ResearchInChina
www.researchinchina.com

Automotive AI Algorithm and Foundation Model Application Research Report, 2023

Sept. 2023

Large AI model research: NOA and foundation model facilitate a disruption in the ADAS industry

Recently some events upset OEMs and small- and medium-sized ADAS companies, as the autonomous driving industry changes faster than most people expected.

In large automotive forums in 2022, ADAS companies focused on introducing their driving-parking integrated solutions, and many of them also carried a large stock for the boom of the market in 2023. In 2023, under the pressure of cost reduction, OEMs have yet to apply driving-parking integration on large scale. On the contrary, driven by Huawei, Haomo.ai, Baidu and emerging carmakers, the theme of competition and promotion has directly shifted to highway NOA and urban NOA in 2023.

Reveled by a self-media and confirmed by channels, the highway NOA project an OEM in Southwest China commissioned multiple medium-sized Tier 1 suppliers to jointly develop has fallen flat, and the OEM thus turns to other Tier 1 suppliers DJI and Huawei to carry out this project. Huawei that originally worked on NOA for high-end models starts launching low- and mid-end solutions to compete with small- and medium-sized ADAS Tier 1 suppliers this year.

Meanwhile, emerging carmakers begin to race in number of cities where they introduce NOA. In late August 2023, Tesla demonstrated the autonomous driving performance of FSD V12, the world's first end-to-end AI-driven autonomous driving system, in a live stream. Elon Musk said that FSD V12 is completely enabled by AI and has no lines of code in it to run for recognition of roads and pedestrians, and the neural networks are used throughout. The C++ code of FSD V12 has been reduced by 10 times from over 20,000 lines to 2,000 lines. 99% of Tesla's decisions are made by neural networks to give visual inputs and control outputs. FSD V12 just works like a human brain. In addition, the powerful capabilities of FSD V12 are trained using a mass of "video data" and enabled by 10,000 H100 GPUs.

The autonomous driving performance of FSD V12 is amazing. With the support of large AI models, autonomous driving is at a critical point. It is said that FSD V12 will enter China in 2024, and will be largely introduced into vehicles in 2025 after being trained on Chinese roads. On September 12, 2023, Yu Chengdong said that AITO's urban NOA can be available in major cities across China in late 2023 (according to Huawei, the meaning of being available in cities is not that it is available to all roads in all cities, but that vehicles need to run on structured roads with clear road boundaries). Tesla and Huawei, two benchmarks of autonomous driving, weigh heavily on other automakers and ADAS Tier 1 suppliers.

Where will the ADAS industry go? How do OEMs and ADAS companies deal with the challenges posed by large AI models and NOA? The Automotive AI Algorithm and Foundation Model Application Research Report, 2023 combs through the development history of ADAS algorithm and large AI model, and explores the development trends of large AI models in automotive.

What changes will end-to-end autonomous driving bring?

Autonomous driving algorithm systems are divided into two categories: end-to-end autonomous driving and modular autonomous driving. The modular autonomous driving system has three layers: environmental perception layer, decision and planning layer, and control and execution layer. In the modular autonomous driving system, different teams are responsible for different modules for better division of labor and cooperation, thus improving development efficiency. The shortcoming is that the entire system is very complex and large, and requires manual design of hundreds or thousands of modules.

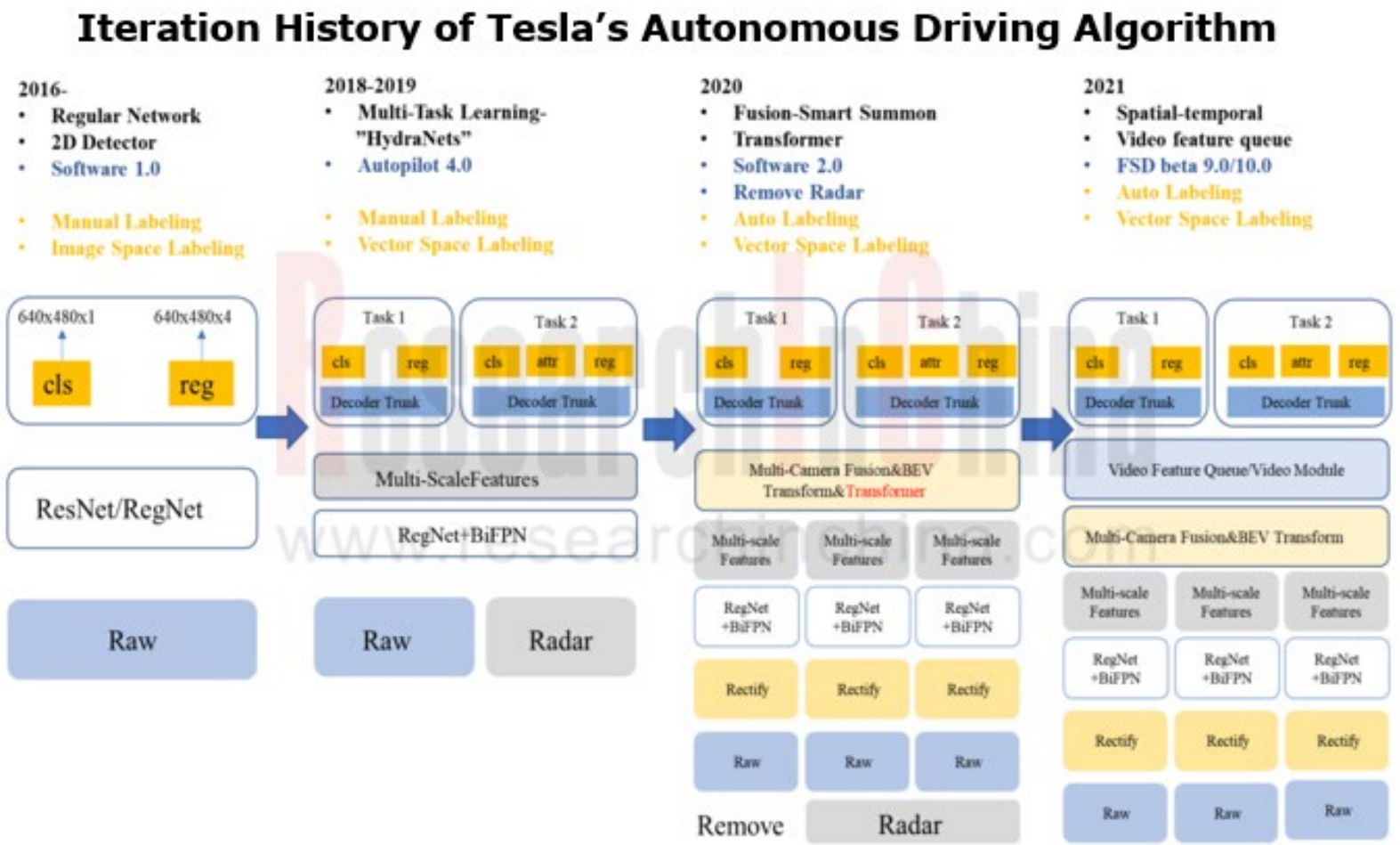
End-to-end autonomous driving means that the vehicle directly sends the information (raw image data, raw point cloud data, etc.) collected by sensors to a unified deep learning neural network, and the neural network processes it and then directly outputs the driving commands (steering wheel angle, steering wheel speed, accelerator pedal opening, brake pedal opening, etc.) of the autonomous vehicle. In end-to-end autonomous driving, there are no complicated rules in manual design, and with only a very small amount of human training data, the deep learning neural network can learn to drive, regardless of HD map coverage.

When autonomous driving evolves to the urban NOA stage, modular autonomous driving algorithms will no longer meet the needs, and end-to-end autonomous driving algorithms will begin to become mainstream. Furthermore most of the accumulated modular autonomous driving algorithms cannot be migrated to end-to-end autonomous driving, requiring a fresh start. Therefore it is not too late for BYD to start great efforts on development of autonomous driving just in 2023. BYD said that combining foundation model technologies such as BEV perception is an opportunity for BYD to overtake at the bend in advanced intelligent driving; and BYD is developing some distinctive ADAS functions by combining intelligent driving with the Yisifang platform.

BYD's concept of lane change to overtake is to skip modular autonomous driving and directly step into end-to-end autonomous driving based on large AI models.

Early autonomous driving perception algorithms were based on conventional computer vision technology. After 2010, as deep learning technology develops, neural network is introduced into autonomous driving perception algorithms, bringing a qualitative improvement in the perception effect of autonomous vehicles. Neural network models applied at the perception layer fall into two categories: small models like CNN and RNN, and Transformer foundation model.

Transformer is a neural network model based on the attention mechanism. This model was proposed in Google’s 2017 paper titled "Attention Is All You Need". Transformer is superior to RNN for its ability to perform parallel computing and handle long sequence inputs. Compared to CNN, Transformer reserves location information and solves the problem of depending on long-distance features. So Transformer has become one of the most popular models in natural language processing. Tesla is the first to introduce Transformer into autonomous driving algorithms, and other emerging carmakers and new brands of conventional automakers follow suit.



Source: Tesla

Iteration History of Mobileye

Source: Mobileye

Why are large AI models needed?

Needed by urban NOA: currently OEMs are expanding from highway NOA to urban NOA. The expansion from highway scenario to urban scenario means vehicles face far more long tail problems (corner cases). Highway scenario is relatively closed in specific road sections, with highly standardized traffic environments, and highway driving rules clearly define driving behaviors of vehicles; traffic participants are simple, not involving pedestrians, and the driving status is also more predictable. All of which makes highway the first scenario to implement NOA. Yet in urban scenarios, complex roads and road conditions (traffic light intersections), multiple traffic participants (pedestrians, low-speed two wheelers), and high scenario heterogeneity (road conditions between cities and even between road sections differ greatly) combine to lead to a surge in corner cases in autonomous driving. The implementation of urban NOA therefore requires higher generalization capabilities of autonomous driving models. Considering commercial application cost, we believe that the application of large AI models to improve generalization capabilities and reduce/control vehicle hardware cost is the key to the evolution of autonomous driving algorithms.

Needed to get rid of HD maps and lower cost: before 2022, Chinese OEMs generally implemented urban NOA using the HD map + single vehicle perception solution. Yet in the implementation process, they found that HD maps pose three big problems: 1) Inability to achieve real-time updates; 2) Regulatory risks; 3) High cost. The upgrading of autonomous driving perception algorithms to the BEV+Transformer architecture helps urban NOA cast off HD maps.

BEV perception model is far more competent to cope with extreme weather conditions: in the post fusion model, the resolution of data/video streams collected by cameras will be much lower in the case of extreme weather conditions such as rain and snow, making it difficult to meet the criteria of acceptability judged by the cameras, so the results transmitted to the backend for planning and control slump. Unlike the post fusion model, the process of converting different views into the BEV in image collection by cameras involves feature level fusion. For example, in extreme weather conditions, some photon information still reflects the situation of the obstacles ahead, which can be used for subsequent planning and control. Under the framework of feature level fusion, the perception model makes far more use of data.

Large AI models have a promising future in intelligent cockpit

Large AI models not only find successful application in autonomous driving, but also have a promising future in intelligent cockpit. Ge Yuming, the head of the C-V2X Working Group of the MIT2020 (5G) Promotion Group, says that foundation models have three impacts on intelligent cockpits:

Firstly, the context understanding capability of foundation models can enhance voice assistant's ability to understand and respond to passengers' voice and semantics, and enables such functions as continuous dialogue, memory dialogue, and active interaction;
Secondly, foundation models can enable vehicle assistants with multimodal understanding and perception capabilities to reduce driver's interaction pressure;

Thirdly, in terms of maps, the accuracy of vehicle navigation capabilities in route optimization and judgment will improve with the application of foundation models.

Source: Huawei

How to deal with the challenges posed by large AI models?

Big data and computing power are important prerequisites for the application of large AI models. The Transformer model requires mileage data of 100 million kilometers or more for qualitative change from quantitative change. The raw data collected by sensors also needs to be annotated before being used for training algorithm models, and automatic annotation tools can greatly improve data processing speed. Since 2018, Tesla's data annotation has gradually developed from 2D manual annotation to 4D spatial automatic annotation; Chinese providers like Xpeng and Haomo.ai have also announced automatic annotation tools, bringing much higher annotation efficiency. In addition to real data, simulation scenes are an important solution to the problem of insufficient data for training foundation models.

Generative AI is expected to significantly enhance the generalization capabilities of simulation scenes, and help OEMs use more simulation scene data, thereby improving the iteration speed of autonomous driving models and shortening the development cycle.

High computing power is another important condition for Transformer model training, and supercomputing centers thus have become an important infrastructure for autonomous driving providers. Tesla's AI computing center Dojo uses a total of 14,000 NVIDIA GPUs to train AI models, increasing network training speed by 30%. Among Chinese providers, Xpeng and Alibaba jointly funded and created "Fuyao", an autonomous driving AI computing center which allows for 170 times faster training of autonomous driving algorithm models.

Except for a few OEMs like NIO, Xpeng and Li Auto, who already work on foundation model application and have ample funds, it is difficult for other OEMs to invest simultaneously in big data and high computing power, supercomputing center, and AI chip self-development as Tesla does. Leveraging the power of large AI model providers to perfect application of foundation models is a relatively pragmatic approach.

If you cannot win in China, you cannot win globally

For small- and medium-sized ADAS Tier 1 suppliers, it is almost impossible to independently launch NOA solutions that can rival Huawei NCA, and the huge investment in intelligent computing centers is an insurmountable mountain. Meanwhile mainstream automakers start turning to independent development of autonomous driving, leaving ever less scope for small- and medium-sized ADAS Tier 1 suppliers. The industry integration is inevitable. They need to quicken the pace of going public to become an integrator, or seek opportunities to be acquired.

In the industry disruption, there are both risks and opportunities. Huawei and other IT giants have high staff cost, and OEMs are also unwilling to be constrained by these IT tycoons, so small- and medium-sized ADAS Tier 1 suppliers are not without survival space. In the next step, resource integration is critical for small- and medium-sized ADAS Tier 1 suppliers. In the trend for cockpit-driving integration and cross-domain integration, close partnerships with large AI model companies (Unisound, SenseTime, AISpeech, etc.), listed cockpit companies and chassis domain companies among others are all options.

Also large AI models bring opportunities to later AI chip companies. Not all the high-compute chips of conventional intelligent driving SoC and cockpit SoC vendors meet the needs of large AI models. Emerging AI chip companies can precisely redesign high-compute chips required by foundation models, and grow rapidly by dint of customization demand from OEMs.

According to Yu Kai, founder and CEO of Horizon Robotics, China runs ahead of foreign countries for more than five years in vehicle intelligence application. NOA and foundation models will help to further widen the gap between Chinese companies and foreign Tier 1 suppliers. After establishing their foothold, China's local Tier 1 suppliers are expected to have opportunities to partner with foreign OEMs and Tier 1 giants.

As Mobileye's CEO Amnon Shashua said, **"If you cannot win in China, you cannot win globally."**

Table of Content (1)

1 Classification and Development History of Autonomous Driving Algorithms
1.1 Classification of Autonomous Driving Systems
1.2 End-to-end Autonomous Driving and Software 2.0
1.3 End-to-end Autonomous Driving Model Case: UniAD
1.4 Development History of Baidu AD Algorithm
1.4.1 Development History of Baidu AD Algorithm: Model 1.0
1.4.2 Development History of Baidu AD Algorithm: Perception 1.0
1.4.3 Development History of Baidu AD Algorithm: Perception 2.0
1.4.4 Development History of Baidu AD Algorithm: Large Perception Model
1.4.5 Development History of Baidu AD Algorithm: Foundation Model Application Cases
1.5 Development History of Tesla AD Algorithm
1.5.1 Development History of Tesla AD Algorithm: Entering the Phase of "More Stress on Perception and Less Stress on Maps"
1.5.2 Development History of Tesla AD Algorithm: Occupancy Network
1.5.3 Development History of Tesla AD Algorithm: FSD Beta V12
1.5.4 Tesla Dojo Supercomputer

2. Common Autonomous Driving AI Algorithms and Models
2.1 Neural Network Models
2.1.1 DNN
2.1.2 CNN
2.1.3 RNN
2.1.4 Transformer
2.1.5 Occupancy Network
2.1.6 Shortcomings of AI Algorithms (1)
2.1.7 Shortcomings of AI Algorithms (2)
2.1.8 Shortcomings of AI Algorithms (3)
2.2 Conventional Autonomous Driving AI Algorithms (Small Models)
2.2.1 2D Object Detection of Early CNN
2.2.2 Core Typical Algorithms of Early CNN for Intelligent Driving
2.2.3 3D Bounding Box
2.2.4 3D Bounding Box Is Inseparable from LiDAR
2.2.5 6D-VISION
2.2.6 Beyond Object Detection: Semantic Segmentation
2.2.7 Road Semantic Segmentation and Moving Object Semantic Segmentation
2.3 Transformer and BEV (Foundation Models)
2.3.1 Transformer Diagram
2.3.2 Three Common Transformer Models
2.3.3 The Base of Foundation Models is Transformer
2.3.4 Why Foundation Models Are Needed
2.3.5 No Code, Only NAS
2.3.6 End-to-end, No Manual Rules Added
2.3.7 Transformer-based End-to-end Object Detection
2.3.8 BEV+Transformer Is "Feature Level Fusion"
2.3.9 Chinese Vehicle Models That Have Been and Will Be Equipped with Large AI Models
2.4 Algorithm Application in Different Scenarios
2.4.1 NVIDIA's Traffic Sign and Signal Light Classification Algorithm
2.4.2 Tesla's Lane Line Model
2.4.3 Li Auto's Complex Intersection Algorithm
2.4.4 Planning & Control: System Classification
2.4.5 Planning & Control: Motion Planning Algorithm
2.4.6 Planning & Control: Tesla's Planning and Control Algorithm
2.5 AI Algorithm's Requirements for Chips
2.5.1 AI Algorithm's Requirements for Chips
2.5.2 AI Algorithm's Requirements for Chips (1)

Table of Content (2)

2.5.3 AI Algorithm's Requirements for Chips (2)
2.5.4 AI Algorithm's Requirements for Chips (3)
2.5.5 AI Algorithm's Requirements for Chips (4)
2.5.6 Memory Chips Are Very Important in the Era of Foundation Models
2.5.7 How To Break the Storage Bottleneck (1)
2.5.8 How To Break the Storage Bottleneck (2)
2.5.9 How To Break the Storage Bottleneck (3)
2.5.10 Emerging Carmakers Learn from Tesla and Self-develop Intelligent Driving Chips

3. Overview of Large AI Models and Intelligent Computing Centers
3.1 Overview of Large AI Model
3.1.1 Development History of Large AI Model
3.1.2 Role of Large AI Models in Development of Artificial Intelligence
3.1.3 Large AI Model Business Models
3.1.4 Challenges in Implementation of Large AI Models and Development Trends
3.1.5 Foundation Model Architecture + Continuous Iteration of Parameter Scale
3.2 Application of Large AI Models in Vehicles
3.2.1 Application Directions of Foundation Models in Vehicles
3.2.2 Application of Foundation Models in Intelligent Cockpits
3.2.3 Application of Foundation Models in Intelligent Driving
3.2.4 Challenges in Application of Large AI Models in Vehicles
3.2.5 Impacts of Foundation Model Application on Sensors
3.3 Intelligent Computing Center
3.3.1 Overview of Intelligent Computing Center
3.3.2 Development History of China Intelligent Computing Center
3.3.3 Era of Intelligent Computing Center 2.0
3.3.4 Construction of Intelligent Computing Centers
3.3.5 Intelligent Computing Center Industry Chain
3.3.6 Overall Architecture Diagram of Intelligent Computing Center
3.3.7 Reasons for Establishing Intelligent Computing Centers for Autonomous Driving
3.3.8 Cost of Building An Autonomous Driving Intelligent Computing Center
3.3.9 Problems in Building An Autonomous Driving Intelligent Computing Center
3.3.10 Large AI Models and Computing Power Configurations of Autonomous Driving Companies
3.3.11 Foundation Model Introduction Modes of OEMs
3.3.12 Summary on Progresses in Foundation Models and Intelligent Computing Centers in Automotive Industry (Suppliers)
3.3.13 Summary on Progresses in Foundation Models and Intelligent Computing Centers in Automotive Industry (OEMs)

4. Tesla's Algorithm and Foundation Model Application
4.1 Tesla's Algorithm Fuses CNN and Transformer
4.1.1 Development History of Visual Perception Framework
4.1.2 Vision Algorithm Architecture
4.1.3 Vision Algorithm Architecture (Including NeRF)
4.1.4 Skeleton, Neck, and Head of Vision Algorithm
4.1.5 Vision System Core - HydraNet
4.1.6 2D to 3D Image Conversion
4.1.7 Use Transformer in the Head Part
4.1.8 Comparison between Swin Transformer and Conventional CNN Backbone Network
4.1.9 Swin Transformer Consumes the Lowest Computing Power
4.1.10 Backbone Network - RegNet
4.1.11 Visual BiFPN
4.2 Transformer Converts 2D into 3D
4.2.1 Transformer, BEV and Vector Space Expression
4.2.2 Image-to-BEV Transformer
4.2.3 Occupancy Network

Table of Content (3)

4.2.4 DETR 3D Architecture
4.2.5 Transformer Model
4.2.6 3D Object Detection
4.2.7 NeRF and Implicit Neural Network
4.3 Occupancy Network, Semantic Segmentation and Spatiotemporal Sequences
4.3.1 Vision Framework
4.3.2 3D Object Detection and 3D Semantic Segmentation
4.3.3 Deconvolution
4.3.4 Video Neural Network Architecture
4.3.5 Feature Queue
4.3.6 Feature Queue - Temporal Sequence
4.3.7 Feature Queue - Spatial Sequence
4.3.8 Spatial Feature
4.4 LaneGCN and Search Tree
4.4.1 Lane Neural Network
4.4.2 Vector Map
4.4.3 LaneGCN Architecture
4.4.4 AR Model
4.4.5 Trajectory Planning MCTS
4.4.6 Core Ideas of MCTS Algorithm Optimization
4.5 Data Closed Loop and Data Engine
4.5.1 Shadow Mode
4.5.2 Data Engine
4.5.3 Data Engine Cases

5. AI Algorithm and Foundation Model Providers
5.1 Haomo.ai
5.1.1 Profile
5.1.2 Data Intelligence System - MANA System
5.1.3 Intelligent Computing Center - MANA OASIS
5.1.4 Research and Application of Foundation Models
5.1.5 MANA's Five Major Models
5.1.6 Human Driving Self-supervised Cognitive Model and Multimodal Model
5.1.6 Video Self-supervised Model
5.1.6 3D Reconstruction Model
5.1.6 Dynamic Environment Model
5.1.7 Sources of Data
5.1.8 Help from Five Major Models and Intelligent Computing Center
5.1.9 Launch of DriveGPT
5.1.10 Comparison between DriveGPT and ChapGPT
5.2 QCraft
5.2.1 Profile
5.2.2 Hyper-converged Perception Framework
5.2.3 Feature and Temporal Fusion Model
5.2.4 OmniNet Model Facilitates the Implementation of Mass Production Solutions
5.2.5 Prediction Algorithm Model
5.2.6 Autonomous Driving R&D Tool Chain
5.3 Baidu
5.3.1 Overview of Baidu AI Cloud
5.3.2 Overview of Baidu Apollo
5.3.3 ERNIE Model
5.3.4 Application of ERNIE Model in Automotive Industry
5.3.5 ERNIE Model Improves Baidu's Perception Algorithm Capabilities
5.3.6 Intelligent Computing Center
5.3.7 Baidu Released ERNIE Bot, and Quite A Few OEMs Have Joined the Ecosystem
5.4 Inspur
5.4.1 Profile
5.4.2 Three Highlights of Huaihai Intelligent Computing Center

Table of Content (4)

5.4.3 Autonomous Driving Computing Framework - AutoDRRT
5.5 SenseTime
5.5.1 Profile
5.5.2 SenseNova Model System
5.5.3 Intelligent Computing Center (AIDC)
5.5.4 Application of Foundation Models in Cockpits
5.5.5 SenseAuto Empower
5.5.6 UniAD Model
5.6 Huawei
5.6.1 Pangu Model 3.0
5.6.2 Pre-annotated Foundation Model
5.6.3 Scene Generation Foundation Model and Foundation Model Cost Reduction
5.6.4 Data Closed Loop Tool Chain
5.7 Unisound
5.7.1 Launched Shanhai Model
5.7.2 Application of AI in Cockpits
5.8 iFLYTEK
5.8.1 Released Spark Cognitive Model
5.8.2 Spark Cognitive Model Is Applied to Intelligent Cockpits
5.8.3 Investment in Large Cognitive Model and Value Realization Mode
5.9 AISpeech
5.9.1 Released A Large Language Model and Sign Agreements with Multiple Automakers
5.9.2 DFM-2 Model
5.9.3 Foundation Model Development Planning
5.10 Megvii Technology
5.10.1 Autonomous Driving Solution
5.10.2 Autonomous Driving Algorithm Model
5.11 Others
5.11.1 Banma Zhixing Introduces Qwen Model
5.11.2 iFLYTEK released Spark Cognitive Model
5.11.3 ThunderSoft's Rubik Model
5.11.4 Volcengine
5.11.5 Horizon Robotics Deploys Foundation Models at the Terminal Side

6. OEMs' Application of Foundation Models
6.1 Xpeng
6.1.1 Profile
6.1.2 Transformer Model
6.1.3 Data Processing
6.1.4 Fuyao Intelligent Computing Center
6.2 Li Auto
6.2.1 Foundation Model Layout
6.2.2 Application of Foundation Models in Autonomous Driving
6.2.3 NPN and TIN
6.2.4 Dynamic BEV
6.3 Geely
6.3.1 Profile
6.3.2 Xingrui Intelligent Computing Center
6.3.3 Leading Technologies of Xingrui Intelligent Computing Center
6.3.4 Capabilities of Xingrui Intelligent Computing Center
6.3.5 Geely-Baidu ERNIE Large Model
6.4 GM
6.4.1 Consider Launching A Vehicle Voice Assistant Based on ChatGPT
6.4.2 AI Cooperation with Google
6.5 BYD
6.5.1 Utilize BEV Perception and Other Foundation Models to Overtake at the Bend
6.5.2 Multi-sensor Multi-task Fusion Perception

Table of Content (5)

6.5.3 Large Data-driven Model That Will Realize the Whole Process of Perception, Prediction, Decision and Planning
6.6 Other Automakers
6.6.1 Great Wall Motor’s Large AI Model Application and Layout
6.6.2 GAC Launches A Large AI Model Platform
6.6.3 Chery EXEED STERRA ES Carries A Large Cognitive Model
6.6.4 SAIC-GM-Wuling
6.6.5 Changan Automobile
6.6.6 Mercedes-Benz Applies ChatGPT

Contact

Beijing Headquarters

TEL: 13718845418
FAX: report@researchinchina.com

Website: ResearchInChina

WeChat: Zuosiqiche

Chengdu Branch

TEL: 028-68738514
FAX: 028-86930659

zuosiqiche