

自動運転の作り方

一般通過自動運転おじさん



@hakuturu583

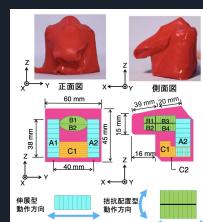


hakuturu583

自己紹介

- 2012
 - 京都大学工学部電気電子情報に落ちて阪大応用理工に入学する
- 2014
 - 大学時代にVRがやりたくてUnityでゲームプログラミングを始める
- 2016-2018
 - 研究室でロボットを始める
 - 柔らかいロボット機構の制御を研究
 - ロボカップでロボットサッカーにハマる
 - 海洋ロボットの開発に着手
 - 今も趣味で自律航行システムを開発中
- 2018-
 - 某自動運転企業に就職

「知能」をどうプログラムで表現できるか
が基本的な興味です

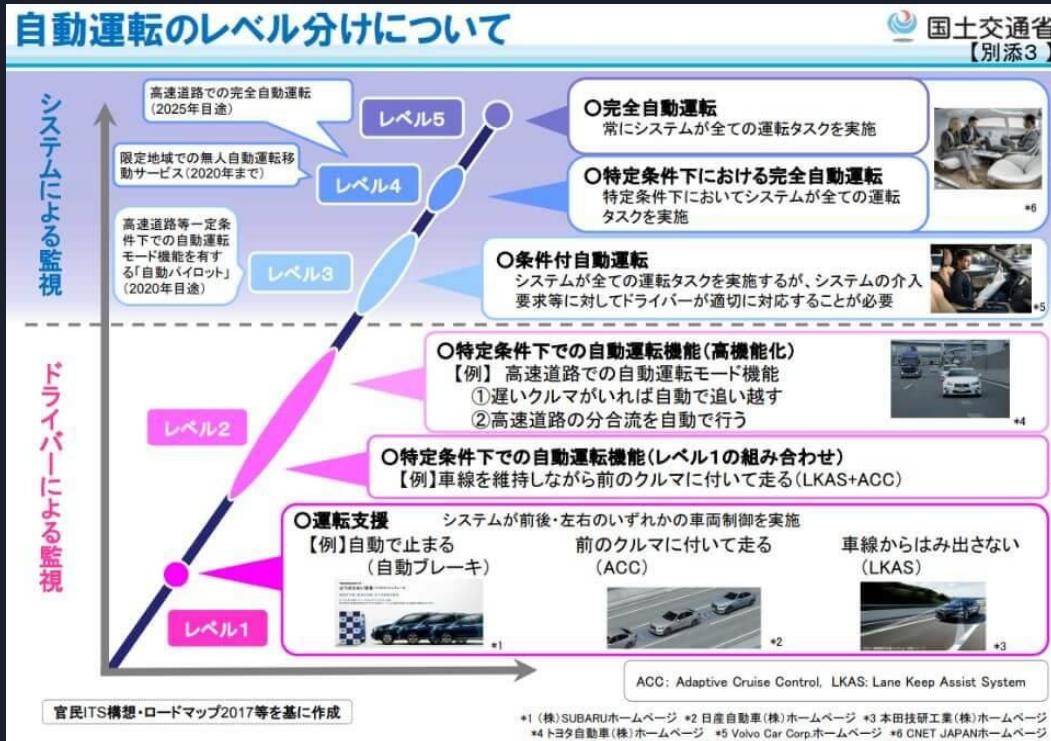


本日のゴール

- 「#自動運転完全に理解した」と呟いて帰ってもらう
- 自動運転の各種流派とアーキテクチャを覚える
- 自動運転の研究と開発にギャップがあることを実感してもらう



そもそも自動運転って？



いわゆる
「自動運転」は
Lv4以上

Lv2以下の車
はドライバーは
あくまで
「人間」

自動運転技術研究開発の歴史

1980年代



NavLabによる
アメリカ横断

2000年代



DARPA Urban
Challenge

2016



世界最強
waymo爆誕

2020



Tesla FSD
リリース

2025



Wayveや中国勢に
よりE2Eがコモデ
ティ化し始める

高精度に自己位置 + ルールベースプランナー

E2Eの勃興

突然ですがクイズです

E2Eとルールベース、
優れてるのはどっち？

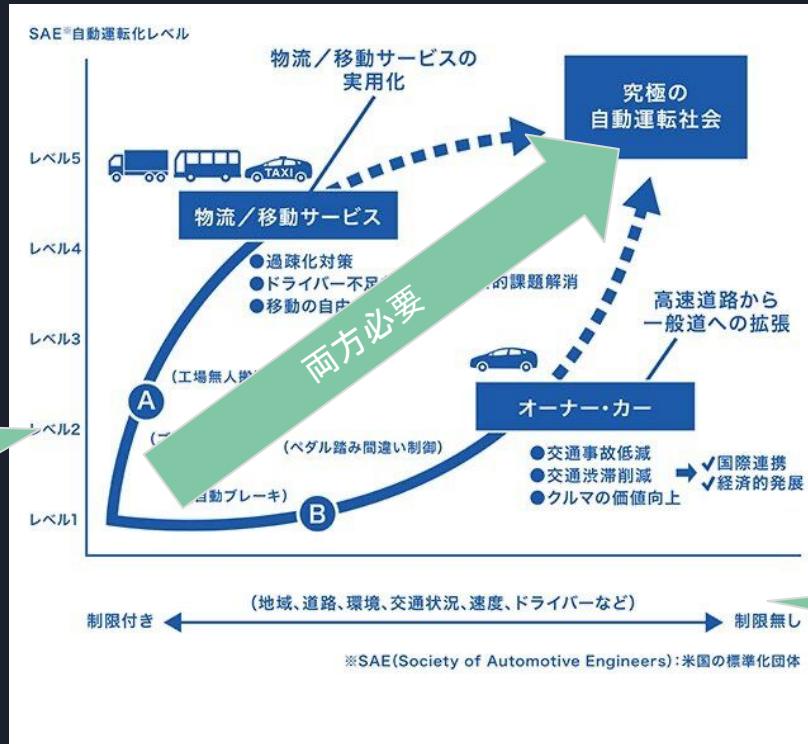
答え



自動運転の性能

E2E手法は
こちらは苦手

安全性
(どれだけ
事故を起こさ
ないか)



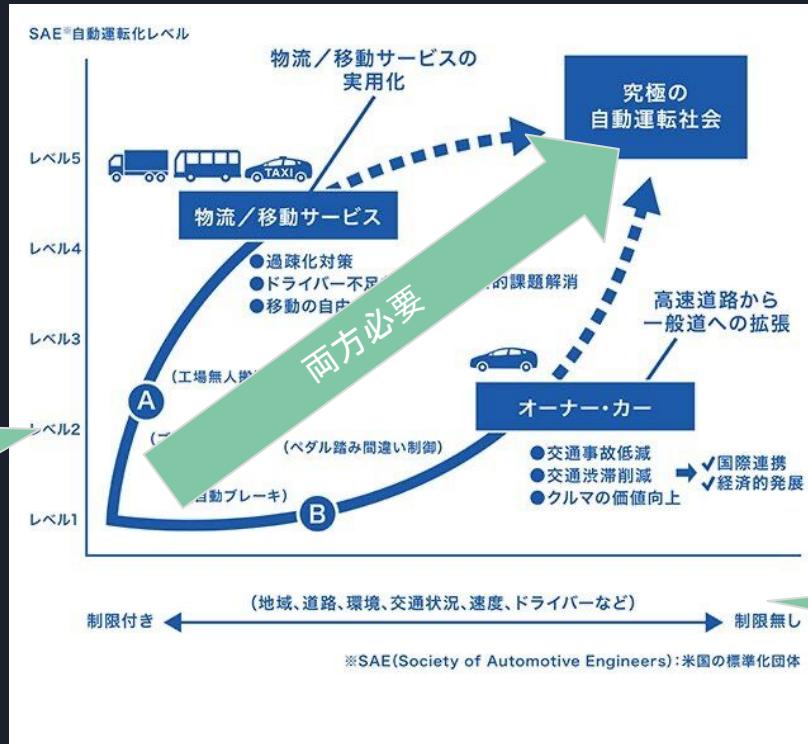
E2E手法は
こちらは得意

可用性
(どんな条件
で使えるか)

自動運転の研究と開発のギャップ

アカデミックなアウトプットになりにくい。
(Oではない)

安全性
(どれだけ事故を起こさないか)



しかし産業界から見ると可用性だけ高くても...

DeepでPonしたら論文書けるので大人気

可用性
(どんな条件で使えるか)

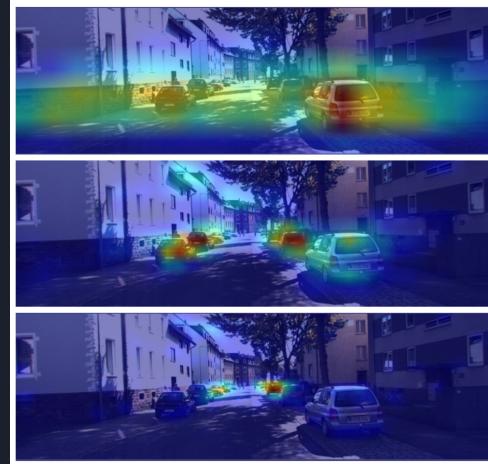
ルールベース/E2Eの比較

	E2E	ルールベース	どっちの性能に寄与
解釈可能性	✗	○	安全性
説明可能性	○	○	寄与しない
柔軟なシーン解釈	○	✗	可用性
デバッグ難易度	✗	○	安全性
デグレ可能性	✗	▲	安全性

解釈可能性と説明可能性



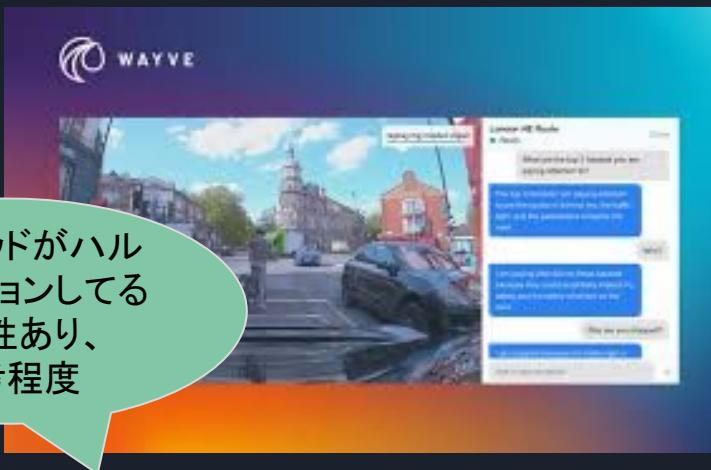
説明可能性: 言語ヘッドを生やして人間から見てわかりやすく状況説明させるなど



解釈可能性: Grad-CAMなどの手法でなぜその推論が出たのか数理的にわかること

解釈可能性と説明可能性

言語ヘッドがハル
シネーションして
る可能性あり、
参考程度



説明可能性: 言語ヘッドを生やして人間から見
てわかりやすく状況説明させるなど



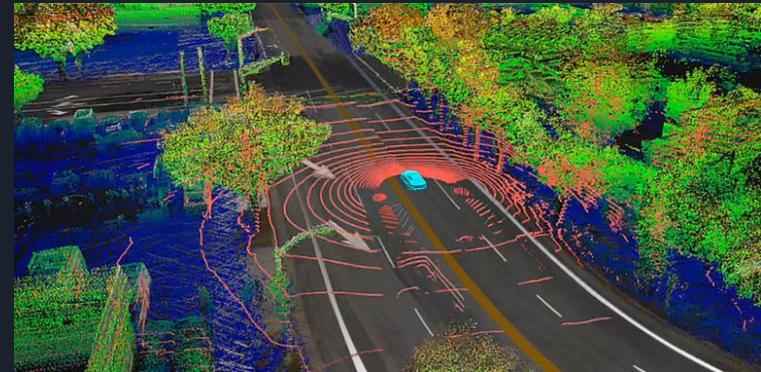
解釈可能性: Grad-CAMなどの手法でなぜその
推論が出たのか数理的にわかること

数理的な根拠はあ
るが、パッと見て何
が原因だったのか
わかるかと言われ
ると。。。。

ルールベース自動運転

ルールベース自動運転のアーキテクチャ

- DARPA Urban Challengeに源流をもつアーキテクチャ
- NDT Matchingを使えば誤差数cmで自己位置推定が可能なことを利用して信号機の位置や白線位置をあらかじめ記録した地図を用いることで白線認識という不安定な要素に頼らず安定した自動運転が可能
- ルールベースのモジュールを上から下に繋げて「正しいデータをやりとりし続ける」方法であるため解釈可能性が高い。しかし融通が効かない



SENSING → LOCALIZATION → PERCEPTION → PLANNING → CONTROL

RULE-BASED AUTONOMOUS DRIVING SYSTEM

Sensing in ルールベース自動運転

- 様々なセンサーのデータを取得し、前処理を行う部分
- センサ山盛りで同期は結構大変
- 正直ここはE2Eにしてもほぼ変わらない
- 労力かかるけど論文にかけないがち



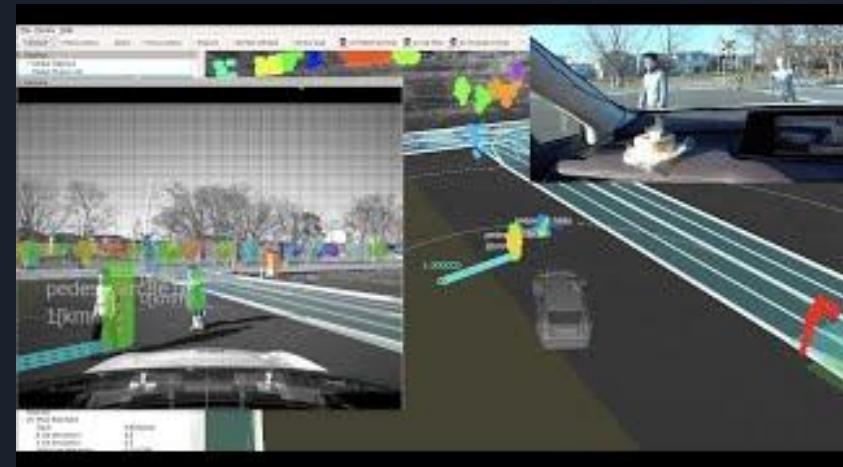
Localization in ルールベース自動運転

- NDT Matchingという手法を使って LiDAR点群と事前にSLAMで作成した地図を正規分布の集合として表現、マッチングする
- NDT Matchingは最適化ベースの手法なので初期値が必要。初期値にはGPSのデータや以前の自己位置推定結果を利用
- LiDARは基本10Hzであり制御に使える頻度で自己位置が取れないのでEKFとの組合せを推奨



Perception in ルールベース自動運転

- 深層学習による画像/点群物体認識によって周囲の環境を認識
- 白線の位置などは自己位置推定結果と高精度地図を付き合わせるため認識しないことがほとんど
- センサ統合を深層学習でやるEarly Fusionと主にルールベースで行うLate Fusionが存在する
- Early Fusionは性能が良いがセンサ配置を変えると学習やり直し。Late Fusionは性能は低いがセンサ配置変えても再学習不要なことが多い



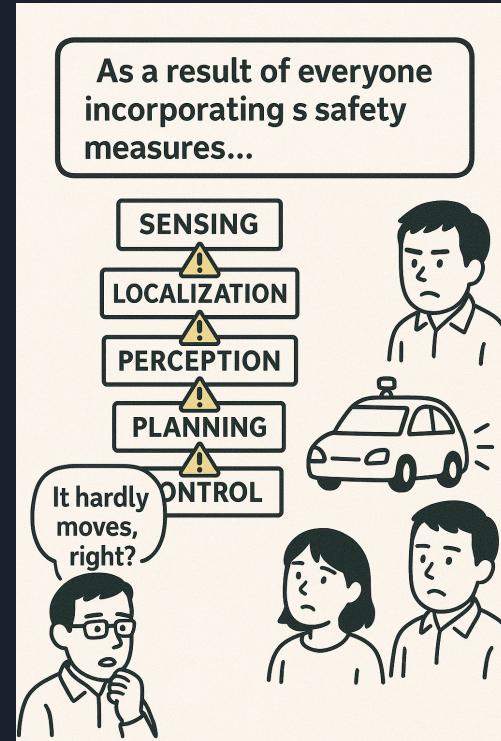
Planning in ルールベース自動運転

- 物体認識結果と高精度地図の情報をもとに車の経路と速度計画を決める
- 経路/速度計画のやり方には色々ある
- この部分だけ機械学習を導入するケースも
- グローバルパスプランニング問題はカーナビがあれば解決するのでMobile Robotと違って明示的にそういったモジュールはない
- 作った経路を追従するのは性能出したければMPC、ざっくりでよければPure Pursuit



ルールベース自動運転の利点と欠点

- 利点
 - 解釈可能である
 - デバッグができる
 - デグレしにくい
- 欠点
 - モジュールごとに安全対策が入る
 - どこか一箇所でも安全対策に引っ掛かってたらエラーになるので可用性が上がらない
 - 各モジュールで安全がわに倒れた実装がなされた結果めちゃくちゃ保守的な動きをして逆に危ないことも
- 課題
 - 道路交通法って完璧に守ると危ない！



E2E自動運転



Q. E2E自動運転とは？

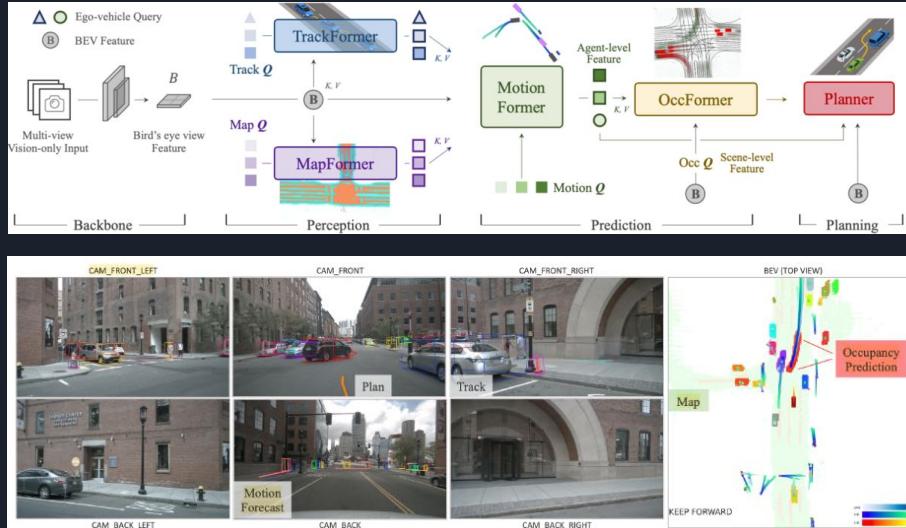
A. センサ情報を受けて運動コマンドを出力する全ての関数が微分可能である自動運転

→ 微分可能だから深層学習により全体最適になることが期待できる

E2E自動運転流派 その1 BEV

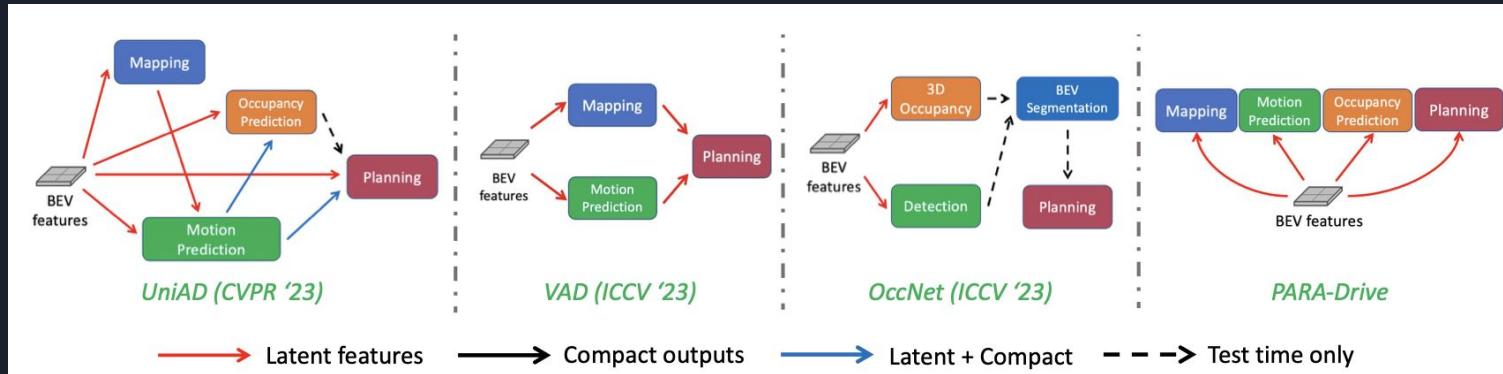


- BEV = Bird eye view
- マルチセンサーの情報を一旦鳥瞰図か3D Occupancy Gridに変換し深層学習で作ったプランナーで運動コマンド
- 世の中で動いてるE2E自動運転は大体このタイプ？言語モデルはまだまだ枯れてなさすぎる印象
- 多分一番有名なのはCVPR 2023 Best PaperのUniAD
- Sensor Fusionのノウハウが最も洗練、統合されている流派



Planning Oriented Autonomous Driving

BEV型の歴史



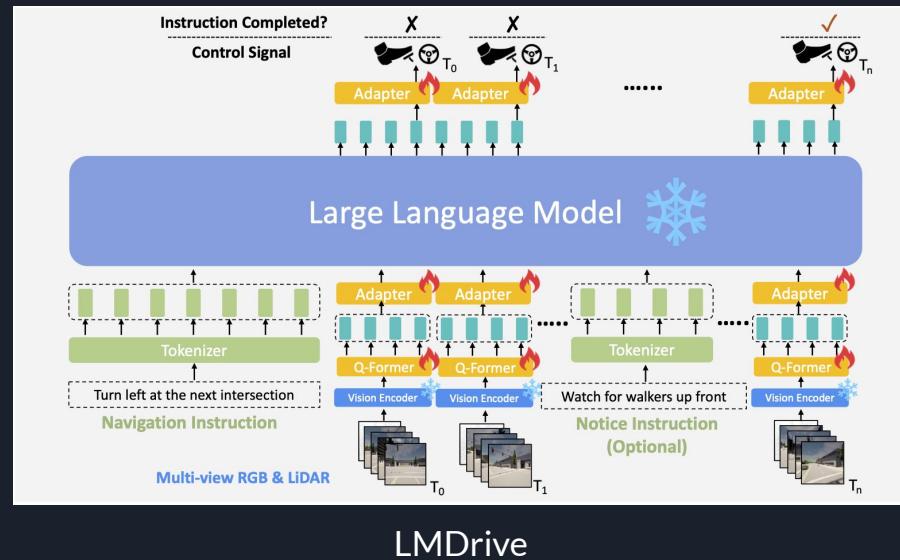
PARA-Drive

- 基本的にはひたすらモデルをダイエットしてきた歴史
- UniADでとりあえず動くようになり
- VADで地図表現を工夫して大幅軽量化
- OccNetで「あれ？マッピング以外もサブタスクとして扱えるのでは？」隣
- PARA-DriveでもうPlanning以外全部サブタスクでよくないかという過激な結論に

E2E自動運転流派 その2 VLA



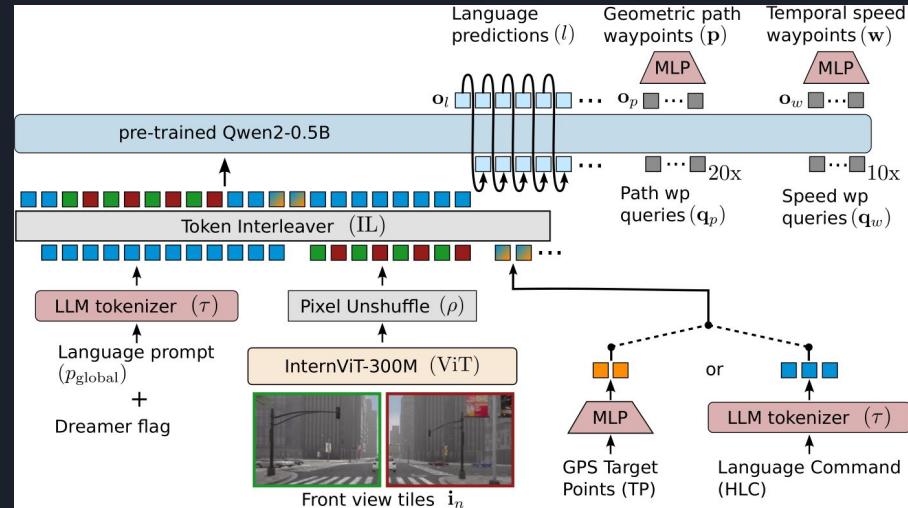
- 言語モデルに画像ぶちこん運動コマンド取ればいいだろという思想
- 研究のトレンドは現在BEVからVLAにシフト
- 言語モデルの知識を取り込むことで、エッジケース対応や複雑なインタラクションを効率よく捌けることを期待
- センサ配置を変更すると再学習のコスト高いがち
- 多分最初に出た研究としてはLMDrive?



VLA型の一例 SimLingo



- 言語モデル内部に「仮想フラグ」を設けて危ういシナリオを仮想的に実行
- 結果として、自動運転スタックは「これはやっちゃんいけなんいんだ」を言語モデル内部に情報として蓄えられる
- 言語モデルはNeuralなガードレール機構を内包し動作する

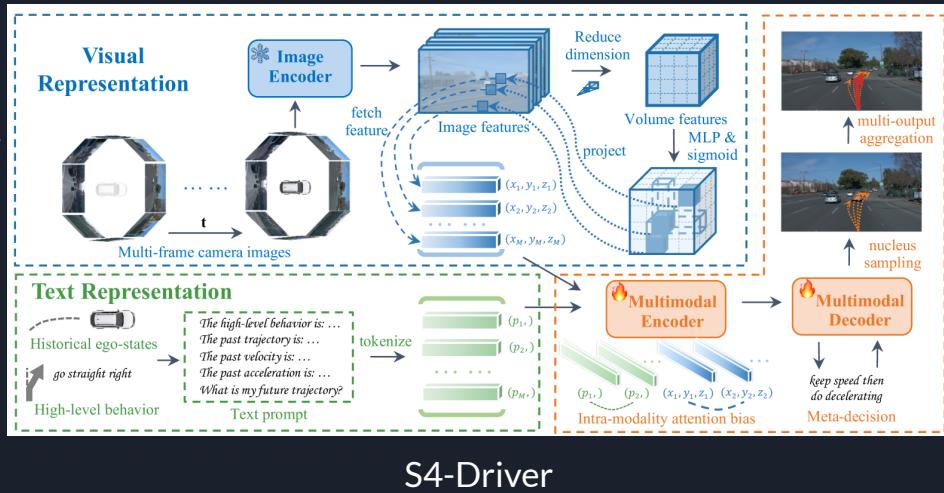


SimLingo

VLA型の一例 SimLingo



- Self-Supervisedな言語モデルベース自動運転システム
- 画像のパッチから計算したEmbeddingをVoxelに埋め込み
- 埋め込み結果を言語モデルベースのEncoder/Decoderを通してTrajectoryを得る
- Ablation Studyによると、経路の表現にはSpecial Tokenを使わなかつた方が性能がいいらしい....

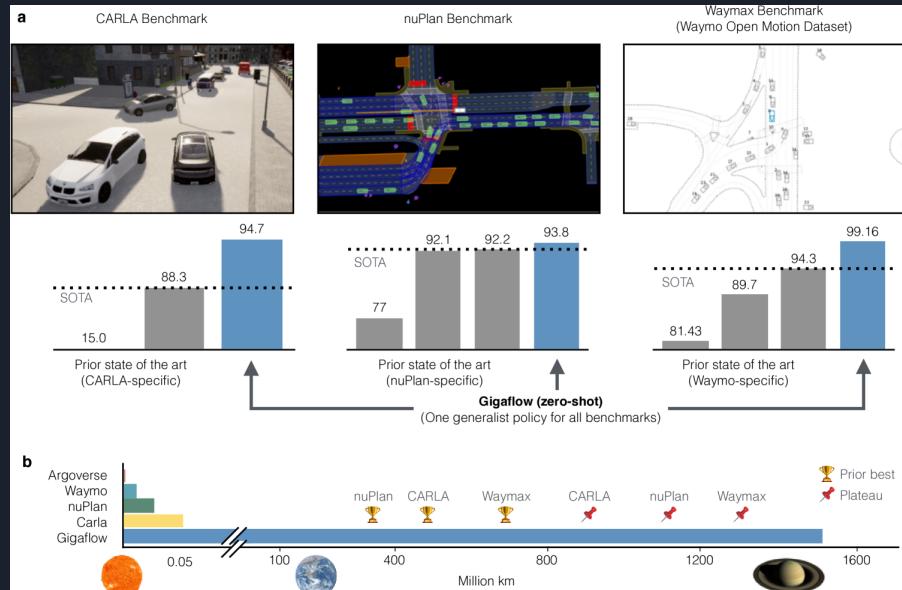


S4-Driver

E2E自動運転流派 その3 強化学習



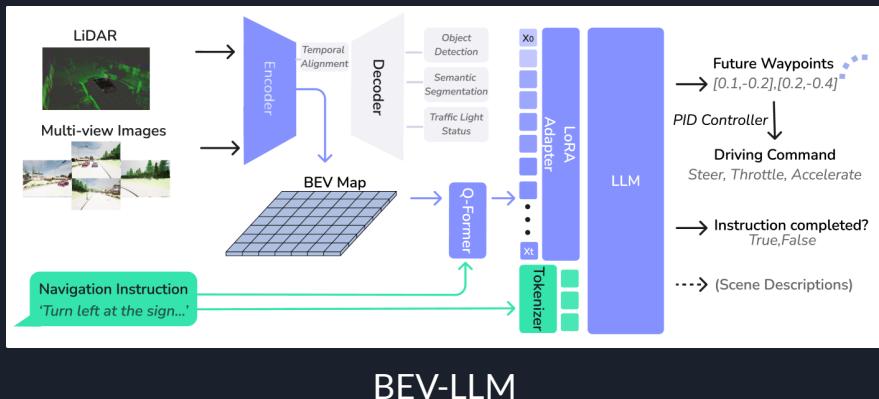
- 古の時代に滅んだと思っていた古代魔法、CVPRに行ったら意外と元気にしてて過激派がいっぱい
- おおん？土星まで走り込めば勝手に自動運転学べるやろというパワー系脳筋ゴリラ
- 高速で学習ぶん回すためにシミュレータは簡素化されており、実機応用に課題
- それはそうと、なんか最近君LLMの刺身のツマみみたいな扱い受けてない？



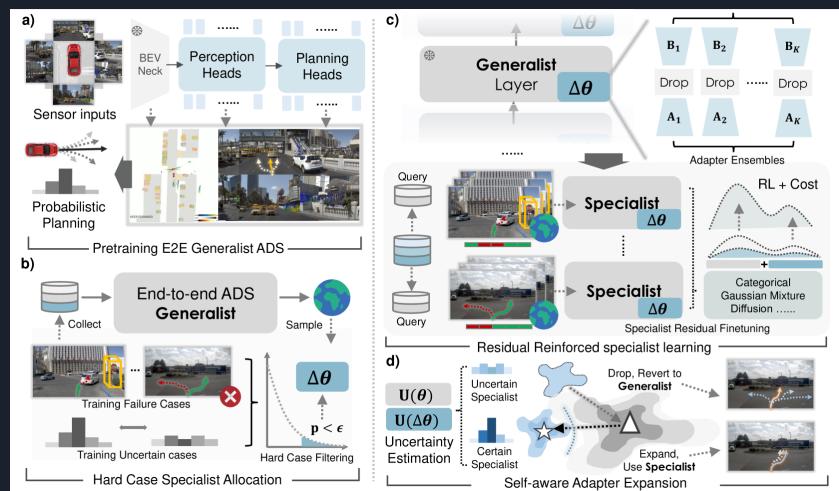
予想される今後の研究のトレンド



- VLA/BEVのハイブリッド
 - BEVでセンサ統合
 - LLMの知識を反映
 - このほかにも、事前学習タスクを工夫する研究も



- 強化学習の活用
 - 苦手ケースをLoRAで対応するのを強化学習で実現



まとめ

E2E自動運転はルールベースの単純な上位互換ではない

- 自動運転の性能には安全性と可用性の2軸あり、両方高める必要がある
- E2Eは可用性に、ルールベースは安全性に効いてくる

産業界と、研究界にかなり乖離がある

- アカデミックな研究は可用性だけで評価されがち
- 産業界で期待されているのでは？と思う研究分野は下記あたり
 - 深層学習の数理モデル
 - 苦手ケースをどう探して潰すか