数据驱动的端到端自动驾驶技术:演进历程、现状与未来挑战*

王瑶　白文静　裴世康　赵文博　徐月云

(国汽(北京)智能网联汽车研究院有限公司,北京 100176)

【摘要】系统梳理了自动驾驶技术自20世纪20年代至2024年的发展历程,分阶段剖析了技术起源与萌芽期、理论奠基与探索期、技术实践突破期、快速发展与范式转型期4个发展阶段及各阶段特点,并聚焦于感知-决策-控制模块化范式、端到端一体化范式、车路云一体化的协同与竞争。结果表明,当前自动驾驶技术呈现"模块化主导、端到端突破"的双轨发展态势,数据闭环与车路云一体化成为解决长尾场景与提升系统泛化能力的关键路径,然而,数据质量不足、算法可解释性欠缺、算力需求激增及政策适配滞后仍是主要挑战。由此可预测,大模型赋能的认知推理、神经符号混合架构、轻量化部署及车路云深度协同等技术将推动自动驾驶向安全、高效、普惠的智能出行生态转型。

关键词:端到端自动驾驶　模块化自动驾驶　数据闭环　车路云一体化

中图分类号:U463.6　　**文献标志码:**A　　**DOI:** 10.20104/j.cnki.1674-6546.20250143

Data-Driven End-to-End Autonomous Driving Technology: Evolution, Current Status, and Future Challenges

Wang Yao, Bai Wenjing, Pei Shikang, Zhao Wenbo, Xu Yueyun

(China Intelligent and Connected Vehicles (Beijing) Research Institute Co., Ltd., Beijing 100176)

【**Abstract**】This paper systematically reviews the developmental trajectory of autonomous driving systems from the 1920s to 2024, analyzing four key phases' s characteristics—technological origins and infancy, theoretical foundations and exploration, practical breakthroughs, and rapid development with paradigm shifts. It focuses on the modular "perception-decision-control" driving, the end-to-end driving and the synergy and competition within the vehicle-road-cloud collaborative. Research indicates that the current autonomous driving landscape exhibits a "dual-track development pattern dominated by modular driving yet driven by end-to-end driving breakthroughs", where data closed-loops and vehicle-road-cloud integration serve as critical pathways to address long-tail scenarios and enhance system generalization. However, insufficient data quality, lack of algorithmic interpretability, surging computational demands and policy adaptation delays remain the main challenges. Predictions suggest that advancements in large model-powered cognitive reasoning, neuro-symbolic hybrid architectures, lightweight deployment, and deep vehicle-road-cloud collaboration will propel autonomous driving toward a safer, more efficient, and inclusive intelligent mobility ecosystem. This study provides a systematic reference for understanding the technological landscape, optimizing development strategies, and formulating industry policies.

Key words: End to end autonomous driving, Modular autonomous driving, Data closed-loops, Vehicle-road-cloud integration

【引用格式】王瑶, 白文静, 裴世康, 等. 数据驱动的端到端自动驾驶技术:演进历程、现状与未来挑战[J]. 汽车工程师, 2026(5): 1-19+33.

WANG Y, BAI W J, PEI S K, et al. Data-Driven End-to-End Autonomous Driving Technology: Evolution, Current Status, and Future Challenges[J]. Automotive Engineer, 2026(5): 1-19+33.

*基金项目:国家重点研发计划项目(2022YFB2503401)。

1 前言

自动驾驶系统（Autonomous Driving System，ADS）作为人工智能与交通系统融合的重要成果，早在20世纪初已现雏形[1]，持续引发了全球范围的广泛关注与深入研究。2016年，国际自动机工程师学会（SAE International）标准SAE J3016_201609《与道路机动车辆驾驶自动化系统有关的术语的分类和定义》将驾驶自动化（Driving Automation）分为0级（L0）~5级（L5）6个等级，即无自动化到完全自动化。

现阶段汽车自动化驾驶技术成熟度以L2及以下等级为主，难以满足L3级场景的安全需求。因此，根据SAE J3016、GB/T 40429—2021《汽车驾驶自动化分级》和行业共识，本文认为，无论是车辆的特定驾驶场景（如示范区自动驾驶出租车（Robotaxi）、矿山、港口），还是全域驾驶场景，需要人类驾驶员接手的驾驶系统均为驾驶辅助系统，人类驾驶员可完全脱手的驾驶系统则称为自动驾驶系统。二者核心内容、升级思路与法律责任不同：驾驶辅助系统的核心是环境感知，旨在帮助人类驾驶员提高车辆驾驶安全性与舒适性，自动驾驶系统的核心是人工智能（Artificial Intelligence，AI），旨在使机器执行人类驾驶员的任务；驾驶辅助系统的升级思路是功能开发，在某个特定场景范围内实现功能，自动驾驶系统的升级思路是驾驶能力提升至人类驾驶员的水平；自动驾驶意味着自动驾驶任务的主要责任从人类驾驶员转移到自动驾驶系统，制造商或运营单位必须对自动驾驶系统的一切行为承担法律责任。辅助驾驶是自动驾驶发展的必要阶段，以最终实现“车辆完全自主决策”的自动驾驶为目标。

汽车驾驶自动化技术的发展经历了由基础控制、视觉导航、多传感器融合、人工规则驱动到当前数据驱动的端到端智能体与大模型驱动的演化，但关于其系统性技术脉络与关键转折点的研究相对稀缺。

因此，本文从历史和技术两个维度梳理自动驾驶系统自20世纪初至今的主要演进历程，最终聚焦于数据驱动范式下端到端自动驾驶技术的突破，包括无线遥控、电磁引导、视觉导航和即时定位与建图（Simultaneous Localization And Mapping，SLAM）的关键阶段，系统总结控制理论（如卡尔曼滤波、动态规划、模型预测控制等）与AI算法（如深度神经网络、强化学习、视觉语言模型、视觉-语言-动作模型等）对技术体系的影响，并探讨当前自动驾驶系统数据闭环、车路云一体化智能网联汽车及二者在解决“长尾问题”、实现系统泛化中的核心作用，厘清自动驾驶系统的发展现状与挑战。最后，对未来基于大模型驱动的自动驾驶系统和ADS的可解释性、安全性和智能化提出展望和建设性建议。

2 端到端自动驾驶技术发展历程

2.1 技术起源与萌芽期（1920~1950年代）

自动驾驶系统的早期探索可追溯至1925年霍迪纳无线电控制（Houdina Radio Control）公司展示的一款名为“美国奇迹”（American Wonder）的无线电遥控汽车，通过后车发射的无线电信号控制转向盘、制动器和加速器，实现了对车辆的远程操作。这一事件被广泛认为是自动驾驶技术的起点[2]。随后，1939年，通用汽车展示了Futurama概念车，通过地面埋设的导线生成磁场指引车辆行驶，这是自动驾驶系统的早期构想之一。这一项目本质上是一种基于特定基础设施的自动化方案，凸显了“受控环境”下自动驾驶系统的初步设想[3]。Futurama项目实际上是对霍迪纳公司技术的商业化延伸，将实验室中的遥控技术转化为对未来交通系统的设想。1942年，德国研发了歌利亚（Goliath）遥控炸弹车，通过有线或无线电控制执行排雷任务[4]，其核心技术包括远程操控系统和路径跟踪算法，尽管未完全实现自主控制，但其为军事自动化提供了宝贵的实践经验。歌利亚的遥控技术为战后民用自动驾驶系统研究提供了技术原型，被认为是现代无人地面车辆（Unmanned Ground Vehicle，UGV）技术的先驱之一，其设计理念对后续遥控武器系统的发展起到了基础性作用[5]。类似地，1953年，巴雷特电子（Barrett Electronics）公司开发的电磁感应引导系统通过地面埋设的导线传输电流生成磁场，车辆通过感应磁场实现路径跟踪[6]。该系统最初应用于仓库货物运输，是工业自动化物流领域的重要里程碑，为后续自动导引车（Automated Guided Vehicle，AGV）技术奠定了基础[7]。

除单车自动驾驶系统研究外，1958年，通用汽车公司与美国无线电公司（Radio Corporation of America，RCA）合作，组装了一套车辆侦测与引导系统，并在内布拉斯加州林肯市郊区的400英尺（约合121.92 m）长的改造高速公路上进行了测试，成功实

现了前后车距保持与自动转向功能。此后，俄亥俄州立大学通信与控制系统实验室于1966年启动了无人驾驶汽车项目，英国交通与道路研究实验室于20世纪90年代开展了无人驾驶测试，美国也在建设试验性电子控制高速公路，英国运输与道路研究实验室进行了车辆试验等[2]。这些技术发展思路均以霍迪纳公司的技术为基础。

20世纪20~50年代的前沿探索者均属依赖预置轨道或磁性标记导航的“地面轨道”，其核心特征是基于电缆、金属传感器和磁场感应器等有线基础设施形成控制环境，如表1所示，从而实现车辆的自动启动、加速、转向与停止等被动控制。这一时期的自动驾驶系统大多未完全脱离人工控制或环境约束，更多地依赖外部硬件或基础设施来引导车辆。因此，早期的自动驾驶系统研究更多侧重于“自动化引导”，而非“自主决策”。

表1　自动驾驶技术起源里程碑性事件

事件	时间	关键人物	里程碑性质
无线电遥控汽车展示	1925年	Francis P. Houdin	远程无线电控制车辆，自动驾驶的开端
Futurama概念车诞生	1939年	Rudolf E. Kálmán	无线电遥控汽车的商业化延伸
遥控炸弹车应用	1942年		无线电遥控车的军事应用
电磁感应引导系统开发	1953年	Barrett Electronics	车辆通过感应磁场实现路径跟踪，AGV技术发展的起源
自动化高速公路建设	1950年代		智能化高速公路的初实践

2.2　理论奠基与探索期（1960~1980年代）

2.2.1　自动驾驶理论奠基

1960~1980年代，自动驾驶系统领域的大量理论基础涌现，如表2所示。1960年代初，基于状态方程的数学模型，以及以最优控制（Optimal Control）理论和卡尔曼滤波（Kalman Filter）为核心的新型控制系统分析与设计方法逐步建立。最优控制理论的核心是最大值原理，该原理通过数学优化方法确定系统的最优控制策略。在自动驾驶系统的早期研究中，最优控制理论主要用于车辆轨迹规划与能耗优化。例如，松弛控制理论用于处理路径规划中的非线性问题，并结合动态模型实现了多目标（如路径长度、曲率等）优化。

表2　1960~1980年代的自动驾驶技术理论基础

理论	核心贡献时间	关键人物/机构	自动驾驶应用场景
最优控制	1960~1970年代	Pontryagin、黄琳	轨迹规划、能耗优化
卡尔曼滤波	1960年代	Rudolf E. Kálmán	状态估计、定位
动态规划	1960~1980年代	Richard Bellman	路径规划、多目标优化
模糊逻辑控制	1980年代	Lotfi Zadeh	非线性控制、决策系统
MAS理论	1980~1981年	Reid G. Smith、Robert M. Ma	多车协同
多智能体系统	1980年代	Vicsek、Borkar/Varaiya	多车协同、交通流优化
视觉导航理论	1987~1994年	Ernst Dickmann	动态视觉模型
PID控制	1970~1980年代	Ziegler–Nichols	速度跟踪、转向控制
神经网络	1991年	Dean A. Pomerleau	端到端驾驶控制

卡尔曼滤波于1960年提出，是一种基于贝叶斯估计的线性高斯模型预测–校正算法，广泛应用于动态系统的状态估计。卡尔曼滤波在自动驾驶系统中的重要应用包括车辆定位与运动状态估计，如通过融合速度、加速度等数据实时预测车辆位置，从而辅助动态环境下的避障决策[8]。此外，卡尔曼滤波还在背景建模与运动目标检测中发挥了重要作用，为环境感知技术的进一步发展奠定了基础[9]。

1961年，动态规划（Dynamic Programming）被应用于航天器制导中的最优路径规划问题，通过问题分解和逐阶段决策策略，为路径优化提供了系统性解决方案。动态规划的优势在于处理带约束的最优路径问题，尤其对复杂环境下的自动驾驶系统路径规划具有重要意义。该技术推动了交通、物流、机器人等领域的路径优化应用，特别是在自动驾驶路径规划中的速度解耦、多目标优化与资源分配方面发挥了关键作用。

1965年提出的模糊逻辑控制（Fuzzy Logic Control，FLC）理论通过语言规则（如“快速减速”“轻微转向”）模拟人类决策，处理不确定性系统问题[10]。

模糊逻辑控制在1980年代广泛应用于车辆轨迹跟踪与鲁棒性控制。同期，基于任务分配的核心协议和协作演化的多智能体系统（Multi-Agent System，MAS）理论框架逐步成型，强调分布式智能体的协作与通信，对早期的多车协同研究具有重要影响。例如，通过分布式一致性算法实现车队的速度同步[11]。与此同时，模型预测控制（Model Predictive Control，MPC）算法、比例-积分-微分（Proportional-Integral-Derivative，PID）控制理论为后期自动驾驶系统框架和工业应用提供了理论支撑。

1980年代是自动驾驶系统从“地面轨道”向“视觉导航”转型的关键时期。大卫·马尔（David Marr）于1982年提出的视觉计算理论构建了利用二维图像重建三维场景的框架，为后续立体视觉与深度感知奠定了基础[12]。1987年，德国慕尼黑联邦国防军大学的恩斯特·迪克曼斯（Ernst Dickmanns）团队提出“动态视觉”（Dynamic Vision）概念，研发了基于视觉的自动驾驶载货汽车，并首次将卡尔曼滤波器应用于道路边界跟踪与车辆状态估计[13]。次年，他们结合实时图像处理与车辆控制验证了单目摄像头在动态场景下的可行性[14]。团队基于动态视觉开发的机器人载货汽车实现了完全自主驾驶，最高车速达到60 km/h，超越了同期美国NavLab（Navigational Laboratory）项目的30 km/h。此外，该团队研发的“4D视觉模型”解决了高速行驶中的实时视觉处理问题，为后续高速自动驾驶系统奠定了基础[15-17]。同期，梅赛德斯-奔驰开发的VaMoRs原型车基于视觉感知与实时决策算法首次实现了高速公路场景下的车道保持功能，验证了人工智能在高速自动驾驶系统中的重要作用[18]。

1991年，美国卡内基梅隆大学的神经网络自主陆地车辆（Autonomous Land Vehicle In a Neural Network，ALVINN）项目首次提出端到端神经网络架构，通过直接从摄像头输入图像生成转向指令，突破了传统基于规则或分模块的自动驾驶系统的设计方式，成为现代深度学习方法的先驱[19-20]。

2.2.2　自动驾驶理论应用探索

随着计算机技术与传感器技术的进步，多个国家开始尝试将自动驾驶系统应用于车辆控制，并在1960~1980年代探索自动驾驶的初步雏形，如表3所示。例如：1970年代日本丰田首次将毫米波雷达应用于试验性车辆，检测前方车辆的相对速度和距离[21]；1975年，德国博世公司尝试使用红外传感器监控载货汽车的盲区[22]。

表3　1960~1980年代的自动驾驶实践探索

项目	时间	关键人物/机构	探索成果
Shakey机器人	1966年	美国斯坦福研究所	分层控制框架
斯坦福车	1960年代	美国斯坦福大学	形成“感知-规划-控制”的三阶段流程
车辆自主导航	1979年	日本东京机械工程实验室	全球首个车道跟随自动驾驶系统
NavLab项目	1984年	美国卡内基梅隆大学	多传感器融合
车辆自主驾驶	1980年代	德国慕尼黑大学	机器人载货汽车完全自主驾驶
VaMoRs原型车	1980年代	德国航空航天工程师、梅赛德斯-奔驰	首次实现高速公路场景下的车道保持功能
中国智能小车	1989年	中国国防科技大学	中国早期智能汽车的代表性成果之一

1966年，首个自动驾驶系统模块化开发思路的探索出现。斯坦福研究所的Shakey机器人采用分层架构设计，成为首个自主机器人。其核心设计包括计划执行层、符号规划系统层、中间动作层和低级动作层，通过马尔可夫表与三角表结构实现稳健执行，启发了后续T-R程序和分层状态机技术（如SMACH库）[23-25]。Shakey的规划系统斯坦福研究所问题求解器（Stanford Research Institute Problem Solver，STRIPS）与A*算法生成路径的创新，推动了模块化架构在自动驾驶中的应用。基于Shakey的设计，斯坦福大学进一步通过“斯坦福车”将感知、决策与控制整合到具体系统中，形成了“感知-规划-控制”的三阶段流程[26]，推动了自动驾驶的系统化发展。

动态视觉探索方面，早在Ernst Dickmanns团队之前，日本在1970年代便开始尝试通过摄像头与图像处理技术实现车辆导航。1979年，由Sadayuki Tsugawa领导的团队在东京机械工程实验室开发了全球首个车道跟随自动驾驶系统，该研究被视为早期自动驾驶系统的重要里程碑。该系统通过双摄像头实时捕捉道路两侧的护栏图像，并通过图像处理技术实现车辆定位与路径规划。与现代的SLAM技术不同，早期系统依赖于预存磁带数据辅助进行实时图像分析，导致环境感知与决策分离，车辆采

用半自主导航[13]。尽管如此，该研究为SLAM技术的发展提供了重要启示。

1984年，卡内基梅隆大学的NavLab项目开创了多传感器融合技术的先河。NavLab项目是美国国防部高级研究计划局（Defense Advanced Research Projects Agency，DARPA）自主陆地车辆（Autonomous Land Vehicle，ALV）计划的核心项目之一，其从单一摄像头扩展到多传感器融合平台，结合视觉传感器（多摄像头）、激光雷达、GPS、惯性导航系统等，为自动驾驶系统领域的多传感器融合技术奠定了基础[27]。同一时期，DARPA的ALV项目在多传感器探索、立体视觉感知、地形建模与路径规划等关键技术方面取得了重要突破[28-31]。

1989年，中国国防科技大学研制出了首辆智能小车，配备了计算机、精确定位、路径规划、运动控制及辅助驾驶仪等系统，是中国早期智能汽车的代表性成果之一[32]。

这一时期的自动驾驶系统发展以理论创新与工程验证为核心：在理论层，建立了最优控制、卡尔曼滤波、动态规划、分层架构、神经网络、模糊逻辑等；在实践层，通过原型测试验证了视觉导航、高速控制的可行性、多传感器融合与跨领域技术整合。这些成果为理论提供了可行的工程实现支撑，也为之后的自动驾驶系统爆发式发展奠定了架构和技术基础。

2.3 技术实践突破期（1990~2000年代）

1990~2000年代，自动驾驶系统领域在传感器技术、神经网络、车路协同、仿真方面迎来了革命性突破，如表4所示。

表4 1990~2000年代的自动驾驶实践突破

突破性项目	时间	关键人物/机构	突破性成果
车载摄像头应用	1990年代		量产车配置引入摄像头，奠定自动驾驶系统发展的基础
毫米波雷达应用	1990年代		毫米波雷达被集成至自动驾驶系统
VITA-1、VITA-2	1994年	奔驰	动态视觉系统在公开道路的验证
NavLab5	1995年	卡内基梅隆大学	开放道路中首次实现端到端自主驾驶，数据驱动范式的首次验证
ARGO	1997年	帕尔马大学	双目视觉系统在公开道路的验证
PATH项目	1997年		首次实现车路协同的规模化应用
仿真技术	1990年代		代表性交通流仿真软件PTV Vissim、Paramics、SUMO；车辆动力学模型CarSim、IPG CarMaker
DARPA挑战赛	2004~2007年	DARPA	多传感器融合与“感知-决策-控制”标准技术栈的形成，完成自动驾驶系统实验室原型向工程化系统的转变；激光雷达工程化应用起点

2.3.1 传感器技术突破

1991年，丰田Soarer成为首款搭载车载摄像头的商用汽车。尽管车载摄像头的概念最早可追溯至1956年别克的Centurion概念车，但由于技术不成熟和成本较高，该功能未能实现商业化应用。直到1991年，丰田Soarer首次将车载摄像头作为量产配置引入市场，标志着该技术的实用化。Soarer搭载的单目视觉系统主要用于环境感知的基础功能，为车载摄像头在自动驾驶系统中的广泛应用奠定了基础[33]。此后，单目视觉技术逐步发展，360°全景环视系统的概念应运而生，标志着摄像头传感器技术进入快速发展及应用阶段。例如，2007年12月，日产发布了全球首款全景环视系统（Around View Monitor，AVM），通过车身周围4个高分辨率超广角摄像头和测距雷达将车身鸟瞰图和侧后方图像实时呈现给驾驶员，为其提供驾驶视角辅助。随后，多个厂商推出了360°环视系统方案，如本田的多视角摄像系统（2008年）、阿尔派的TOPVIEW系统（2009年）和富士通的多角度视野（Multi-Angle Vision）系统（2010年）。

随着毫米波射频芯片技术的发展，雷达体积显著缩小，Ernst Dickmanns的研究结合毫米波雷达、超声波雷达和视觉传感器实现了车道保持功能[34]。1990年代，毫米波雷达开始集成到车辆的前向碰撞预警系统中。日本车企率先推出基于24 GHz雷达的自适应巡航（Adaptive Cruise Control，ACC）系统，通过实时测距实现车速自动调节。与此同时，毫米波雷达在恶劣天气条件下（如雨雪、雾霾等）表现出的优势逐渐凸显，2000年代初，得益于77 GHz雷达优于光学传感器的穿透能力[35]，欧洲车企将其应用于自动紧急制动（Autonomous Emergency Braking，AEB）系统。

激光雷达在1990年代主要应用于航空航天与

测绘等非车载领域，其高精度和主动探测能力为自动驾驶系统的应用提供了基础[36]。例如，1992年三菱Debonair的激光雷达预警系统、1995年三菱Diamante的ACC系统及1999年奔驰S级的Distronic系统都标志着激光雷达在自动驾驶系统中的初步应用，但DARPA挑战赛成为激光雷达在自动驾驶系统领域突破的真正催化剂。

2.3.2 神经网络突破

开放道路试验促进了自动驾驶系统的发展。1994年，基于Ernst Dickmanns团队的动态视觉系统成果，奔驰搭建了VITA-1和VITA-2无人驾驶原型车，并进行了大量公路测试[37]。尽管仍有工程师坐在驾驶员位置，但车辆的自动驾驶能力已初步实现，包括车道跟踪、障碍物避让及自动超车等。1995年，卡内基梅隆大学的NavLab5完成了“无手穿越美国”试验，标志着神经网络在开放道路上实现端到端自主驾驶，突破了传统的“感知-决策-控制”模块化架构。该试验中，车辆配备了视觉摄像头、GPS、陀螺仪和转向编码器，通过神经网络控制转向盘，验证了数据驱动方法在自动驾驶系统中的应用可行性[38]。这一成果推动了模块化架构与学习方法相结合的混合模式发展。1997年，意大利帕尔马大学的Alberto Broggi团队提出了基于双目立体视觉的一般障碍物和车道检测（Generic Obstacle and Lane Detection，GOLD）算法，并在ARGO原型车中实现应用。ARGO系统能够实时处理道路障碍物和车道线检测数据，并在1998年完成了2 000 km的高速公路测试，实现94%的无人驾驶里程，且在复杂路况下的检测准确性显著提升[39]，为后续视觉算法的发展提供了重要参考。

2.3.3 车路协同突破

在单车自动驾驶系统发展的同时，1997年，美国加州先进公共交通与公路研究合作伙伴（Partners for Advanced Transit and Highways，PATH）项目启动了“自动公路系统”，首次实现了车路协同的规模化应用[40]。该系统基于车路协同架构，通过车辆间（Vehicle to Vehicle，V2V）通信和车辆与基础设施（Vehicle to Infrastructure，V2I）通信实现编队行驶、协同避障与动态路径规划，为智能网联汽车（Intelligent and Connect Vehicle，ICV）和协同自适应巡航控制技术的后续发展奠定了基础[41]。

2.3.4 仿真技术突破

自动驾驶仿真一直是自动驾驶技术研发与验证体系中不可或缺的重要环节。它通过构建虚拟的道路场景、车辆动力学模型以及传感器模型感知环境，实现各种复杂工况下自动驾驶系统功能和安全性能的评估与验证。在2000年之前，自动驾驶仿真雏形出现，主要集中在交通流仿真和车辆动力学模型仿真领域。此阶段的仿真技术重点在于基于物理规则的建模方法，以简化的模型模拟交通流及车辆行为[42]。受计算能力限制，这些早期技术所需计算资源较少，在模拟的精确性和多样性上存在不足。代表性交通流仿真软件包括PTV Vissim（1992）、Paramics（1994）和SUMO（开源、可扩展）。这些工具主要用于研究交通系统的流量特性、优化交通信号，以及为交通政策制定提供支持。交通流仿真侧重于车辆交互及宏观交通模式。车辆动力学模型如CarSim和IPG CarMaker（均创建于1990年代），着眼于模拟单辆车的动力学行为。这些工具用于模拟车辆在不同操控条件下的响应，帮助工程师在开发阶段进行车辆动力学参数的优化。此时的自动驾驶仿真依赖真实数据且场景单一，未涉及感知与决策模块。

综合来看，2004~2007年，DARPA挑战赛推动了现代自动驾驶技术体系的形成。通过真实场景验证，DARPA挑战赛验证了模块化架构的可行性，促进了学术界和工业界将实验室算法转化为工程化系统。斯坦福大学的Stanley[43]（2005年冠军）和卡内基梅隆大学的Boss[44]（2007年冠军）代表了两种典型的模块化架构范式。Stanley采用“感知-决策-控制”分层架构，通过多传感器（激光雷达、立体视觉和GPS）融合技术构建高精度环境地图，验证了多模态数据融合在应对复杂沙漠地形中的关键作用[45]。决策层利用概率图模型（马尔可夫决策过程）处理动态路径规划，结合机器学习技术提升了自主导航的实时性[46]。控制层则开发了鲁棒性强的非线性控制器，实现了精确的轨迹跟踪[47]。卡内基梅隆大学的Boss系统创新性地采用行为树模型，将驾驶任务分解为“变道-超车-停车”等原子行为节点，进一步推动了分层规划与行为树技术的发展。

DARPA挑战赛标志着自动驾驶从实验室原型向工程化系统的转变，其模块化设计理念至今仍为行业主流架构。此阶段自动驾驶系统领域传感器革命与竞赛驱动并行，以“算法智能化、架构模块化、场景复杂化”为核心，机器学习首次承担感知任务，自动驾驶系统领域出现了革命性应用。车路协

同与车联网（Vehicle to Everything，V2X）技术的雏形、决策算法的智能化升级以及多传感器融合技术的工程化验证，为2010年之后的自动驾驶系统商业化浪潮奠定了基础。

2.4 快速发展与范式转型期（2010年之后）

2.4.1 模块化自动驾驶范式确立

在21世纪初期多项重大技术取得突破后，自动驾驶系统模块化范式迅速发展，并实现了驾驶辅助系统模块化范式的产业化应用。

2012年，谷歌发布了配备激光雷达、摄像头和毫米波雷达的自动驾驶原型车，是自动驾驶系统的初步量产尝试。该原型车在2014年成功完成了超过100×10^4英里（$160.934\ 4\times10^4$ km）的测试里程[48]，突破性验证了多传感器融合架构量产的可靠性，标志着该架构可实现工程化生产。2013年，汽车开放系统架构（AUTomotive Open System ARchitecture，AUTOSAR）规范提出了层次化、模块化的汽车电子架构，强调通过功能分离和接口标准化实现硬件的可扩展性。对于自动驾驶系统而言，AUTOSAR为其“感知-决策-控制”功能模块提供了底层支撑平台，确保通过标准化接口、通信协议和安全机制实现各功能模块的高效协同与可靠运行。AUTOSAR为模块化设计的工业化应用奠定了理论基础[49]。

自动驾驶系统模块化范式主要包括感知、决策控制3个模块[50]。感知模块负责从环境中收集信息，识别并理解周围的物体和情况，如行人、其他车辆、交通标志等。决策模块负责根据感知模块提供的信息作出驾驶行为的规划和决策，如选择车速、转向、加速、制动等。控制模块负责将决策结果转化为具体的驾驶行为，如控制车辆的转向、加速和制动。

随着自动驾驶系统模块化范式逐渐明晰，神经网络成为主流的技术手段。

在感知（Perception）模块，深度神经网络（Deep Neural Networks，DNN）和卷积神经网络（Convolutional Neural Networks，CNN）发挥重要作用。CNN能够处理来自摄像头的图像数据，进行实时物体检测和分类[51-52]。通过训练大量图像数据，神经网络能够从视觉传感器（如摄像头）中识别物体、障碍物、道路标线、交通标志等。同时，CNN及其变种（如U-Net）可用于语义分割[53]，将图像中的像素分类，从而帮助自动驾驶系统理解复杂的场景。在处理激光雷达（Light Detection And Ranging，LiDAR）和毫米波雷达的数据时，神经网络也能通过点云数据的聚类、分类等方式帮助自动驾驶系统识别周围环境。

在决策（Decision）模块，神经网络通常采用强化学习、模仿学习等方法进行复杂决策任务的优化。强化学习是一种常用的神经网络训练方法，自动驾驶系统可以通过模拟环境进行训练，使得神经网络根据环境状态（如与周围车辆的距离、交通信号灯状态等）学习如何采取最佳动作。通过与环境交互，系统可以评估不同决策的效果，并通过奖励反馈不断调整策略，以优化整体性能。模仿学习是一种通过观察专家驾驶行为（如人工驾驶员或历史数据）来训练神经网络的方式。神经网络可以学习专家的决策模式，进而应用于实际驾驶中。

在控制（Control）模块，神经网络通常通过控制策略的学习和优化来实现控制。例如，神经网络可以与传统的MPC方法结合，提升车辆的实时控制能力。例如，通过神经网络预测车辆在未来时刻的状态或环境变化，帮助MPC算法作出更准确的控制决策。

随着神经网络算法在自动驾驶中的应用，以及800万像素摄像头、固态激光雷达等高级传感器的普及，车辆环境感知能力迅速提升，驾驶辅助系统模块化范式实现产业化应用。截至目前，众多企业发布了其模块化辅助驾驶系统。例如，福特Co-Pilot 360、奔驰驾驶领航（Drive Pilot）、通用超级巡航（Super/Ultra Cruise）、特斯拉自动辅助驾驶（Autopilot）、沃尔沃领航辅助（Pilot Assist）、比亚迪DiLink、蔚来NIO Pilot、小鹏XPILOT、吉利千里浩瀚智驾系统、广汽ADiGO GSD等。这些系统基本普及了自适应巡航控制、车道保持辅助、交通拥堵辅助功能，并集成巡航辅助、主动变道辅助等基础行车功能，以及智能泊车辅助、地下停车场停车等基础泊车功能。进阶的驾驶辅助系统可实现高速领航辅助、城区领航辅助等复杂的行车功能，以及记忆泊车辅助、代客泊车辅助等复杂的泊车功能。这些功能极大提升了驾驶安全性和便利性，尽管仍依赖驾驶员监督，但显著降低了人为操作失误的风险。

2.4.2 端到端自动驾驶范式出现

2015年前后，自动驾驶范式出现重大转型：一方面，模块化自动驾驶系统普遍依赖于已知数据的监督训练或基于人工经验的规则设计，以应对特定场景需求，面对未知环境时，这类系统的有效性难

以保证，传统模块化自动驾驶系统路径的持续推进面临挑战；另一方面，神经网络虽然在自动驾驶中的子任务上取得了显著成功，但在处理整个自动驾驶系统的驾驶任务时无法高效决策和实时控制，出现了自动驾驶系统复杂性与协调性不足的问题。为解决这些问题，端到端（End-to-End，E2E）学习逐渐成为自动驾驶技术的重要方法。端到端方法通过神经网络将输入（如传感器数据）直接映射到输出（如控制指令），不再需要将问题分解为多个子任务，如表5所示。

表5　端到端自动驾驶范式转型关键突破

范式转型代表	时间	代表性人物/机构	突破性内容
DAVE-2	2016年	Bojarski	端到端范式的原型与演示
OpenPilot	2017年	Comma.ai	首个商业化端到端自动驾驶产品
BEV+ Transformer	2020年	特斯拉	实现多摄像头数据融合
BEV	2021年	特斯拉	实现感知模块的端到端化
BEVFormer	2022年		纯视觉自动驾驶系统新范式
UniAD	2023年	OpenDriveLab	感知、预测、规划模块联合优化设计的端到端
DriveGPT4	2023年		首次将LLM整合到端到端自动驾驶框架中，模型更具解释性
GPT-4V	2023年		VLM使自动驾驶系统的决策更加拟人化，提升系统的安全性与可靠性
EMMA	2024年	Waymo	VLA不仅关注环境建模与决策，还涉及规划与控制

在端到端自动驾驶范式探索中，英伟达Bojarski等于2016年基于AlVINN系统，通过扩展摄像头数量和训练数据规模，提出了端到端驾驶模型DAVE-2。该模型直接将单目摄像头图像映射为转向控制指令，从而支持复杂工况，如砂石路面或夜间驾驶等工况。试验结果证明，仅使用不足100 h的覆盖高速公路、普通道路及多种气象条件（如晴天、多云、雨天）的训练数据即可实现98%的自动驾驶覆盖率[54]。DAVE-2通过端到端全流程优化，彻底取代了传统的“感知-决策-控制”模块化范式，其模型能够自主学习车道线等道路语义特征，标志着自动驾驶系统深度神经网络的突破性应用。自2016年英伟达首次公开端到端自动驾驶深度学习框架以来，业界基于该范式的原型与演示层出不穷，充分验证了端到端方法在真实驾驶环境中的可行性与优越性。2017年，Wayve.AI仅用1天时间训练端到端模型，即可应对复杂城市驾驶场景，展现了端到端技术在快速学习和适应能力方面的潜力。同年，Comma.ai推出了OpenPilot软件，最初为L2级辅助驾驶系统，后逐步发展为端到端神经网络模型[55]，成为首个商业化端到端自动驾驶产品[56]。2019年，特斯拉通过硬件升级推出Autopilot 2.0系统，通过影子模式收集真实驾驶数据，为端到端技术迭代奠定了基础。2020年，特斯拉采用鸟瞰图（Bird's Eye View，BEV）+Transformer架构实现了多摄像头数据融合。2021年，特斯拉又通过多传感器融合实现了感知模块的端到端化，为后续端到端技术的全面应用奠定了基础。2022年，鸟瞰图视角转换方法（Bird's-Eye-View Transformer，BEVFormer）解决了3D感知难题[57]，不仅带动了自动驾驶感知领域的研究热潮，也为纯视觉自动驾驶系统和多模态感知算法的发展提供了新范式。同一期间，理想汽车成为全世界第一个在量产车上实现BEV 3D视觉感知算法和3D激光雷达前融合的企业。2023年，特斯拉正式发布完全自动驾驶（Full Self Driving，FSD）v12版本，采用端到端架构，实现了感知、决策、规划的一体化，显著提升了驾驶体验的拟人化与安全性，推动了全球车企加速端到端技术的研发。同年，OpenDriveLab发布了首个感知决策一体化的自动驾驶通用大模型UniAD[58]。该模型深度依赖神经网络，尤其是Transformer架构和端到端联合优化的设计理念，整合了感知、预测和规划模块，进一步提升了复杂场景下的环境理解与决策能力。

随着计算能力的提高和数据集的扩展，端到端学习模式逐渐应用于大模型开发。端到端学习的核心是通过一个统一的模型直接学习从输入到输出的映射，而大模型提供了实现这一目标的强大计算能力和学习能力。大模型能够更好地进行多模态数据融合以提高系统的感知能力。同时，大模型在处理复杂任务时可提供更强的数据处理能力和泛化能力。当前，自动驾驶领域的典型大模型有大语言模型（Large Language Model，LLM）、视觉语言模

型（Visual Language Models，VLM）、视觉-语言-动作模型（Vision Language Action，VLA）。

LLM通过大规模预训练和微调，将神经网络与自然语言任务紧密结合，具备强大的语言理解与生成能力，旨在实现对人类语言的理解和生成。2023年，DriveGPT4首次将大语言模型整合到端到端自动驾驶框架中，突破了传统的规则或黑箱模型在可解释性上的瓶颈[59]。引入LLM后，端到端系统的泛化能力得到提升，能够更有效处理长尾场景，尤其是应对罕见和复杂的驾驶情境。例如，DiMA（Distilling Multi-modal large language models for Autonomous driving）通过知识蒸馏将多模态LLM的知识转移至视觉规划器，显著降低了长尾场景中的轨迹误差[60]。同时，LLM的零样本学习能力使得系统能够在无标注数据的情况下成功泛化到新的任务和场景。

相较于LLM，VLM将自然语言处理与计算机视觉（Computer Vision，CV）能力结合，突破了文本与视觉信息的界限，实现了多模态数据的融合[61]。在自动驾驶领域，VLM能够识别道路环境、交通标志、行人和车辆等，通过大量训练理解交通场景中的复杂交互事件，从而为自动驾驶系统提供决策支持。VLM的引入使得自动驾驶系统的决策更加拟人化，能够根据当前环境作出合理的驾驶决策，提升了系统的安全性与可靠性。VLM的代表性应用包括GPT-4V在驾驶场景中的视觉问答与理解[62-64]、LLaVA-AD（Large Language and Vision Assistant-Active Deduction）在视觉语言导航和交通标志理解中的应用[65]、BLIP-2（Bootstrapping Language Image Pre-training 2）结合文本描述优化驾驶决策[66]以及Flamingo在复杂驾驶场景中的推理能力[67]等。在国内，理想成为首家将VLM概念引入自动驾驶的车企，其在车端芯片上部署了双系统方案，其中系统1为端到端模型，负责视觉感知并执行操作，系统2为VLM，提供复杂环境理解、导航地图解读和交通规则理解，辅助系统1进行规划决策[68]。

然而，VLM需要在算力占用和功能实现之间进行权衡，VLA在自动驾驶领域的应用显得更加自然。VLA模型最早源于机器人领域。2023年7月，谷歌DeepMind发布了全球首个控制机器人的视觉语言动作模型RT-2[69]。此后，VLA模型概念迅速扩展至智能驾驶领域。VLA在VLM基础上发展而来，能够从传感器数据中提取丰富的环境信息，借助语言模型理解人类指令，并生成可解释的决策过程，最终将多模态信息转化为具体的驾驶操作指令。即与VLM专注于环境建模不同，VLA不仅关注环境建模，还涉及规划与控制。此外，VLA能够理解并响应训练数据集中未出现的命令，基于底层语言模型的思维链进行推理，从而作出合理决策[69]。据行业专家预测，现阶段端到端系统能够推理出未来7 s内的驾驶情境，而基于VLA的端到端系统能够推理至10~30 s，这使得其在应对复杂长尾场景（如施工现场、潮汐车道、交通指挥手势等）时的表现更为出色。此外，VLA系统具备可解释性，能够通过车内屏幕向用户解释“为什么这样驾驶”，增强了用户对自动驾驶系统的信任。VLA使得驾驶决策过程可追溯，弥补了传统端到端系统的“黑箱”问题，提升了可解释性和监管可信度。在国外，Waymo的用于自动驾驶的端到端多模态模型（End-to-end Multimodal Model for Autonomous driving，EMMA）通过VLA理解路牌、导航信息和指挥手势，已在自动驾驶出租车中实现应用[70]。Wayve公司通过LINGO-2结合大规模Transformer架构与端到端的多模态学习能力，展示了如何在无需高精度地图的情况下进行决策。在国内，理想汽车发布MindVLA大模型架构，将车辆从运输工具升级为“具备空间认知与逻辑推理能力的智能体”。通过22亿参数模型（部署于车端Orin-X芯片）实现3D空间理解、逻辑推理与行为生成三大能力跃迁。

2.4.3 仿真测试快速发展

不论是模块化自动驾驶还是端到端自动驾驶范式，面对ADAS功能、ADS的商业化，仿真测试需求均愈加明显。自动驾驶的核心痛点之一是成本，而仿真测试能将大量的自动驾驶开发和测试成本转化为图形处理单元（Graphics Processing Unit，GPU）的物料成本和工程师的知识经验成本，进而大幅缓解该痛点。

仿真场景是整个仿真体系的开端。早期场景构建主要依赖人工规则设计，以标准化测试场景为主，覆盖范围有限，缺乏动态交互性。随着传感器技术的发展，自然驾驶数据成为场景库的主要来源，可通过聚类分析提取典型场景（如高速公路跟车、换道）。当前，利用AI技术合成极端天气、光照等罕见场景的方法更为普遍，以解决真实数据边界场景不足问题。从仿真软件看，受限于计算资源和3D建模技术，2000~2015年间高保真仿真器开发停

滞，仅PreScan、rFpro、VI-grade等获大企业资助的项目留存。2017年，开源生态CARLA发布，通过提供开源数字资产和传感器模拟、支持强化学习训练，成为学界标准平台。同期，AirSim（微软）、LGSVL（LG）推动了高保真感知仿真。当前，LGSVL等软件已兼容Autoware、Apollo，支持端到端测试，亦有NVIDIA DRIVE Sim等支持分布式仿真加速。从仿真硬件来看，2010年前主要针对传感器/控制器单元进行单模块硬件在环（Hardware-in-the-Loop，HiL）测试，如基于LabVIEW构建无人机自动驾驶仪HiL测试。2010~2020年间更侧重于系统级HiL测试与整车在环（Vehicle-in-the-Loop，ViL）测试，通过多传感器信号融合，支持ADAS测试、实车测试、车路协同验证。当前，在数字孪生等技术驱动下，硬件平台可同步模拟驾驶员行为（驾驶员在环）、交通流、环境扰动等。此外，仿真场景、仿真软硬件的拟真度、关联性、标准化的仿真方法与流程等都是仿真评价关注的重点。

3 端到端自动驾驶技术发展现状

3.1 自动驾驶技术发展路线

从技术成熟度角度看，目前自动驾驶领域的成熟模型仍为模块化系统，而端到端+VLM已成为主流热点，VLA则代表了自动驾驶技术的最新发展方向，行业形成了“模块化系统主导、端到端系统突破”的双轨发展态势，如表6所示。

表6 自动驾驶技术范式对比

维度	端到端自动驾驶技术	模块化自动驾驶技术
数据需求	千亿级里程数据，强调整体分布均衡性	分模块优化，数据量级低但需针对性标注（如障碍物检测需高精度3D框数据）
数据利用方式	端到端联合优化，数据驱动全局策略生成	分模块独立训练，依赖人工规则衔接各环节
长尾问题解决	通过数据泛化自主覆盖长尾场景	需人工设计规则或增加特定场景数据
迭代成本	初期数据闭环投入高，后期边际成本递减	单模块优化成本低，但系统级协同成本高

模块化自动驾驶系统的感知层、决策规划层和控制执行层相互独立，设计方案依赖于人工规则，且信息传递过程中容易积累误差。因此，模块化自动驾驶系统主要适用于L0~L2阶段等要求简单、风险可预测的场景。而端到端算法基于数据驱动，能够实现信息的无损传递，系统整体针对最终任务进行优化。这使得端到端系统具备更高的迭代效率、较低的维护成本以及更广泛的应用场景。因此，端到端系统可在驾驶自动化L3、L4、L5级以及复杂场景中发挥重要作用。

当前，端到端系统根据感知模块与决策模块的集成程度可分为“模块化端到端”与“一体化端到端”。模块化端到端将感知与决策规划模块视为独立部分，分别应用神经网络，但模块间依然依赖人工设计的数据接口。这种设计使得各模块可以独立开发和优化，便于针对特定功能进行模块替换升级，增加了系统灵活性与可替换性，且易于定位问题，单一模块故障不影响全局，其参数调优和错误分析更具针对性。但模块化端到端存在模块之间的接口和协调增加系统集成复杂度、多级数据处理可能增加响应时间等缺点。典型的模块化端到端包括感知端到端[71-72]、预测端到端[73-74]、规划端到端[75]。从行业实践看，华为乾崑ADS 3.0由通用障碍物感知（General Obstacle Detection，GOD）网络与预测决策规控（Prediction Decision Planning，PDP）网络组成，GOD网络负责感知障碍物，PDP网络则根据感知信息作出决策。类似地，小鹏AI天玑系统包括感知神经网络XNet、规划控制大模型XPlanner和语言大模型XBrain。特斯拉FSD V12保留感知与规划模块，但采用端到端方法训练路径规划。模块化端到端去除了对高精度地图的依赖，大幅提升了自动驾驶系统的灵活性和适应性。相较之下，一体化端到端将感知与决策控制模块打通，采用统一的深度学习模型，通过直接输入感知数据生成车辆的控制输出，能够更高效、直接地处理信息并作出反应。这种架构通过消除模块划分，减少了信息损失，且对数据标注的需求较低，同时具有更强的泛化能力。但一体化端到端微调难度增加，系统修改和故障分析困难，必要时需要重新训练整个网络。此外，一体化端到端需要更多数据进行训练以涵盖广泛场景，典型代表包括商汤绝影R-UniAD方案、轻舟智航基于地平线单征程6M芯片的端到端方案、Momenta R6飞轮大模型、理想MindVLA、特斯拉FSD V13等。

另外，要推动端到端自动驾驶技术的发展，数据驱动是核心。完全自主决策的自动驾驶一方面

需要现实世界中的极端场景（Corner Case）完全被解决，另一方面需要数据驱动学习的自动驾驶决策系统，从而实现训练数据越多，越能够接近人类的驾驶行为风格，甚至超越人类的决策能力，从而提升驾驶安全性。如特斯拉FSD V12通过全球超过500万辆车的实时行驶数据（覆盖复杂路口、恶劣天气等长尾场景）实现“类人类驾驶”的决策能力，其模型参数量级已突破百亿，远超模块化架构的算法组合。再如，相较于传统模块化架构依赖人工规则应对细分场景（如无保护左转、施工路段绕行），端到端模型通过海量真实数据学习场景间的隐含关联，可自主推导出通用驾驶策略。Waymo的ChauffeurNet端到端模型在未针对“动物闯入”场景进行专项训练的情况下，仍能通过数据泛化生成合理避让动作。上述功能都离不开数据驱动，作为自动驾驶系统学习和改进的基石，数据驱动已经被业界认为是实现自动驾驶的热门方向。

数据驱动的自动驾驶旨在通过大量的数据采集、处理和分析，使自动驾驶系统能够自主学习和适应各种复杂的驾驶场景。因此，数据驱动的端到端自动驾驶技术需构建“数据采集→场景挖掘→自动标注→模型训练→仿真验证”的全链条闭环，从而重新定义自动驾驶的进化范式。

3.2 自动驾驶数据闭环现状

从技术相关性看，自动驾驶数据闭环[76]是当前行业发展端到端自动驾驶技术的攻坚重点，旨在通过持续的数据采集与处理、模型训练与迭代，形成动态优化的技术体系。其核心在于将实际驾驶中产生的数据实时反馈并用于算法优化，解决长尾问题和极端场景的挑战，从而提升系统的可靠性、适应性、实时性、持续优化性。典型的数据闭环流程包括数据采集、数据预处理与标注、模型训练与验证、空中下载（Over-The-Air，OTA）更新。

3.2.1 数据采集

数据采集通过多模态传感器（激光雷达、摄像头、毫米波雷达等）获取道路环境信息，同时记录车辆的姿态、速度、加速度及驾驶者行为数据，涵盖高速、城区及复杂交通场景。以往，多数自动驾驶企业依赖高成本的专用数据采集车辆，配备激光雷达、摄像头、毫米波雷达等传感器。然而，单一企业实现全面的高质量数据采集较为困难，因而出现了车企与第三方机构共建数据平台的合作模式。这种模式降低了成本，但也面临数据兼容性和共享可行性等挑战。目前，车企通过大量量产车进行数据采集，借助特斯拉提出的影子模式，将异常接管、交通违规、危险事故等高价值场景回传云端，以满足训练端到端模型所需的数据量。尽管如此，量产车的合规性和传输成本仍限制了其在模型冷启动时的应用，因此仍需依赖研发型采集车获取基础数据。

随着越来越多的极端场景由“未知”转为“已知”，量产车挖掘极端场景的能力逐渐减弱，难度逐渐提高。研究表明，数据量增加可提升自动驾驶模型的训练效果，但达到一定量后，效果愈发减弱[77]，因此亟需高价值场景的生成方法，如基于仿真的合成引擎、生成对抗网络（Generative Adversarial Network，GAN）与扩散模型、事件驱动模拟（Scenario-based Simulation）、神经辐射场（Neural Radiance Fields，NeRF）三维重建、3D高斯泼溅（3D Gaussian Splatting，3DGS）、世界模型（World Model）[78]等。其中，基于仿真的合成引擎的代表技术，如CARLA、LGSVL Simulator、虚幻引擎（Unreal Engine）、Unity等，通过物理建模、三维建模和脚本控制生成特定交通状况、极端天气、复杂交互等“稀缺高难”场景。GAN与扩散模型的代表技术如风格生成对抗网络（Style Generative Adversarial Network，StyleGAN）、潜在扩散模型（Stable Diffusion），可生成逼真的图像、三维点云，通过条件输入（如特定交通事件）来定制场景。事件驱动模拟以真实事故或极端案例库为基础，抽象出“危险高价值”事件模板，用仿真或AI生成周边环境及参与者行为。NeRF的代表性技术如MARS（Modular And Realistic Simulator）[79]、EmerNeRF[80]，直接利用光线追踪对原始图像进行处理，即NeRF仅需要视觉传感器即可实现高质量的未知视角合成及场景重建[81]，但NeRF计算效率低、内存开销大。3DGS的代表性技术包括GaussianPro[82]、解耦混合高斯分布（Decoupled Hybrid Gaussian Splatting，DHGS）[83]、TrafficBots[84]等，3DGS通过透明度混合（Alpha Blending）实现了高效光栅化与高质量渲染，在计算效率与视觉保真度之间取得了较好的平衡。世界模型的代表性技术包括DriveDreamer[85-86]、GAIA-2（Generative AI for Autonomy）[87]、EMM-AU（Enhanced Multi-Modal Accident video Understanding）[88]，不仅能够合成新场景，还能解释大量的传感器数据、预测潜在的未来场景并弥补信息差距。整体而言，仿真合成引擎、

事件驱动模拟兼具批量性和可控性，是当前自动驾驶高风险长尾场景测试的主力。NeRF、3DGS强调基于真实采集数据的高保真三维重建，适合还原罕见真实事件。世界模型与生成模型（GAN/扩散）更偏向“虚拟想象力”，能补充稀缺数据分布，适合探索未发现的边界情况。以华为乾崑智驾ADS 4为例，采用面向未来自动驾驶时代的世界引擎（World Engine，WE）+世界行为模型（World Action model，WA）架构（WEWA架构），端到端时延可降低50%，通行效率提升20%，紧急制动率降低30%。同时，位于云端的乾崑智驾世界引擎可利用AI生成难例扩散模型，所提供的高质量、高密度的难例场景密度是真实世界的1 000倍，实现用AI训练AI。此外，数据传输、加密、存储及网络优化等技术日益成熟，满足了快速提高的数据要求。

3.2.2　数据预处理与标注

数据预处理通过算法对采集的数据进行校准、时间同步、清洗、降噪和标注，确保数据的一致性和可用性，并提高后续处理效率。由于车端数据来自多种传感器，每种设备输出数据的格式不同，因此数据预处理的首要步骤是将其统一格式化。其中，时间戳需要统一格式以同步不同源的数据，传感器读数需要转换成相同的量纲或单位系统。数据清洗可识别并补全因技术问题或干扰导致的缺失数据。针对处理后的数据，需标记图像或雷达点云中的物体类型和运动状态。通常，采集回传的数据量较大，需经压缩后存储，并根据不同任务需求解析。随着云计算、边缘计算及AI算法的发展，数据处理效率大幅提升。云计算提供强大的存储与计算能力，支持大规模数据存储与复杂模型优化；边缘计算降低了数据传输延迟，提升了数据隐私保护能力；AI算法，尤其是无监督和半监督学习，使得数据清洗和算法迭代更加高效。

自动驾驶数据标注技术早期依赖人工标注静态目标，传统3D目标检测与车道线检测标注依赖单帧点云或图像数据，存在标注成本高、时序一致性弱的问题。激光雷达的普及与端到端技术的崛起催生了多模态时序标注需求。为应对海量点云数据与跨帧一致性难题，行业提出4D-BEV（3D空间+时间维度）自动化标注框架。AI自动标注技术通过机器学习与深度学习自动识别并标记大规模数据集中的对象，显著提高了标注速度与准确性，并减少了人工标注中的错误。4D-BEV标注通过引入时间维度与俯视空间表征，构建了感知标注的新范式：将点云序列转换为BEV空间的4D体素张量（时间×3D空间），通过预训练的3D检测器与多目标跟踪器生成初始物体轨迹。BEV视角与下游规划模块的空间一致性显著提升了标注的实用性。随着LLM与VLM的应用，标注技术也在探索语义理解驱动[89]。例如，利用自然语言指令生成复杂标注（如“标注左侧被部分遮挡的自行车”）。此外，随着端到端自动驾驶技术的发展，现有的包括3D目标检测、车道线检测、路面标志语义分割等感知模块的真值的标注需转向面向一体化算法的标注，如规划标注。

3.2.3　模型训练与验证

模型训练与验证过程通常通过分布式计算集群进行深度学习训练，并通过仿真测试验证模型性能。经过标注的数据将被用于模型训练，过程包括预训练与微调。预训练通过自学习大量无标签数据构建底层模型，微调则根据特定任务优化模型参数，以提高其在新数据上的性能。由于自动驾驶数据量庞大，单机训练效率低，通常采用分布式训练方式。高性能计算（High Performance Computing，HPC）中心可在处理大规模数据和高算力需求方面发挥重要作用，部分车企已建设或规划了自己的高性能计算中心。高性能计算中心不仅可提供强大的计算能力，还能有效降低测试和优化的时间成本。

模型训练完成后使用场景库进行测试，如虚拟仿真测试。仿真测试借助NeRF、3DGS等新兴技术生成丰富的测试场景，以减少成本、提高安全性，并通过XiL（X-in-the-Loop）自动化测试平台进行验证。根据被测对象的不同，测试平台可进行模型在环（Model-in-the-Loop，MiL）、软件在环（Software-in-the-Loop，SiL）、HiL、驾驶员在环（Driver-in-the-Loop，DiL）和ViL等核心测试。自动驾驶仿真平台需兼容车辆动力学、环境感知传感器、交通场景等仿真工具。然而，目前行业内的仿真工具种类繁多、数据格式不统一，且仿真可信度尚需进一步验证。此外，现有的量产算法仍采用传统模块化结构，测试时可针对模块进行仿真，如构造显示模块的输入，接收显示模块的输出进行闭环控制以及结果评价。而端到端系统的仿真则需要重构，端到端的智能驾驶仿真必须从传感器输出数据切入[90]，提供一种置信度高、一致性好、覆盖全面、故障注入便

捷的全链路闭环仿真方案，NeRF、3DGS、世界模型等场景生成技术所构造的仿真环境是进行端到端仿真的有效手段。

3.2.4 OTA更新

数据闭环的最后环节是极端场景挖掘与迭代，即将训练好的模型部署到车端，车端继续收集数据，进行新一轮训练，最终通过OTA升级模型。OTA通过云端系统远程更新车辆电子控制单元（Electronic Control Unit，ECU）软件，实现自动化更新。随着OTA技术的成熟，其应用已从售后升级扩展至更多场景，并呈爆发式增长。高效的云服务为OTA的优化升级提供助力，并推动数据闭环的持续高效运转，以提供优质的终端体验。

3.3 车路云一体化智能网联汽车系统建设现状

无论是自动驾驶“模块化系统主导、端到端系统突破”发展路线，还是自动驾驶数据闭环，均是从单车智能视角出发的。然而，单车智能在视野、视距、视效和信息获取等方面存在一定的“感知局限性”，并面临算力物理极限与安全性威胁等挑战，因此，车路协同作为一种具有可实现性的弥补性解决方案，逐渐成为当前的研究重点。

车路云一体化智能网联汽车系统由智能网联汽车、智能化路侧基础设施、云控基础平台三大主体构成，并通过通信网、相关支撑平台以及其他交通参与者的协同作用，形成一个综合性的智能系统。该系统基于车路云一体化融合的架构，具备分层解耦和跨域共享两大技术特征。单车智能与车路云一体化智能网联汽车并非彼此独立的技术路线，而是相辅相成、相互促进的关系。具体而言，单车智能聚焦于解决局部问题，而车路云一体化则着眼于系统层面的优化，二者相互依托，推动整体智能化水平的提升。在车路云一体化智能网联汽车的应用场景中，车路云一体化典型功能包括：超视距感知与安全预警（如闯红灯预警、交叉口碰撞预警、左转辅助、盲区变道预警、逆向超车预警、紧急制动预警等）；“鬼探头”问题的应对（行人/两轮车快速穿越）；网联自适应巡航控制（如车辆直接获取交通灯信息，顺畅起步与停车）；车车通信（避免二次事故）；L4级低速无人驾驶（适用于机场、矿山、港口等特定场景）。这些场景展示了车路云一体化智能网联汽车在提升道路安全性、优化交通效率以及推动自动驾驶发展中的重要作用。

从车路云一体化智能网联汽车系统的三大主体出发，除单车外，实现高效的路侧全息感知是关键。车路协同应用场景的路端信息来源于路侧感知系统，该系统的精度、可靠性和实时性直接决定了场景的应用效果。为提升路侧感知系统的性能，行业正在研发专用激光雷达、高可信边缘计算单元，并开发覆盖设备层、融合感知层、全域跟踪层和事件检测层的全流程感知算法。云端方面，需要建立智能网联云控基础平台，汇聚来自车端、路端及第三方平台的大量数据，将原始数据转化为服务数据，支持云控服务、运营管理及测试评价等平台应用。云端研究主要集中于云端与边缘端的动态协同，如通过分布式计算优化资源分配，推动云控平台“逻辑协同、物理分散”架构设计，实现算力资源的动态弹性扩展。

此外，蜂窝车联网（Cellular V2X，C-V2X）通信技术是实现实时数据交互的核心，能够有效支持车与车、车与路之间的协同通信。当前，C-V2X研究的重点包括5G/6G网络下低时延通信、高可靠传输协议设计、车-路-云多节点协同的通信资源调度算法，以及大规模设备接入的稳定性等。整体而言，C-V2X通信技术相对成熟，相关技术标准、产品标准和测试标准已趋于完善。同时，基于高精度地图和全球导航卫星系统（Global Navigation Satellite System，GNSS）授时的统一时空底座建设至关重要。GNSS授时为时间基准，高精度地图则作为空间底座。在车端，自动驾驶系统依赖OpenDRIVE地图理解道路静态信息，并利用V2X通信动态感知图层进行协同感知信息的映射；在路端，激光雷达、毫米波雷达和摄像头等传感器需要依靠点云图、瓦片图和矢量图进行标定；在云端，数字孪生系统通过瓦片图和矢量图实现三维场景建模，映射多方获取的动、静态信息。通过“统一图源，多端共用”的方式，可以实现车路云三端空间基准的统一。

纵观全球，美国、欧盟、日本、韩国、中国都在加速C-V2X车端和路侧设施部署，推动智能交通系统建设。各国在车路协同应用场景的设计与开发中，主要聚焦于智能网联汽车的协同感知、协同决策与协同控制需求[91]，通过车端、路端、云端的信息交互，采用特定消息格式，并利用C-V2X通信技术将信息传递至车辆。此过程中，信息通过人机交互界面对驾乘人员进行预警，或通过自动驾驶域控制器影响自动驾驶的感知、决策、规划与控制系统。此外，系统还具备按需定制个性化场景的能力。为此，需要

开发高效的协同决策算法，以解决车路云系统在动态交通场景下的实时响应问题，并推动关键数据集的构建。例如，OPV2V（Open dataset for Perception with V2V communication）数据集作为首个专注于V2V协同感知的模拟数据集，为相关领域的研究奠定了数据基础[92]。V2XSet和V2X-Sim[93]则进一步扩展了对各种V2X场景的研究[94]，涵盖了更复杂的协同感知场景。然而，模拟数据集与实际场景之间的差异促使研究人员开发出更接近现实的数据集。DAIR-V2X[95]和V2X-Real[96]等大型现实世界数据集的出现，推动了协同感知研究的发展。TUMTraf-V2X数据集[97]通过在复杂交叉路口收集数据，并提供精确的GPS和惯性测量单元（Inertial Measurement Unit，IMU）数据标注，进一步促进了多智能体在真实复杂交通场景中有效协作的应用研究。

我国各地区结合智能网联汽车发展状况，依托地区优势、特色资源，到2024年8月，共建设了17个智能网联汽车示范区、7个国家级车联网先导区、16个双智城市、20个车路云一体化试点城市，开放测试示范道路超过22 000 km，发放测试牌照超过5 200张，累计测试总里程8 800×10^4 km。

在车路云一体化架构的支持下，更广泛的数据闭环成为可能，如通过路端、云端为车端算法补充高价值场景数据，云端整合多城市、多路段、多车数据，构建分钟级地图更新与全域交通态势地图，并支持大规模仿真训练。

4 端到端自动驾驶技术发展挑战与展望

端到端自动驾驶是数据驱动的模型，有望实现全场景自动驾驶。与人工智能算法相同，其在数据、算法、算力三个核心要素方面同样面临一系列复杂的挑战。

在数据层面，端到端算法高度依赖于大规模高质量数据集，数据的数量、标注质量与分布特性直接决定了模型训练效果。优质数据集的建设与企业的自研实力、产品综合竞争力、资金投入和智能驾驶车型销量密切相关，这种关联性进一步加剧了车企间的马太效应。与此同时，长尾场景数据的稀缺导致自动驾驶系统在应对特殊情境时仍存在决策层面的挑战。现有的仿真技术尚难以全面复现复杂的动态交互场景，限制了算法验证与优化的能力。未来，数据层面的关键发展方向包括：

a. 加强多模态数据的融合，扩大长期、全球数据的收集规模，提升系统在复杂环境下的感知与决策能力；

b. 提升数据筛选和管理能力，为自动驾驶系统提供高质量数据；

c. 优化现实数据与模拟数据的结合，解决数据分布不均的问题，提高模型的实际应用可靠性；

d. 探索数据合成技术等低成本数据规模扩展方法，提升系统在极端环境下的适应性；

e. 突破车端与路侧视角数据的转换，便于车端融合路侧、云端数据，从而快速获取真实世界高价值场景数据，优化端到端自动驾驶算法。

在算法层面，随着DeepSeek等新型算法的涌现，自动驾驶技术逐步从依赖硬件堆叠与算力竞赛转向通过算法创新提高算力利用效率。端到端模型通过非线性复杂映射直接关联输入、输出，导致决策过程难以直观理解，模型的可解释性与调试优化难度成为其重要挑战，安全性验证也因此受限。未来，算法的发展将聚焦在以下方面：

a. 引入LLM、VLM、VLA等大模型，增强系统的场景理解与推理能力，尤其是在处理长尾事件时；

b. 开发具备可解释性的神经符号混合架构，并结合数字孪生技术实现决策过程可视化，成为满足ISO 21448预期功能安全标准要求的重要方向；

c. 优化能耗与实时性，采用轻量化技术（如知识蒸馏）支持边缘计算环境下的低延迟推理，开发生物启发式架构（如脉冲神经网络与神经形态芯片），以降低功耗、提升实时性；

d. 构建驾驶员意图预测模型，完善L3级驾驶员接管策略设置，提升人机协同效率；

e. 完成数据标注“双轨制”迁移，将传统2D/3D感知标注向专家驾驶轨迹的时空连续性标注、复杂场景的决策边界标注、物理约束的数学表达标注等规划标注迁移，同时，需发展时空一致、连续的自动标注技术（如4D标注）以适应端到端技术的成熟；

f. 探索高效、准确的车路云协同感知、决策、控制算法，补足单车智能的劣势，随着技术的演进，未来自动驾驶算法有望形成“端到端+多模态大模型冗余系统”的双轨并行模式，这一趋势要求智能驾驶团队具备卓越的模型框架设计与快速迭代能力，能够高效地将前沿理论转化为量产技术。

在算力层面，道路交通环境的复杂性和训练数据体量的持续增长，使得自动驾驶模型对算力的需

求不断攀升。尽管专用芯片和固态激光雷达在性能与成本的平衡上取得了进展，但高精度感知与实时决策仍需更强大的硬件支撑。未来，算力的发展可聚焦于建设高性能计算中心提供超级规模算力，以及车端与云端算力的协同提升。例如，部署VLA模型对芯片算力的要求可能需提升至$2\,000\times10^{12}$次/s（英伟达DRIVE Thor级别）。可以预见，在芯片整合与垂直整合能力上具备优势的企业未来将在VLA等大模型应用中抢占先机。

除面临技术层面的挑战外，端到端自动驾驶还要解决自动驾驶汽车生态、商业、政策方面的发展难题。

在生态层面，V2X标准体系建设与数字孪生仿真技术的进步为自动驾驶场景验证与算法优化提供了有力支撑。然而，跨领域协作、标准统一以及数据融合仍是亟待突破的瓶颈。尤其是在高速移动环境中，全路段高精度定位、多源异构数据的实时融合与复杂交通事件识别等问题，制约着系统性能的进一步提升。基于大模型的车云协同正引领智能驾驶步入正向闭环演进模式。通过5G网络，路测数据可实时上传至云端训练平台，优化后的模型反哺车端系统，加速了模型在真实场景下的迭代与进化。未来，车路云一体化的发展应重点围绕以下路径推进：通过典型案例城市示范，推动技术规模化落地；补齐细分技术环节，完善产业链布局；构建高效、低延迟、智能化的数据闭环系统；建立车路云数据共享统一的标准和管理平台，通过数据共享机制协同解决端到端自动驾驶的数据隐私、安全以及利益分配等问题。

在商业层面，随着高阶智能驾驶技术的规模化生产与应用，市场技术路径逐渐分化，成本与政策的博弈日益成为制约商业化落地的核心因素。以零跑汽车为代表的车企正引领10~15万元市场的智能化变革，推动“智驾平权”时代到来。国内亦有如北京亦庄高级别自动驾驶示范区建立的“车路云一体化”试点，实现了路口通信延迟降低至20 ms，单车智能硬件成本下降40%的实际效果。从全球视角看，图森未来在美国实现无人载货汽车日均运营里程超800 km，西井科技在厦门港完成24 h无人集装箱转运，提升了燃油效率并降低了能耗。未来，端到端自动驾驶技术商业发展的关键方向包括：

a. 通过数据共享平台实现多源异构数据交易；

b. 国产方案通过软、硬件协同优化降低算力需求与成本，助力自动化驾驶系统普及至10万元级车型；

c. 车企、科技公司与出行平台形成“算法-制造-运营”闭环（如广汽+滴滴、小马智行+优步），推动数据共享与场景落地；

d. 我国在低成本激光雷达、视觉算法等领域的创新，正在定义全球自动驾驶“新成本曲线”，有望主导国际标准，并在自动驾驶出租车、物流、特种作业等领域率先实现规模化盈利，并推动全球交通生态向智能化加速转型。

在政策层面，国家层面支持力度持续增强，目前已有50余个城市发布自动驾驶相关法规，为技术试点与规模化应用奠定了政策基础。随着《智能网联汽车道路测试与示范应用管理规范（试行）》等政策陆续出台，行业正加速迈向“规模化落地”阶段。“车路云一体化”建设正进入关键期，在提升交通效率、丰富模型训练数据、加快自动驾驶功能落地等方面发挥重要作用。未来，政策上核心着力点聚焦于：解决全国各地自动驾驶测试和运营标准存在差异的问题，以避免影响车路云一体化智能网联汽车系统规模化部署；立法完善责任认定机制，明确L3及以上级别自动驾驶汽车的事故责任划分，避免商业化落地受限；助力构建海内外共性数据库（如政策、路况信息），降低我国汽车出海的全球化风险。

5 结束语

本文对端到端自动驾驶技术的发展历程进行了回顾，梳理出4个关键的历史阶段：1920年代至1950年代的技术起源与萌芽期、1960年代至1980年代的理论奠基与探索期、1990年代至2000年代的技术实践突破期，以及2010年代至今的快速发展与范式转型期。每个阶段都以特定的技术进步和挑战为特征，从最初的遥控车辆到当今复杂的人工智能驱动系统，展现了自动驾驶技术从概念到实践的演变轨迹。技术起源与萌芽期以“地面轨道派”为特征，依赖无线电遥控、电磁引导等预置基础设施实现被动控制，虽未脱离人工干预，但为自动化物流与军事应用提供了雏形。理论奠基与探索期涌现出卡尔曼滤波、动态规划、模糊逻辑控制等核心理论，同时动态视觉与多传感器融合技术取得突破，模块化自动驾驶架构初现雏形。技术实践突破期以传感器革命与车路协同为标志，毫米波雷达、激光雷达的工程化应用与DARPA挑战赛验证了模

块化架构的可行性，数据驱动的端到端自动驾驶范式初现端倪。快速发展与范式转型期实现了模块化自动驾驶架构的量产，同时，以深度学习和大模型为驱动力的端到端自动驾驶新范式通过统一模型实现感知-决策-控制一体化，LLM、VLM与VLA的引入显著提升了系统推理能力与可解释性，车路云一体化则通过协同感知与算力动态分配突破单车智能的物理局限。

本文指出，自动驾驶技术当前呈现“模块化主导、端到端突破”的双轨发展态势。模块化架构凭借其可解释性与较高的工程成熟度主导驾驶自动化等级L2及以下场景，而端到端架构通过数据闭环与多模态大模型赋能驾驶自动化等级L3及以上的复杂场景，二者在功能互补中推动技术迭代。数据闭环通过高价值场景生成与数据合成技术优化长尾问题，车路云一体化则从系统层面提升安全性、效率与泛化能力。然而，数据质量不足、算法黑箱化、算力需求激增及政策适配滞后仍是当前面临的主要挑战。未来，自动驾驶技术的发展将依赖于大模型赋能的认知推理、硬件与算法的协同优化，以及法规框架的适应性调整。模块化与端到端架构将在不同应用场景中形成互补，大模型驱动的推理能力、生物启发式架构的能效优化，以及车路云一体化体系的完善，将共同推动智能交通系统向安全、高效、可持续方向转型。

自动驾驶的最终竞争，不仅是技术的比拼，更是对安全性、商业逻辑与社会接受度的综合考验。企业需在硬件成本控制、算法鲁棒性提升与政策合规性之间实现动态平衡，而用户的实际体验与市场选择将最终决定行业格局。期待不远的将来，L4级自动驾驶的规模化商用与L5级原型技术的突破将共同书写智能出行革命的新篇章。

参 考 文 献

[1] GANDIA R M, ANTONIALLI F, CAVAZZA B H, et al. Autonomous Vehicles: Scientometric and Bibliometric Review[J]. Transport Reviews, 2019, 39(1): 9-28.

[2] BIMBRAW K. Autonomous Cars: Past, Present and Future a Review of the Developments in the Last Century, the Present Scenario and the Expected Future of Autonomous Vehicle Technology[C]// 2015 12th International Conference on Informatics in Control, Automation and Robotics (ICINCO). Colmar, France: IEEE, 2015: 191-198.

[3] ERJAVEC J. Automotive Technology[M]. Boston, Massachusetts, USA: Thomson Delmar Learning, 2005.

[4] SINGER P W. Wired for War[M]. London, UK: Penguin Press, 2009.

[5] KUSHWAHA J, GUPTA A, VISHWAKARMA D K, et al. RAKSHAK-The Multipurpose Unmanned Ground Vehicle [J]. International Research Journal on Advanced Engineering Hub (IRJAEH), 2024, 2(6): 1816-1820.

[6] PANDEY R, SINGH A. Utilization of AGVs and Machines in FMS Environment[J]. Journal of Material Science & Engineering, 2016, 5(4).

[7] CHEN Z H, CHEN X Y, WANG S M. Research on an AGV Path Planning Method[C]// Proceedings of the 2022 International Conference on Artificial Life and Robotics. Online, Japan: Alife Robotics Corp. Ltd., 2022: 1058-1062.

[8] 王奕凡. 卡尔曼滤波在自动驾驶的应用[J]. 软件, 2021, 42(1): 127-129.
WANG Y F. The Application of Kalman Filter in Automatic Driving[J]. Software, 2021, 42(1): 127-129.

[9] 夏伟才, 曾致远. 一种基于卡尔曼滤波的背景更新算法[J]. 计算机技术与发展, 2007(10): 134-136+143.
XIA W C, ZENG Z Y. Background Update Algorithm Based on Kalman Filtering[J]. Computer Technology and Development, 2007(10): 134-136+143.

[10] ZADEH L A. Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems [M]. Singapore: World Scientific, 1996.

[11] FENG X L, YANG X L. Fast Convergent Average Consensus of Multiagent Systems Based on Community Detection Algorithm[J]. Advances in Difference Equations, 2018, 2018(1): 1-13.

[12] MARR D. Vision[M]. San Francisco, CA, USA: W. H. Freeman, 1982.

[13] HU A. Neural World Models for Computer Vision[D]. Cambridge: University of Cambridge, 2022.

[14] DICKMANNS E D, GRAEFE V. Dynamic Monocular Machine Vision[J]. Machine Vision and Applications, 1988, 1(4): 223-240.

[15] DICKMANNS E D. The 4D-Approach to Dynamic Machine Vision[C]// Proceedings of the 33rd IEEE Conference on Decision and Control. Lake Buena Vista, FL, USA: IEEE, 1994: 3770-3775.

[16] DICKMANNS E D. Dynamic Vision for Perception and Control of Motion[M]. London: Springer Verlag London Limited, 2007.

[17] AUDIBERT R B, LEMOS H, AVELAR P, et al. On the Evolution of A. I. and Machine Learning: Towards Measuring and Understanding Impact, Influence, and Leadership at Premier A.I. Conferences[EB/OL]. (2022-05-26)[2025-10-30]. https://arxiv.org/abs/2205.13131.

[18] RANE V, POOJARI H, SHARMA P, et al. LiDAR Based Self-Driving Car[J]. International Journal for Research in Applied Science and Engineering Technology, 2022, 10 (IV): 261-270.

[19] POMERLEAU D A. Efficient Training of Artificial Neural Networks for Autonomous Navigation[J]. Neural Computation, 1991, 3(1): 88-97.

[20] POMERLEAU D A. Neural Network Perception for Mobile Robot Guidance[M]. New York, NY, USA: Springer, 1993.

[21] ZHOU T H, YANG M M, JIANG K, et al. MMW Radar-Based Technologies in Autonomous Driving: A Review[J]. Sensors, 2020, 20(24).

[22] KHANAM F T Z, CHAHL L A, CHAHL J S, et al. Noncontact Sensing of Contagion[J]. Journal of Imaging, 2021, 7(2).

[23] FIKES R E, NILSSON N J. Strips: A New Approach to the Application of Theorem Proving to Problem Solving[J]. Artificial Intelligence, 1971, 2(3/4): 189-208.

[24] FERREIN A. Robot Controllers for Highly Dynamic Environments with Real-Time Constraints[J]. Künstliche Intelligenz, 2010, 24: 175-178.

[25] KUIPERS B, FEIGENBAUM E A, HART P E, et al. Shakey: From Conception to History[J]. The AI Magazine, 2017, 38(1): 88-103.

[26] COX I J, WILFONG G T. Autonomous Robot Vehicles[M]. New York, NY, USA: Springer-Verlag, 1990.

[27] SHAFER S A, STENTZ A, THORPE C E. An Architecture for Sensor Fusion in a Mobile Robot[C]// Proceedings of the 1986 IEEE International Conference on Robotics and Automation. San Francisco, CA, USA: IEEE, 1986: 2002-2011.

[28] LOWRIE J W, THOMAS M, GREMBAN K, et al. The Autonomous Land Vehicle (ALV) Preliminary Road-Following Demonstration[C]// Cambridge Symposium. Cambridge, United States: Society of Photo-Optical Instrumentation Engineers, 1985.

[29] TURK M A, MORGENTHALER D G, GREMBAN K D, et al. VITS-A Vision System for Autonomous Land Vehicle Navigation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1988, 10(3): 342-361.

[30] 刘欣. ALV立体视觉导航技术研究[D]. 长沙: 国防科技大学, 2005.
LIU X. Research on ALV Stereoscopic Vision Navigation Technology[D]. Changsha: National University of Defense Technology, 2005.

[31] 杜歆. 用于导航的立体视觉系统[D]. 杭州: 浙江大学, 2003.
DU X. Stereo Vision System for Navigation[D]. Hangzhou: Zhejiang University, 2003.

[32] 杨华. 致力于推动中国智能汽车发展——记清华大学“THMR课题组”负责人何克忠教授[J]. 中国高校科技, 2012(增刊1): 55.
YANG H. Dedicated to Advancing the Development of Intelligent Vehicles in China - A Profile of Professor He Kezhong, Head of Tsinghua University's “THMR Research Group”[J]. Chinese University Science & Technology, 2012 (S1): 55.

[33] LIU J H. An Overview of ADAS System Sensors in the Automotive Electronics Sector[J]. Highlights in Science, Engineering and Technology, 2024, 111: 52-57.

[34] KIM G Z, EOM J S, PARK Y W. Alien Pulse Rejection in Concurrent Firing LIDAR[J]. Remote Sensing, 2022, 14(5).

[35] BROOKER G M. Mutual Interference of Millimeter-Wave Radar Systems[J]. IEEE Transactions on Electromagnetic Compatibility, 2007, 49(1): 170-181.

[36] 张春蓉. 激光雷达在自动驾驶中的技术应用[J]. 产业与科技论坛, 2021, 20(21): 35-36.
ZHANG C R. The Technological Application of LiDAR in Autonomous Driving[J]. Industrial & Science Tribune, 2021, 20(21): 35-36.

[37] DICKMANNS E D. Vehicles Capable of Dynamic Vision[J]. Proceedings of the Fifteenth International Joint Conference on Artifical Intelligence, 1997, 2: 1577-1592.

[38] MOHAMMED A S, AMAMOU A, AYEVIDE F K, et al. The Perception System of Intelligent Ground Vehicles in All Weather Conditions: A Systematic Literature Review[J]. Sensors, 2020, 20(22).

[39] BROGGI A, BERTOZZI M, FASCIOLI A. ARGO and the MilleMiglia in Automatico Tour[J]. IEEE Intelligent Systems and Their Applications, 1999, 14(1): 55-64.

[40] SAHAGUN P. Platooning of Autonomous Vehicles Operating on an Automated Highway System (AHS) [D]. Barcelona: Universitat Politècnica de Catalunya, 2019.

[41] XIE S T, LI Z H, ARVIN F, et al. A Review of Multi-Vehicle Cooperative Control System in Intelligent Transportation[J]. International Journal of Automotive Manufacturing and Materials, 2023, 2(3).

[42] LI Y Y, YUAN W, ZHANG S A, et al. Choose Your Simulator Wisely: A Review on Open-Source Simulators for Autonomous Driving[J]. IEEE Transactions on Intelligent Vehicles, 2024, 9(5): 4861-4876.

[43] THRUN S. A Personal Account of the Development of Stanley, the Robot That Won the DARPA Grand Challenge [J]. AI Magazine: Artificial Intelligence, 2006, 27(4): 69-82.

[44] URMSON C, ANHALT J, BAGNELL D, et al. Autonomous Driving in Urban Environments: Boss and the Urban Challenge[J]. Journal of Field Robotics, 2008, 25(8): 425-466.

[45] THRUN S. Winning the Darpa Grand Challenge: A Robot Race Through the Mojave Desert[J]. 21st IEEE/ASE International Conference on Automated Software Engineering. Tokyo, Japan: IEEE, 2006: 11.

[46] THRUN S, MONTEMERLO M, DAHLKAMP H, et al. Stanley: The Robot That Won the DARPA Grand Challenge [J]. Journal of Field Robotics, 2006, 23(9): 661-692.

[47] 王鑫, 凌铭, 饶启鹏, 等. 基于改进Stanley算法的无人车路径跟踪融合算法研究[J]. 汽车技术, 2022(7): 25-31.
WANG X, LING M, RAO Q P, et al. Research on Fusion Algorithm of Unmanned Vehicle Path Tracking Based on Improved Stanley Algorithm[J]. Automobile Technology, 2022(7): 25-31.

[48] QAYYUM A, USAMA M, QADIR J, et al. Securing Connected & Autonomous Vehicles: Challenges Posed by Adversarial Machine Learning and the Way Forward[J]. IEEE Communications Surveys & Tutorials, 2020, 22(2): 998-1026.

[49] BERNARDESCHI C, DI NATALE M, DINI G, et al. Verifying Data Secure Flow in AUTOSAR Models[J]. Journal of Computer Virology and Hacking Techniques, 2018, 14: 269-289.

[50] PENDLETON S D, ANDERSEN H, DU X X, et al. Perception, Planning, Control, and Coordination for Autonomous Vehicles[J]. Machines, 2017, 5(1).

[51] CONTRERAS M, JAIN A, BHATT N P, et al. A Survey on 3D Object Detection in Real Time for Autonomous Driving [J]. Frontiers in Robotics and AI, 2024, 11.

[52] LUO W H, XING J L, MILAN A, et al. Multiple Object Tracking: A Literature Review[J]. Artificial Intelligence, 2021, 293.

[53] HE Y, YU H S, LIU X Y, et al. Deep Learning Based 3D Segmentation: A Survey[EB/OL]. (2024-09-17)[2025-10-30]. https://arxiv.org/abs/2103.05423.

[54] BOJARSKI M, TESTA D D, DWORAKOWSKI D, et al. End to End Learning for Self-Driving Cars[EB/OL]. (2016-04-25)[2025-10-30]. https://arxiv.org/abs/1604.07316.

[55] CHEN L, TANG T T, CAI Z T, et al. Level 2 Autonomous Driving on a Single Device: Diving into the Devils of Openpilot[EB/OL]. (2022-06-16) [2025-10-30]. https://arxiv.org/abs/2206.08176.

[56] TAMPUU A, MATIISEN T, SEMIKIN M, et al. A Survey of End-to-End Driving: Architectures and Training Methods [J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 33(4): 1364-1384.

[57] LI Z Q, WANG W H, LI H Y, et al. BEVFormer: Learning Bird's-Eye-View Representation from Multi-Camera Images via Spatiotemporal Transformers[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2025, 47(3): 2020-2036.

[58] HU Y H, YANG J Z, CHEN L, et al. Planning-Oriented Autonomous Driving[C]// 2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Vancouver, BC, Canada: IEEE, 2023: 17853-17862.

[59] XU Z H, ZHANG Y J, XIE E Z, et al. DriveGPT4: Interpretable End-to-End Autonomous Driving via Large Language Model[J]. IEEE Robotics and Automation Letters, 2024, 9(10): 8186-8193.

[60] HEGDE D, YASARLA R, CAI H, et al. Distilling Multi-Modal Large Language Models for Autonomous Driving[EB/OL]. (2025-01-16) [2025-10-30]. https://arxiv. org/abs/2501.09757.

[61] ZHOU X C, LIU M Y, YURTSEVER E, et al. Vision Language Models in Autonomous Driving: A Survey and Outlook[J/OL]. IEEE Transactions on Intelligent Vehicles. (2024-05-16) [2025-10-30]. https://ieeexplore. ieee. org/abstract/document/10531702.

[62] YANG Z Y, LI L J, LIN K, et al. The Dawn of LMMs: Preliminary Explorations with GPT-4V(ision) [EB/OL]. (2023-10-11) [2025-10-30]. https://arxiv. org/abs/2309.17421.

[63] WEN L C, YANG X M, FU D C, et al. On the Road with GPT-4V(ision): Early Explorations of Visual-Language Model on Autonomous Driving[EB/OL]. (2023-11-28) [2025-10-30]. https://arxiv.org/abs/2311.05332.

[64] LI Z X. GPT-4V Explorations: Mining Autonomous Driving [EB/OL]. (2024-01-24)[2025-10-30]. https://arxiv.org/abs/2406.16817.

[65] LIU H T, LI C Y, WU Q Y, et al. Visual Instruction Tuning [C]// Proceedings of the 37th International Conference on Neural Information Processing Systems. New Orleans, LA, USA: Curran Associates Inc., 2023: 34892-34916.

[66] LI J N, LI D X, XIONG C M, et al. BLIP: Bootstrapping Language-Image Pre-Training for Unified Vision-Language Understanding and Generation[J]. Proceedings of Machine Learning Research, 2022, 162: 12888-12900.

[67] ALAYRAC J B, BARR I, BARREIRA R, et al. Flamingo: A Visual Language Model for Few-Shot Learning[C]// 36th Conference on Neural Information Processing Systems (NeurIPS 2022). New Orleans, Louisiana, USA: Neural

Information Processing Systems Foundation Inc., 2022.

[68] TIAN X Y, GU J R, LI B L, et al. DriveVLM: The Convergence of Autonomous Driving and Large Vision-Language Models[J]. Proceedings of Machine Learning Research, 2025, 270: 4698-4726.

[69] ZITKOVICH B, YU T H, XU S C, et al. RT-2: Vision-Language-Action Models Transfer Web Knowledge to Robotic Control[J]. Proceedings of Machine Learning Research, 2023, 229: 2165-2183.

[70] HWANG J J, XU R S, LIN H, et al. EMMA: End-to-End Multimodal Model for Autonomous Driving[EB/OL]. (2025-09-23)[2025-10-30]. https://arxiv.org/abs/2410.23262.

[71] LIN X W, PEI Z X, LIN T W, et al. Sparse4D V3: Advancing End-to-End 3D Detection and Tracking[EB/OL]. (2023-11-02) [2025-10-30]. https://arxiv. org/abs/2311.11722.

[72] CHEN J C, WU Y F, TIAN J Q, et al. MapTracker: Tracking with Strided Memory Fusion for Consistent Vector HD Mapping[C]// 18th European Conference on Computer Vision (ECCV 2024). Milan, Italy: Springer, 2024: 90-107.

[73] HU A, MUREZ Z, MOHAN N, et al. FIERY - Future Instance Prediction in Bird's-Eye View from Surround Monocular Cameras[J]. 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Montreal, QC, Canada: IEEE, 2021: 15253-15262.

[74] JIANG B, CHEN S Y, WANG X G, et al. Perceive, Interact, Predict: Learning Dynamic and Static Clues for End-to-End Motion Prediction[EB/OL]. (2022-12-05) [2025-10-30]. https://arxiv.org/abs/2212.02181.

[75] CHEN S Y, JIANG B, GAO H, et al. VADv2: End-to-End Vectorized Autonomous Driving Via Probabilistic Planning [EB/OL]. (2024-02-20)[2025-10-30]. https://arxiv.org/abs/2402.13243.

[76] LI L C, SHAO W, DONG W, et al. Data-Centric Evolution in Autonomous Driving: A Comprehensive Survey of Big Data System, Data Mining, and Closed-Loop Technologies [EB/OL]. (2024-01-26)[2025-10-30]. https://arxiv.org/abs/2401.12888.

[77] ZHENG Y P, YANG P X, XIA Z P, et al. Preliminary Investigation into Data Scaling Laws for Imitation Learning-Based End-to-End Autonomous Driving[EB/OL]. (2025-10-13)[2025-10-30]. https://arxiv.org/abs/2412.02689.

[78] GUAN Y C, LIAO H C, LI Z N, et al. World Models for Autonomous Driving: An Initial Survey[J/OL]. IEEE Transactions on Intelligent Vehicles. (2024-05-08)[2025-10-30]. https://ieeexplore. ieee. org/abstract/document/10522953.

[79] WU Z R, LIU T Y, LUO L Y, et al. MARS: An Instance-Aware, Modular and Realistic Simulator for Autonomous Driving[C]// CAAI International Conference on Artificial Intelligence (CICAI 2023). Singapore: Springer, 2024: 3-15.

[80] YANG J W, IVANOVIC B, LITANY O, et al. EmerNeRF: Emergent Spatial-Temporal Scene Decomposition Via Self-Supervision[C]// The 12th International Conference on Learning Representations (ICLR 2024). Vienna, Austria: ICLR, 2024.

[81] 成欢，王硕，李孟，等. 面向自动驾驶场景的神经辐射场综述[J]. 图学学报, 2023, 44(6): 1091-1103.
CHENG H, WANG S, LI M, et al. A Review of Neural Radiance Field for Autonomous Driving Scene[J]. Journal of Graphics, 2023, 44(6): 1091-1103.

[82] CHENG K, LONG X X, YANG K Z, et al. GaussianPro: 3D Gaussian Splatting with Progressive Propagation[C]// Proceedings of the 41st International Conference on Machine Learning (ICML 2024). Vienna Austria: JMLR. org, 2024: 8123-8140.

[83] SHI X, CHEN L L, WEI P, et al. DHGS: Decoupled Hybrid Gaussian Splatting for Driving Scene[C]// 2025 IEEE International Conference on Multimedia and Expo Workshops (ICMEW). Nantes, France: IEEE, 2025: 1-6.

[84] ZHANG Z J, LINIGER A, DAI D X, et al. TrafficBots: Towards World Models for Autonomous Driving Simulation and Motion Prediction[J]. 2023 IEEE International Conference on Robotics and Automation (ICRA). London, United Kingdom: IEEE, 2023: 1522-1529.

[85] WANG X F, ZHU Z, HUANG G, et al. DriveDreamer: Towards Real-World-Driven World Models for Autonomous Driving[C]// 18th European Conference on Computer Vision (ECCV 2024). Milan, Italy: Springer, 2024: 55-72.

[86] ZHAO G S, NI C J, WANG X F, et al. DriveDreamer4D: World Models Are Effective Data Machines for 4D Driving Scene Representation[C]// 2025 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Nashville, TN, USA: IEEE, 2025: 12015-12026.

[87] RUSSELL L, HU A, BERTONI L, et al. GAIA-2: A Controllable Multi-View Generative World Model for Autonomous Driving[EB/OL]. (2025-03-26)[2025-10-30]. https://arxiv.org/abs/2503.20523v1.

[88] LI C, ZHOU K Y, LIU T, et al. AVD2: Accident Video Diffusion for Accident Video Description[EB/OL]. (2025-03-04)[2025-10-30]. https://arxiv.org/abs/2502.14801.

（下转第33页）

Prioritized Experience Replay[J]. Digital Transportation and Safety, 2025, 4(3): 170-174.

[15] 张奔. 基于深度强化学习的车辆车道保持算法研究[D]. 南京: 南京理工大学, 2021.
ZHANG B. Research on Vehicle Lane Keeping Algorithm Based on Deep Reinforcement Learning[D]. Nanjing: Nanjing University of Science and Technology, 2021.

[16] TERVEN J. Deep Reinforcement Learning: A Chronological Overview and Methods[J]. AI, 2025, 6(3).

[17] IBRAHIM S, MOSTAFA M, JNADI A, et al. Comprehensive Overview of Reward Engineering and Shaping in Advancing Reinforcement Learning Applications[J]. IEEE Access, 2024, 12: 175473-175500.

[18] LU C B, SHI L X, CHEN Z W, et al. Overcoming the Curse of Dimensionality in Reinforcement Learning Through Approximate Factorization[EB/OL]. (2024-11-12) [2025-12-11]. https://arxiv.org/abs/2411.07591.

[19] 贺伊琳, 宋若旸, 马建. 基于强化学习DDPG的智能车辆轨迹跟踪控制[J]. 中国公路学报, 2021, 34(11): 335-348.
HE Y L, SONG R Y, MA J. Intelligent Vehicle Trajectory Tracking Control Based on Reinforcement Learning DDPG [J]. China Journal of Highway and Transport, 2021, 34(11): 335-348.

[20] 李文礼, 邱凡珂, 廖达明, 等. 基于深度强化学习的高速公路换道跟踪控制模型[J]. 汽车安全与节能学报, 2022, 13(4): 750-759.
LI W L, QIU F K, LIAO D M, et al. Highway Lane Change Tracking Control Model Based on Deep Reinforcement Learning[J]. Journal of Automotive Safety and Energy Efficiency, 2022, 13(4): 750-759.

[21] 孙腾超, 陈焕明. 基于深度强化学习的自主换道控制模型[J]. 农业装备与车辆工程, 2024, 62(4): 30-34.
SUN T C, CHEN H M. Autonomous Lane Changing Control Model Based on Deep Reinforcement Learning[J]. Agricultural Equipment & Vehicle Engineering, 2024, 62 (4): 30-34.

（责任编辑　斛　畔）

修改稿收到日期为2025年12月11日。

（上接第19页）

[89] TIAN Y H, QIN H, XIA F, et al. Syntax-Driven Approach for Semantic Role Labeling[C]// International Conference on Language Resources and Evaluation. Marseille, France: European Language Resources Association, 2022: 7129-7139.

[90] CHEN L, WU P H, CHITTA K, et al. End-to-End Autonomous Driving: Challenges and Frontiers[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2024, 46(12): 10164-10183.

[91] 岳超, 钟佳儒, 宁启立, 等. 新技术形势下“车能路云”融合发展战略研究[J]. 中国工程科学, 2024, 26(1): 45-58.
YUE C, ZHONG J R, NING Q L, et al. Development Strategy of Vehicle-Energy-Road- Cloud Collaboration in the New Technological Situation[J]. Strategic Study of CAE, 2024, 26(1): 45-58.

[92] XU R S, XIANG H, XIA X, et al. OPV2V: An Open Benchmark Dataset and Fusion Pipeline for Perception with Vehicle-to-Vehicle Communication[C]// 2022 International Conference on Robotics and Automation (ICRA). Philadelphia, PA, USA: IEEE, 2021: 2583-2589.

[93] LI Y M, MA D K, AN Z Y, et al. V2X-Sim: Multi-Agent Collaborative Perception Dataset and Benchmark for Autonomous Driving[J]. IEEE Robotics and Automation Letters, 2022, 7(4): 10914-10921.

[94] XU R S, XIANG H, TU Z Z, et al. V2X-ViT: Vehicle-to-Everything Cooperative Perception with Vision Transformer [C]// 17th European Conference on Computer Vision (ECCV 2022). Tel Aviv, Israel: Springer, 2022: 107-124.

[95] YU H B, LUO Y Z, SHU M, et al. DAIR-V2X: A Large-Scale Dataset for Vehicle-Infrastructure Cooperative 3D Object Detection[C]// 2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, LA, USA: IEEE, 2022: 21329-21338.

[96] XIANG H, ZHENG Z L, XIA X, et al. V2X-Real: A Large-Scale Dataset for Vehicle-to-Everything Cooperative Perception[EB/OL]. (2024-12-16) [2025-10-30]. https://arxiv.org/abs/2403.16034v2.

[97] ZIMMER W, WARDANA G A, SRITHARAN S, et al. TUMTraf V2X Cooperative Perception Dataset[C]// 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Seattle, WA, USA: IEEE, 2024: 22668-22677.

（责任编辑　斛　畔）

修改稿收到日期为2025年7月14日。