应用人工智能技术的自动驾驶系统安全测评方法研究

陈 贞[1]，李京泰[2]，郭 煌[3]，徐晓庆[1]

（1. 北京镝石数据科技有限公司，北京 100176；2. 工业和信息化部 装备工业发展中心，北京 100846；
3. 北京赛目科技股份有限公司，北京 100080）

摘 要：汽车智能化和网联化的快速发展促进了人工智能技术的探索与商业化应用，但人工智能技术的应用带来的自动驾驶安全风险也日益突显。因此，建立应用人工智能技术的自动驾驶系统安全测试与评估方法，成为平衡技术发展与安全风险的重要基础。从系统安全验证的角度，融合考虑人工智能系统的生命周期、安全要求、验证与确认方法及持续的风险评估和安全分析，提出了一种覆盖设计开发、测试评估、部署和运行3个阶段的安全评估方法。同时，提出了保障系统安全的开发、设计、测试、优化等措施，旨在为应用人工智能技术的自动驾驶系统测试与安全评估方法提供参考。

关键词：人工智能；自动驾驶；智能网联汽车；测试方法；安全评估

中图分类号：U461.91 **文献标志码：**A **DOI：**10.3969/j.issn.2095–1469.2025.04.03

A Safety Assessment Method for AI-Powered Automated Driving Systems

CHEN Zhen[1]，LI Jingtai[2]，GUO Huang[3]，XU Xiaoqing[1]

（1. Beijing Dishi Data Technology Co.，Ltd.，Beijing 100176，China；
2. Equipment Industry Development Center，Ministry of Industry and Information Technology，Beijing 100846，China；
3. Beijing Saimo Technology Co.，Ltd.，Beijing 100080，China）

Abstract: The rapid development of connected and intelligent vehicles is accelerating the exploration and commercialization of artificial intelligence (AI) technologies. Yet the broader and deeper application of AI in automated driving also brings increasingly prominent safety risks. Thus, developing safety testing and assessment methods for AI-applied automated driving systems is crucial for balancing technological innovation with safety concerns. From a system-safety perspective, this paper proposes a safety assessment method covering three stages: design and development, testing and evaluation, and deployment and operation. The method integrates the life cycle of AI system, safety requirements, verification and validation methods, and continuous risk assessment and safety analysis. Furthermore, the measures for development, design, testing, and optimization to ensure system safety are proposed, providing a reference for future testing and safety assessment of AI-based automated driving systems.

Keywords: artificial intelligence; automated driving; intelligent and connected vehicle; testing methods; safety assessment

收稿日期：2024–05–29 **改稿日期：**2024–07–16 **网络首发日期：**2024–07–19

基金项目：新一代人工智能国家科技重大专项（2022ZD0116311）

参考文献引用格式：

陈贞，李京泰，郭煌，等.应用人工智能技术的自动驾驶系统安全测评方法研究[J].汽车工程学报，2025，15(4)：457–467.

CHEN Zhen，LI Jingtai，GUO Huang，et al.A Safety Assessment Method for AI-Powered Automated Driving Systems[J].Chinese Journal of Automotive Engineering，2025，15(4)：457–467.(in Chinese)

随着汽车产业向智能网联汽车转型升级和自动驾驶商业应用探索加速，人工智能（Artificial Intelligence，AI）技术在汽车领域的应用不断深入。AI技术可以应用于自动驾驶系统的环境感知、决策规划等多个环节，建立基于AI的视觉模型、决策规划模型，可实现基于传感器输入的环境识别和建模、意图识别和行为预测、车辆轨迹规划等[1-2]，如图1所示。AI具备学习和泛化能力，通过分析海量驾驶场景和行为，并借助高性能训练算力集群，识别公路、城市道路等条件下的危险场景和边缘场景，可以实现自动驾驶系统的快速迭代和端到端优化，加速长尾效应的收敛[3-4]。然而，AI的应用也可能会给车辆带来新的功能安全、预期功能安全、网络安全、数据安全等安全风险，同时AI的“黑盒效应”[5]使自动驾驶系统存在模型不透明、决策过程不可解释等问题，导致验证评估过程复杂且困难，增加了对自动驾驶安全测评的难度。因此，建立适当的安全测试与评估方法，是支撑AI快速商业落地，保障车用AI安全运行的重要前提。当前，AI在自动驾驶系统应用中的安全测评方法仍在研究和探索中，现有的实证研究相对有限，这在一定程度上限制了AI在自动驾驶系统领域的商业化应用。

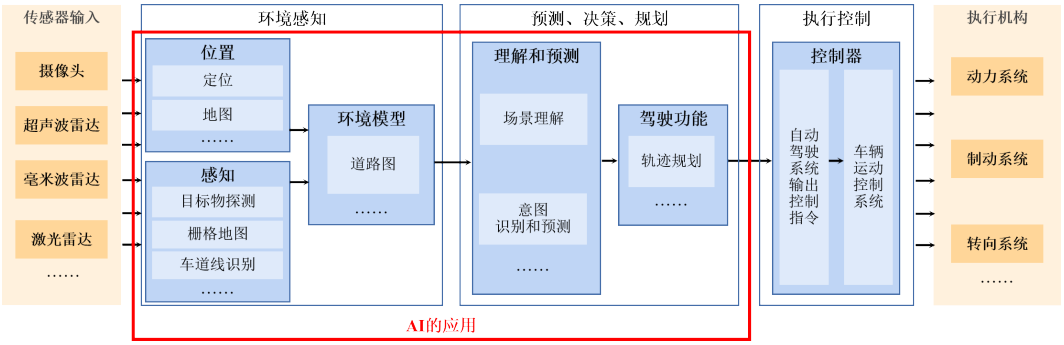

图1 AI在自动驾驶系统中的应用

2024年3月21日，联合国通过一项关于AI管理问题的决议[6]，该决议强调AI的监管和治理办法，包括在AI系统的生命周期中，加强制定和实施有效的保障措施，进行风险及影响评估，并强调有效的数据管理、生成和获取等方面的重要性。2024年5月21日，欧盟理事会批准了《人工智能法案》（EU AI Act）[7]，首次为AI技术制定了全面而详细的监管制度，该法案提出对AI采取基于风险的监管规则，强调了AI的鲁棒性、安全性、数据治理、透明度等原则。联合国世界车辆法规协调论坛自动驾驶和网联车辆工作组（WP.29/GRVA）开展了AI对GRVA负责的相关法规适用性分析[8-9]，主要涉及与UN R156软件升级及管理的关系，以及从测试验证的角度关注测试结果的可重复性及一致性、数据管理、可解释性、透明度、可靠性等问题，为形成道路车辆背景下的AI管理指导意见提供支撑。国际标准化组织道路车辆电子电气部件及通用系统工作组（ISO/TC22/SC32）正在研究制定道路车辆安全和人工智能标准[10]，对AI从开发到部署的生命周期提出适当的安全要求，并对AI的验证和确认以及系统的安全性评估提出相应的目标和约束。2023年7月，国家互联网信息办公室等部门联合发布《生成式人工智能服务管理暂行办法》[11]，对生成式人工智能服务的透明度、准确性和可靠性提出要求，并规定数据标注工作的质量和规范性。

本文基于国内外对于AI在自动驾驶系统应用中的安全测试评估方法的研究，以应用AI的自动驾驶系统为研究对象，从系统安全验证的角度出发，提出了一种涵盖AI安全生命周期的设计、开发、运行等阶段的系统安全测试与评估方法，为应用AI技术的自动驾驶系统性能与安全测评提供参考。本文所指的AI系统是具备如学习、推理、改进等能力的系统，通过分析大量数据、应用模型等

实现其目标，例如应用于自动驾驶系统的感知模型、决策规划大模型、端到端大模型等作为组件集成到自动驾驶系统。

1　AI 系统安全测试与评估要点

当前，采用基于场景的模拟仿真、封闭场地、实际道路测试等“多支柱”测试评估方法对自动驾驶系统进行安全测评已成为行业共识，并正在探索其在量产自动驾驶产品的测评与型式批准中的应用[12-15]。AI 技术的广泛应用为自动驾驶带来了革命性的变化和新期望，由于 AI 系统在信息获取、认知、知识形成、决策结果等方面存在不确定性，为应用 AI 的自动驾驶系统安全测评带来了新的挑战。本文基于 AI 系统的安全生命周期，提出了一种针对应用 AI 的自动驾驶系统的安全测试与评估方法，并提出需要重点关注的关键评估项，如图 2 所示。

图 2　应用 AI 的自动驾驶系统安全测试与评估框架

1.1　AI 系统安全测试与评估关注要点

（1）建立完整的 AI 安全生命周期流程[10]。为保障 AI 系统安全生命周期的安全性和可靠性[16]，汽车生产企业针对应用 AI 的自动驾驶系统建立安全管理体系，并在安全生命周期的各个阶段采取相应的技术措施，确保系统安全和残余风险可接受。

（2）明确合理的 AI 系统安全要求[10]。为确保 AI 系统的测试和安全评估有据可依，汽车生产企业在设计开发前明确分配给 AI 系统的安全要求，并确认 AI 系统安全要求的合理性和一致性，以便基于安全要求进行 AI 系统的验证和确认。

（3）完备且有效的验证和确认方法[10]。为确认应用 AI 的自动驾驶系统的行为能力和安全性，开展对 AI 训练模型的可信度、数据完整性和准确性等方面的评估，并确认 AI 系统在生命周期内从设计到使用的所有活动的完备性和有效性，以确保测试和监测结果符合企业的产品安全声明。

（4）持续的风险评估和安全分析[10]。持续的安全监测、风险评估和安全分析、风险缓解和避免覆盖了应用 AI 的自动驾驶系统生命周期。为确保系统的安全运行，缓解或避免风险，做好应急处置，需要建立风险监测、识别、分析和处理流程，以持续保障产品生产的一致性和安全性。

1.2　AI 系统安全生命周期

为确认系统残余风险可接受，并证明系统不会导致车辆级别的不合理安全风险，需要采用系统性的评估方法对 AI 安全生命周期进行评估。AI 安全生命周期包括设计开发阶段、测试评估阶段和运行保障阶段，涵盖安全要求的推导、系统开发设计、验证和确认、部署和安全监测等活动[17]，如图 3 所示。

（1）安全要求的推导。在安全要求的推导过程中，需要充分考虑 AI 系统安全要求的规范性、完整性和有效性。在设计开发阶段，根据自动驾驶安全要求及功能的设计运行条件，明确分配给 AI 系统的安全要求，并在开发测试验证等过程中根据实际运行表现迭代其安全要求。

（2）系统开发设计。为确保功能实现并满足安全要求，在产品设计开发阶段，企业结合自动驾驶系统和 AI 系统的安全要求，选择合适的技术路线和架构方案、设计控制失效风险的措施、规划安全监测和持续保障计划，以确保设计开发过程具有可

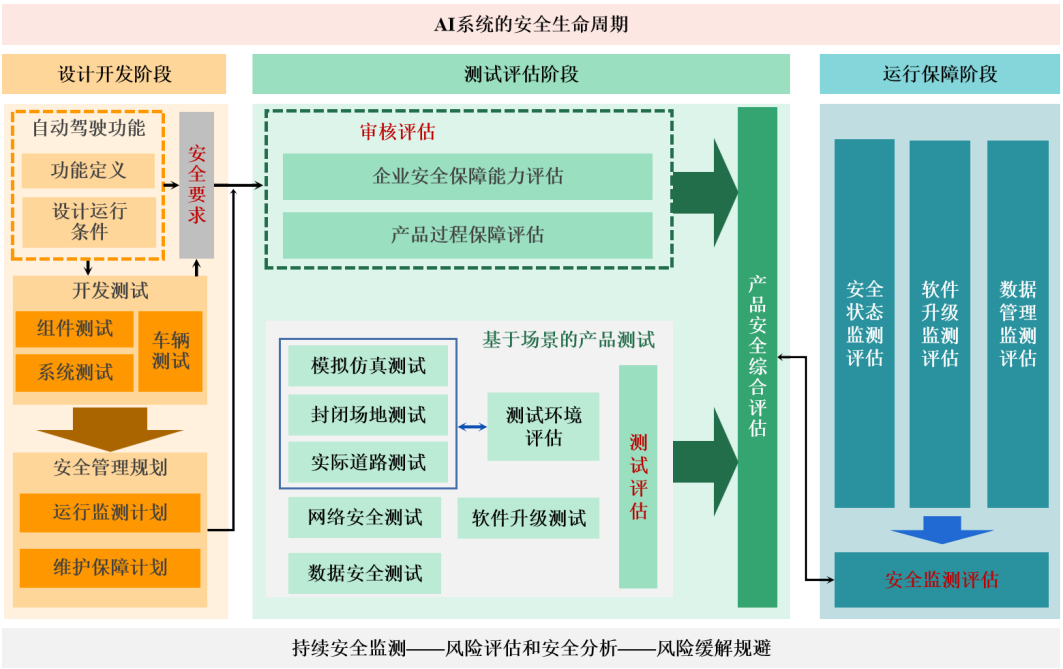

图 3 应用 AI 的自动驾驶系统与 AI 系统安全生命周期

追溯性和可持续性。形成一个性能稳定的初始版本后，基于组件测试、系统测试、车辆测试等结果进行更新迭代，形成用于整车安全性能确认评估、符合安全要求并且残余风险可接受的版本。

（3）验证和确认。在验证和确认阶段，需要基于安全要求对产品进行全面审核评估。开展 AI 训练模型的可信度、使用的训练数据集、数据完整性和准确性、仿真工具链置信度及测试结果可信度等方面的评估，以确保数据收集规范、准确、完整，且开发流程安全可靠。对自动驾驶系统产品进行基于场景的仿真、封闭场地、实际道路等“多支柱”测试，同时在实车测试时反馈安全监测情况，完成产品的安全性综合评估分析。对系统安全性评估时，需要考虑 AI 系统安全要求的合理性和一致性，确认 AI 生命周期活动的完备性和有效性。开展 AI 系统的验证和确认，确保符合汽车生产企业声明的安全要求，不会造成不合理的安全风险。

（4）部署和安全监测。在市场部署后的运行保障阶段，为了控制并消减 AI 系统的残余风险，采取持续的安全监测、数据管理及缓解和避免风险等措施，进行相应的软件升级活动。运行保障阶段的监测评估，有助于确保自动驾驶安全测评的有效性，同时可完善自动驾驶安全测试评估方法。

1.3 AI 系统安全要求

AI 系统的安全要求指 AI 系统级的技术安全要求，涵盖指标要求、系统设计、数据收集、数据训练、性能评估等方面[8]，如图 4 所示。AI 系统的安全要求需要与更高层系统（如 ADS 系统感知模块系统）的要求相协调，以避免冲突。AI 系统的安全要求既是衡量 AI 系统性能的标准，也是评估性能不足的依据。安全要求应具备可追溯性，需要制定具体的定量或定性指标。定量指标与分配给 AI 系统的安全要求应具有一致性，定性指标也应具备完整性和有效性。受 AI 特性影响，可能会导致其在运行期间产生不确定性，通过单一定量或定性指标有时难以充分证明 AI 系统符合安全要求。因此，需要结合安全监测结果等多种证据来证明 AI 系统的安全稳定。

在开发过程中，可根据需要调整和更新 AI 系统的安全要求，迭代方向旨在实现全局最优的性能指标，并确保系统级风险水平和误差在可接受范围内。如果 AI 系统的错误率过高，不符合最初的安

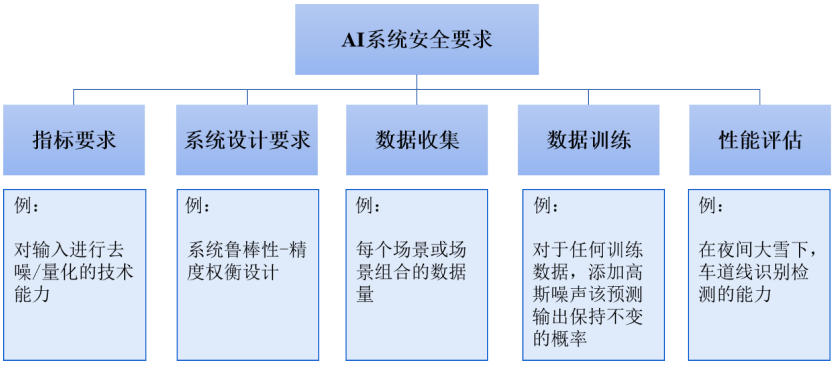

图 4　AI 系统安全要求示例

全要求，可围绕鲁棒性、可解释性、预测一致性等系统安全属性来检查原因，并根据具体原因调整其安全要求。

1.4　AI 系统验证和确认

为确保 AI 系统符合其安全要求，并在集成于上层系统时满足相应的预期用途，且不存在非预期的功能或性能不足，需要对 AI 系统开展有效的验证和确认流程。AI 系统的验证和确认需要做到以下 4 个方面。

（1）验证训练模型的可信度，确保其具有一致性和可追溯性。根据自动驾驶系统任务特点，可采取预训练、奖励训练、监督微调、强化学习等方式进行模型的训练和微调。

（2）确保数据在其生命周期内的准确性、机密性和完整性。定期审查数据质量和标注正确性，提升数据准确性，减小数据偏差；建立数据管理库，记录数据集的更改，确保机密性和完整性。

（3）确保测试验证对设计运行范围的覆盖度，充分验证 AI 系统性能。通过实车采集数据和仿真建模等方式，组合搭建更多复杂场景，构建设计开发过程阶段的测试场景集。

（4）评估测试阶段的 AI 系统性能，进行 AI 系统本身的独立性能分析、AI 系统的组件级分析测试和车辆级测试，以确保符合 AI 系统安全要求。

1.5　AI 系统风险评估和安全分析

风险评估和安全分析用于识别可能违反安全技术要求的相关风险问题，并评价系统残余风险水平，以持续保障产品的稳定运行和生产一致性。在任何阶段发现 AI 系统的测试或运行结果出现错误时，均需进行风险评估和安全分析。风险评估和安全分析可以分为两个部分。

（1）风险评估。评估系统测试或运行结果错误对系统安全要求的影响，判断由此造成的风险是否可接受。如果风险不可接受，则需要进行安全分析。

（2）安全分析。识别导致系统功能或性能不足的因素，并根据安全分析的结果采取风险缓解或避免措施。

导致系统安全风险的原因可能包括 AI 系统架构设计不合理、其他车辆系统或部件的错误输入或输出、训练数据的安全性、准确性或完整性等问题。通过深入分析 AI 架构各层与各节点，可以理解和评估架构设计及其可能带来的安全风险；审查数据集，识别训练数据是否存在覆盖偏差、不完整或质量等问题；考虑系统的复杂耦合性，分析 AI 组件，评估模型性能并进行失效分析；对应用 AI 的自动驾驶系统进行整体性安全分析，识别 AI 系统对上层系统带来的安全风险。

2　应用 AI 的自动驾驶系统生命周期管理及安全测评方法

2.1　设计开发阶段

由于 AI 技术演进迅速且正在扩大应用范围，当前难以对所有可能应用于自动驾驶系统中的 AI 系统的可接受残余风险水平提出具体指标要求。因此，加强设计开发阶段的安全要求验证和评估，通过软件升级迭代优化显得尤为重要。如图 5 所示，

基于V型开发流程，在需求设计阶段，根据自动驾驶功能定义和设计运行条件，进行需求定义和分析，完成系统架构设计、单元设计，选取合适数据集，开发、训练优化模型，通过组件测试、系统测试、整车测试，根据残余风险水平持续优化完善自动驾驶系统，最终释放基线版本。通过采用覆盖整个生命周期的分级分层、持续提升置信度等安全方法，可以合理进行自动驾驶产品的开发和迭代更新，从而有效控制设计不足、性能局限等带来的风险。

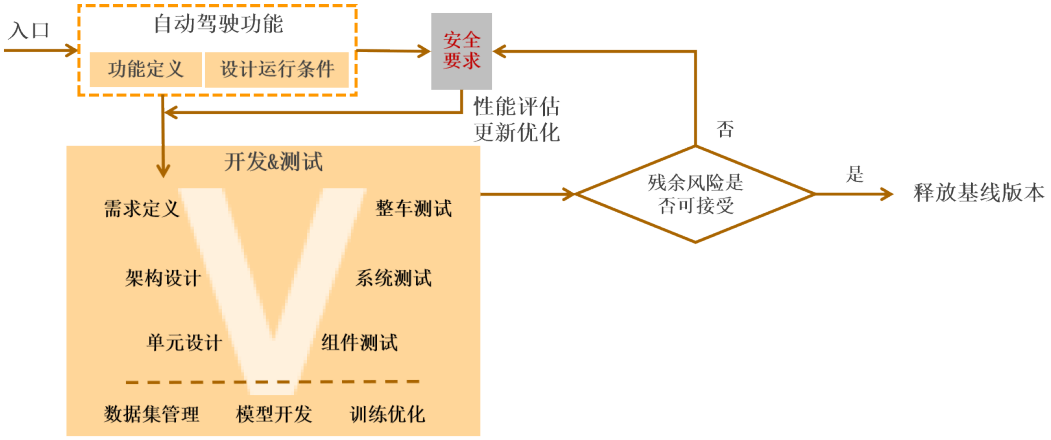

图5 应用AI的自动驾驶系统设计开发阶段主要流程

2.1.1 安全要求的重新定义

在设计开发过程中，在受AI算法性能、训练数据数量和质量、安全要求证明的置信度水平等局限性的情况下，为了控制AI系统可能出现的残余风险，可以对AI组件、AI系统、集成AI的自动驾驶系统、整车功能及性能进行测试，确保测试输入足够充分。根据测试结果反馈，持续优化迭代AI系统，以满足自动驾驶系统级的技术安全概念和分配给AI系统的安全要求。如不满足系统安全要求，可结合自动驾驶系统级的功能和安全概念、相应的解决安全相关风险的技术措施等，重新定义AI系统安全要求，直至获得最终的安全目标和验收标准。

2.1.2 系统安全保障措施

基于推理的算法使AI系统的需求和实现过程难以具备较细颗粒度的可追溯性，因此，可以选择模型可信度分析等保障AI系统安全性的措施，使释放的最终版本残余风险可接受并满足安全要求。

（1）模型可信度分析。通过模型校准的方法确认模型设计的可信度，最小化和缓解训练模型的不足。

（2）因果链分析。确认AI系统输入输出的因果关系，从而提升AI系统的可信度和模型可解释性。

（3）安全管理规划。制定运行保障阶段的安全监测和维护保障计划，采取必要的监测和运行保障措施，以对AI产品进行合理有效的安全管理。

2.2 测试评估阶段

AI系统的输入通常是多模态数据，如图像、视频等，传统的测试场景参数的输入覆盖可能导致约束数量不足。此外，AI系统的复杂架构和行为不可预测等特性，对安全性分析和能力评估带来了一定挑战。因此，对于应用AI的自动驾驶系统，有必要考虑AI特性、数据管理[18]等方面对AI系统的安全要求，充分的网络安全保护能力，以及因AI技术应用需要考虑新的攻击点或漏洞及相应的风险防御措施。基于“多支柱”测试评估方法，考虑AI系统安全要求和ADS安全要求，基于自动驾驶功能和动态场景收集等构建场景库，开展仿真测试、封闭场地测试和实际道路测试，并进行相应的安全评估和安全监测[19]，实现对应用AI的自动驾驶系统的测试评估分析，如图6所示。

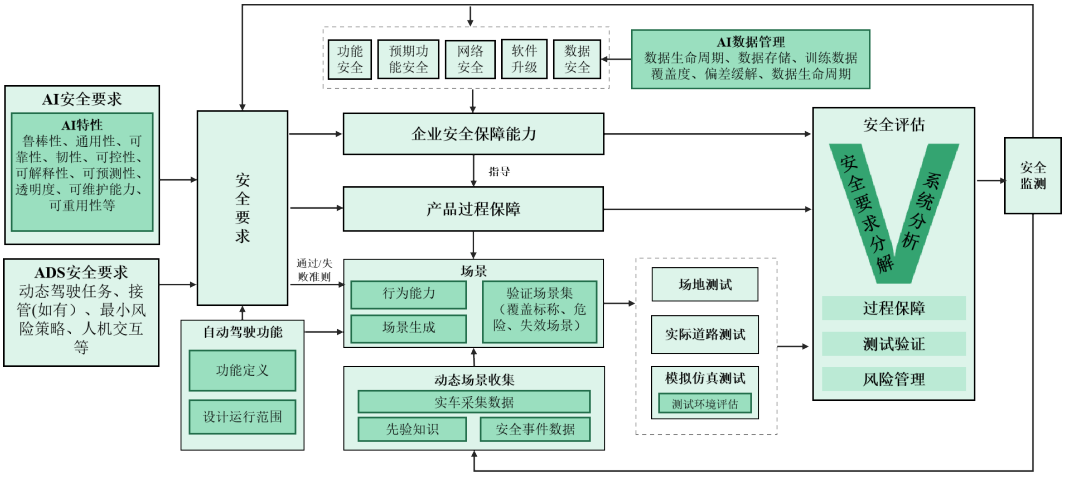

图 6　应用 AI 的自动驾驶系统测试评估框架

2.2.1　场景生成

基于场景的测试方法，需要根据自动驾驶功能及其设计运行范围，对相关道路使用者的行为进行假设，结合 ADS 行为能力、场景识别等分析方法生成包含实际参数的场景[20]。为了构建高覆盖度、合理可预见的测试场景库，除了依靠先验知识进行建模以外，还可利用地图、实车测试时自车状态、传感器采集到的数据等动态场景收集信息源构造高拟真度道路场景。同时，还可以实现对极端天气、其他道路使用者的不可预测行为、非结构化道路等复杂场景的建模，从而降低测试成本和测试安全风险。采用基于机理建模和数据驱动的测试场景生成方法，形成可解释性强、生成效率高、场景覆盖全面的测试场景库。在测试过程中，通过组合测试条件并利用测试工具链的泛化能力，可以实现设计运行范围内和边界上行为能力测试的相对完整性。

2.2.2　测试方法

测试方法包括仿真测试、封闭场地测试和实际道路测试。由于传统的场景枚举测试较难满足应用 AI 的自动驾驶系统对所有合理可预见场景的测试验证需求，因此，场景多样性强、安全性高、泛化能力强、生成效率高的仿真测试成为评测 AI 系统的重要支柱。

（1）仿真测试。构造覆盖度全面的测试场景，生成设计运行范围内、边界上及边缘场景的测试用例。基于相应的测试通过/失败准则，验证应用 AI 的自动驾驶系统的性能和安全性[21]。

（2）封闭场地测试。通过构造安全危险场景，验证应用 AI 的自动驾驶系统的基本功能和对危险场景的响应能力[22]。

（3）实际道路测试。基于仿真和封闭场地测试结果，确认应用 AI 的自动驾驶系统能力具备一定可靠性后，开展实际道路测试，验证残余风险是否达到可接受水平。

2.2.3　审核评估

在审核评估过程中，除了需要充分考虑测试报告结果外，还需要考虑实际运行环境和模型训练环境的差异性。由企业根据对应用 AI 的自动驾驶系统的生命周期过程，整理提炼并充分说明针对应用 AI 的自动驾驶系统是否建立健全的设计开发流程、测试验证体系和持续保障一致性的机制和方法。在设计开发过程中，需要说明对自动驾驶系统的安全要求设计的考虑、AI 系统架构设计的选择、AI 模型训练过程的可信度验证内容等。在测试验证阶段，基于场景的测试方法形成充分的测试验证报告，提供安全要求的实现证据，并说明自动驾驶系统实际部署和运行过程中，企业持续的迭代优化保障系统安全的能力和 AI 安全生命周期内的安全保障措施。在对 AI 模型安全进行分析的过程中，重点关注 AI 系统的鲁棒性、通用性、可靠性、韧性、可控性、可解释性、可预测性、透明度、可维护能

力、可重用性[23]等水平，如自动驾驶系统在长尾场景下的泛化能力、在极端复杂场景下的鲁棒性等。在安全评估阶段，基于系统工程方法[24]，对安全要求展开需求分解和定义，基于企业安全保障能力和产品的过程保障、测试验证报告结果，以及风险管理措施和计划，全面综合地确认自动驾驶系统的性能和安全水平。

2.3 运行保障阶段

由于真实环境复杂多变，AI模型训练环境与实际运行环境存在差异，即使经过开发设计和测试评估阶段的验证，也难以保证自动驾驶系统能应对实际场景中的所有风险。因此，可以通过采取运行过程中的安全监测等必要技术手段，持续监测自动驾驶系统的运行稳定性、可靠性、鲁棒性等，提升并优化系统性能，确保系统能力与分配的安全目标保持一致。具体技术手段包括。

（1）充分告知。由于车辆的实际使用者可能缺乏对AI局限性的了解，存在因误用或滥用系统带来的安全隐患。因此，为避免用户的滥用或误用，需要对用户进行相关培训，充分告知系统的性能限制、正确使用系统的方法以及适用场景[25]。

（2）安全监测。采取适当的技术手段监测AI系统运行过程中的行为，监测内容可以包括AI系统是否在其设计运行范围内工作、AI模型是否产生潜在异常或错误输出、AI系统是否直接参与安全事件或间接造成伤害等。

（3）风险缓解或避免。如监测到异常或危险行为，需进行如1.5节所述的风险评估和安全分析，判断是否需要采取相应的缓解或避免风险的技术措施，包括更新训练数据集、完善系统架构设计、对AI系统进行重新训练、更新或完善监测机制和安全机制、修改设计运行范围等，可能涉及对AI系统的重新开发、验证和确认。若风险较高并且没有及时的解决方案，采取措施限制AI系统的使用并及时告知用户需要采取的技术手段，确保系统造成的风险可控。最终，将系统升级更新为置信度水平提升、符合安全要求和生产一致性的新版本，如图7所示。

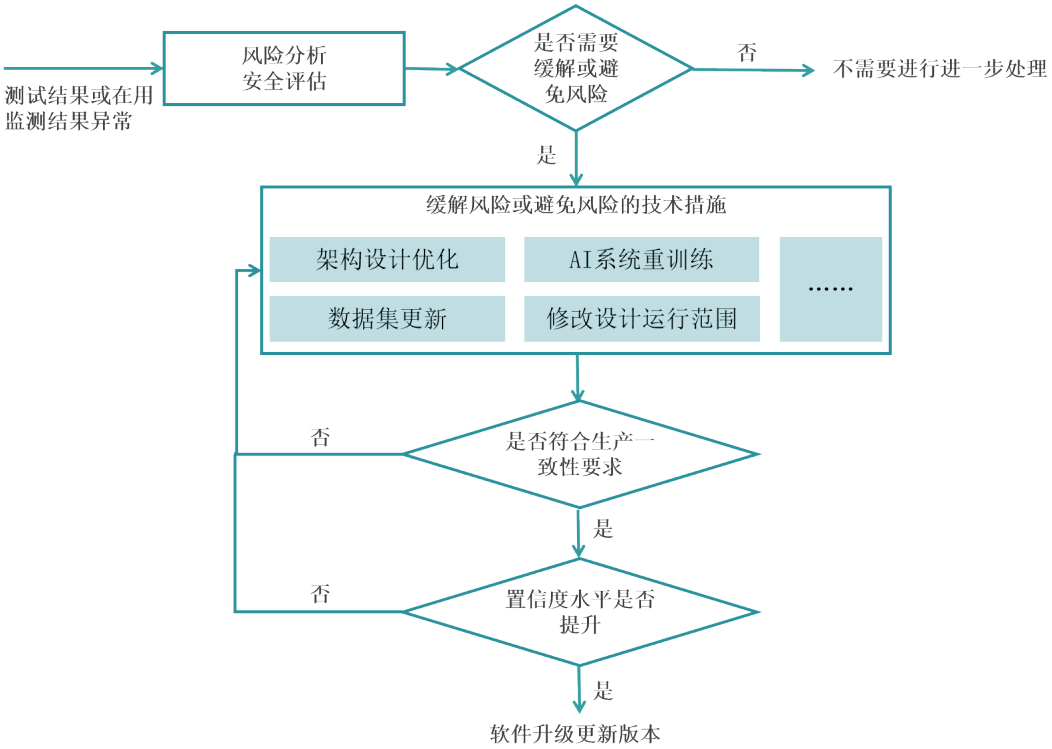

图7 缓解或避免风险主要流程

3　结语

本文基于AI技术在自动驾驶系统中的应用现状和趋势，从系统安全验证的角度出发，针对应用AI的自动驾驶系统，考虑AI的安全生命周期、AI系统的安全目标和风险/安全分析，探索性地提出了一种安全测试与评估方法。同时，本文提出了提升AI系统可解释性、可靠性及安全性等的相应措施，相关技术措施的采用可以作为安全评估内容的一部分，支撑证明AI系统安全管理体系的有效性。

AI、自动驾驶等技术已成为推动汽车产业转型升级和提升竞争力的关键。端到端神经网络技术的AI大模型逐渐应用于智能驾驶系统[26]，为智能网联汽车发展带来更多安全和舒适性提升的可能。AI在车辆路径规划、交通行为预测、行为决策、横纵向控制等方面可以发挥快速迭代优化等优势能力[27]。但是AI的规模化商业应用还需要考虑诸多问题，如数据管理[28]、算法管理[29]、科技伦理[30-31]等问题。未来需要进一步研究应用AI技术自动驾驶汽车的功能安全、预期功能安全、网络安全、数据安全、软件升级等安全测评方法和工具，如相关测试工具链的研发、测试场景集的设计、数据管理与评估方法的构建等方面，支撑相关法律法规、技术标准、评测工具链的建立，促进自动驾驶技术和安全测评体系的协调统一发展。

参考文献（References）

[1] 刘旖菲，胡学敏，陈国文，等．视觉感知的端到端自动驾驶运动规划综述[J]．中国图象图形学报，2021(1)：49-66.
LIU Yifei, HU Xuemin, CHEN Guowen, et al. Review of End-to-End Motion Planning for Autonomous Driving with Visual Perception [J]. Journal of Image and Graphics, 2021(1):49-66.(in Chinese)

[2] OMEIZA D, WEBB H, JIROTKA M, et al. Explanations in Autonomous Driving: A Survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 23(8): 10142-10162.

[3] HUANG Junjie, HUANG Guan, ZHU Zheng, et al. BEVDet: High-Performance Multi-Camera 3D Object Detection in Bird-Eye-View [Z]. arXiv: 2112.11790, 2021.

[4] 覃熊艳，张雄飞，张剑平．关于AI大模型在自动驾驶中的应用研究[J]．汽车与驾驶维修，2024(3)：29-32.
QIN Xiongyan, ZHANG Xiongfei, ZHANG Jianping. Research on the Application of AI Large Models in Automated Driving [J]. Auto Driving & Service, 2024(3):29-32.(in Chinese)

[5] CASTELVECCHI D. Can We Open the Black Box of AI? [J]. Nature, 2016, 538: 20-23.

[6] United Nations General Assembly. Seizing the Opportunities of Safe, Secure and Trustworthy Artificial Intelligence Systems for Sustainable Development [Z]. United Nations: United Nations General Assembly, 2024.

[7] European Commission. Artificial Intelligence Act [Z]. Europe: European Commission, 2024.

[8] United Nations Economic Commission for Europe. Artificial Intelligence and Vehicle Regulations[Z]. Europe: United Nations Economic Commission for Europe, 2021.

[9] United Nations Economic Commission for Europe. Proposal for a Draft Resolution with Guidance on Artificial Intelligence in the Context of Road Vehicles Under Review by GRVA [Z]. Europe: United Nations Economic Commission for Europe, 2023.

[10] Road Vehicles—Safety and Artificial Intelligence: ISO/CD PAS 8800: 2024[S]. Geneva: International Organization for Standardization, 2024.

[11] 国家互联网信息办公室，国家发展和改革委员会，教育部，等．生成式人工智能服务管理暂行管理办法[Z]．北京：国家互联网信息办公室，2023.
Cyberspace Administration of China, National Development and Reform Commission, Ministry of Education of the People's Republic of China, et al. Interim Measures for the Management of Generative Artificial Intelligence Services [Z]. Beijing: Cyberspace Administration of China, 2023.(in Chinese)

[12] United Nations Economic Commission for Europe. Proposal for a Second Iteration of the New Assessment/Test Method for Automated Driving[Z]. Europe: United Nations Economic Commission for Europe, 2022.

[13] Representatives of China, European Union, Japan and the United States of America. Proposal for Amendments to ECE/TRANS/WP. 29/2019/34 Framework Document on

Automated/Autonomous Vehicles (Levels 3 and Higher) [Z]. Europe: United Nations Economic Commission for Europe, 2019.

[14] 刘法旺,徐晓庆,陈贞,等.搭载自动驾驶功能的智能网联汽车安全测试与评估方法研究[J].汽车工程学报,2022:12(3):221-227.
LIU Fawang, XU Xiaoqing, CHEN Zhen, et al. Research on Safety Test and Assessment Method of Intelligent and Connected Vehicle with Autonomous Driving Function [J]. Chinese Journal of Automotive Engineering, 2022, 12(3):221-227.(in Chinese)

[15] European Commission. Regulation (EU) 2019/2144 of the European Parliament and of the Council as Regards Uniform Procedures and Technical Specifications for the Type-Approval of the Automated Driving System (ADS) of Fully Automated Vehicles [Z]. Europe: European Commission, 2022.

[16] European Commission. Ethics Guidelines for Trustworthy AI[Z]. Europe: European Commission, 2019.

[17] Economic and Social Council. Proposal for a Draft Resolution with Guidance on Artificial Intelligence in the Context of Road Vehicle [Z]. Europe: Economic and Social Council, 2023.

[18] 李国良,周煊赫.面向AI的数据管理技术综述[J].软件学报,2021,32(1):21-40.
LI Guoliang, ZHOU Xuanhe. Survey of Data Management Techniques for Artificial Intelligence [J]. Journal of Software, 2021, 32(1):21-40.(in Chinese)

[19] Organisation Internationale des Constructeurs d'Automobiles, European Association of Automotive Suppliers. Presentation for FRAV First Session [Z]. Europe: United Nations Economic Commission for Europe, 2019.

[20] 刘法旺,曹建永,张志强,等.基于场景的智能网联汽车"三支柱"安全测试评估方法研究[J].汽车工程学报,2023,13(1):1-7.
LIU Fawang, CAO Jianyong, ZHANG Zhiqiang, et al. A Scenario-Based "Three-Pillar" Safety Testing and Assessment Method for Intelligent and Connected Vehicles[J]. Chinese Journal of Automotive Engineering, 2023, 13(1):1-7.(in Chinese)

[21] 全国汽车标准化技术委员会.自动驾驶功能仿真测试方法标准化需求研究[Z].天津:全国汽车标准化技术委员会,2020.
National Technical Committee of Auto Standardization. Research on Standardization Requirements for Simulation Testing Methods of Automated Driving Function [Z]. Tianjin: National Technical Committee of Auto Standardization, 2020.(in Chinese)

[22] United Nations Economic Commission for Europe. Guidelines and Recommendations for Automated Driving System Safety Requirements, Assessments and Test Methods to Inform Regulatory Development [Z]. Europe: United Nations Economic Commission for Europe, 2024.

[23] 全国网络安全标准化技术委员会.人工智能安全标准化白皮书[Z].北京:全国网络安全标准化技术委员会,2023.
National Technical Committee 260 on Cybersecurity of Standardization Administration of China. White Paper on Artificial Intelligence Security Standardization [Z]. Beijing: National Technical Committee 260 on Cybersecurity of Standardization Administration of China, 2023. (in Chinese)

[24] HASKINS C. INCOSE-TP-2003-002-03, INCOSE Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities [M]. International Council on Systems Engineering, 2006.

[25] 毛向阳,尚世亮,崔海峰,等.自动驾驶汽车安全影响因素分析与应对措施研究[J].上海汽车,2018(1):33-37.
MAO Xiangyang, SHANG Shiliang, CUI Haifeng, et al. Analysis of Safety Influencing Factors of Automated Vehicle and Research on Countermeasures [J]. Shanghai Auto, 2018(1):33-37.(in Chinese)

[26] 李升波,刘畅,殷玉明,等.汽车端到端自动驾驶系统的关键技术与发展趋势[J].人工智能,2023(5):1-16.
LI Shengbo, LIU Chang, YIN Yuming, et al. Key Technologies and Development Trends of End to End Automated Driving System [J]. AI-View, 2023(5):1-16.(in Chinese)

[27] LAUGIER C. Impact of AI on Autonomous Driving [C]// WRC 2019: IEEE World Robot Conference, Beijing, China. 2019:1-27.

[28] 全国人民代表大会常务委员会.中华人民共和国数据安全法[Z].北京:全国人民代表大会常务委员会,2021.
The National People's Congress of the People's Republic of China. Data Security Law of the People's Republic of China [Z]. Beijing: The National People's Congress of the People's Republic of China, 2021.(in Chinese)

[29] 国家互联网信息办公室,中央宣传部,教育部,等.关于加强互联网信息服务算法综合治理的指导意见[Z].北京:国家互联网信息办公室,2021.
Cyberspace Administration of China, Central Propaganda

Department, Ministry of Education of the People's Republic of China, et al.Guiding Opinions on Strengthening the Comprehensive Governance of Internet Information Service Algorithms [Z].Beijing: Cyberspace Administration of China, 2021.(in Chinese)

[30] 科学技术部,教育部,工业和信息化部.科技伦理审查办法(试行)[Z].北京:科学技术部,2023.
Ministry of Science and Technology of the People's Republic of China, Ministry of Education of the People's Republic of China, Ministry of Industry and Information Technology of the People's Republic of China. Measures for the Review of Science and Technology Ethics (Trial) [Z]. Beijing: Ministry of Science and Technology of the People's Republic of China, 2023.(in Chinese)

[31] Commission Ethics. Ethical Rules for Automated and Connected Vehicular Traffic [Z]. Germany: Ethics Commission, 2017.

作者简介

陈贞（1984–），女，湖北钟祥人，博士，工程师，主要研究方向为智能网联汽车安全测试与评估方法。
E-mail：chenzhen@dystech.cn

通信作者

徐晓庆（1990–），男，山东临沂人，博士，工程师，主要研究方向为智能网联汽车安全测试与评估方法。
E-mail：xuxiaoqing@dystech.cn