International Journal of Frontiers in Engineering Technology
ISSN 2706-655X Vol.6, Issue 4: 25-30, DOI: 10.25236/IJFET.2024.060405

Research on the Prospect of Future Vehicle Use Based on AI Large Model

Biao Wang*, **Xiangyu Chen, Ziyu Gao, Xueli Wan, Ao Yang**

College of Mechanical Engineering, North China University of Science and Technology, Tangshan, 063210, China
Corresponding author: zgmrwb@163.com

Abstract: *Since 2023, Artificial Intelligence Large Model has become the hottest topic of discussion with the explosion of OpenAI. Based on the above problems in the actual use of vehicles and the current research results on the application of AI large models, this paper will briefly describe the application of AI large models in vehicles in the future from the aspects of concept, theoretical basis and practical application. At present, the car mainly uses "imperative interaction", similar to voice remote control, which mainly relies on the vehicle user to provide simple instructions to the vehicle intelligent system through voice. With the characteristics of multimode, multi-attention and autonomous fine-tuning of the AI large model, it can be applied to any individual vehicle, through a certain amount of time of interactive learning. AI large models provide strong support for vehicle intelligence and automation, improve driving safety and comfort, and build "person vehicle-home" connectivity. It will promote the development of the automotive industry in the direction of more intelligent and user-friendly.*

Keywords: *AI Large model, Vehicle traffic, AI autonomous driving*

1. Introduction

Since 2023, Artificial Intelligence Large Model have become the hottest topic of discussion with the popularity of OpenAI. At the same time, based on the rapid development[1] of the chip industry and computer computing power, Artificial General Intelligence (AGI) has begun to become the focus of major technology enterprises. For example, the current most concerned Chat GPT(Chat generative Pre-trained Transformer) can be classified as a linguistic version of AGI.

At present, most of the intelligent in-vehicle assistants are realized by means of language interaction technology and Internet of things technology, which can assist the car owner to complete some simple operations. However, there is still a huge gap between them and the existing Artificial Intelligence Large Model in understanding the syntax and semantics of certain requirements and performing corresponding operations.

Urban traffic is an inevitable problem in the use of vehicles. At present, the common problem of urban traffic is: peak congestion. The actual travel routes of vehicles mostly rely on navigation and driving experience, resulting in low utilization rate of some traffic roads. At the same time, the urban traffic system can be summarized as: diversity, complexity, hybrid, dynamic, coupling[2]. That is, there are a variety of transportation modes in the urban transportation system, involving a variety of aspects, a large time span, and the actual road conditions and travel demand are constantly changing dynamically. Some studies have proposed that new traffic models should be developed with the help of AI large models to meet the future. For example, the MT-GPT multi-mode traffic model proposed by Zhou Zhen et al., which can provide data-driven for multi-directional decision-making tasks in complex multi-mode transportation systems in the future. At the same time, Trans World NG, a traffic simulator proposed[3]by foreign scholars, introduces simulation system into AI large model, which provides strong support for the application of large model in the traffic field.

Based on the above problems in the actual use of vehicles, and the current research results of the application of AI large model, this paper will briefly describe the application of AI large model in vehicles in the future from the aspects of concept, theoretical basis and practical application.

ISSN 2706-655X Vol.6, Issue 4: 25-30, DOI: 10.25236/IJFET.2024.060405

2. The interaction and application of the two models

2.1 Interaction with AI large model

The fact that AI large models, represented by Chat GPT, can now achieve excellent language dialogue is rooted in a major change in machine deep learning. The initial deep learning[4] was "supervised learning" + "go it alone", and then developed into the "pre-trained model" + "task fine-tuning" mode based on neural network. After the release of Chat GPT, it marked that deep learning had developed into the "pre-trained large model" + "prompt generation" mode, that is, training the corresponding model through huge data. In practice, due to the huge data training, the trained large model can realize the transformation and decision-making of different text tasks under the same task module. The rest of the AI large model interaction is mainly summarized as the following aspects[5]:

Natural Language processing (NIP), which is a branch of artificial intelligence and linguistics, its main role is to realize the mutual transformation between computer language and human natural language, and at the same time, the computer can understand, interpret and execute human natural language. Before AI large models were widely known, Natural language processing (NIP) mainly relied on rule-based methods (i.e., using predefined rules to interpret human natural language) and statistical methods (i.e., based on a large number of language database statistical information for language processing). With the continuous development of machine deep learning, neural network models have been applied to natural language processing, and with the release of Chat GPT, natural language processing has begun to use Transformers models. Transformers model can realize multi-task processing, which is the main choice for training AI large model natural language processing.

Autonomous fine-tuning and intention recognition: In the actual application process of the large model, different use scenarios and tasks determine that the large model cannot be designed simply according to the appearance. Based on the continuous innovation and development of deep learning mode, the current deep learning mode can meet the large model in the use process, continuously collect data information independently, so as to carry out independent fine-tuning and identify user intention. And then the exclusive AI large model is realized.

At present, the interaction of large models still refers to the deep learning neural network. Take the YOLOV5 deep learning neural network as an example, and its main working mode is shown in Fig 1.

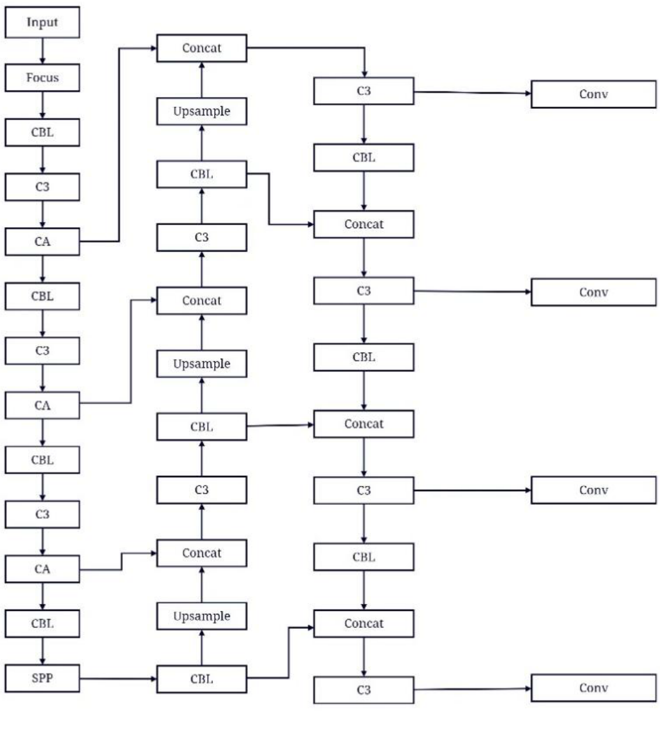

Figure 1: Main working mode of YoloV5

The When an image is fed into YOLO, it goes through a series of convolutional and pooling layers. The main role of these layers is toperform feature extraction on the image, gradually reducing the spatialdimension of the image while increasing the depth of the features. Through the processing of these layers, the network can capture the important information and details in the image. Next, YOLO will divide the input image into a grid of SxS. Each grid cell is responsible for detecting objects that lie within its range. This division allows the network to focus more on local areas of the image, improving the accuracy of object detection. For each grid cell, YOLO predicts B bounding boxes. Each bounding box contains five prediction parameters: (x,y,w,h,confidence). Where (x,y) represents the offset of the box center with respect to the grid cell, (w,h) represents the width and height of the box, and confidence is the confidence of the object predicted by the box. Confidence is an important metric that reflects the product of the IoU (intersection over Union) ratio of the predicted box to the true box and the probability that the box contains any object or not. In addition to bounding box predictions, YOLO also predicts C conditional class probabilities for each bounding box. These probabilities represent the probability that a given box contains each class. In this way, the network can consider multiple categories at the same time, improving the accuracy of object detection. During training, YOLO defines a loss function to measure the difference between the predicted results and the true labels. This loss function is composed of position error, confidence error, and classification error. By optimizing this loss function, the network can continuously adjust its weight parameters to reduce the gap between the predicted results and the true labels. When trained, YOLO can be used for object detection tasks. In the testing phase, to eliminate duplicate detection boxes, YOLO applies a Non-maximum suppression (NMS) algorithm. The algorithm sorts all the boxes based on their confidence and then gradually eliminates the boxes with higher confidence[6-7].

2.2 Application of large model interaction

At present, cars mainly use "imperative interaction", similar to voice remote control, which mainly relies on vehicle users to provide simple instructions to the vehicle intelligent system through voice. This interaction mode can meet some basic needs, but based on the powerful interaction ability of AI large model, if applied to the vehicle, future driving will have travel scene dialogue interaction. That is, when the user has a travel need, the vehicle AI large model can actively interact with each other to plan the trip, at the same time, it performs logical reasoning and strategy suggestions on the user's needs and decisions, and answers some knowledge and other deep interactive functions.

Based on the characteristics of multi-mode, multi-attention and autonomous fine-tuning of AI large model, it can be applied to any individual vehicle, through a certain amount of time of interactive learning, so as to realize that each car has an AI intelligent vehicle housekeeper consistent with the owner. At the same time, under the effect of multi-mode and multi-attention of AI large model, taking the group chat UAV group of Northwestern Polytech University as reference, a unified AI large model of urban road can be established, and the unified large model of road can be connected to the large model of individual vehicles. The two types of large models exchange data with each other, realizing the autonomous and intuitive information interaction between drivers and vehicles, and between vehicles. Then, the separate large models on each vehicle can understand each other's next driving intention, and then efficiently complete the task.

3. Vertical application of large models

3.1 Vehicle vertical large model

At present, the most popular Chat GPT belongs to the general language large model according to the division method in the field of artificial intelligence, that is, it does not optimize for a specific task, but learns general problems and patterns through large-scale data training. Taking Chat GPT as an example, the general language large model has great advantages in the field of natural language processing. It can complete text generation, machine translation, question answering and other language tasks. General-purpose large models in other domains such as DALLE, a multi-modal large model developed by Open AI, combine common trained models and certain kinds of diffusion models are able to convert text descriptions into corresponding images. Sora, a large visual model, can understand the direct relationship between the meaning of natural text and the generated video, and can successfully generate 60-second video according to the laws of the objective world. According to the division method in the field of artificial intelligence, that is, AI large models are divided according to the data and model level, there are also vertical large models in addition to the general large model. Different from the general large model, the vertical large model is based on the general model, through additional training and optimization for the data of a specific industry or field, so as to obtain higher accuracy and specialization

ISSN 2706-655X Vol.6, Issue 4: 25-30, DOI: 10.25236/IJFET.2024.060405

in the field. For example, a vertical large model in the finance domain would focus on financial terminology, trading rules, and related regulations, while a model in the healthcare domain would focus on medical knowledge and case analysis. Compared with the general large model, the vertical large model pays more attention to the combination of deep learning and domain knowledge to meet the needs of specific scenarios. At the same time, the cost of debugging and deployment is lower and the cost performance is higher, which can better meet the personalized needs and improve the use efficiency. vehicles is shown in Fig 2.

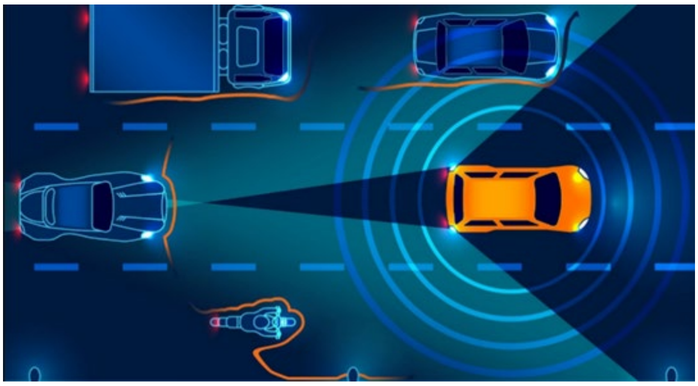

Figure 2: Automatic driving of vehicles

It involves four continuous tasks, such as perception, prediction, planning and execution. At present, HU has proposed a full-stack controllable end-to-end autonomous driving scheme based on Transformer decoder, called UniAD, which enables the large vehicle model to integrate the above four continuous tasks into the same neural network system. The joint operation of four tasks can be realized at the same time. In addition, the test results of the large model system in the formal environment are in good agreement with the actual test situation.

In addition to autonomous driving, it is a hot field of vehicle vertical large model application to involve large models in the travel transportation network. To integrate the large model into the transportation network is to make the model understand, manage and predict the huge and complex transportation system. Based on the connectivity of large models and traffic network models, it has been proposed to combine machine learning technology with multivariate data fusion in series, and use the characteristics of multi-attention and multi-level learning of large models to understand and infer road realtime information. At the same time, Gaussian regression process in existing traffic models is used to complete the prediction and execution tasks in road traffic systems. At the same time, some researchers have proposed a large model of vehicle vertical traffic based on generative transformers, which combines the prediction of actual road conditions with virtual dynamic simulation. The large model can better help drivers complete travel tasks by capturing the behavior and needs of vehicle drivers and combining simulation capabilities.

In February 2024, Zhou Zhen et al. proposed the latest vehicle vertical Transportation large-scale model -- MT-GPT (Multi-modal Transportation Generative Pre-trained Transformer), which is a multimodal transportation large-scale model. It can provide data-driven and practical decision-making suggestions for multi-direction and multi-granularity decision problems in multi-model transportation system tasks. At the same time, MT-GPT can provide perception and decision support for multi-spatio-temporal scale urban transportation planning, network design, transportation infrastructure construction and traffic management in the aspects of digitalization, informatization and intelligence [8-10]. In the MT-GPT traffic large model, the concept of "point, line and plane" structure is adopted, in which a single point is specifically represented as an intersection or a road section in the multi-mode traffic large-scale model, which is also the smallest structural unit in road traffic. As a real spatial individual, a single point is usually characterized by limited topological information and relatively low representation dimension. A branch route is formed by coupling several single points together through an ordered topological sequence. Due to the spatio-temporal correlation among multiple single points in a branch route, the traffic state in a branch route will propagate and dissipate in a linear and directed manner. For example, surveillance cameras and millimeter-wave radars can provide different raw data formats, so multi-source data needs to be preprocessed to further pre-match in time and space, so as to construct the dynamic evolution process of traffic.

3.2 Difficulties in the practical application of large models

Although large models have shown great ability in dealing with complex tasks and data, there are still a series of technical, economic, and social difficulties to be overcome in the practical application process. First of all, the technical level is one of the key challenges for the implementation of AI large models. Large models usually require a large amount of computing resources and data to train and optimize, which leads to high deployment costs and energy consumption. In addition, large models may have performance bottlenecks when dealing with specific domains or tasks, which need to be customized and optimized for specific scenarios. At the same time, the security of AI models is also an important issue. How to ensure that the model will not produce harmful output or be maliciously exploited is an urgent problem to be solved.

Secondly, the economic level also poses a certain pressure on the landing application of large models. The development and deployment cost of large models is high, which is a big burden for many enterprises. In addition, as competition intensifies, companies need to constantly invest money and human resources to stay ahead of the technology, which further increases the economic pressure. Therefore, how to reduce the cost of large models and improve their economic benefits is an important driving force to promote their landing application. Finally, challenges at the social level cannot be ignored. The application of large models can have far-reaching effects on society, including job markets, privacy protection, ethics, and so on. For example, large models may replace some traditional occupations and trigger changes in the job market. At the same time, when dealing with personal data, large models may involve privacy leakage, and relevant regulations and regulatory measures need to be strengthened. In addition, the decision making process of large models often lacks transparency, which may trigger ethical and moral controversies. Therefore, in the process of promoting the application of large models, these social factors need to be fully considered to ensure that they are in line with social values and legal and regulatory requirements. To sum up, the practical application of large models faces technical, economic, and social challenges. To overcome these challenges, the paper need to strengthen interdisciplinary cooperation, promote technological innovation and regulation construction, and actively explore business models and cooperation opportunities to jointly promote the healthy development of AI technology.

4. Conclusions

At present, more and more fields are developing and applying large models, and its development and application in various industries has become an irreversible trend. The successful cases of large models such as Chat GPT, Pangu and Hongmeng clearly demonstrate their powerful ability in dealing with complex decision-making tasks. The realization of these achievements is inseparable from the support of massive data, powerful computing resources, and advanced algorithm technology. With the continuous improvement of scientific and technological productivity, these hardware and software resources will become more abundant, laying a solid foundation for the further application of AI large models.

However, despite the great potential of large-scale models in theory, there are still many challenges in practical applications, such as data privacy protection, model interpretation, and computational efficiency. However, with the continuous evolution of The Times and the continuous improvement of technology, these challenges will be gradually overcome, which will promote the real landing and wide application of AI large model technology in the automotive field. This process will further promote the automotive industry to the direction of intelligence and automation, and bring more convenience and innovation to human society.

References

[1] Geman O .DAHP–TOPSIS-Based Channel Decision Model for Co-Operative CR-Enabled Internet on Vehicle (CR-IoV)[J].Sustainability, 2021, 13.DOI:10.3390/su132413966.
[2] Zhang Y , Gao M , Zheng R ,et al.Self-Controlled In-Vehicle Traffic Light Based on The 4th Generation Mobile Communication[C]//2021 IEEE International Conference on Mechatronics (ICM).IEEE, 2021.DOI:10.1109/ICM46511.2021.9385614.
[3] Wang Z , Xu A , Zhang Y ,et al.Research on Electric Vehicle Charging Scheduling Strategy Based on Graph Model[J].Journal of Physics Conference Series, 2020, 1673:012063.DOI:10.1088/1742-6596/1673/1/012063.
[4] Zeng,Qiang,Wen,et al.The interactive effect on injury severity of driver-vehicle units in two-vehicle crashes.[J].Journal of Safety Research, 2016.DOI:10.1016/j.jsr.2016.10.005.

ISSN 2706-655X Vol.6, Issue 4: 25-30, DOI: 10.25236/IJFET.2024.060405

[5] Matsumoto S , Saito M .Adaptive Identification Method for Vehicle Driving Model Capable of Driving with Large Acceleration Changes and Steering[J].Journal of Advanced Computatioanl Intelligence and Intelligent Informatics, 2023, 27(4 TN.163):609-615.
[6] Li X , Epureanu B I .AI-based competition of autonomous vehicle fleets with application to fleet modularity[J].European Journal of Operational Research, 2020, 287.DOI:10.1016/j.ejor.2020.05.020.
[7] Shimada T , Nishigaki T , Onoda T .Prediction of Departure and Travel Time of Individual Vehicle Based on Past Usage Patterns by Using Support Vector Machine[J].Proceedings of the Annual Conference of JSAI, 2020:4I2GS204-4I2GS204.DOI:10.11517/pjsai.JSAI2020.0_4I2GS204.
[8] Zhu A H , Fu C , Yang J W ,et al.Research on the wheel wear of metro vehicles based on the time-varying passenger flow[J].Industrial Lubrication and Tribology, 2019, ahead-of-print(ahead-of-print).DOI:10.1108/ILT-01-2019-0014.
[9] Fridman L , Brown D E , Glazer M ,et al.MIT Advanced Vehicle Technology Study: Large-Scale Naturalistic Driving Study of Driver Behavior and Interaction with Automation[J].IEEE Access, 2017.DOI:10.1109/ACCESS.2019.2926040.
[10] Bourdine A V , Bukashkin S A , Buzov A V ,et al.Concept of distributed corporative wireless vehicle voice networks based on radio-over-fiber technique[C]//XIII International Scientific & Technical Conference on Optical Technologies in Telecommunications.2016.DOI:10.1117/12.2234577.