



ResearchInChina
www.researchinchina.com

Automotive AI Foundation Model Technology and Application Trends Report, 2023-2024

Feb. 2024

End-to-End autonomous driving foundation models boom.

Since 2023 ever more vehicle models have begun to be connected with foundation models, and an increasing number of Tier1s have launched automotive foundation model solutions. Especially Tesla's big progress of FSD V12 and the launch of SORA have accelerated implementation of AI foundation models in cockpits and intelligent driving.

End-to-End autonomous driving foundation models boom.

In February 2023, Tesla FSD v12.2.1, which adopts an end-to-end autonomous driving model, began to be pushed in the United States, not just to employees and testers. According to the feedback from the first customers, FSD V12 is quite powerful, allowing ordinary people who previously did not believe in and use autonomous driving to dare to use FSD. For example, Tesla FSD V12 can bypass puddles on roads. A Tesla engineer commented: this kind of driving approach is difficult to implement with explicit code, but Tesla's end-to-end approach makes it almost effortlessly.

The development of AI foundation models for autonomous driving can be divided into four phases.

Phase 1.0 uses a foundation model (Transformer) at the perception level.

Phase 2.0 is modularization, with foundation models used in perception, planning & control and decision.

Phase 3.0 is end-to-end foundation models (one "end" is raw data from sensors, and the other "end" directly outputs driving actions).

Phase 4.0 is about heading from vertical AI to artificial general intelligence (AGI's world model).

Most companies are now in Phase 2.0

Most companies are now in Phase 2.0, while Tesla FSD V12 is already in Phase 3.0. Other OEMs and Tier1s have followed up with the end-to-end foundation model FSD V12. On January 30, 2024, Xpeng Motor announced that its end-to-end model will be fully available to vehicles in the next step. It is known that NIO and Li Auto will also launch "end-to-end based" autonomous driving models in 2024.

FSD V12's driving decisions are generated by an AI algorithm. It uses end-to-end neural networks trained with massive video data to replace more than 300,000 lines of C++ code. FSD V12 provides a new path that needs to be verified. If it is feasible, it will have a disruptive impact on the industry.

On February 16, OpenAI introduced text-to-video model SORA, signaling the wide adoption of AI video applications. SORA not only supports generation of up to 60-second videos from texts or images, but it well outperforms previous technologies in capabilities of video generation, complex scenario and character generation, and physical world simulation.

Through vision both SORA and FSD V12 enable AI to understand and even simulate the real physical world. Elon Musk believes that FSD 12 and Sora are just two of the fruits of AI's ability to recognize and understand the world through vision, and FSD is ultimately used for driving behaviors, and Sora is used to generate videos.

The high popularity of SORA is further evidence of the rationality of FSD V12. Musk said "Tesla generative video from last year".



Source: Twitter

AI foundation models evolve rapidly, bringing new opportunities.

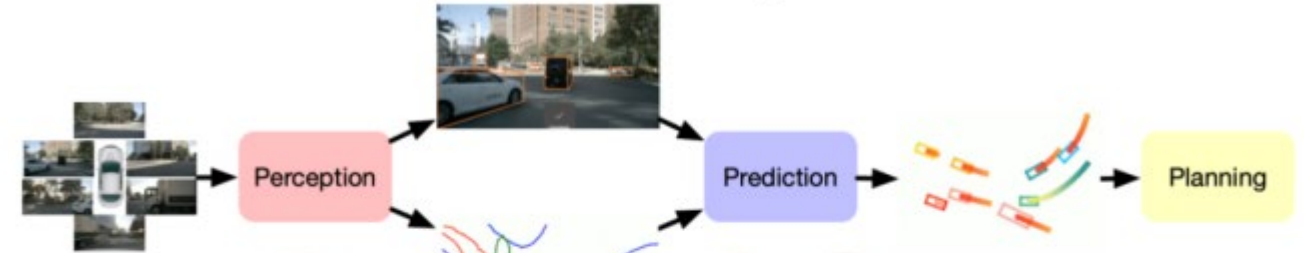
AI foundation models evolve rapidly, bringing new opportunities.

In recent three years foundation models for autonomous driving have undergone several evolutions, and the autonomous driving systems of leading automakers must be rewritten almost every year, which also provides entry opportunities for late entrants.

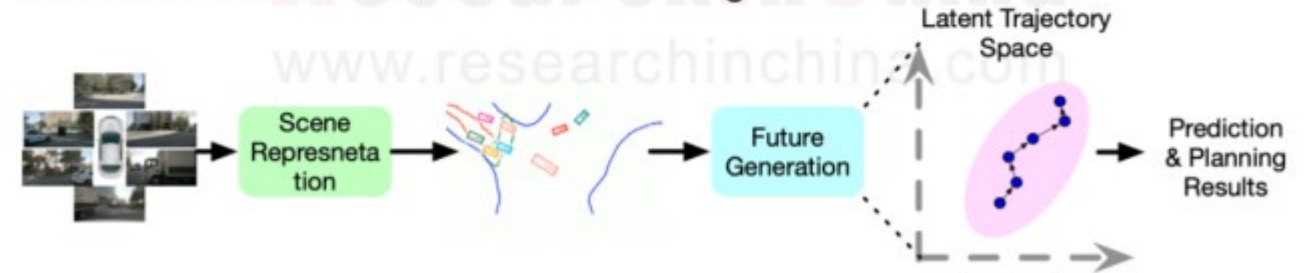
At CVPR 2023, UniAD, an end-to-end autonomous driving algorithm jointly released by SenseTime, OpenDriveLab and Horizon Robotics, won the 2023 Best Paper.

In early 2024, Waytous' technical team and the Institute of Automation Chinese Academy of Sciences jointly proposed GenAD, the industry's first generative end-to-end autonomous driving model which combines generative AI and end-to-end autonomous driving technology. This technology is a disruption to UniAD progressive process end-to-end solution, and explores a new end-to-end autonomous driving mode. The key is to using generative AI to predict temporal evolution of the vehicle and surroundings in past scenarios.

Conventional End-to-End Autonomous Driving



Generative End-to-End Autonomous Driving



Source: Waytous

In February 2024, Horizon Robotics and Huazhong University of Science and Technology proposed VADv2, an end-to-end driving model based on probabilistic planning. VADv2 takes multi-view image sequences as input in a streaming manner, transforms sensor data into environmental token embeddings, outputs the probabilistic distribution of action, and samples one action to control the vehicle. Using only camera sensors, VADv2 achieves state-of-the-art closed-loop performance in CARLA Town05 benchmark test, much better than all existing approaches. It runs stably in a fully end-to-end manner, even without rule-based wrapper.

Method	Modality	Reference	Driving Score \uparrow	Route Completion \uparrow	Infraction Score \uparrow
CILRS [9]	C	CVPR 19	7.8	10.3	0.75
LBC [6]	C	CoRL 20	12.3	31.9	0.66
Roach [54]	C	ICCV 21	41.6	96.4	0.43
Transfuser [†] [40]	C+L	TPAMI 22	31.0	47.5	0.77
ST-P3 [18]	C	ECCV 22	11.5	83.2	-
VAD [23]	C	ICCV 23	30.3	75.2	-
ThinkTwice [21]	C+L	CVPR 23	70.9	95.5	0.75
MILE [16]	C	NeurIPS 22	61.1	97.4	0.63
Interfuser [45]	C	CoRL 22	68.3	95.0	-
DriveAdapter+TCP [20]	C+L	ICCV 23	71.9	97.3	0.74
DriveMLM [49]	C+L	arXiv	76.1	98.1	0.78
VADv2	C	Ours	85.1	98.4	0.87

Table 1. Closed-loop evaluation on Town05 Long benchmark.

Source: Horizon Robotics

VADv2 achieved a Drive Score of 85.1

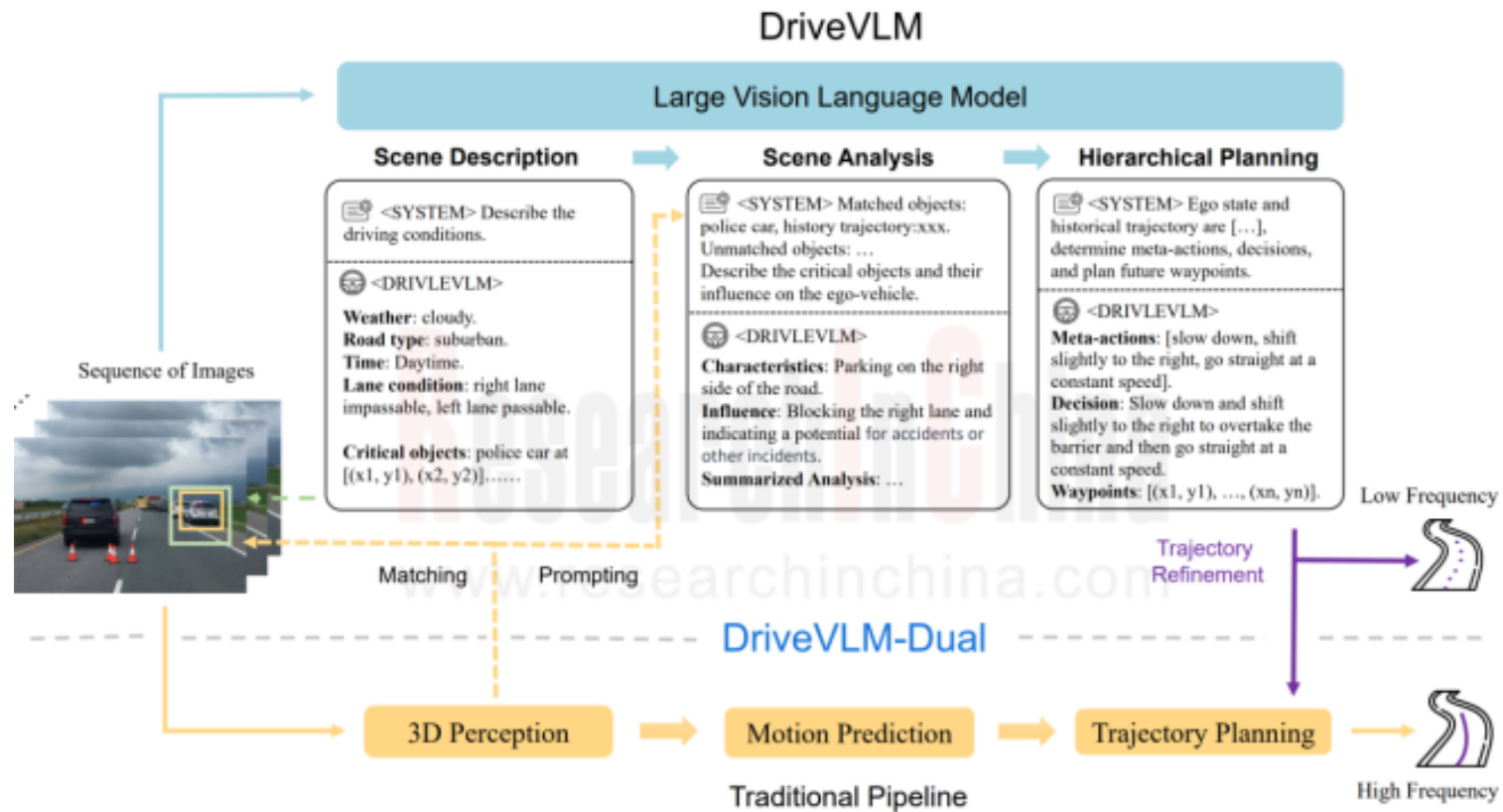
On the Town05 Long benchmark, VADv2 achieved a Drive Score of 85.1, a Route Completion of 98.4, and an Infraction Score of 0.87, as shown in Tab. 1. Compared to the previous state-of-the-art method, VADv2 achieves a higher Route Completion while significantly improving Drive Score by 9.0. It is worth noting that VADv2 only utilizes cameras as perception input, while DriveMLM utilizes both cameras and LiDAR. Furthermore, compared to the previous best method which only relies on cameras, VADv2 demonstrates even greater advantages, with a remarkable increase in Drive Score of up to 16.8.

Method	Modality	Driving Score ↑	Route Completion ↑
CILRS [9]	C	7.5	13.4
LBC [6]	C	31.0	55.0
Transfuser [40]	C+L	54.5	78.4
ST-P3 [18]	C	55.1	86.7
VAD [23]	C	64.3	87.3
VADv2	C	89.7	93.0

Table 2. Closed-loop evaluation on Town05 Short benchmark.

Source: Horizon Robotics

Also in February 2024, the Institute for Interdisciplinary Information Sciences at Tsinghua University and Li Auto introduced DriveVLM (its whole process shown in the figure below). A range of images are processed by a large visual language model (VLM) to perform specific chain of thought (CoT) reasoning to produce driving planning results. This large VLM includes a visual encoder and a large language model (LLM).



Source: Interdisciplinary Information Sciences at Tsinghua University

AI and cloud companies attract attention as foundation models emerge

Due to limitations of VLMs in spatial reasoning and high computing requirements, DriveVLM team proposed DriveVLM-Dual, a hybrid system that combines advantages of DriveVLM and conventional autonomous driving pipelines. DriveVLM-Dual optionally combines DriveVLM with conventional 3D perception and planning modules, such as 3D object detector, occupancy network, and motion planner, allowing the system to achieve 3D localization and high-frequency planning. This dual-system design, similar to slow and fast thinking processes of human brain, can effectively adapt to changing complexity of driving scenarios.

AI and cloud companies attract attention as foundation models emerge.

As AI foundation models emerge, computing power, algorithm and data are indispensable. AI companies (iFLYTEK, SenseTime, Megvii, etc.) that are good at algorithms and have a large reserve of computing power, and cloud computing companies (Inspur, Volcengine, Tencent Cloud, etc.) with powerful intelligent computing centers, come under a spotlight of OEMs.

In the field of AI Foundation Model, SenseTime has deployed cockpit multimodal foundation model SenseChat-Vision, Artificial Intelligence Data Center (AIDC, with computing power of 6000P), and autonomous driving foundation model DriveMLM. In early 2024, SenseTime launched DriveMLM and achieved good results on CARLA, the most authoritative list of closed-loop test. DriveMLM is an intermediate solution between modular and end-to-end solutions and is interpretable.

Method	Type	DS↑	RC↑	IS↑	MPI↑
Roach [81]	DD	43.6	80.4	0.54	-
Interfuser [57]	DD	68.3	95.0	0.72	0.70
ThinkTwice [27]	DD	70.9	95.5	0.75	0.40
Apollo [3]	FSM	71.4	92.2	0.80	0.76
DriveMLM	LLM	76.1	98.1	0.78	0.96

Source: SenseTime

AI and cloud companies attract attention as foundation models emerge

Due to limitations of VLMs in spatial reasoning and high computing requirements, DriveVLM team proposed DriveVLM-Dual, a hybrid system that combines advantages of DriveVLM and conventional autonomous driving pipelines. DriveVLM-Dual optionally combines DriveVLM with conventional 3D perception and planning modules, such as 3D object detector, occupancy network, and motion planner, allowing the system to achieve 3D localization and high-frequency planning. This dual-system design, similar to slow and fast thinking processes of human brain, can effectively adapt to changing complexity of driving scenarios.

AI and cloud companies attract attention as foundation models emerge.

As AI foundation models emerge, computing power, algorithm and data are indispensable. AI companies (iFLYTEK, SenseTime, Megvii, etc.) that are good at algorithms and have a large reserve of computing power, and cloud computing companies (Inspur, Volcengine, Tencent Cloud, etc.) with powerful intelligent computing centers, come under a spotlight of OEMs.

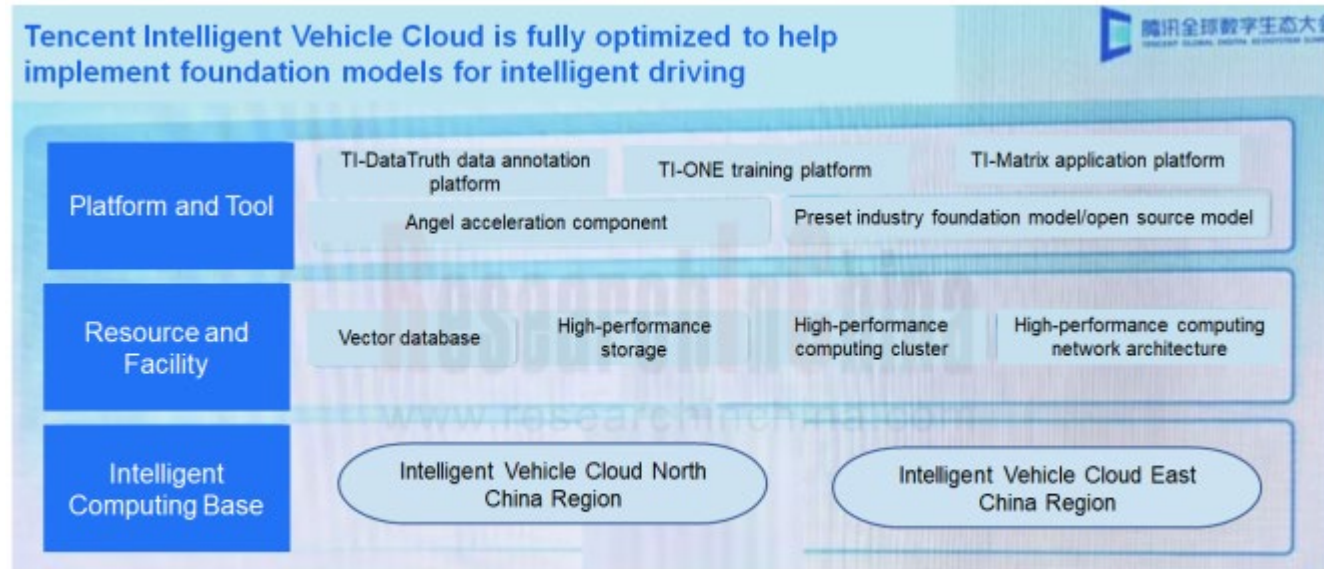
In the field of AI Foundation Model, SenseTime has deployed cockpit multimodal foundation model SenseChat-Vision, Artificial Intelligence Data Center (AIDC, with computing power of 6000P), and autonomous driving foundation model DriveMLM. In early 2024, SenseTime launched DriveMLM and achieved good results on CARLA, the most authoritative list of closed-loop test. DriveMLM is an intermediate solution between modular and end-to-end solutions and is interpretable.

Method	Type	DS↑	RC↑	IS↑	MPI↑
Roach [81]	DD	43.6	80.4	0.54	-
Interfuser [57]	DD	68.3	95.0	0.72	0.70
ThinkTwice [27]	DD	70.9	95.5	0.75	0.40
Apollo [3]	FSM	71.4	92.2	0.80	0.76
DriveMLM	LLM	76.1	98.1	0.78	0.96

Source: SenseTime

AI and cloud companies attract attention as foundation models emerge

For collection of autonomous driving corner cases, Volcengine and Haomo.ai work together to use foundation models to generate scenarios and improve annotation efficiency. The cloud service capabilities provided by Volcengine help Haomo.ai to improve the overall pre-annotation efficiency of DriveGPT by 10 times.



Source: Tencent

In 2023, Tencent released upgraded products and solutions in Intelligent Vehicle Cloud, Intelligent Driving Cloud Map, Intelligent Cockpit and other fields. In terms of computing power, Tencent Intelligent Vehicle Cloud enables 3.2Tbps bandwidth, 3 times higher computing performance, 10 times higher communication performance, and an over 60% increase in computing cluster GPU utilization, providing high-bandwidth, low-latency intelligent computing power support for training foundation models for intelligent driving. As for training acceleration, Tencent Intelligent Vehicle Cloud combines Angel Training Acceleration Framework, with training speed twice and reasoning speed 1.3 times faster than the industry's mainstream frameworks. Currently Bosch, NIO, NVIDIA, Mercedes-Benz, and WeRide among others are users of Tencent Intelligent Vehicle Cloud. In 2024, Tencent will further strengthen construction of AI foundation models.

Table of Content (1)

1 Classification of Autonomous Driving (AD) Algorithms and Common Algorithm Models

- 1.1 AD System Classification and Software 2.0
- 1.2 Baidu AD Algorithm Development History
 - 1.2.1 Model 1.0
 - 1.2.2 Perception 1.0
 - 1.2.3 Perception 2.0
 - 1.2.4 Perception Foundation Model
 - 1.2.5 Foundation Model Application Case
- 1.3 Tesla AD Algorithm Development History
 - 1.3.1 Development History (1)
 - 1.3.2 Development History (2)
 - 1.3.3 Development History (3)
 - 1.3.4 Development History (4)
 - 1.3.5 Dojo Supercomputing Center
- 1.4 Neural Network Model
 - 1.4.1 DNN
 - 1.4.2 CNN
 - 1.4.3 RNN
 - 1.4.4 Transformer and Foundation Model
 - 1.4.5 Occupancy Network
 - 1.4.6 Disadvantage of AI Foundation Model (1)
 - 1.4.7 Disadvantage of AI Foundation Model (1)
- 1.5 Traditional AD AI Algorithms (Small Model)
- 1.6 Transformer and BEV (Foundation Model)
 - 1.6.1 Transformer Schematic
 - 1.6.2 Three Common Transformers
 - 1.6.3 Foundation Model's Root is Transformers
 - 1.6.4 Why Need Foundation Model

- 1.6.5 No Code, Only NAS
- 1.6.6 End-to-end without Manual Rules
- 1.6.7 Transformer-based End-to-end Target Detection
- 1.6.8 BEV + Transformer is "Feature-level Fusion"
- 1.7 End-to-end Foundation Model Cases
 - 1.7.1 Case 1 (1)
 - 1.7.2 Case 2 (2)
 - 1.7.3 Case 3 (3)

2 Overview of AI Foundation Model and Intelligent Computing Center

- 2.1 AI Foundation Model
 - 2.1.1 Introduction
 - 2.1.2 Development Background
 - 2.1.3 Development History
 - 2.1.4 Position
 - 2.1.5 Business Mode
 - 2.1.6 Challenges and Future Development Trends of Implementation
 - 2.1.7 Continuous Iteration of Technical Architecture + Parameter Size
- 2.2 Application of AI Foundation Model in Automotive
 - 2.2.1 Application Direction
 - 2.2.2 Application in Intelligent Cockpit (1)
 - 2.2.3 Application in Intelligent Cockpit (2)
 - 2.2.4 Application in Intelligent Cockpit (3)
 - 2.2.5 Application Challenges
 - 2.2.6 Vehicle Models in China That Already Have and Will Soon Carry AI Foundation Models
- 2.3 Autonomous Driving (AD) Multimodal Basic Foundation Model
 - 2.3.1 Application Scenarios
 - 2.3.2 Typical Applications

Table of Content (2)

- 2.3.3 Typical Applications and Limitations Analysis
- 2.3.4 Major Adaptable Scenarios and Applications
- 2.3.5 Adaptation Scenario Cases
- 2.3.6 AD World Model and Video Generation
- 2.3.7 AD VFM: Adaptable Scenarios and Main Applications
- 2.3.8 AD: Simulation Data Generation and Scenario Reconstruction
- 2.4 Intelligent Computing Center
 - 2.4.1 Introduction
 - 2.4.2 Development History in China
 - 2.4.3 2.0 Era
 - 2.4.4 Construction
 - 2.4.5 Industry Chain
 - 2.4.6 Overall Architecture Diagram
 - 2.4.7 Development Trend (1)
 - 2.4.8 Development Trend (2)
 - 2.4.9 Development Trend (3)
 - 2.4.10 Development Trend (4)
 - 2.4.11 Reasons for Building AD Intelligent Computing Center
 - 2.4.12 Costs of Building AD Intelligent Computing Center
 - 2.4.13 Problems in Building AD Intelligent Computing Center
 - 2.4.14 AI Foundation Model and Computing Power Configuration of AD Companies
 - 2.4.15 Modes of OEMs Introducing Foundation Models
 - 2.4.16 Summary of Foundation Model and Intelligent Computing Center Progress in Automotive Industry (Suppliers)
 - 2.4.17 Summary of Foundation Model and Intelligent Computing Center Progress in Automotive Industry (OEMs)

3 Tesla Algorithm and Foundation Model Analysis

- 3.1 Algorithm Fusion of CNN and Transformer

- 3.1.1 Development of Visual Perception Frameworks
- 3.1.2 Visual Algorithm Architecture
- 3.1.3 Visual Algorithm Architecture (with NeRF)
- 3.1.4 Visual Algorithm Skeleton, Neck, and Head
- 3.1.5 Visual System Core: HydraNet
- 3.1.6 2D to 3D Images
- 3.1.7 Comparison of Swin Transformer with Traditional CNN Backbone
- 3.1.8 Backbone RegNet
- 3.1.9 Visual BiFPN
- 3.2 Transformer Turns 2D into 3D
 - 3.2.1 Transformer, BEV and Vector Space Expression
 - 3.2.2 Image-to-BEV Transformer
 - 3.2.3 Occupancy Network
 - 3.2.4 DETR 3D Architecture
 - 3.2.5 Transformer Model
 - 3.2.6 3D Target Detection
- 3.3 Occupancy Network, Semantic Segmentation and Time-space Sequence
 - 3.3.1 Vision Framework
 - 3.3.2 3D Target Detection and 3D Semantic Segmentation
 - 3.3.3 Deconvolution
 - 3.3.4 Video Neural Network Architecture
 - 3.3.5 Feature Queue
 - 3.3.6 Feature Queue - Time
 - 3.3.7 Feature Queue - Space
- 3.4 LaneGCN and Search Tree
 - 3.4.1 Lane Neural Network
 - 3.4.2 Vector Map
 - 3.4.3 LaneGCN Architecture
 - 3.4.4 AR Model

Table of Content (3)

- 3.4.5 MCTS for Trajectory Planning
- 3.4.6 MCTS Algorithm for Optimizing Core Ideas
- 3.5 Data Closed Loop and Data Engine
 - 3.5.1 Shadow Mode
 - 3.5.2 Data Engine
 - 3.5.3 Data Engine Case

4 AI Algorithms and Foundation Model Providers

- 4.1 Haomo.ai
 - 4.1.1 Profile
 - 4.1.2 Data Intelligence System
 - 4.1.3 Intelligent Computing Center
 - 4.1.4 Research and Application of Foundation Models
 - 4.1.5 Five Models of MANA
 - 4.1.6 Self-supervised Foundation Model
 - 4.1.7 Dynamic Environment Foundation Model
 - 4.1.8 Sources of Data
 - 4.1.9 Assistance of Five Foundation Models and Intelligent Computing Center to Haomo.ai
 - 4.1.10 Launch of DriveGPT
 - 4.1.11 Comparison between DriveGPT and ChatGPT
 - 4.1.12 2023 Haomo.ai AI Day: AD Semantic Perception Foundation Model
 - 4.1.13 2023 Haomo.ai AI Day: AD 3.0 Era
 - 4.1.14 2023 Haomo.ai AI Day: Introducing External Foundation Model
 - 4.1.15 2023 Haomo.ai AI Day: Automatic Annotation
- 4.2 QCraft
 - 4.2.1 Profile
 - 4.2.2 Ultra-fusion Perception Framework
 - 4.2.3 Feature and Timing Fusion Foundation Model

- 4.2.4 OmniNet Foundation Model Promotes the Implementation of Production Solutions
- 4.2.5 Predictive Algorithm Model
- 4.2.6 AD R&D Toolchain - QCraft Matrix
- 4.3 Baidu
 - 4.3.1 Baidu Apollo and Foundation Model
 - 4.3.2 AI Base of Baidu Intelligent Cloud
 - 4.3.3 ERNIE Foundation Model
 - 4.3.4 Application of ERNIE Foundation Model in Automotive Industry
 - 4.3.5 ERNIE Foundation Model Improves Baidu's Perception Algorithm Capabilities
 - 4.3.6 Baidu Intelligent Computing Center
 - 4.3.7 Baidu Officially Launched ERNIE Bot
- 4.4 Inspur
 - 4.4.1 Profile
 - 4.4.2 Three Highlights of Huaihai Intelligent Computing Center
 - 4.4.3 Yuan 2.0 Foundation Model
 - 4.4.4 AD Computing Framework AutoDRRT
 - 4.4.5 Inspur Helps XX Automotive Intelligent Computing Center Building
- 4.5 SenseTime
 - 4.5.1 Profile
 - 4.5.2 Sensenova Foundation Model Base
 - 4.5.3 Sensenova · SenseChat Multimodal Foundation Model
 - 4.5.4 Sensenova Foundation Model System
 - 4.5.5 Artificial Intelligence Data Center (AIDC)
 - 4.5.6 Foundation Model Application in Cockpit
 - 4.5.7 SenseAuto Empower
 - 4.5.8 UniAD Foundation Model
 - 4.5.9 AD Foundation Model

Table of Content (4)

- 4.6 Huawei
 - 4.6.1 Pangu Model 3.0
 - 4.6.2 Pre-annotation Foundation Model
 - 4.6.3 Scenario Generation Foundation Model and Foundation Model Cost Reduction
 - 4.6.4 Data Closed Loop Toolchain
 - 4.6.5 Computing Power Scale
- 4.7 Unisound
 - 4.7.1 Launched Shanghai Foundation Model
 - 4.7.2 AI Application in Cockpit
- 4.8 iFLYTEK
 - 4.8.1 Released "Spark" Cognitive Foundation Model
 - 4.8.2 "Spark" Cognitive Foundation Model Applied to Intelligent Cockpit
 - 4.8.3 Investment in Cognitive Foundation Model and Value Realization Approach
 - 4.8.4 "Spark" Foundation Model V3.0 Upgrades Cockpit, Sound and Intelligent Driving Functions
- 4.9 AISpeech
 - 4.9.1 Released Language Foundation Model and Signed with Several Auto Companies
 - 4.9.2 DFM-2 Foundation Model
 - 4.9.3 Foundation Model Development Plan
 - 4.9.4 Automotive Scenarios of Foundation Model
- 4.10 Megvii Technology
 - 4.10.1 AD Solution
 - 4.10.2 AD Algorithm Model
 - 4.10.3 General World Model for AD
- 4.11 Volcengine
 - 4.11.1 Profile
 - 4.11.2 Launched Volcano Ark Foundation Model
 - 4.11.3 Foundation Model Implementation

- 4.12 Tencent Cloud
 - 4.12.1 Profile
 - 4.12.2 Tencent Cloud Helps to Implement Intelligent Driving Foundation Model
 - 4.12.3 Tencent Cloud's Project Cases
- 4.13 Other Companies
 - 4.13.1 Banma Zhixing
 - 4.13.2 ThunderSoft
 - 4.13.3 Horizon Robotics' End-side Deployment of Foundation Model

5 Foundation Model of OEMs

- 5.1 Xpeng Motor
 - 5.1.1 Profile
 - 5.1.2 Transformer Foundation Model
 - 5.1.3 Data Processing
 - 5.1.4 Fuyao Intelligent Computing Center
 - 5.1.5 Foundation Model Application
 - 5.1.6 XBrain Intelligent Driving Foundation Model
 - 5.1.7 XGPT Foundation Model
- 5.2 Li Auto
 - 5.2.1 Profile
 - 5.2.2 Application of Foundation Model in Li Auto AD
 - 5.2.3 NPN and TIN
 - 5.2.4 Dynamic BEV
 - 5.2.5 DriveVLM-Dual: Visual Language Model (VLM) + Autonomous Driving Modular Pipeline
 - 5.2.6 Foundation Model Installation
 - 5.2.7 Smart Space 3.0 Connected to Mind GPT
 - 5.2.8 Foundation Model Mind GPT Progress
 - 5.2.9 In-Cabin Interaction Capabilities with Mind GPT

Table of Content (5)

- 5.3 Geely
 - 5.3.1 Profile
 - 5.3.2 Geely Xingrui Computing Center
 - 5.3.3 Leading Technologies of Geely Xingrui Computing Center
 - 5.3.4 Capabilities of Geely Xingrui Computing Center
 - 5.3.5 Geely Xingrui AI Foundation Model
 - 5.3.6 Geely-Baidu ERNIE Foundation Model
 - 5.3.7 Geely OCC Installation
- 5.4 BYD
 - 5.4.1 AI Foundation Model Layout
 - 5.4.2 Xuanji Architecture
 - 5.4.3 Xuanji Architecture: Integration of Xuanji AI Foundation Model in Vehicles
 - 5.4.4 Xuanji Architecture: Quad-Chain AI Foundation Model Layout
 - 5.4.5 Xuanji Architecture: BEV Perception Model
 - 5.4.6 Xuanji Architecture: Decision Planning Foundation Model
 - 5.4.7 Vehicle Intelligence AI Foundation Model Layout
 - 5.4.8 AI Foundation Model Development Plan
- 5.5 GM
 - 5.5.1 Launch of In-vehicle Voice Assistant Based on ChatGPT Technology
 - 5.5.2 GM and Google Cooperate on AI
- 5.6 Changan Automobile
 - 5.6.1 Layout of AI Foundation Model and Intelligent Computing Center
 - 5.6.2 Deepal GPT
- 5.7 Other Auto Enterprises
 - 5.7.1 GWM: All-round Layout of AI Foundation Model
 - 5.7.2 Chery: EXEED STERRA ES Equipped with Cognitive Foundation Model
 - 5.7.3 GAC
 - 5.7.4 SAIC-GM-Wuling
 - 5.7.5 Mercedes-Benz

- 5.7.6 Volkswagen
- 5.7.7 Stellantis
- 5.7.8 PSA

6 Application Trends of Sora and AI Foundation Model in Automotive

- 6.1 Analysis of Sora Text-to-Video Foundation Model
 - 6.1.1 Basic Functions
 - 6.1.2 Application Cases
 - 6.1.3 Case Analysis
 - 6.1.4 Basic Principles and Social Value
 - 6.1.5 Introduction of Basic Principles
 - 6.1.6 Advantages and Limitations
 - 6.1.7 Prospects and Future
- 6.2 Explanation of Sora's Underlying Algorithm Architecture
 - 6.2.1 Sora Text-to-Video Foundation Model: Basic System Introduction
 - 6.2.2 Interpreting Sora Module: Spacetime latent patches
 - 6.2.3 Interpreting Sora Module: Video compression network
 - 6.2.4 Interpreting Sora Module: Video compression network
 - 6.2.5 Interpreting Sora Module: Scaling transformers
- 6.3 Generative World Model and Intelligent Vehicle Industry
 - 6.3.1 Impacts of Sora on Intelligent Vehicle Industry and Forecast
 - 6.3.2 Comparison of Video Generation Capabilities between Sora and Tesla FSD-GWM
 - 6.3.3 Wayve Generative World Model
 - 6.3.4 Wayve Generative World Model System Architecture
 - 6.3.5 Multimodal Generative World Model of KIT University
- 6.4 Application Trends of AI Foundation Model in Automotive
 - 6.4.1 Trend 1
 - 6.4.2 Trend 1: Case

Table of Content (6)

- 6.4.3 Trend 2
- 6.4.4 Trend 3
- 6.4.5 Trend 4
- 6.4.6 Trend 5
- 6.5 AI Foundation Model Requirements for Chips
 - 6.5.1 Requirement 1
 - 6.5.2 Requirement 2
 - 6.5.3 Requirement 3
 - 6.5.4 Requirement 4
 - 6.5.5 Requirement 5
 - 6.5.6 Storage Chips Are Very Important in Foundation Model Era
 - 6.5.7 How to Solve Storage Bottleneck Problems (1)
 - 6.5.8 How to Solve Storage Bottleneck Problems (2)
 - 6.5.9 How to Solve Storage Bottleneck Problems (3)
 - 6.5.10 Emerging Carmakers Learn Tesla and Self-develop NPUs



Beijing Headquarters

TEL: 13718845418

Email: report@researchinchina.com

Website: [ResearchInChina](http://ResearchInChina.com)

WeChat: Zuosiqiche



Chengdu Branch

TEL: 028-68738514

FAX: 028-86930659

