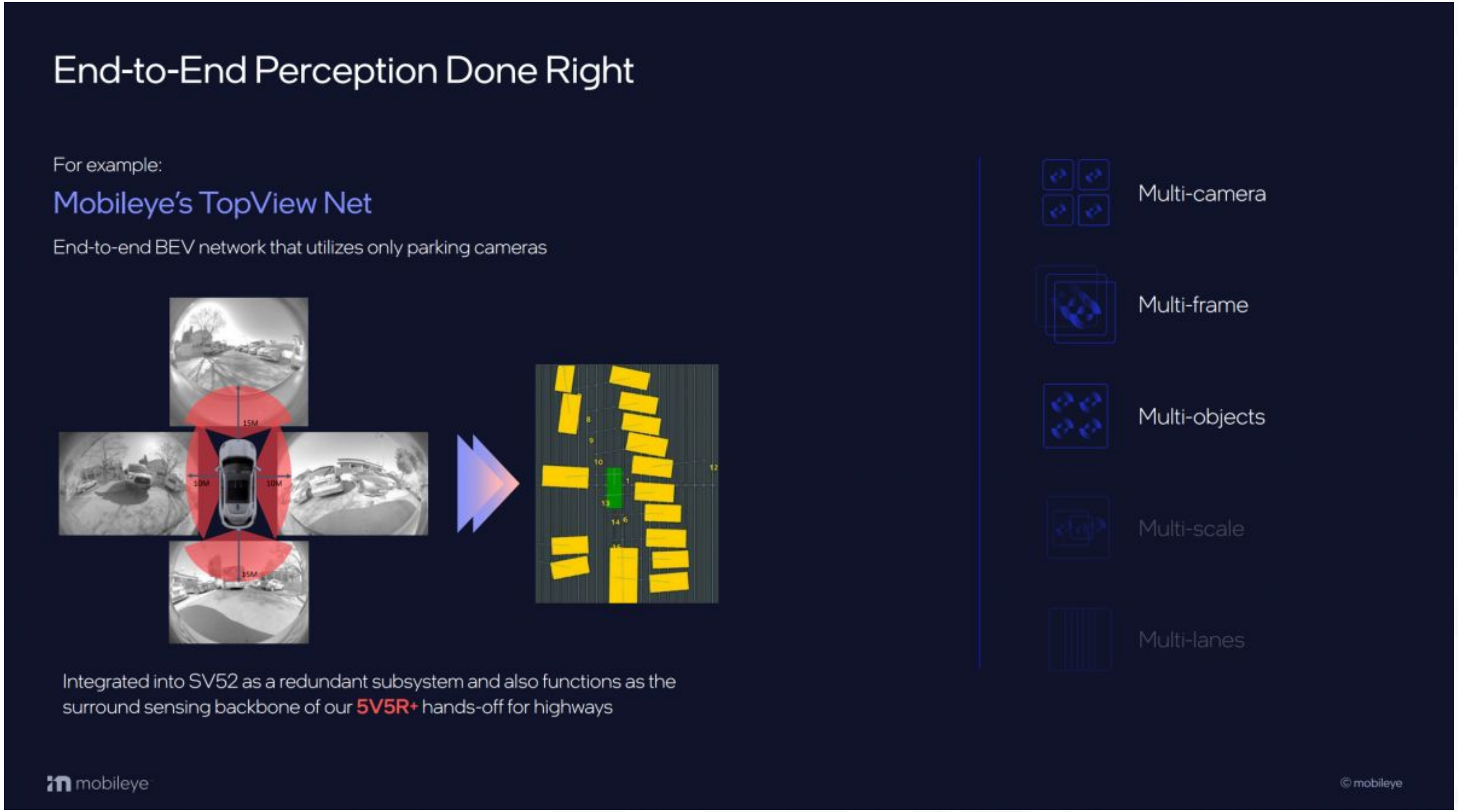


https://www.mobileye.com/ces-2024/

# Mobileye - End-to-end Perception Done Right

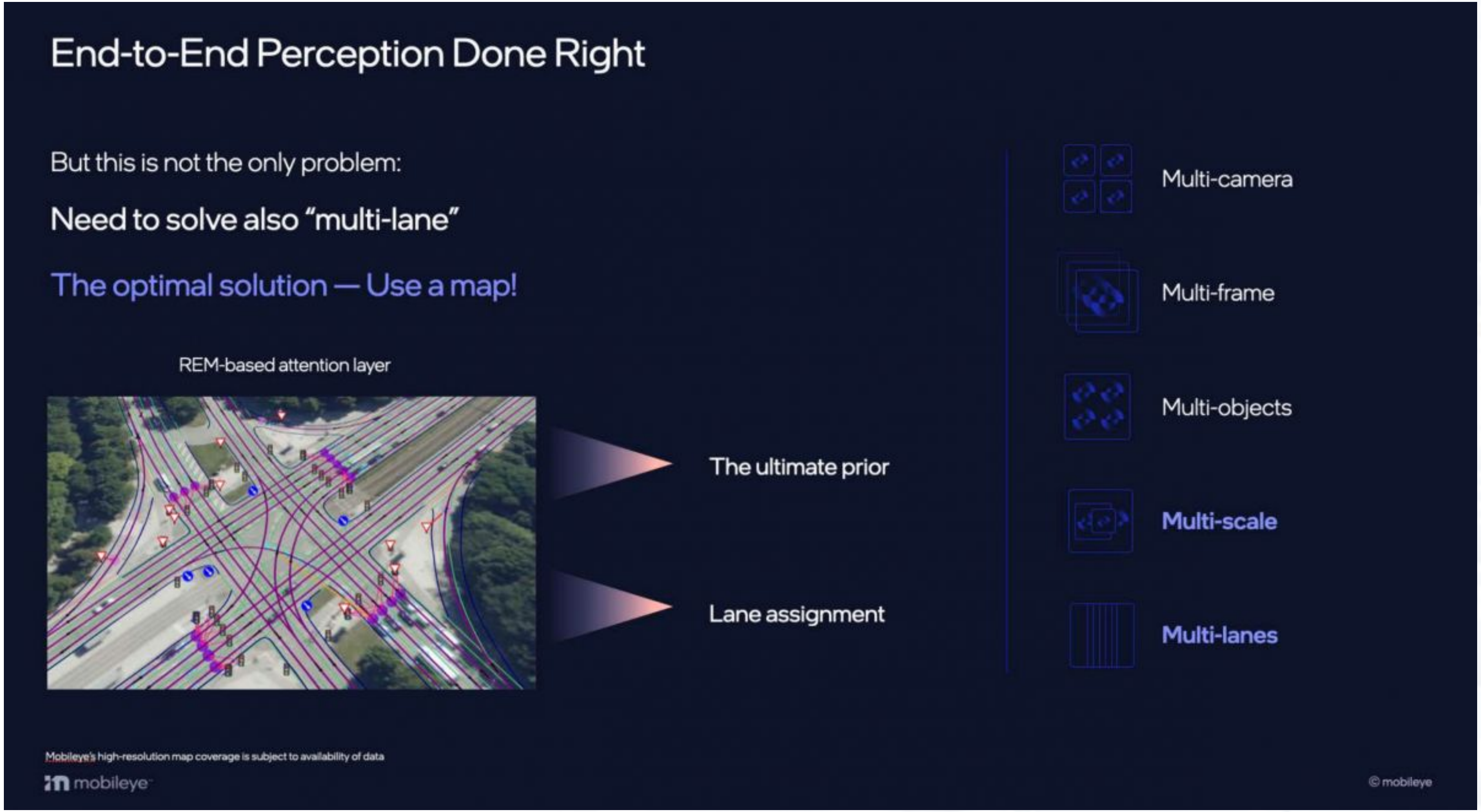


# Mobileye - End-to-end Perception Done Right



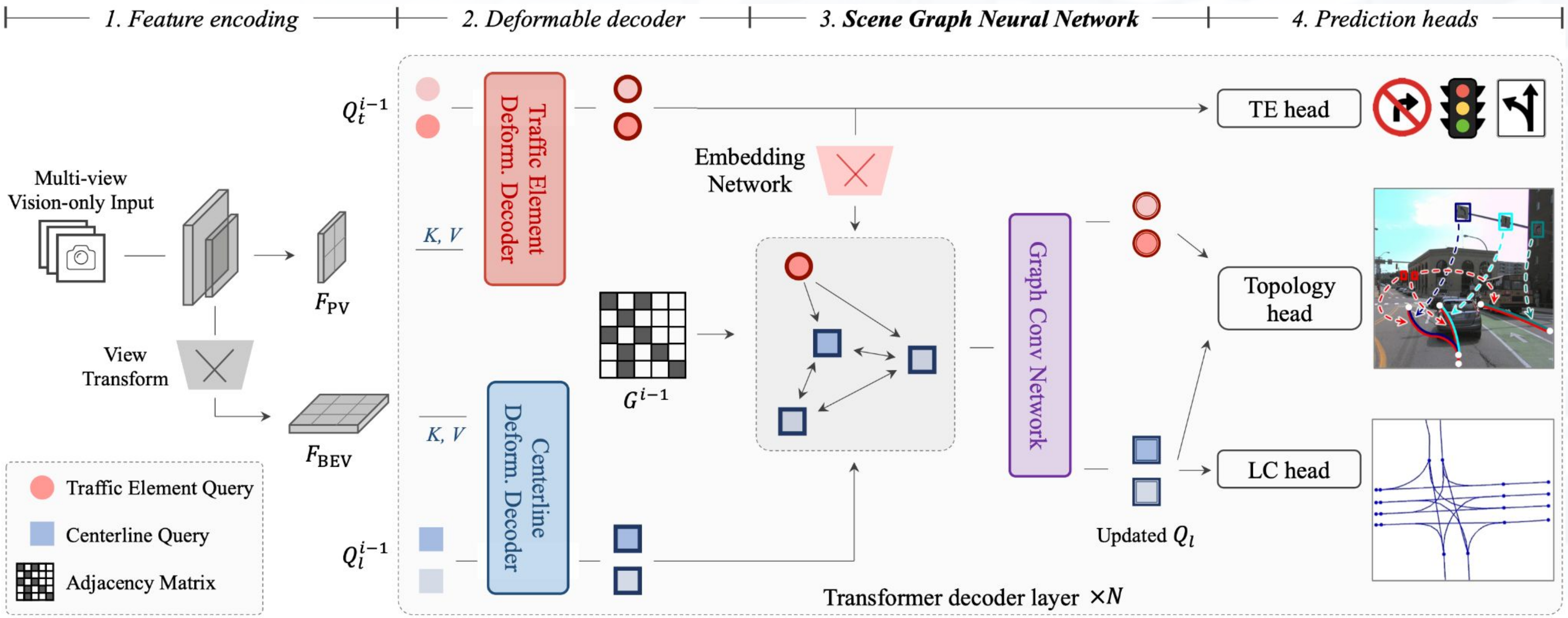
# Mobileye - End-to-end Perception Done Right - Multi-scale


# Mobileye - End-to-end Perception Done Right - Multi-lanes

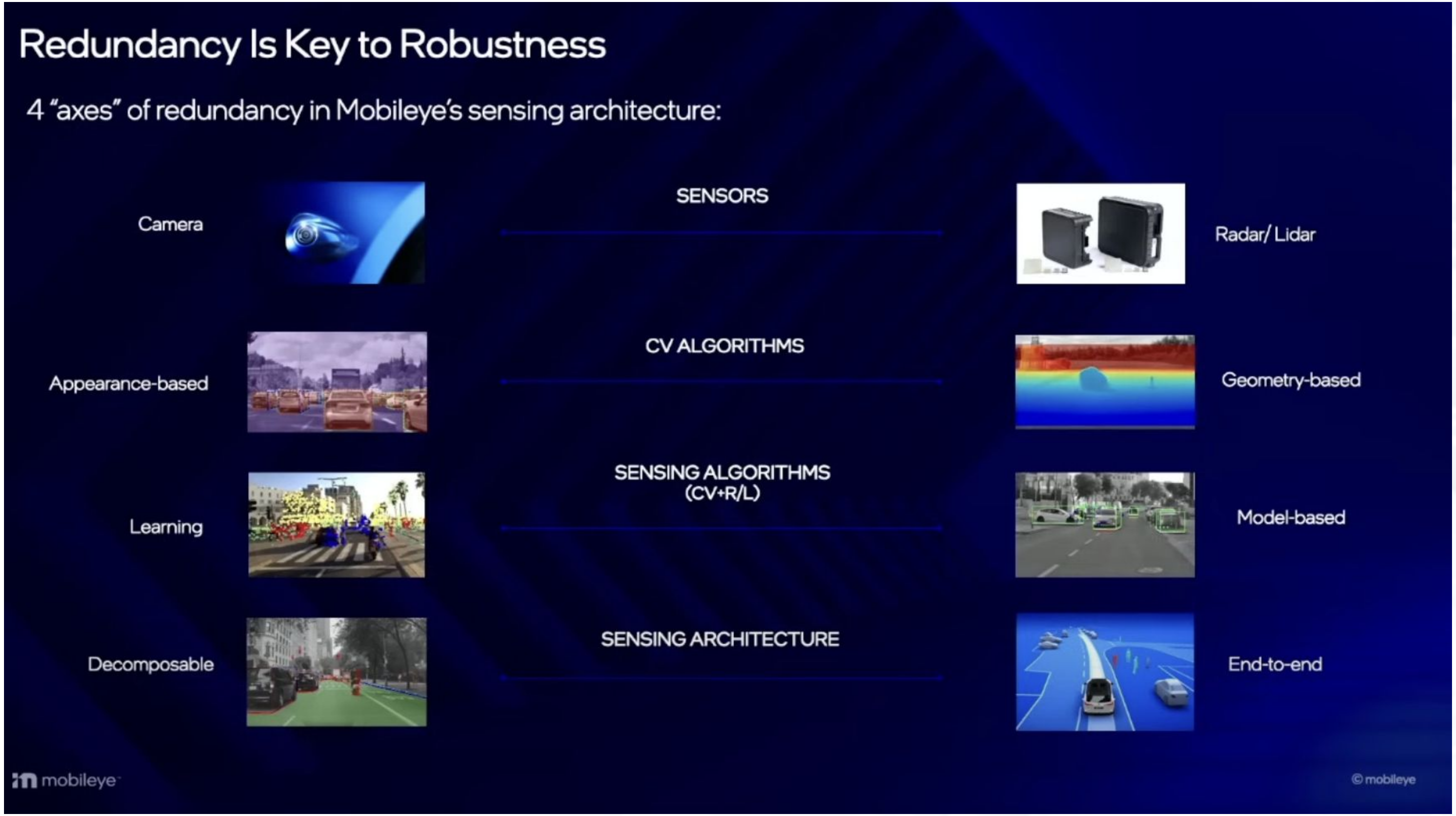


# Mobileye - 车道线 / 道路结构认知

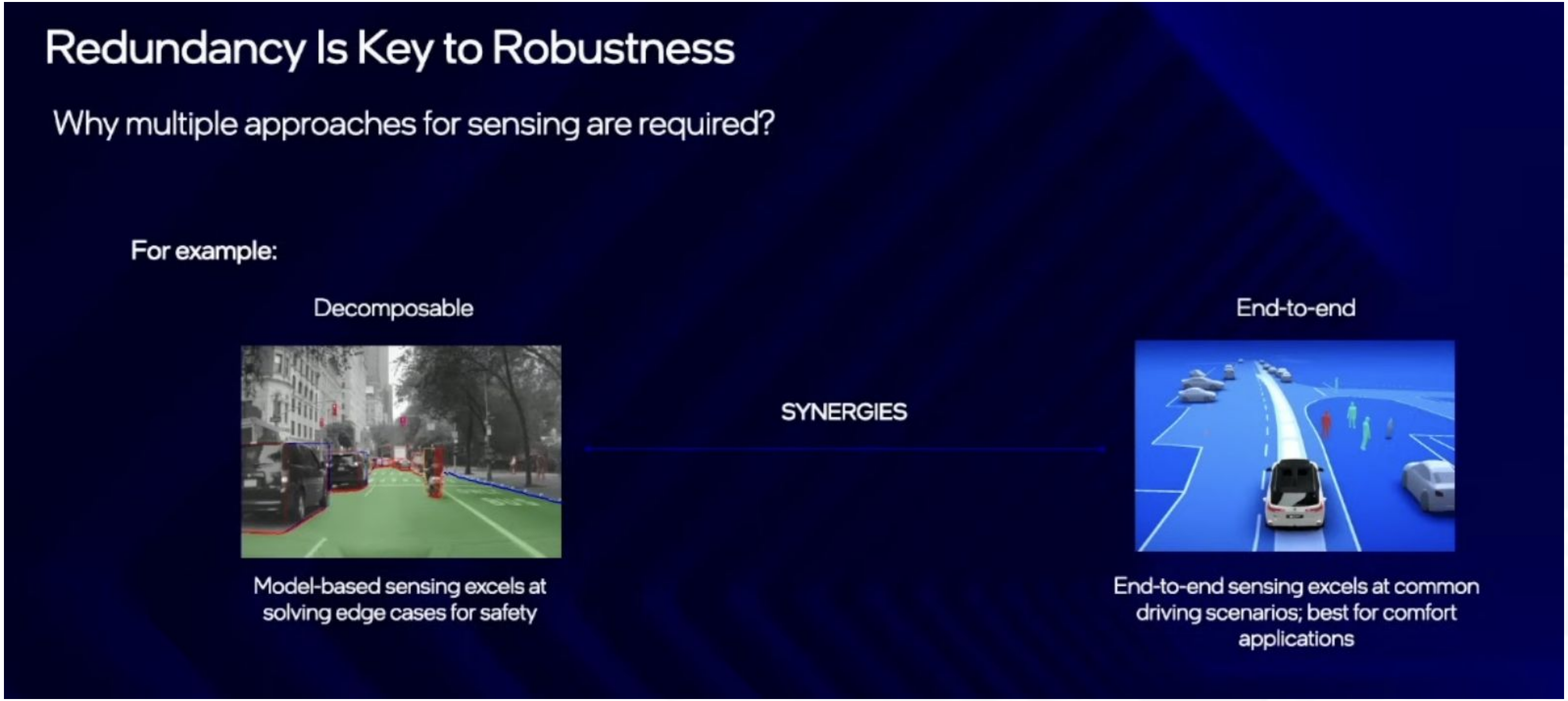
- TopoNet: https://arxiv.org/abs/2304.05277
- From landline detection to topology reasoning



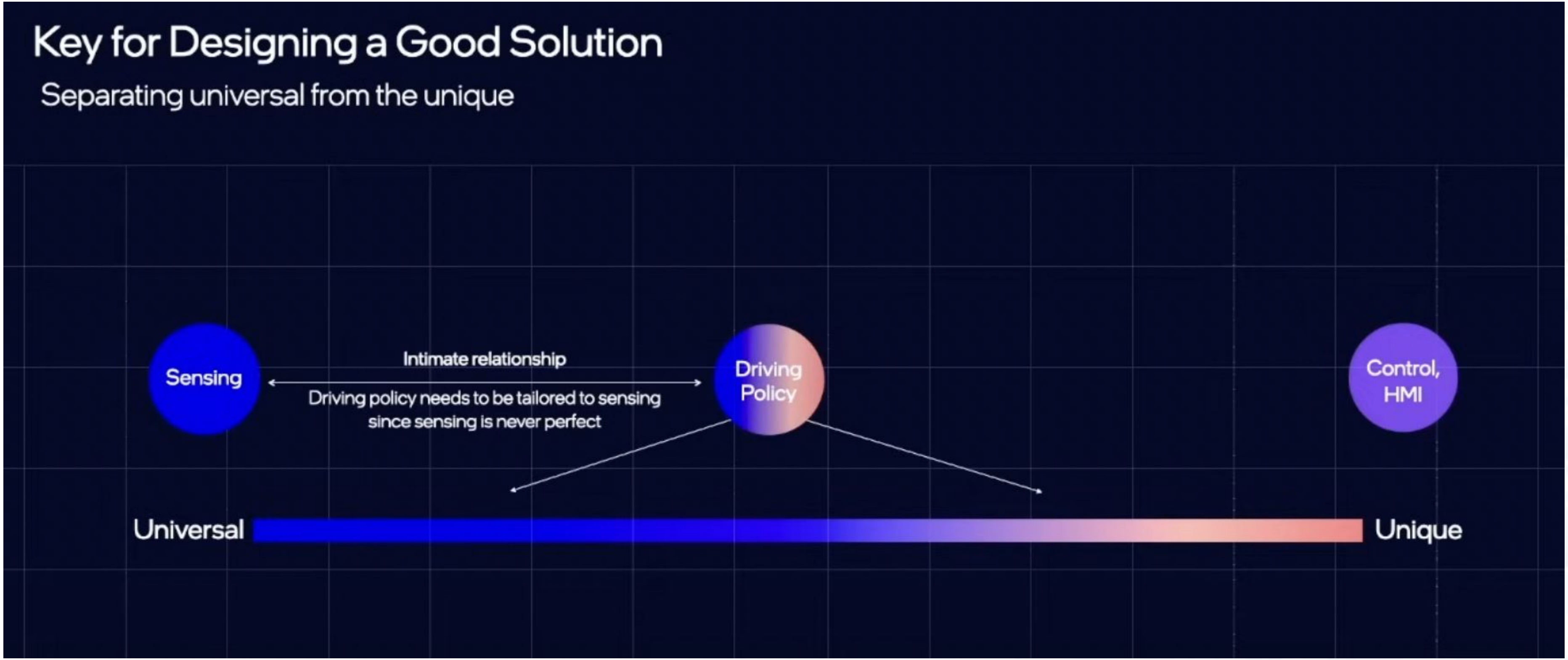
# Mobileye - Redundancy



# Mobileye - Redundancy



# Mobileye - How to scale while empowering OEM to own driving?



# Mobileye - How to scale while empowering OEM to own driving?

## Key for Designing a Good Solution

Separating universal from the unique

### Universal (OEM-agnostic)

**Facts**

- Perception of the surroundings (objects, road users, etc.)

**Semi-facts (predicting the future)**

- Intentions of road users

**Uncertainties**

- Lack of visibility, occlusions, error bars, etc.

**Optimization**

- Efficient data structures (e.g., "find all lanes at distance *d* from a query point")
- Optimization engines (e.g., "given desired offset per each road user, and lateral limiters, optimize a trajectory")

### Unique (OEM-specific)

**Discrete driving decisions**

- Lane changes, overtakes, yield or take-way, negotiation, etc.

**Continuous longitudinal planning**

- Acceleration and braking profiles/ jerk limiters
- Margins (keeping distance, headway, etc.)

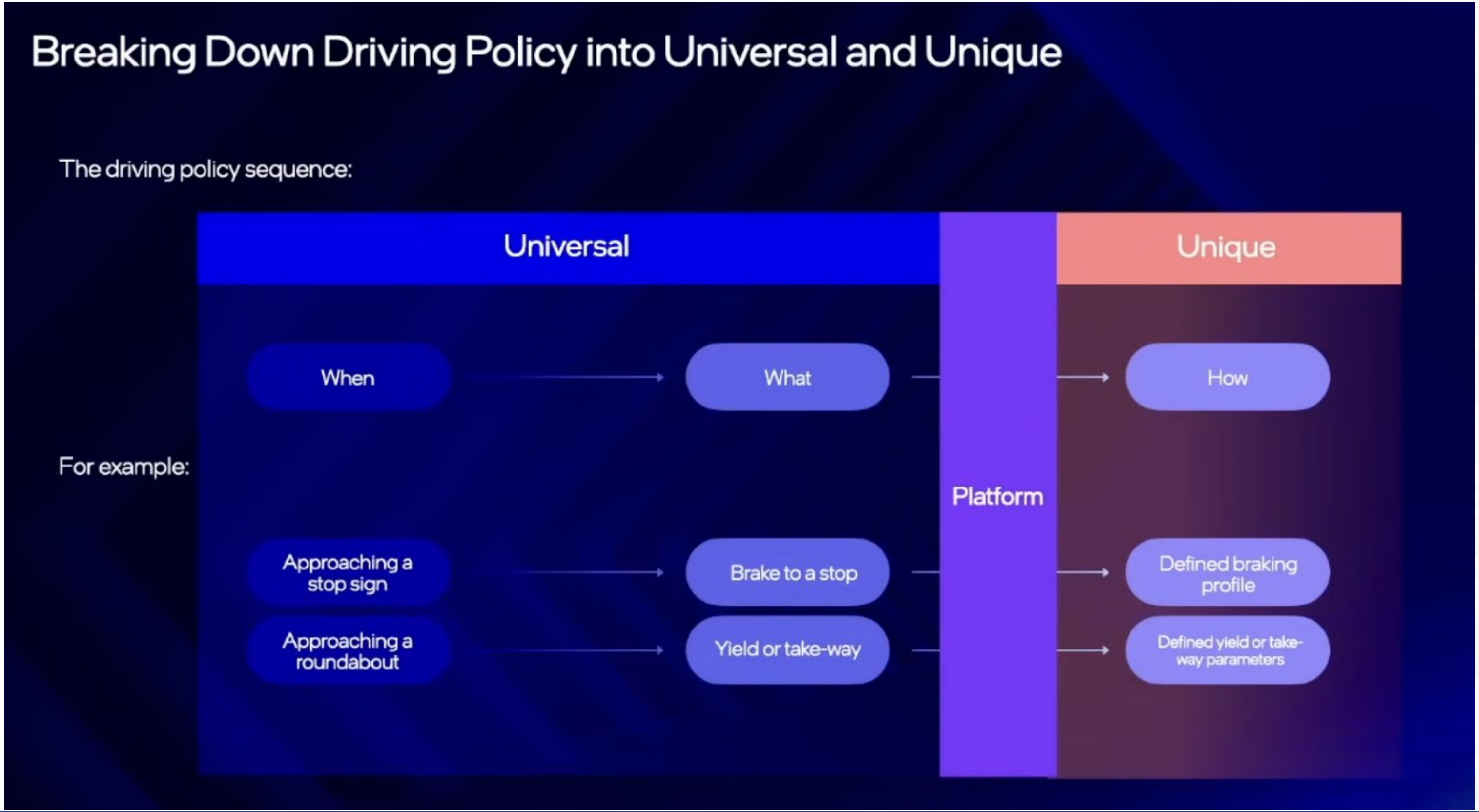
**Lateral planning**

- Lateral acceleration and velocity
- Offset parameters per road user

**Control**

**HMI**

# Mobileye - How to scale while empowering OEM to own driving?



OpenDriveLab

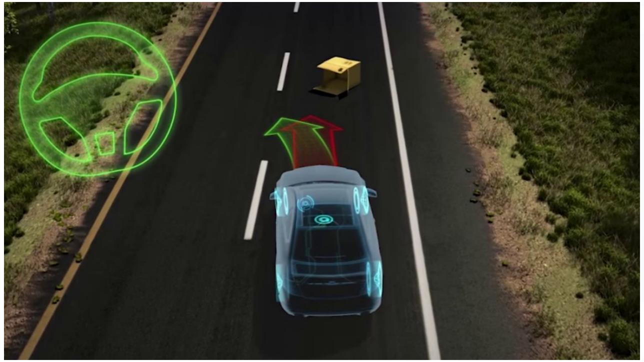
上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

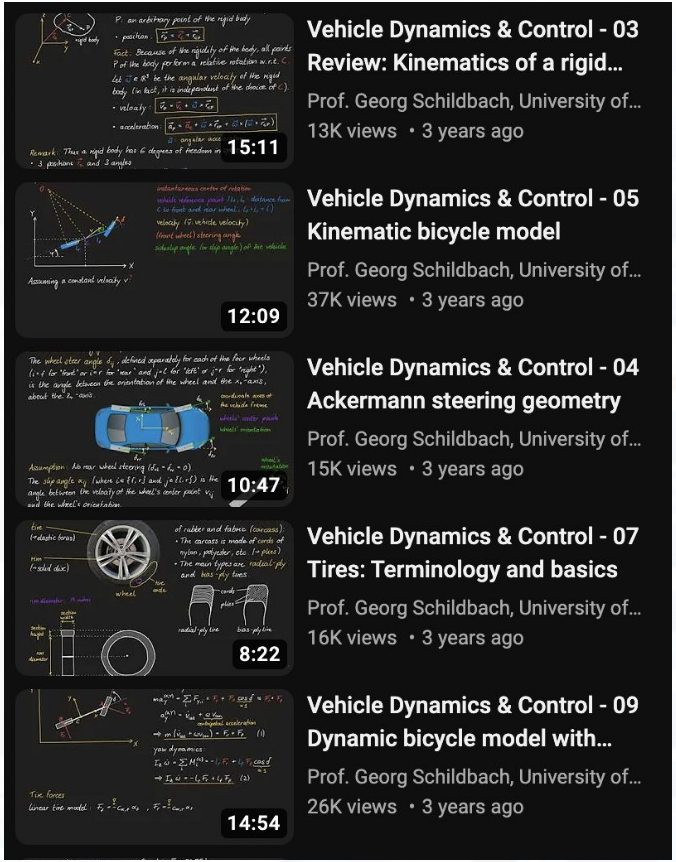
# Vehicle Dynamics

Software Algorithm

# 车辆动力学

- Credit from https://www.youtube.com/watch?v=Cg0L_HZYxP4 and Andreas Geiger
- Knowledge of **vehicle dynamics** enables accurate vehicle control

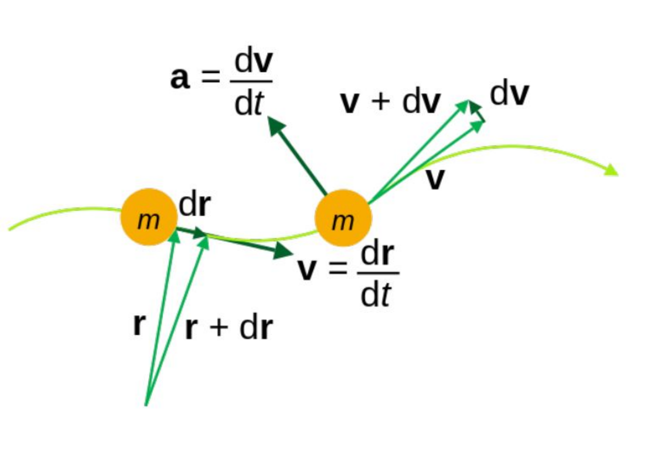




# 车辆动力学 - Kinematics（动力学） vs Kinetics（运动学，揭示规律）

**Kinematics:**

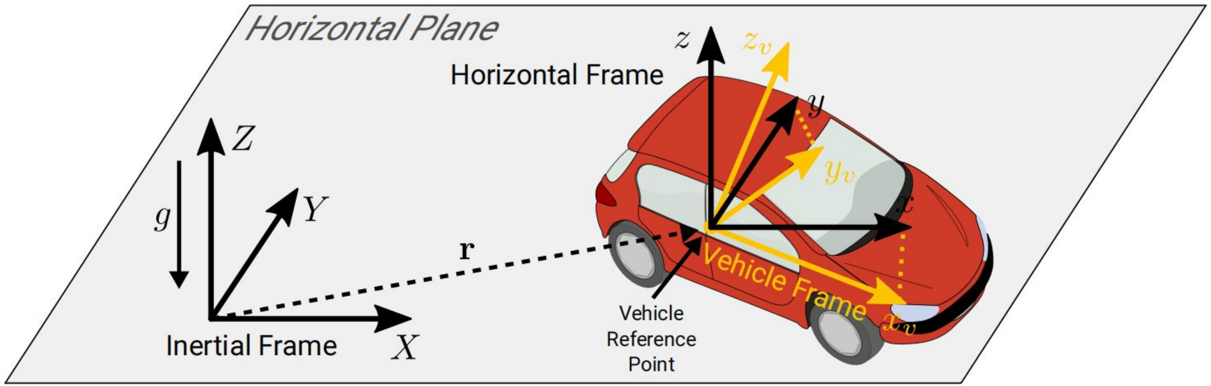
- Greek origin: “motion”, “moving”
- Describes motion of points and bodies
- Considers position, velocity, acceleration, ..
- Examples: Celestial bodies, particle systems, robotic arm, human skeleton




**Kinetics:**

- Describes causes of motion
- Effects of forces/moments
- Newton’s laws, e.g., $F = ma$

# 车辆动力学 - Coordinate Systems



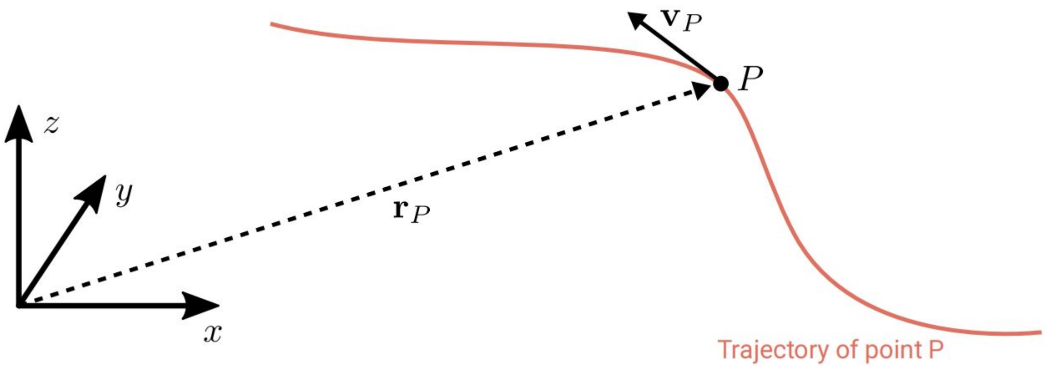

- **Inertial Frame:** Fixed to earth with vertical $Z$-axis and $X/Y$ horizontal plane
- **Vehicle Frame:** Attached to vehicle at fixed reference point; $x_v$ points towards the front, $y_v$ to the side and $z_v$ to the top of the vehicle (ISO 8855)
- **Horizontal Frame:** Origin at vehicle reference point (like vehicle frame) but $x$- and $y$-axes are projections of $x_v$- and $y_v$-axes onto the $X/Y$ horizontal plane

# 车辆动力学 - Kinematics of a Point

The **position** $\mathbf{r}_P(t) \in \mathbb{R}^3$ of point $P$ at time $t \in \mathbb{R}$ is given by 3 coordinates.

**Velocity** and **acceleration** are the first and second derivatives of the position $\mathbf{r}_P(t)$.

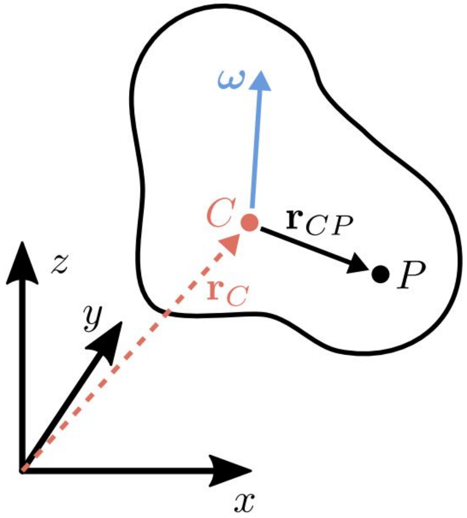
$$\mathbf{r}_P(t) = \begin{pmatrix} x(t) \\ y(t) \\ z(t) \end{pmatrix} \qquad \mathbf{v}_P(t) = \dot{\mathbf{r}}_P(t) = \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \\ \dot{z}(t) \end{pmatrix} \qquad \mathbf{a}_P(t) = \ddot{\mathbf{r}}_P(t) = \begin{pmatrix} \ddot{x}(t) \\ \ddot{y}(t) \\ \ddot{z}(t) \end{pmatrix}$$



# 车辆动力学 - Kinematics of a Rigid Body

A **rigid body** refers to a collection of infinitely many infinitesimally small mass points which are rigidly connected, i.e., their relative position remains unchanged over time. It's **motion** can be compactly described by the motion of an (arbitrary) reference point $C$ of the body plus the relative motion of all other points $P$ with respect to $C$.

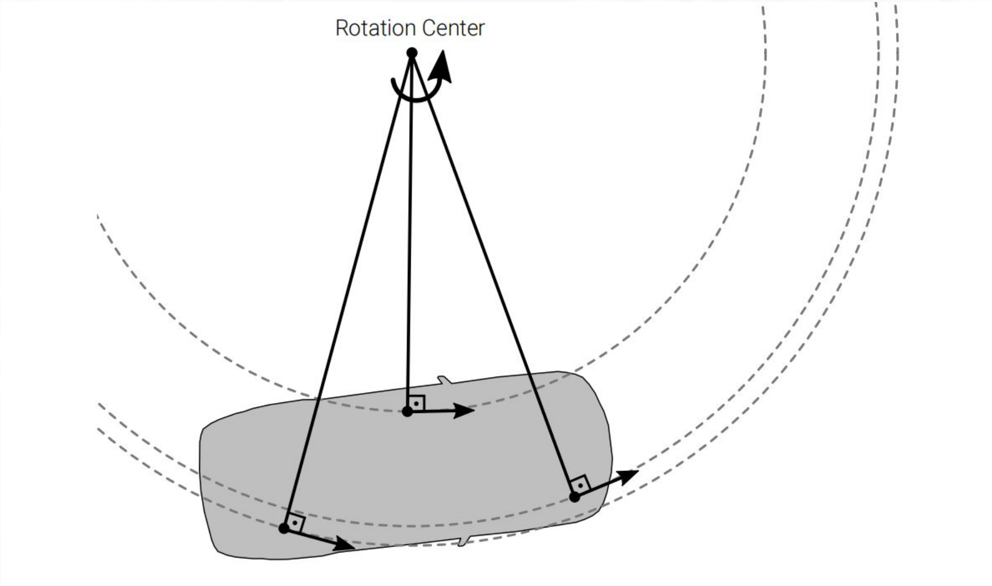
- $C$: Reference point fixed to rigid body
- $P$: Arbitrary point on rigid body
- $\boldsymbol{\omega}$: Angular velocity of rigid body
- Position: $\mathbf{r}_P = \mathbf{r}_C + \mathbf{r}_{CP}$
- Velocity: $\mathbf{v}_P = \mathbf{v}_C + \boldsymbol{\omega} \times \mathbf{r}_{CP}$
- Due to rigidity, points $P$ can only rotate wrt. $C$
- Thus a rigid body has 6 DoF (3 pos., 3 rot.)



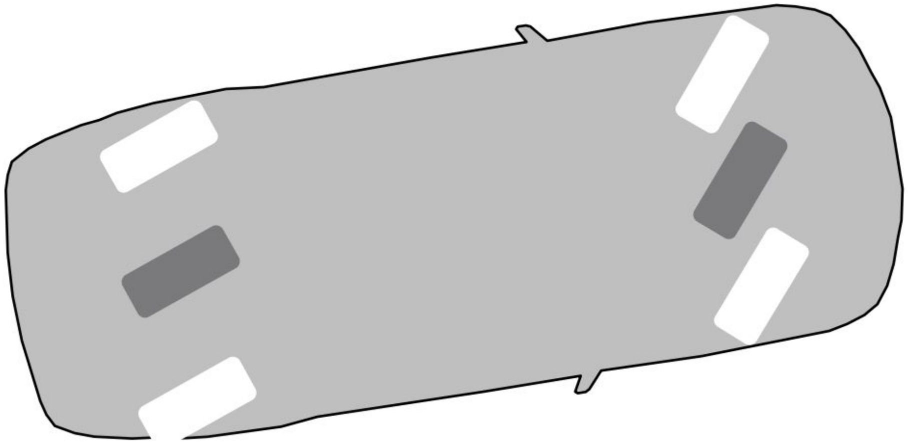
# 车辆动力学 - Kinematic Bicycle Model (1/n)

Rigid body motion:

Different points on the rigid body move along different circular trajectories

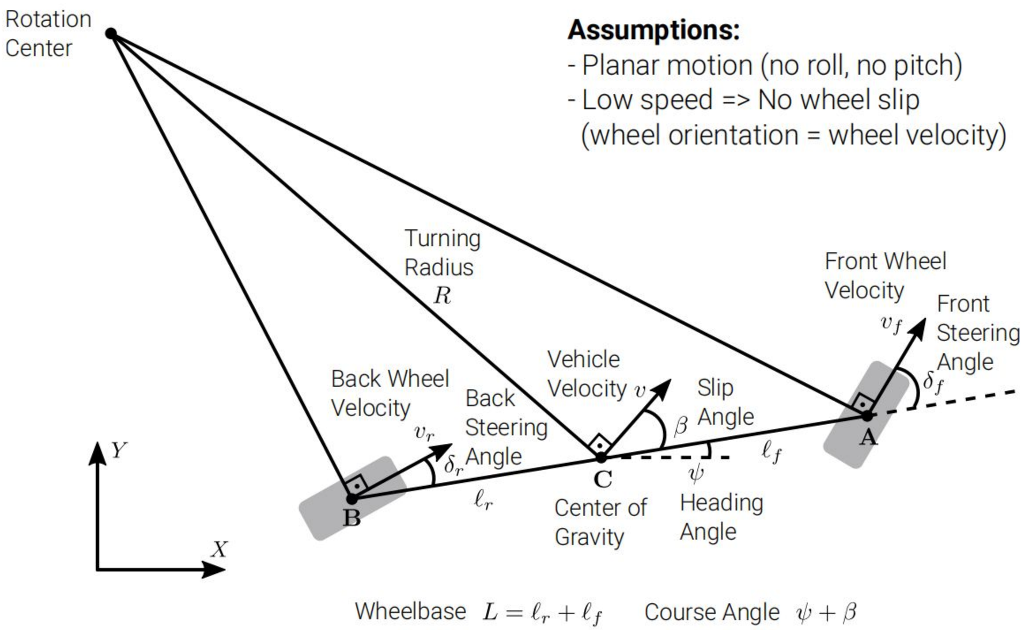


# 车辆动力学 - Kinematic Bicycle Model (1/n)



- The **kinematic bicycle model** approximates the 4 wheels with 2 imaginary wheels

# 车辆动力学 - Kinematic Bicycle Model (1/n)

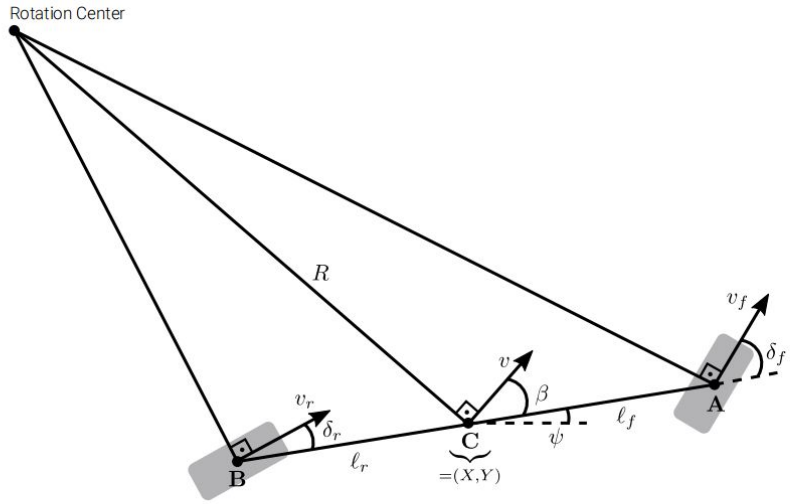



**Assumptions:**
- Planar motion (no roll, no pitch)
- Low speed => No wheel slip
  (wheel orientation = wheel velocity)

- The **kinematic bicycle model** approximates the 4 wheels with 2 imaginary wheels

# 车辆动力学 - Kinematic Bicycle Model (1/n)

## Model




## Motion Equations

$$\dot{X} = v\cos(\psi + \beta)$$

$$\dot{Y} = v\sin(\psi + \beta)$$

$$\dot{\psi} = \frac{v\cos(\beta)}{\ell_f + \ell_r}(\tan(\delta_f) - \tan(\delta_r))$$

$$\beta = \tan^{-1}\left(\frac{\ell_f\tan(\delta_r) + \ell_r\tan(\delta_f)}{\ell_f + \ell_r}\right)$$

(proof as exercise)

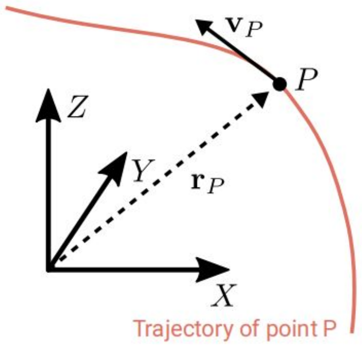
# 车辆动力学 - Dynamics of a Rigid Body

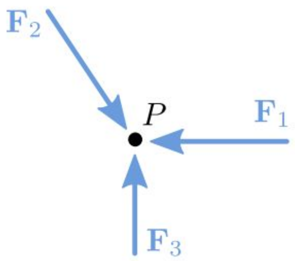
**Translatory Motion of a Point:**

- Consider **point** $P$ with mass $m$ in $\mathbb{R}^3$
- Let $\mathbf{r}_P(t) \in \mathbb{R}^3$ be its **position** in an inertial reference frame
- Let $\mathbf{v}_P(t)$ denote its **velocity** and $\mathbf{a}_P(t)$ its **acceleration**
- The **linear momentum** of $P$ is defined as $\mathbf{p}_P(t) = m\mathbf{v}_P(t)$
- By **Newton's second law** we have

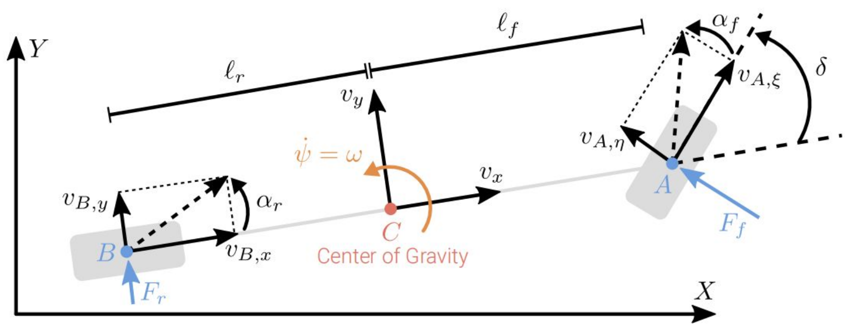
$$\frac{d}{dt}\mathbf{p}_P(t) = m\,\mathbf{a}_P(t) = \mathbf{F}_{net}(t) = \sum_i \mathbf{F}_i(t)$$

where $\mathbf{F}_i(t)$ represent all forces acting on the point mass $P$



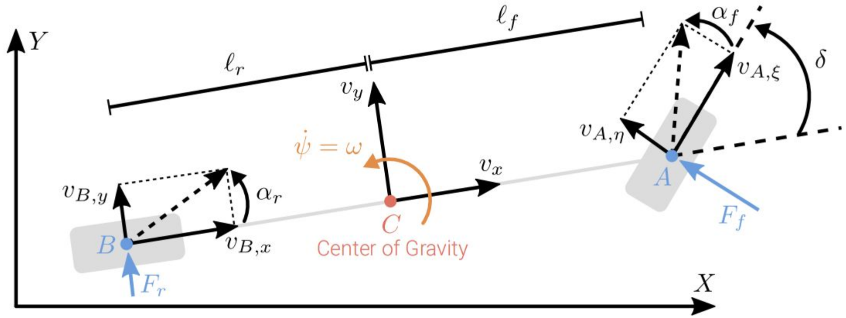
# 车辆动力学 - Dynamic Bicycle Model (1/2)




**Assumptions:**

- The vehicle's motion is restricted to the X/Y plane
- The vehicle is considered as a rigid body
- Only lateral tire forces, generated by a linear tire model
- Small steering angle $\delta$: $\sin\delta \approx \delta \quad \tan\delta \approx \delta \quad \cos\delta \approx 1$
- Constant longitudinal velocity $v_x$

# 车辆动力学 - Dynamic Bicycle Model (2/2)




**State Space Representation:**

$$\begin{bmatrix} \dot{v}_y \\ \dot{\psi} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} -\frac{c_r + c_f}{m v_x} & 0 & \frac{c_r l_r - c_f l_f}{m v_x} - v_x \\ 0 & 0 & 1 \\ \frac{l_r c_r - l_f c_f}{I_z v_x} & 0 & -\frac{l_f^2 c_f + l_r^2 c_r}{I_z v_x} \end{bmatrix} \underbrace{\begin{bmatrix} v_y \\ \psi \\ \omega \end{bmatrix}}_{\text{State}} + \begin{bmatrix} \frac{c_f}{m} \\ 0 \\ \frac{c_f}{I_z} l_f \end{bmatrix} \underbrace{\delta}_{\text{Input}}$$

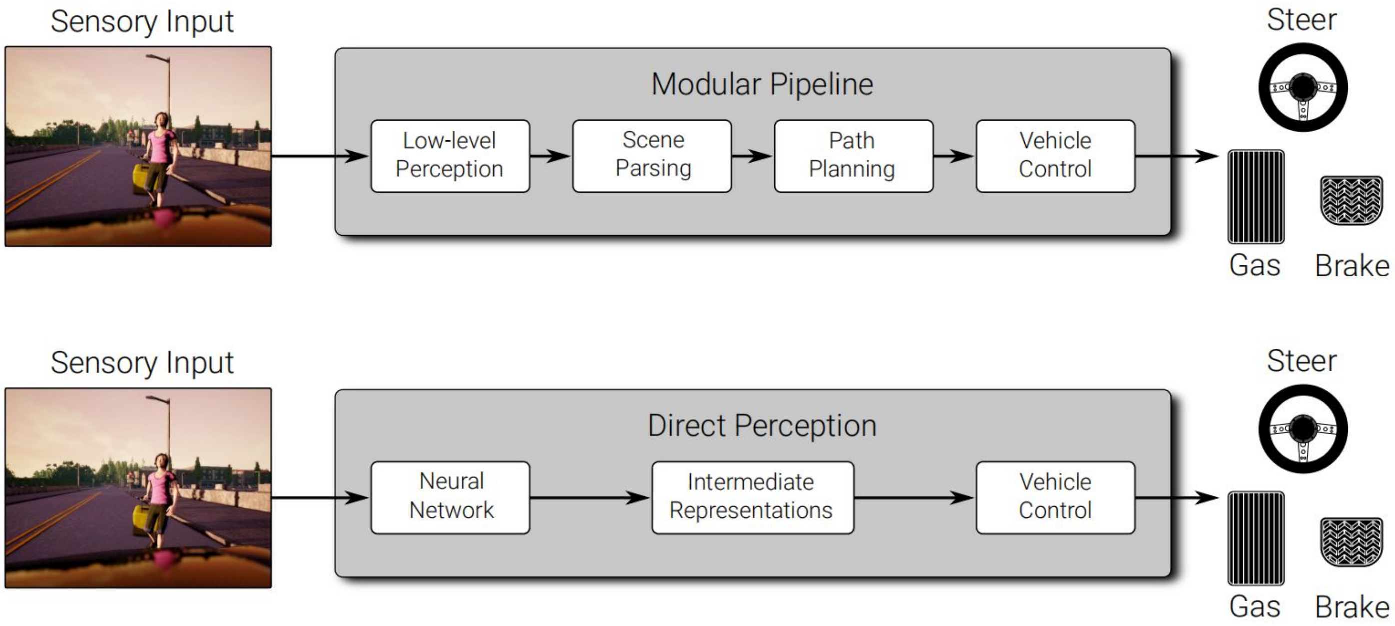
Can be augmented by the global position to a nonlinear state space model

# 车辆动力学 - Summary

- A vehicle can be modeled as a rigid body
- It is subject to holonomic and non-holonomic constraints
- The bicycle model approximates the vehicle using 2 wheels
- The kinematic bicycle model assumes no **wheel slip** (low speeds) **车轮滑转**
- However, modeling tires requires to consider slip
- **Sliding friction** 滑动摩擦 is smaller than static friction
- We want to operate in the static friction area of the force curve
- The circle of forces tells us that lat. and long. forces are dependent
- **The dynamic bicycle model** takes into account tire forces and wheel slip

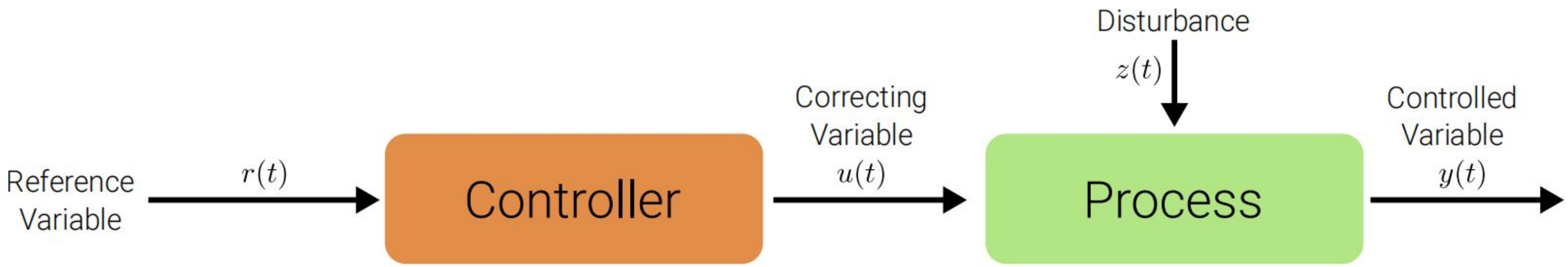
# Vehicle Dynamics meets Vehicle Control (1/3)

Approaches to Self-Driving



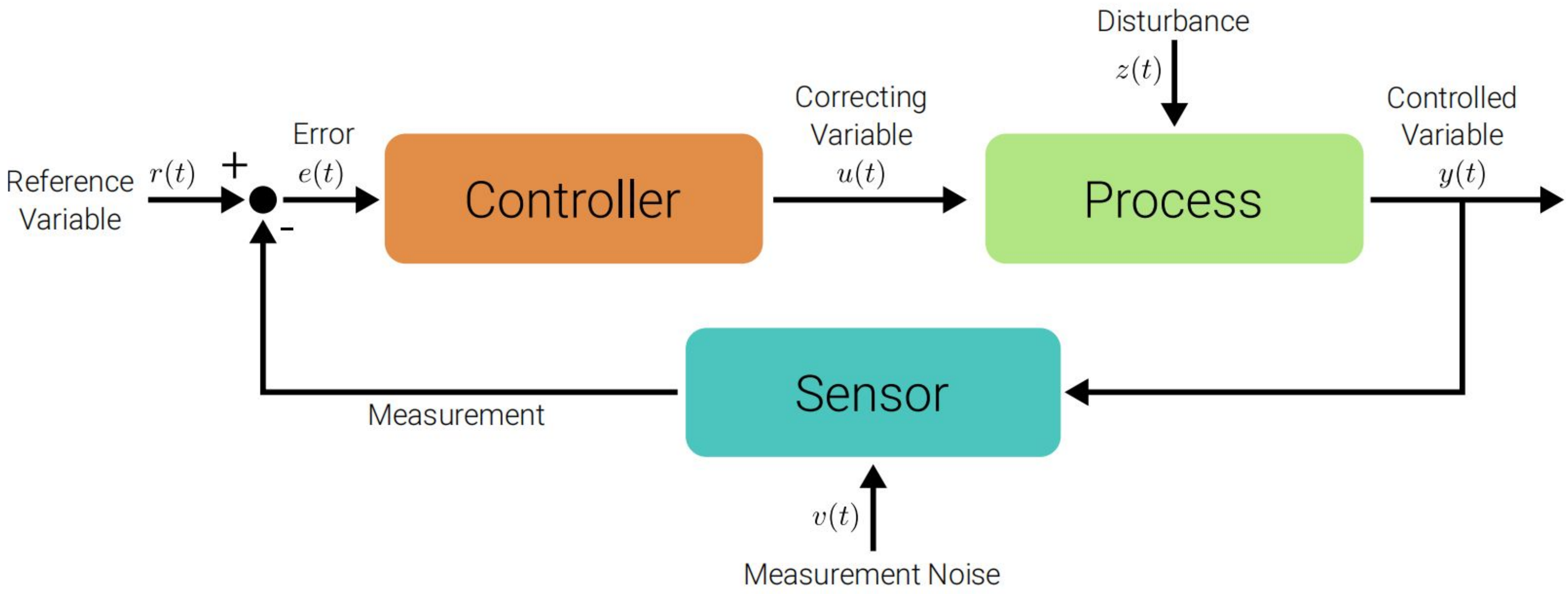

- Self-driving cars and driver assistance systems require **vehicle control**

# Vehicle Dynamics meets Vehicle Control (2/3) - open loop




- Requires **precise knowledge** of the plant and the influence factors
- **No feedback** about the controlled variable
- Cannot handle unknown disturbances, resulting in **drift**

# Vehicle Dynamics meets Vehicle Control (3/3) - closed loop




- Exploit feedback loop to minimize error between reference and measurement

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

# Datasets / Benchmarks / Ecosystem

Dr. Hongyang Li

Shanghai AI Lab and OpenDriveLab
Mar 13 2024

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

# 自动驾驶数据集综述

4 Datasets/Benchmarks/Ecosystem

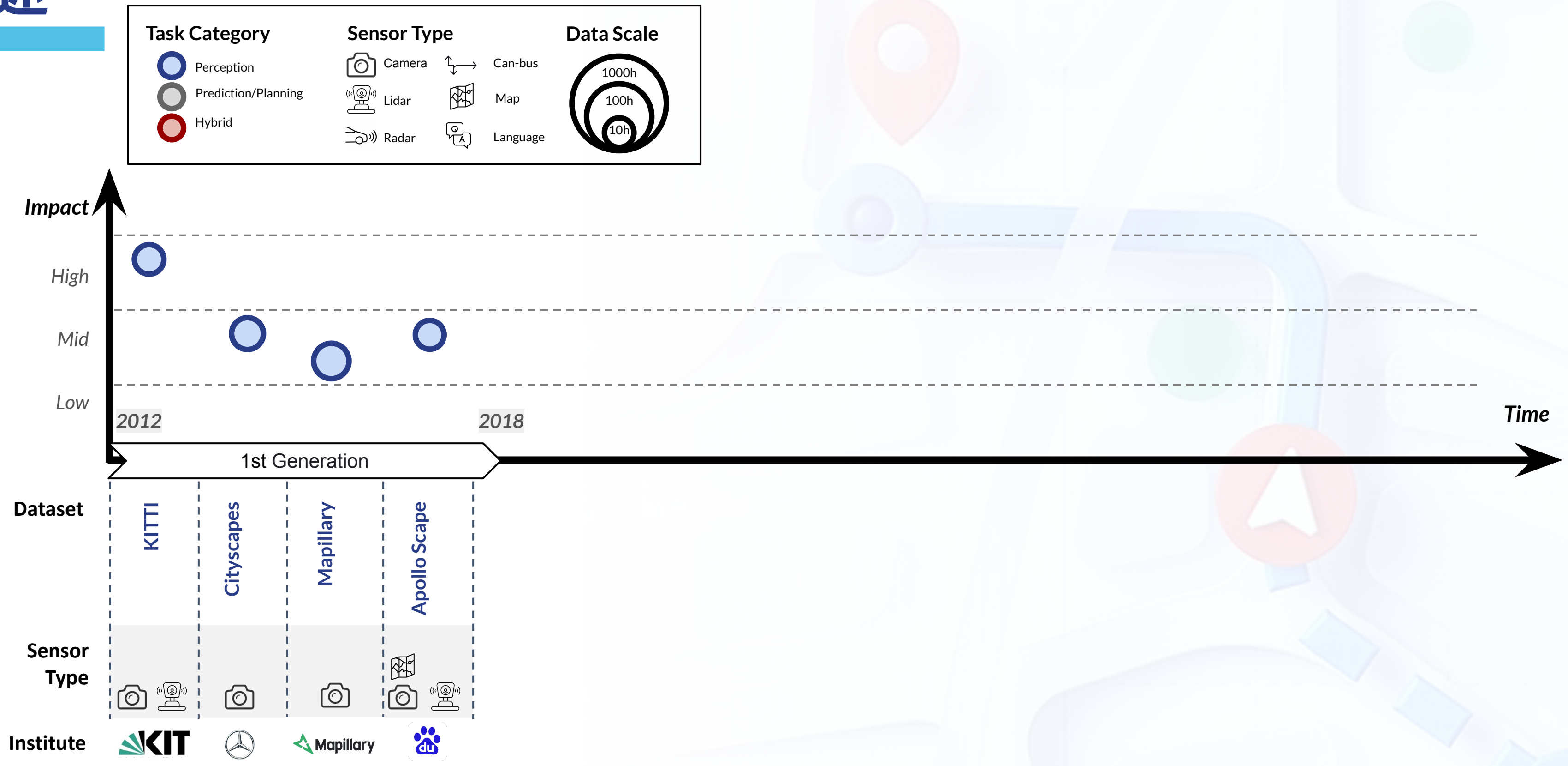
# 概述

**现有自动驾驶数据集可大致分为两代**：

第一代数据集的传感模态复杂度相对较低、数据集规模相对较小，且大多**局限于感知级任务**，以发布于 2012 年的 KITTI 为代表。

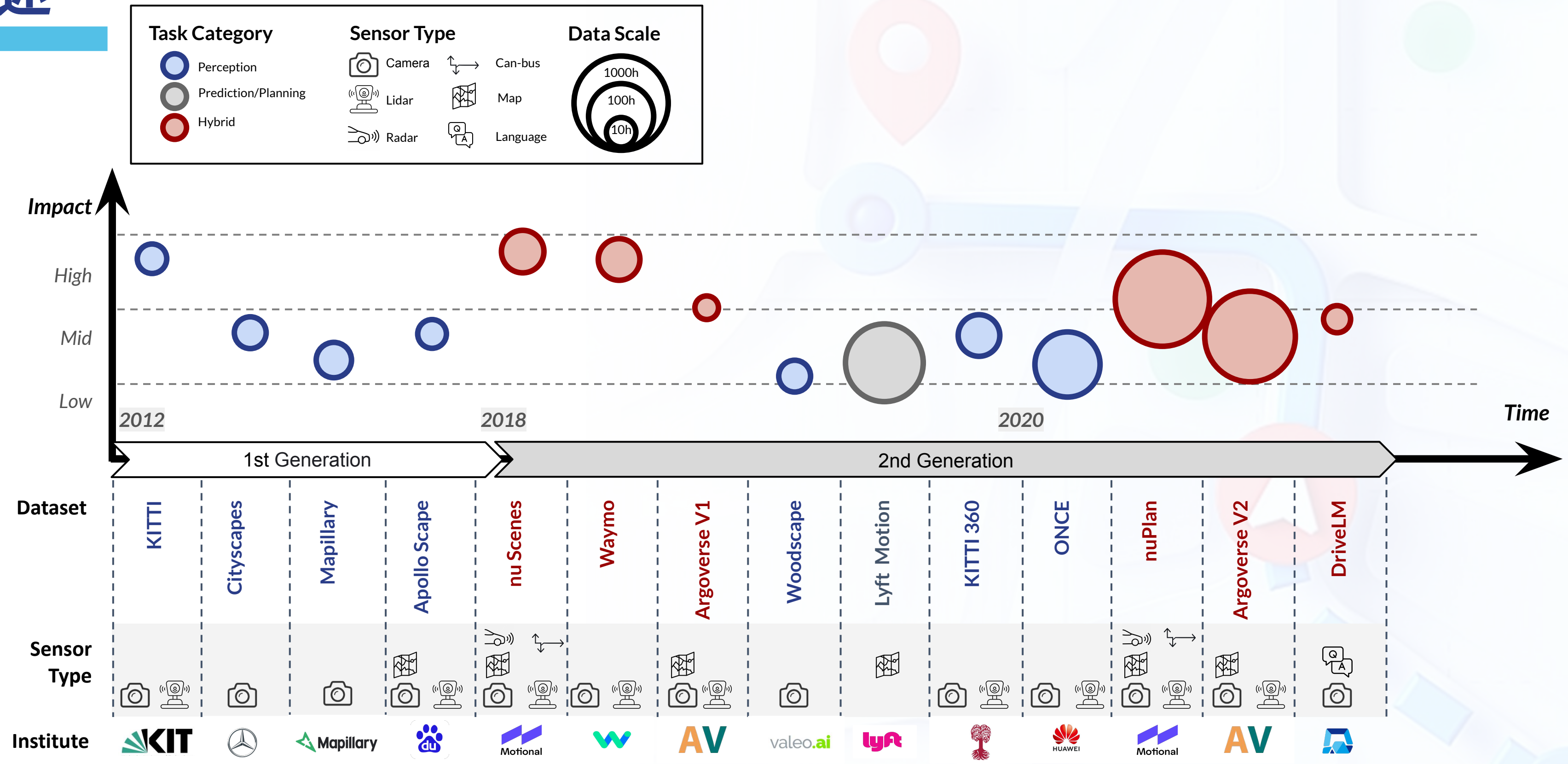
相比于第一代数据集，第二代数据集的特征为传感模态复杂度较高、数据集规模与多样性较丰富、所设置任务**从感知扩展到预测、规控上**，以 2019 年前后提出的 nuScenes、Waymo 为代表。

# 概述



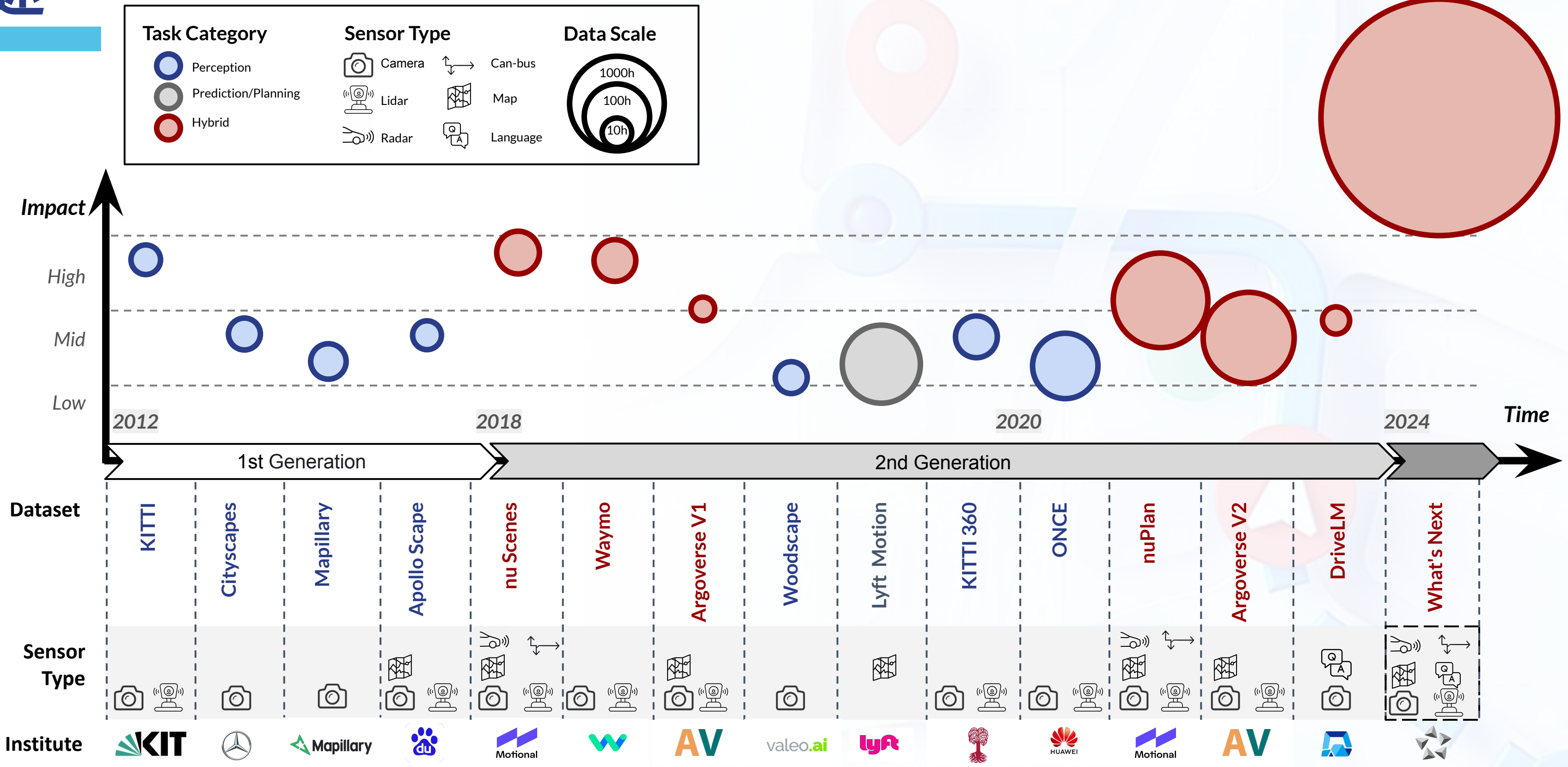
Task Category
Perception
Prediction/Planning
Hybrid
Sensor Type
Camera
Can-bus
Lidar
Map
Radar
Language
Data Scale
1000h
100h
10h
Impact
High
Mid
Low
2012
2018
Time
1st Generation
Dataset
KITTI
Cityscapes
Mapillary
Apollo Scape
Sensor Type
Institute
KIT
Mapillary
du

# 概述



Task Category
Perception
Prediction/Planning
Hybrid
Sensor Type
Camera
Can-bus
Lidar
Map
Radar
Language
Data Scale
1000h
100h
10h
Impact
High
Mid
Low
2012
2018
2020
Time
1st Generation
2nd Generation
Dataset
KITTI
Cityscapes
Mapillary
Apollo Scape
nu Scenes
Waymo
Argoverse V1
Woodscape
Lyft Motion
KITTI 360
ONCE
nuPlan
Argoverse V2
DriveLM
Sensor Type
Institute
KIT
Mapillary
Motional
AV
valeo.ai
lyft
HUAWEI
Motional
AV

# 概述



Task Category
Perception
Prediction/Planning
Hybrid
Sensor Type
Camera
Lidar
Radar
Can-bus
Map
Language
Data Scale
1000h
100h
10h
Impact
High
Mid
Low
2012
2018
2020
2024
Time
1st Generation
2nd Generation
Dataset
KITTI
Cityscapes
Mapillary
Apollo Scape
nu Scenes
Waymo
Argoverse V1
Woodscape
Lyft Motion
KITTI 360
ONCE
nuPlan
Argoverse V2
DriveLM
What's Next
Sensor Type
Institute

# 感知类数据集

自动驾驶感知类（Perception）任务包括二维与三维物体检测、追踪、可行驶区域检测、栅格检测（Occupancy Detection）等内容。

感知任务类数据集主要由**数据场景的丰富度**，**传感器种类**以及**标注类型**几个要素构成。

# 感知类数据集

## 感知类数据集梳理（部分）

| Dataset | Year | Data Diversity | | | Sensor | | | Annotation | Impact |
|---|---|---|---|---|---|---|---|---|---|
| | | Scenes | Hours | Region | Camera | LiDAR | Other | | |
| KITTI [33] | 2012 | 50 | 6 | EU | Front-view | ✓ | GPS & IMU | 2D BBox &3D BBox | **High** |
| Cityscapes [10] | 2016 | - | - | EU | Front-view | ✗ | | 2D Seg | Mid |
| Lost and Found [12] | 2016 | 112 | - | - | Front-view | ✗ | | 2D Seg | |
| Mapillary [11] | 2016 | - | - | Global | Street-view | ✗ | | 2D Seg | Mid |
| DDD17 [36] | 2017 | 36 | 12 | EU | Front-view | ✗ | GPS & CAN-bus & Event Camera | - | |
| Apolloscape [14] | 2016 | 103 | 2.5 | AS | Front-view | ✗ | GPS & IMU | 3D BBox & 2D Seg | Mid |
| BDD-X [37] | 2018 | 6984 | 77 | NA | Front-view | ✗ | | Language | |
| HDD [38] | 2018 | - | 104 | NA | Front-view | ✓ | GPS & IMU & CAN-bus | 2D BBox | Mid |
| IDD [39] | 2018 | 182 | - | AS | Front-view | ✗ | | 2D Seg | |
| SemanticKITTI [24] | 2019 | 50 | 6 | EU | ✗ | ✓ | | 3D Seg | |
| Woodscape [20] | 2019 | - | - | Global | 360° | ✓ | GPS & IMU & CAN-bus | 3D BBox & 2D Seg | Mid |
| DrivingStereo [25] | 2019 | 42 | - | AS | Front-view | ✓ | | - | |
| Brno-Urban [40] | 2019 | 67 | 10 | EU | Front-view | ✓ | GPS & IMU & Infrared Camera | - | |
| A*3D [41] | 2019 | - | 55 | AS | Front-view | ✓ | | 3D BBox | Mid |
| Talk2Car [42] | 2019 | 850 | 283.3 | NA | Front-view | ✓ | | Language & 3D BBox | |
| Talk2Nav [21] | 2019 | 10714 | - | Sim | 360° | ✗ | | Language | |

# 感知类数据集

## 复合类数据集梳理

| Dataset | Year | Data Diversity | | | Sensor | | | Annotation | Impact |
|---|---|---|---|---|---|---|---|---|---|
| | | Scenes | Hours | Region | Camera | LiDAR | Other | | |
| nuScenes [8] | 2019 | 1000 | 5.5 | AS & NA | 360° | ✓ | GPS & CAN-bus & Radar & HDMap | 3D BBox & 3D Seg | **High** |
| Argoverse V1 [16] | 2019 | 324k | 320 | NA | 360° | ✓ | HDMap | 3D BBox & 3D Seg | **High** |
| Waymo [9] | 2019 | 1000 | 6.4 | NA | 360° | ✓ | | 2D BBox & 3D BBox | **High** |
| KITTI-360 [26] | 2020 | 366 | 2.5 | EU | 360° | ✓ | | 3D BBox& 3D Seg | Mid |
| ONCE [15] | 2021 | - | 144 | AS | 360° | ✓ | | 3D BBox | Mid |
| nuPlan [18] | 2021 | - | 120 | AS & NA | 360° | ✓ | | 3D BBox | **High** |
| Argoverse V2 [19] | 2022 | 1000 | 4 | NA | 360° | ✓ | HDMap | 3D BBox | Mid |
| DriveLM [62] | 2023 | 1000 | 5.5 | AS & NA | 360° | ✗ | | Language | Mid |

# 感知类数据集

影响力最高的几个数据集在创新性都有较大贡献：

- **KITTI** 数据集首次配备了激光雷达与相机传感器，并定义了深度估计、目标检测、目标跟踪等 10 个任务与相应的评价指标
- **Waymo、nuScenes、Argoverse** 等数据集最先配备了环视相机，并且场景规模更加丰富，支持横跨感知、预测及决策规划等多个模块任务
- 其中，**nuScenes** 数据集还配置了高精地图数据，贴近实际量产环境下系统传感器输入

# 建图类数据集

自动驾驶建图任务是对于静态物体的感知和理解，为下游任务构建出对静态环境的准确认知。

静态场景中的交通元素感知不仅包含对**路面元素**如车道线、人形横道的识别，还包括对交通灯、交通标志牌等与**车辆行驶状态相关元素**的识别

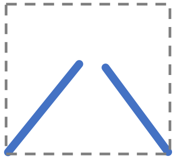
# 建图类数据集

| Dataset | Year | Data Diversity | | Sensor | | Annotation | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Scenes | Frames | Camera | LiDAR | Type | Space | Inst. | Track. |
| Caltech Lanes [73] | 2008 | 4 | 1224/1224 | Front-view Image | ✗ | Laneline | PV | ✓ | ✗ |
| VPG [74] | 2017 | - | 20K/20K | | ✗ | | PV | ✗ | - |
| TUsimple [75] | 2017 | 6.4K | 6.4K/128K | | ✗ | | PV | ✓ | ✗ |
| CULane [76] | 2018 | - | 133K/133K | | ✗ | | PV | ✓ | - |
| ApolloScape [14] | 2018 | 235 | 115K/115K | | ✓ | | PV | ✗ | ✗ |
| LLAMAS [77] | 2019 | 14 | 79K/100K | | ✗ | | PV | ✓ | ✗ |
| 3D Synthetic [78] | 2020 | - | 10K/10K | | ✗ | | PV | ✓ | - |
| CurveLanes [79] | 2020 | - | 150K/150K | | ✗ | | PV | ✓ | - |
| VIL-100 [80] | 2021 | 100 | 10K/10K | | ✗ | | PV | ✓ | ✗ |
| OpenLane-V1 [81] | 2022 | 1K | 200K/200K | | ✗ | | 3D | ✓ | ✓ |
| ONCE-3DLane [82] | 2022 | - | 211K/211K | | ✗ | | 3D | ✓ | - |
| OpenLane-V2 [83] | 2023 | 2K | 72K/72K | Multi-view Image | ✗ | Lane Centerline, Lane Segment | 3D | ✓ | ✓ |

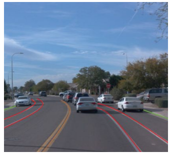
OpenDriveLab

# 建图类数据集

## 道路结构认知领域发展脉络

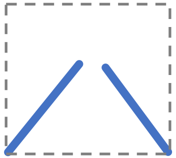


2D Lane Detection

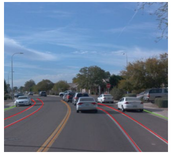


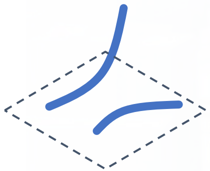
# 建图类数据集

## 道路结构认知领域发展脉络



2D Lane Detection

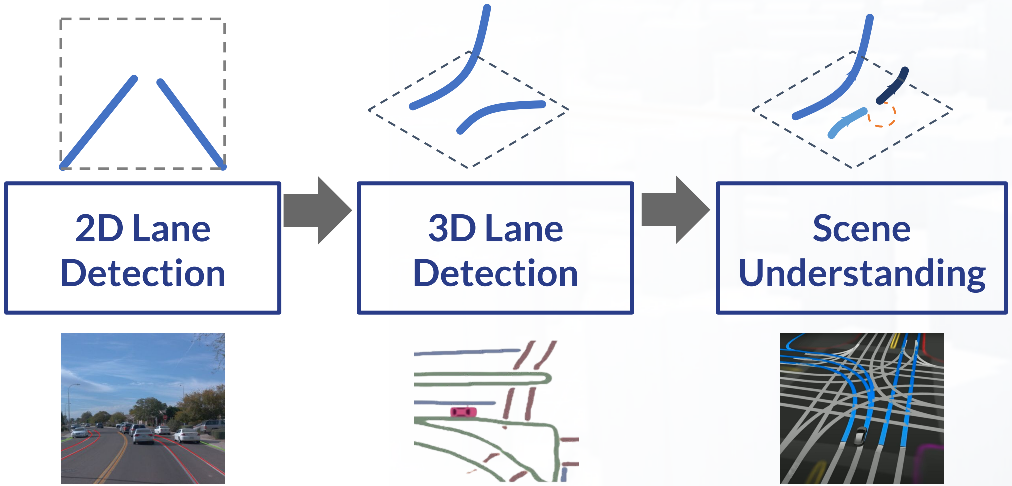




3D Lane Detection

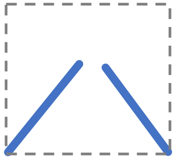
# 建图类数据集

## 道路结构认知领域发展脉络

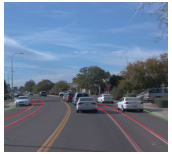


# 建图类数据集

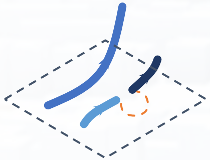
## 道路结构认知领域发展脉络



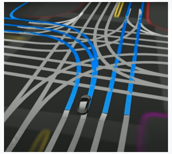
2D Lane Detection



3D Lane Detection



Scene Understanding



HD Map

Driving

“重感知，轻地图” 是未来的发展方向！

# 预测与规划类数据集

传统的模块化规控技术路线会将规控任务具体拆分为不同维度的子任务，具体的可以分为**路网级别**、**道路级别**与**车辆级别**。

# 预测与规划类数据集

传统的模块化规控技术路线会将规控任务具体拆分为不同维度的子任务，具体的可以分为**路网级别**、**道路级别**与**车辆级别**。

- **路网级别:** 根据高清地图提供的静态道路拓扑数据和浮动车辆或卡口提供的实时道路流量数据，可以进行交通流预测和路线规划，为车辆规划从起点到终点的行驶路线，通常以最小化行驶路程或旅行时间为目标。

# 预测与规划类数据集

传统的模块化规控技术路线会将规控任务具体拆分为不同维度的子任务，具体的可以分为**路网级别、道路级别**与**车辆级别**。

- **路网级别:** 根据高清地图提供的静态道路拓扑数据和浮动车辆或卡口提供的实时道路流量数据，可以进行交通流预测和路线规划，为车辆规划从起点到终点的行驶路线，通常以最小化行驶路程或旅行时间为目标。
- **道路级别:** 以完成路线引导的驾驶目的为导向，根据车载相机、激光雷达、毫米波雷达等传感器所提供的自车附近的小范围的道路场景感知信息，预测周围车辆未来数秒时间内的驾驶行为和运动轨迹，进而规划出安全、高效、舒适行驶的自车运动轨迹。

# 预测与规划类数据集

传统的模块化规控技术路线会将规控任务具体拆分为不同维度的子任务，具体的可以分为**路网级别**、**道路级别**与**车辆级别**。

- **路网级别:** 根据高清地图提供的静态道路拓扑数据和浮动车辆或卡口提供的实时道路流量数据，可以进行交通流预测和路线规划，为车辆规划从起点到终点的行驶路线，通常以最小化行驶路程或旅行时间为目标。
- **道路级别:** 以完成路线引导的驾驶目的为导向，根据车载相机、激光雷达、毫米波雷达等传感器所提供的自车附近的小范围的道路场景感知信息，预测周围车辆未来数秒时间内的驾驶行为和运动轨迹，进而规划出安全、高效、舒适行驶的自车运动轨迹。
- **车辆级别:** 结合车辆运动学和动力学模型，获取操纵车辆加减速和转向所对应的油门与刹车踏板行程、方向盘或转向轮转角，以最小操纵代价的前提下驱使车辆完成给定的行驶轨迹。

# 预测与规划类数据集

| Subtask | Input | Output | Evaluation | Dataset | Reference |
|---|---|---|---|---|---|
| Motion Prediction | Surrounding Traffic States | Spatiotemporal Trajectories of Single/Multiple Vehicle(s) | Displacement Error | Argoverse [16]<br>nuScenes [8]<br>Waymo [9]<br>Interaction [144]<br>MONA [148] | [135], [136], [137]<br>[138], [139], [140]<br>[141], [142], [143]<br>[145], [146], [147] |
| Trajectory Planning | Motion States for Ego Vehicles, Scenario Cognition and Prediction | Trajectories for Ego Vehicles | Displacement Error, Safety, Compliance, Comfort | nuPlan [18]<br>CARLA [30]<br>MetaDrive [155]<br>Apollo [159] | [149], [150], [151]<br>[152], [153], [154]<br>[156], [157], [158]<br>[160], [161], [162] |
| Path Planning | Maps for Road Network | Routes Connecting to Nodes and Links | Efficiency, Energy Conservation | OpenStreetMap [163]<br>Transportation Networks [167]<br>DTAlite [171]<br>PeMS [175]<br>New York City Taxi Data [179] | [164], [165], [166]<br>[168], [169], [170]<br>[172], [173], [174]<br>[176], [177], [178]<br>[180], [181], [182] |

OpenDriveLab

# 预测与规划类数据集

规控任务的评价方式可以划分为**开环评价**和**闭环评价**：

- **开环评价**是指以数据集所提供的传感器信息和高阶行车指令信息（前进、左转、右转等）作为输入，来获取出车辆未来轨迹的方法。
- **闭环评价**是指让所构建的模型能够在动态场景中进行校验，模型输出驱动自车进行行驶，周围车辆会根据自车行为做出交互式反应。

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

# 数据算法闭环体系

4 Datasets/Benchmarks/Ecosystem

# 数据算法闭环体系

## 存在的问题

- **长尾问题**，原因在于训练模型的数据量不足而导致存在少量情况未被模型学习，而在模型推理阶段，模型并不能对这些边缘场景给出正确的结果

# 数据算法闭环体系

## 存在的问题

- **长尾问题，**原因在于训练模型的数据量不足而导致存在少量情况未被模型学习，而在模型推理阶段，模型并不能对这些边缘场景给出正确的结果
- 对基于规则的模块，通过**人工设计**各种规则来使模块输出符合人为设计逻辑的结果。该方法耗时耗力，并且难以覆盖所有情况，有可能导致自动驾驶系统在某些未见场景下失效

# 数据算法闭环体系

## 存在的问题

- **长尾问题**，原因在于训练模型的数据量不足而导致存在少量情况未被模型学习，而在模型推理阶段，模型并不能对这些边缘场景给出正确的结果
- 对基于规则的模块，通过**人工设计**各种规则来使模块输出符合人为设计逻辑的结果。该方法耗时耗力，并且难以覆盖所有情况，有可能导致自动驾驶系统在某些未见场景下失效

## 解决方法

→ **海量数据**

# 数据算法闭环体系

**存在的问题**

**解决方法**

- **长尾问题**，原因在于训练模型的数据量不足而导致存在少量情况未被模型学习，而在模型推理阶段，模型并不能对这些边缘场景给出正确的结果
- 对基于规则的模块，通过**人工设计**各种规则来使模块输出符合人为设计逻辑的结果。该方法耗时耗力，并且难以覆盖所有情况，有可能导致自动驾驶系统在某些未见场景下失效

**海量数据**

---

- 训练过程中，**数据噪声**的引入会不可避免地对优化过程产生负面影响，并降低模型性能。

# 数据算法闭环体系

**存在的问题**

**解决方法**

- **长尾问题**，原因在于训练模型的数据量不足而导致存在少量情况未被模型学习，而在模型推理阶段，模型并不能对这些边缘场景给出正确的结果
- 对基于规则的模块，通过**人工设计**各种规则来使模块输出符合人为设计逻辑的结果。该方法耗时耗力，并且难以覆盖所有情况，有可能导致自动驾驶系统在某些未见场景下失效

→ **海量数据**

---

- 训练过程中，**数据噪声**的引入会不可避免地对优化过程产生负面影响，并降低模型性能。

→ **高质量数据**

# 数据算法闭环体系

**现阶段，如何高效地构建海量高质量的数据仍是一个开放性的问题。**学术界与工业界对于构建自动驾驶数据集所采用的方案不尽相同，**数据采集、质量把控、标注技术**等方面都根据各自拥有的平台与技术有所变化。

# 数据算法闭环体系

**现阶段，如何高效地构建海量高质量的数据仍是一个开放性的问题。**学术界与工业界对于构建自动驾驶数据集所采用的方案不尽相同，**数据采集、质量把控、标注技术**等方面都根据各自拥有的平台与技术有所变化。

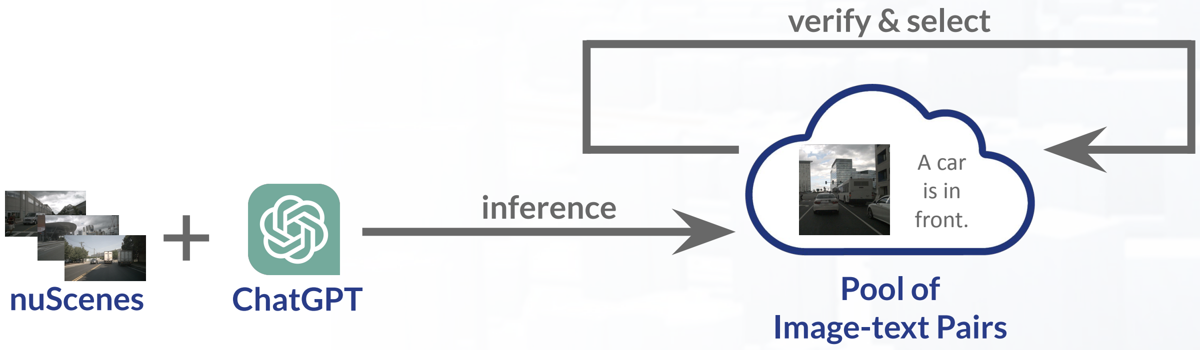
| Solution | Owner | Data Catelog | Data Retrieval | Auto-labeling | Model Training | Simulation | Open-Source |
|---|---|---|---|---|---|---|---|
| Autonomous Driving Data Framework (ADDF) [234] | amazon | Scene Description via Scene Metadata | OpenSearch | Object and Lane Detection via Open-source Models | ✓ | ✓ | ✓ |
| Full Self Driving (FSD) [235] | TESLA | Misprediction Identification, Label Correction, and Selection on Most Valuable Examples | | Static Scene Annotation via Multi-trip Reconstruction | Dojo Supercomputer | Scene Generation in Minutes | ✗ |
| MagLev [236] | NVIDIA | Generating Dataset via Searching, Collection, Labeling, and Export | Elastic Search and Categorization via Active Learning | - | Multi-node Training and Parallel Evaluation | DRIVE Sim | ✗ |
| OpenTrek [237] | Alibaba | - | Multi-modal Retrieval based on Semantics, Images, Labels, Similarity, *etc.* | - | ✓ | ✓ | ✗ |

# 数据算法闭环体系

**应用案例:** 大语言模型通过数据算法闭环体系，生成海量高质量数据。

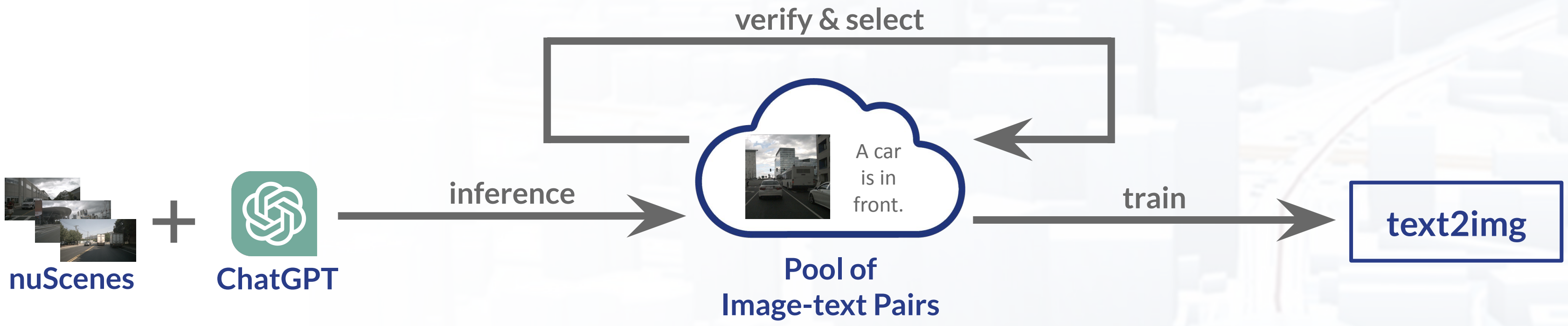
# 数据算法闭环体系

**应用案例:**大语言模型通过数据算法闭环体系，生成海量高质量数据。



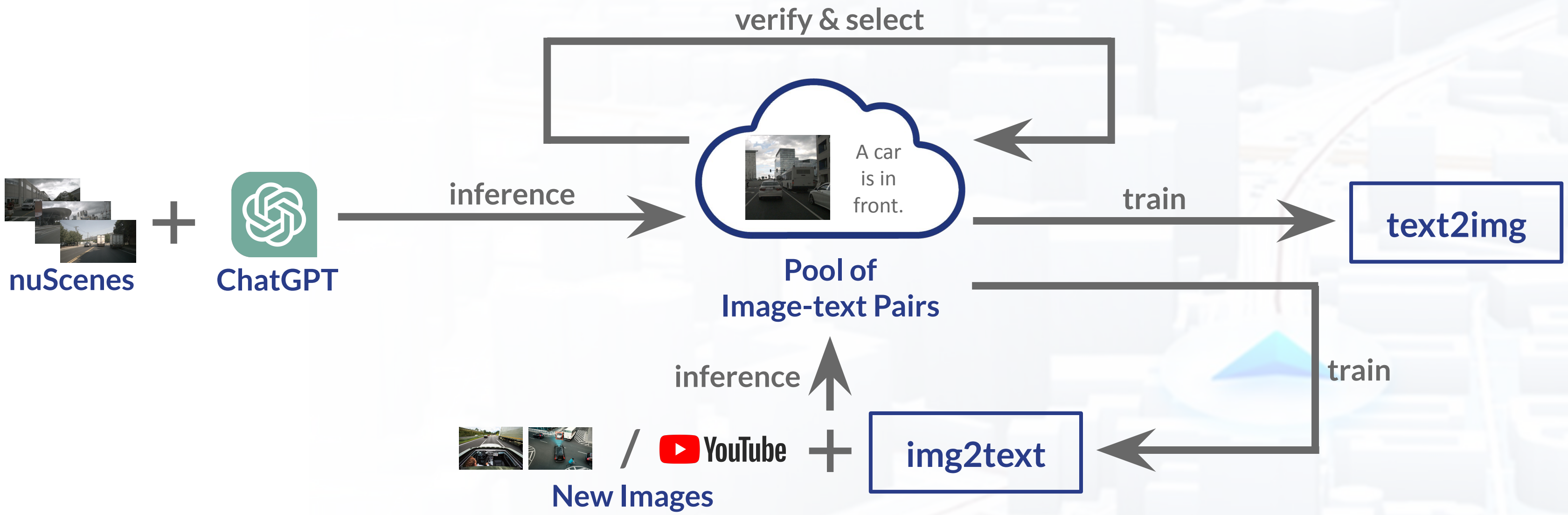
# 数据算法闭环体系

**应用案例:** 大语言模型通过数据算法闭环体系，生成海量高质量数据。



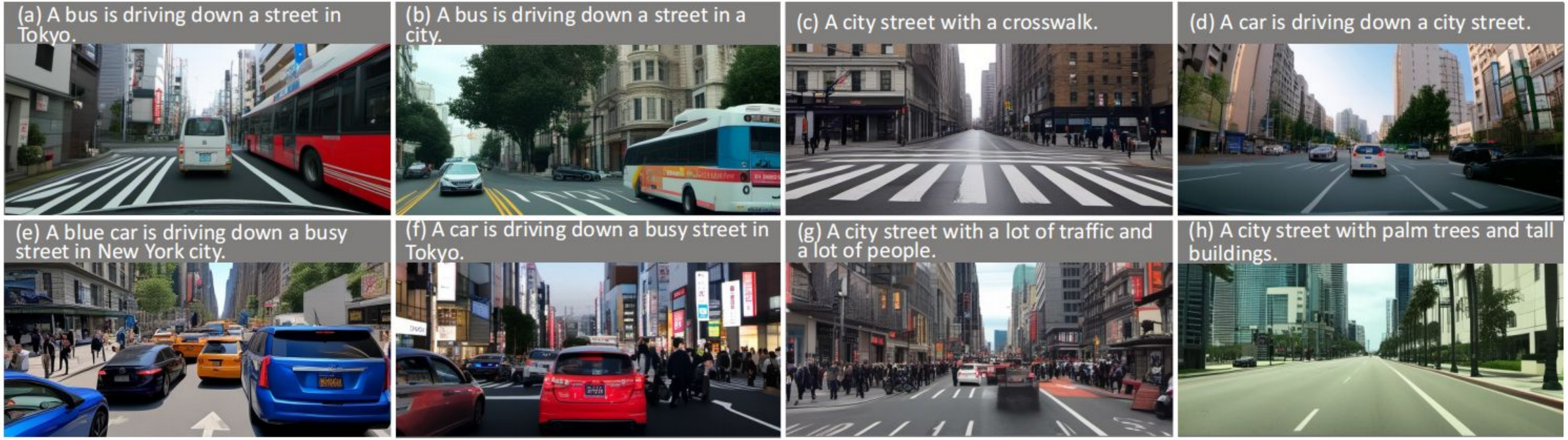
# 数据算法闭环体系

**应用案例:**大语言模型通过数据算法闭环体系，生成海量高质量数据。



# 数据算法闭环体系

**案例模型生成结果:** 通过对语言大模型的运用，研究人员达成通过文本输入生成高质量自动驾驶图片的目标。



OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

# 榜单与生态

4 Datasets/Benchmarks/Ecosystem

# 赛事榜单

随着数据集的发展，近几年自动驾驶国内外挑战赛与榜单也不再局限于感知类任务，正逐步朝**感知决策一体化、端到端、大模型**等方向发展。

模型评测方式也从单一方面的前景物体检测或背景地图描述，变成更加全面的**面向自动驾驶最终目标的评测**。

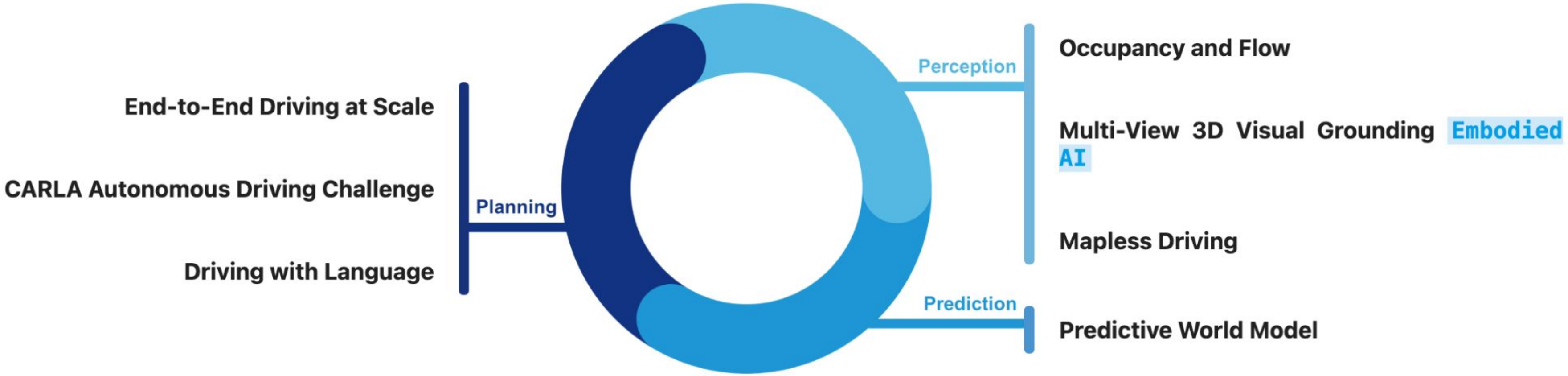
# 赛事榜单

| Title | Host | Year | Task | Evaluation | # Entry | Test Server |
|---|---|---|---|---|---|---|
| Autonomous Driving Challenge [27] | OpenDriveLab | CVPR 2023 | **Perception** / OpenLane Topology | OpenLane-V2 Score (OLS) | 111 | ✓ |
| | | | **Perception** / Online HD Map Construction | mAP | | ✗ |
| | | | **Perception** / 3D Occupancy Prediction | mIoU | | ✓ |
| | | | **Prediction & Planning** / nuPlan Planning | Mean Overall Score | | ✗ |
| Waymo Open Dataset Challenges [9, 28] | Waymo | CVPR 2023 | **Perception** / 2D Video Panoptic Segmentation | weighted Segmentation and Tracking Quality | 35 | ✓ |
| | | | **Perception** / Pose Estimation | Pose Estimation Metric (PEM) | | ✓ |
| | | | **Prediction** / Motion Prediction | Soft mAP | | ✓ |
| | | | **Prediction** / Sim Agents | Realism Meta Metric | | ✓ |
| | | CVPR 2022 | **Prediction** / Motion Prediction | Soft mAP | 128 | ✓ |
| | | | **Prediction** / Occupancy and Flow Prediction | AUC on Joint Occupancy and Flow Metric | | ✓ |
| | | | **Perception** / 3D Semantic Segmentation | mIOU | | ✓ |
| | | | **Perception** / 3D Camera-only Detection | Longitudinal Affinity Weighted LET-3D-AP | | ✓ |
| | | CVPR 2021 | **Prediction** / Motion Prediction | Soft mAP | 115 | ✓ |
| | | | **Predirction** / Interaction Prediction | mAP | | ✓ |
| | | | **Perception** / Real-time 3D Detection | Mean Average Precision with Heading (APH) | | ✓ |
| | | | **Perception** / Real-time 2D Detection | mAP | | ✓ |

# Autonomous Grand Challenge

https://opendrivelab.com/challenge2024/

CVPR 2024 Workshop



# Autonomous Grand Challenge

## 占据栅格与运动估计 / Occupancy and Flow

- **赛题描述：**三维框往往不足以描述一般物体，受机器人学概念的启发，可将感知表征描述成**对栅格化三维空间的占据情况预测**。在这个赛道中，参赛者不仅要给出三维空间的栅格化表示，还须给出栅格的运动预测。
- **基线模型：**OccNet https://github.com/OpenDriveLab/OccNet/tree/occnet
- **评价指标：**占据分数（occupancy score）
- **数据集：**nuScenes，光轮占据栅格仿真数据集。

# Autonomous Grand Challenge

## 无图驾驶 / Mapless Driving

- **赛题描述**：在没有高清地图的情况下，自动驾驶汽车需要高水平的场景理解能力，本赛道旨在探索场景推理能力的极限。将多视角图像和标清地图作为输入信息，神经网络不仅要输出车道和交通元素的感知结果，同时还须输出车道之间、车道和交通元素之间的拓扑关系。
- **基线模型**：LaneSegNet / TopoNet / TopoMLP / MapTR
- **评价指标**：OpenLane-V2 UniScore (OLUS)
- **数据集**：OpenLane-V2 数据集

# Autonomous Grand Challenge

## 预测世界模型 / Predictive World Model

- **赛题描述：**作为现实世界的抽象时空表征，世界模型可以根据当前状态预测未来状态。世界模型的学习过程有可能为自动驾驶提供一个预先训练好的基础模型。在只有视觉输入的情况下，神经网络会输出未来的点云，以证明其对世界的预测能力。
- **基线模型：**ViDAR https://github.com/OpenDriveLab/ViDAR
- **评价指标：**Chamfer Distance
- **数据集：**OpenScene

# Autonomous Grand Challenge

## 大规模端到端驾驶 / End-to-end Driving at scale

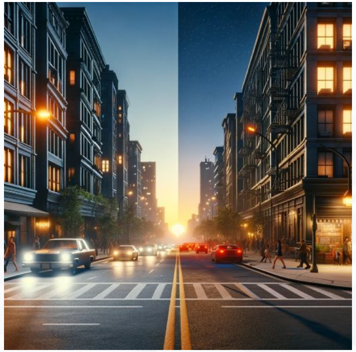
- **赛题描述：**由于之前的数据集规模有限，且**开环和闭环指标之间存在偏差**，因此利用真实数据对传感器运动驾驶政策进行基准测试具有挑战性。在本赛道中，我们使用大规模 OpenScene 数据集，旨在缩小两种评估范式之间的差距。
- **基线模型：**NAVSIM
  https://github.com/autonomousvision/navsim/blob/main/docs/agents.md
- **数据集：**OpenScene

# Autonomous Grand Challenge

## 驾驶结合自然语言 / Driving with language

- **赛题描述：**这项任务结合语言模式，将视觉语言模型（VLM）与自动驾驶系统联系起来。该模型将引入视觉语言模型的推理能力来做出决策，并追求可概括、可解释的驾驶行为。给定多视角图像作为输入，模型需要回答涉及驾驶各个方面的问题。
- **基线模型：**llama-adapter v2（Finetuned with DriveLM）
- **评价指标：**Chamfer Distance
- **数据集：**DriveLM

# Autonomous Grand Challenge



## CARLA 自动驾驶挑战 / CARLA Autonomous Driving Challenge

- **赛题描述：**为了验证 AD 系统的有效性，我们需要一个具有闭环设置的终极规划框架。CARLA AD 排行榜要求代理驾车通过一组预定义的路线。对于每条路线，代理都将在起点处初始化，并通过 GPS 式坐标、地图坐标或路线指示提供的路线描述，定向驶向目的地点。路线可在各种情况下定义，包括高速公路、城区、住宅区和乡村环境。排行榜可评估 AD 代理在各种天气条件下的表现，包括白天、日落、雨天、雾天和夜晚等。
- **评测体系：CARLA Leaderboard 2.0**
  https://leaderboard.carla.org/leaderboard/

# End-of-Lecture

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

# 课程总结

# 总结

- Mobileye 总结
- Dynamics 总结
- 数据集与评测总结
- 挑战赛总结

OpenDriveLab

上海人工智能实验室
Shanghai Artificial Intelligence Laboratory

# 课后部分

# 思考题