

题目编号：XH-202602

面向智能驾驶的大模型应用场景研究 比赛方案

一、发榜单位

东风汽车集团有限公司研发总院

二、题目名称

面向智能驾驶的大模型应用场景研究

三、题目介绍

随着 L3+ 高阶智能驾驶向“人机协同”方向演进，语音控制作为最自然的人机交互方式，已成为突破复杂场景决策效率、提升驾驶体验的核心技术方向。当前智能驾驶语音控制面临“两层核心壁垒”：一是场景化适配不足，传统语音控制仅支持固定指令集，难以理解模糊化、组合式自然语言指令，且无法精准联动视觉、车辆状态数据完成复杂决策（如“雨天绕开前方施工区域并减速”）；二是车规级部署瓶颈，语音解析模型与决策模型的多模态融合架构参数量大，算力需求高，难以适配车载芯片的算力 / 功耗 / 内存约束，且轻量化易导致指令解析精度下降、决策延时超标，制约规模化落地。

本研究以“语音控制场景化应用为基础，轻量化车规级部

署为核心”，设置双赛道递进式竞技模式：基础赛道依托开源仿真平台验证大模型的语音指令解析、多模态融合决策能力，通过“语音输入—车辆响应”的实时联动提升观赏性；挑战赛道聚焦车规级芯片的轻量化部署，突破“语音解析+场景决策”全链路的硬件适配瓶颈，推动语音控制技术从“仿真可行”走向“实车可用”，为智能驾驶人机协同交互提供“场景验证+硬件适配”的全链路解决方案。

四、参赛对象

学生赛道：2026年6月1日以前正式注册的国内全日制非成人教育的普通高等学校在校专科生、本科生、硕士和博士研究生（不含在职研究生），以及全日制职业教育本科、高职高专在校学生，可通过学生赛道申报作品参赛。

高校青年教师在指导学生参赛的同时不得以参赛人员身份参加同一选题比赛。发榜单位及同发榜单位有相关隶属关系单位的青年不得参加本单位选题比赛。

本赛题的参赛对象可以团队或个人形式参赛，每个团队不超过10人，每件作品可由不超过3名指导教师进行指导。可以跨专业、跨学校、跨单位、跨地域组队，但同一团队所有成员均应符合本赛道相关年龄、身份要求。每件作品只可由1所高等院校或科研院所作为参赛主体提交申报。

五、答题要求

1. 基础赛道：语音控制仿真场景验证（必选，晋级挑战赛

道前提)

核心目标：验证大模型对自然语言语音指令的解析精度、多模态融合决策能力，确保语言控制在智能驾驶典型场景中的实用性与可靠性。

场景覆盖要求：基于 CARLA 开源仿真平台，适配智能驾驶语音控制核心场景（含基础操控、复杂避障、应急响应三类子场景），需处理语音指令（自然语言）、视觉数据（图像 / 点云）、车辆状态（速度 / 转向 / 位置）、环境数据（路况 / 天气）多模态输入；

模型架构要求：基于主流开源大模型架构（如 Qwen2.5 - VL、OpenVLA、多模态 Transformer），构建“语音指令解析 + 多模态融合决策”一体化模型，支持自然语言指令（含方言 / 轻微噪音）的语义理解、场景关联与动作生成；

性能要求：在仿真场景中，核心指标需达标——语音指令识别（通过文本输入模糊指令）准确率 $\geq 95\%$ 、指令解析延时 $\leq 50\text{ms}$ 、决策响应全流程延时 $\leq 150\text{ms}$ ，场景任务完成率 $\geq 90\%$ （无碰撞、无违规、指令响应无遗漏）；

数据集要求：基于 NuScenes、Waymo 仿真数据集 + 智能驾驶语音指令专用数据集（含单步 / 组合 / 应急指令、不同口音 / 噪音环境），构建“语音 + 视觉 + 车辆状态”多模态训练 / 测试数据集，保证场景同源性。

2. 挑战赛道：车规级轻量化部署攻关（仅基础赛道达标团

队可参与)

核心目标：在保证语音控制场景应用性能的前提下，完成“语音解析+决策”一体化模型的车规级芯片轻量化部署，突破算力、功耗、延迟约束。

轻量化技术要求：采用结构化 / 非结构化剪枝、INT8 / INT4 / FP8 低比特量化、知识蒸馏、多模态特征联合精简等技术，完成模型改造，需重点优化语音解析模块与决策模块的联动效率，提供明确的轻量化策略与量化分析；

车规级适配要求：针对指定车规级芯片（如 Thor-u，地平线 J6），开展算子适配与 CPU / GPU / NPU / DSP 异构调度优化，解决语音-视觉数据传输瓶颈、算力浪费问题；

性能保持要求：轻量化后模型在车规级芯片（含 HIL 硬件在环测试）中，核心性能较原始模型衰减 $\leq 3\%$ ，模型端到端推理耗时 $\leq 200\text{ms}$ ；

3. 技术应用场景

基础操控场景：通过语音指令实现车辆启动 / 停止、加速 / 减速、左转 / 右转、变道等基础动作；

复杂避障场景：结合视觉数据，通过语音指令完成“绕开前方行人后提速”、“避让施工锥桶并回归原车道”等组合动作；

应急响应场景：在雨天、夜间等极端环境下，响应“突发加塞，紧急避让”、“前方路况危险，减速停车”等模糊应急

指令；

车规级部署场景：适配车载硬件平台，为 L3+ 高阶智能驾驶提供自然、实时、可靠的语音人机交互方案。

六、作品评选标准

（一）基础赛道提交材料需提交以下内容：

1.技术方案的所有内容文档。

2.模型架构：详细说明四大模块（语音解析、视觉理解、语义对齐、动作生成）的设计、数据流、技术选型依据及关键超参数调优过程。

3.多模态融合：阐述语音-视觉-车辆状态数据的对齐方法与动态权重调整机制，以及针对噪声、模糊指令等的鲁棒性优化策略。场景适配：分场景（基础 / 复杂 / 极限）说明决策逻辑、异常处理与容错机制。

4.数据集构建：说明数据构成、预处理、标注规范、数据划分与增强方法。

5.模型相关文件（可执行、可复现）。

源代码包含：完整训练 / 推理代码，需详细注释及环境依赖说明。

模型权重：提交预训练权重与最终训练权重，并提供参数统计报告。

语音处理模块：独立提交 ASR / NLU 模块及方言 / 噪音优化代码与测试用例。

6. 仿真测试全量报告。

测试环境：说明 CARLA 仿真平台配置、硬件与软件环境。

分场景测试数据：提供基础操控、复杂避障、极限应急场景的完整时序数据、核心指标统计及异常记录。

指标计算：给出准确率、对齐精度、延时等指标的计算方法与原始数据。

可视化材料：提供过程视频、指标曲线与决策流程可视化图。

数据集相关佐证材料——说明文档：数据来源、规模统计与质量评估报告。样本示例：随机抽样数据及标注文件示例。

验证报告：数据集适配性与模型泛化能力的验证结果。

（二）挑战赛道（轻量化与车载部署）需提交材料为以下内容（仅基础赛道达标团队可参与）：

1.策略设计：说明剪枝、量化、蒸馏等技术的组合使用、参数设置与迭代优化记录。

2.模块优化：针对 VLA 各模块的差异化轻量化方案、特征压缩创新及精度损失补偿方法。

3.芯片适配：指定芯片型号与架构，详述算子适配、异构调度策略及性能测试报告。

4.轻量化模型与部署代码包，需含有——模型文件：轻量化后权重与结构文件，提供压缩前后对比报告及中间过程文件。

部署代码：适配芯片的推理代码、框架配置文件及预处理优化

代码。算子与调度模块：自定义算子源码与可执行文件、异构调度模块独立代码。

5.车规级芯片实机测试报告,需含有测试平台和核心指标说明。测试平台：芯片配置、硬件参数、软件环境与数据传输链路说明。核心指标：轻量化效果（FLOPs 压缩比、推理时间）。

（三）作品场景覆盖要求

1.基础赛道：CARLA 仿真 VLA 语音控制竞技

所有场景基于 CARLA 开源仿真平台定制，统一仿真镜像（含固定车辆参数、传感器配置、VLA 模型适配接口、语音指令输入模块），保证公平性，通过“语音输入实时投屏 + VLA 多模态处理流程可视化 + 车辆动作同步展示”提升观赏性：

场景 1：基础语音操控工况（入门）

环境：晴天白天、城市主干道（双向 6 车道），视野良好，无动态干扰；

语音指令类型：单步基础指令（如“保持当前车道，提速至 60km/h”、“前方 300 米路口右转”、“向左变道”、“减速至 40km/h”）；

数据输入：语音指令 + 前视摄像头 + 车辆状态数据（VLA 模型实时对齐语义与环境信息）；

任务目标：完成 5km 连续驾驶，无违规、无碰撞，语音指令识别准确率 $\geq 95\%$ ，多模态语义对齐精度 $\geq 98\%$ ，端到端延时 $\leq 150\text{ms}$ ；

考核重点:VLA 模型基础指令解析精度与动作响应稳定性。

场景 2: 复杂避障语音操控工况 (进阶)

环境: 阴天傍晚 (低光照)、城市次干道 + 十字路口 + 公交站, 混合交通流 (私家车 + 公交车 + 非机动车 + 行人);

语音指令类型: 复杂组合指令 (如 “看到前方横穿马路的行人, 减速避让后向左变道超越慢车”、“前方公交站有行人上下车, 靠边减速至 30km/h, 确认安全后继续行驶”);

数据输入: 语音指令 + 多视角摄像头 + 激光雷达 + 车辆状态数据 (VLA 模型联动多模态数据完成决策);

任务目标: 完成 8km 连续驾驶, 违规 ≤ 1 次、无碰撞, 语音指令识别准确率 $\geq 96\%$, 组合指令解析无遗漏, 多模态语义对齐精度 $\geq 98.5\%$;

考核重点:VLA 模型复杂指令理解与多模态融合决策能力。

场景 3: 极限应急语音操控工况 (高阶)

环境: 雨天夜间 (低光 + 路面反光 + 雨雾遮挡)、城市快速路 + 施工路段, 车道收窄, 含突发动态干扰 (突发加塞、施工锥桶、临时横穿行人);

语音指令类型: 模糊化 / 应急式指令 (如 “前方路况危险, 保持安全车速”、“突发车辆加塞, 紧急避让”、“施工路段, 减速并道至左侧车道”);

数据输入: 带轻微环境噪音的语音指令 + 多视角摄像头 (低信噪比) + 车辆状态数据 (动态变化);

任务目标：完成 6km 连续驾驶，无违规、无碰撞、无决策失效，应急场景响应延时 $\leq 120\text{ms}$ ，VLA 模型在极端环境下多模态语义对齐精度 $\geq 97\%$ ；

考核重点：VLA 模型极限环境鲁棒性与应急决策能力。

2.挑战赛道：车规级芯片 + HIL 硬件在环测试（实车适配核心）

测试平台配置：统一提供车规级芯片（如地平线征程 6）+ HIL 测试系统，集成传感器仿真模块、实时操作系统，支持 VLA 模型多模态数据实时输入与指标量化；

测试数据来源：提取基础赛道 3 类场景的 1000 帧典型多模态数据（语音 + 视觉 + 车辆状态），保证测试与仿真场景同源；

核心考核指标：实时采集 VLA 模型推理延时、FLOPs、内存占用、功耗、异构算力利用率，及语音指令识别准确率、多模态语义对齐精度等性能指标，验证轻量化模型的车规级适配能力。

3.挑战赛道：车规级芯片

测试平台配置：车规级芯片，集成传感器仿真模块、实时操作系统，支持 VLA 模型多模态数据实时输入与指标量化；

测试数据来源：提取基础赛道场景的典型多模态数据（语音 + 视觉 + 车辆状态），保证测试与仿真场景同源；

核心考核指标：实时采集 VLA 模型推理延时、FLOPs、内存占用、及文本模糊指令识别准确率等性能指标，验证轻量化

模型的车规级适配能力。

4.作品评审标准

4.1 总分计算规则

仅参与基础赛道：总分 = 基础赛道得分（满分 40 分）+ 主观评分（技术方案完整性 20 分，场景观赏性 10 分），满分 70 分；

晋级挑战赛道：总分 = 基础赛道得分 × 40% + 挑战赛道得分 × 60%，满分 100 分（主观评分融入对应赛道指标）。

4.2 基础赛道得分（40 分，晋级门槛 + 计分基础）

指标类别	具体指标	分值	达标要求 (晋级线)	计分逻辑
场景应用效果	场景任务完成率	15 分	$\geq 90\%$	基础分 10 分；每超过 1%加 0.5 分，上限 15 分。
核心性能指标	语音指令识别准确率	10 分	$\geq 95\%$	基础分 6 分；每超过 0.5%加 0.4 分，上限 10 分。
核心性能指标	多模态语义对齐精度	8 分	$\geq 98\%$	基础分 5 分；每超过 0.5%加 0.3 分，上限 8 分。
核心性能指标	端到端决策响应延	7 分	$\leq 150\text{ms}$	（达标得 4 分，每低 10ms 加 0.3 分）

	时			
--	---	--	--	--

4.3 挑战赛道得分（60 分，仅晋级团队参与计分）

指标类别	具体指标	分值	评分标准
轻量化效果	FLOPs 压缩比	15 分	≤ 0.5 得 15 分， $0.5 < \text{压缩比} \leq 0.7$ 得 10 分， $0.7 < \text{压缩比} \leq 0.9$ 得 5 分
车规级适配性	异构算力利用率	10 分	$\geq 80\%$ 得 10 分， $70\% \leq \text{利用率} < 80\%$ 得 7 分， $60\% \leq \text{利用率} < 70\%$ 得 3 分
功耗与内存	功耗 & 内存占用	10 分	功耗 $\leq 25\text{W}$ 且内存 $\leq 512\text{MB}$ 得 10 分，一项不达标得 5 分，两项不达标得 0 分
性能保持率	核心指标衰减幅度	15 分	$\leq 3\%$ 得 15 分， $3\% < \text{衰减} \leq 5\%$ 得 8 分， $> 5\%$ 得 0 分
技术创新性	轻量化 / 适配策略	10 分	策略具有创新性（如 VLA 多模态特征联合压缩、芯片定制化语义对齐算子），实用性强得 10 分

七、作品提交时间

参赛分线网上报名和提交作品两个阶段。

2026 年 5 月 30 日至 6 月 30 日,各参赛团队通过大赛申报系统进行线上报名（见第八章）。

2026 年 5 月至 9 月上旬，各参赛团队根据题目开展研发攻

关；同期，我司根据组委会活动安排和策划方案组织学生参赛，安排专业人员给予指导，为参赛团队提供支持保障等。

2026年9月5日前，各参赛团队要向发榜单位完成作品提交，具体要求详见本方案第八点第（二）款，并严格遵照发榜单位明确的提交规范执行。

2026年9月30日前，由我司完成初审，确定入围终审擂台赛的晋级作品和团队。

2026年10月，发榜单位安排专门团队提供帮助和指导，各晋级团队完善作品。

2026年11-12月，根据组委会要求组织终审擂台赛，角逐“擂主”。

八、参赛报名及作品提交方式

（一）报名方式

1.参赛选手登录“挑战杯”官网 www.tiaozhanbei.net，在“揭榜挂帅”擂台赛报名入口注册账号，登录大赛申报系统在线填写报名信息。报名信息提交后，下载打印系统生成的报名表。

2.申报人在报名表对应位置加盖所在学校公章。

3.将盖章版报名表扫描件上传至报名系统，等待系统审核。请参赛选手注意查看审核状态，如审核不通过，需重新提交。

4.系统开放报名时间为2026年5月30日—6月30日，逾期后系统将自动关闭报名功能。

（二）作品提交方式

各参赛队伍请将所有作品文件与材料打包为一个压缩文件发送至指定邮箱。

各队伍可以多次通过邮件提交作品，基础赛道提交作品给赛事联系人邮箱时可按要求提交章节六中提及的六项内容，并至少在7月30日之前提交第一版到后文中的联系人邮箱。基础赛道达标后，请在9月30日之前提交挑战赛道最终作品内容到指定邮箱中。

各参赛团队在提交作品时，同步报送1份经报名系统审核通过的参赛报名表，报名表所有信息须与系统内填报内容完全一致。

九、赛事保障

本单位将为此次比赛提供案例赛题说明，支持本次赛事开展，赛题说明将会以网盘链接的方式发给参赛队伍。同时，本单位将组建专业指导团队，指导团队将由出题单位专家组成，赛事办公室设在东风公司研发总院人工智能开发中心。参赛过程中，参赛团队如需本单位提供与项目相关的其他必须帮助，可提前与赛事办公室联系，我们将在许可范围内给予参赛团队帮助。同时，本单位将根据人工智能领域组委会每个阶段的活动策划等要求提供线上线下相结合等帮提解读工作等。

十、设奖情况及奖励措施

（一）设奖情况

原则上设特等奖5个，一等奖5个，二等奖5个，三等奖5

个（将根据实际报名情况，对评奖数量作适当调整）。从特等奖获奖团队中决出 1 个“擂主”，擂主奖金和特等奖奖金不可兼得。

（二）奖励措施

1.本单位将结合项目实际，针对学生赛道，奖励原则为擂主 10 万/个，特等奖 2 万元/个，一等奖 1 万元/个，二等奖 0.5 万元/个，三等奖 0.2 万元/个，以上均为税后奖励；

2.工作成果如获本单位认可，投入应用实践，团队成员可以允许参与项目研发，同时根据项目成果给予额外奖励；

3.特等奖团队成员如就业投递本单位研究岗位，可优先进入二面，并且同等条件下可优先录取；

4.所有获奖的队伍成员可以优先获取本单位的实习实践机会。

（三）奖金发放方式

比赛结束后，比赛专班工作人员与获奖团队取得联系，填写奖金申请表，待获奖团队提供银行卡详细信息后，将奖金按组委会要求进行发放。

十一、比赛专班联系方式

本单位将组建赛事办公室，专家指导团队共两人，进行技术指导和保障，方便参赛团队咨询；赛事服务团队共两人，负责收集比赛作品与组委会对接以及后期相关比赛赛务的协调联络。以下邮箱在报名开始后即可接收咨询邮件和作品提交邮件。

1.专家指导团队

顾问专家：熊老师，联系电话：15196030340

顾问专家：李老师，联系电话：18186245349

负责比赛期间技术指导保障。

2.赛事服务团队

联络专员：潘老师，联系电话：13190039930，联系和作品提交邮箱：panxinze@dfmc.com.cn

联络专员：林老师，联系电话：13597586708，联系和作品提交邮箱：tc-linyao@dfmc.com.cn

负责比赛期间组织服务及后期相关赛务协调联络。

3.联系时间

比赛期间工作日（9:00-17:00）

附：发榜单位简介

东风汽车集团有限公司是以汽车制造、销售、服务和技术研发为主业的商业一类央企，前身是始建于 1969 年的第二汽车制造厂。56 年来，累计产销汽车超 6200 万辆，上缴税费约 7000 亿元。截至 2024 年底，公司总资产 4703.86 亿元，从业人数 11.7 万人。2025 年，东风汽车新能源汽车销量历史性地跨越了百万辆大关，达到 104 万辆，同比增长 21%；自主品牌销量超过 150 万辆，同比增长超过 9%。主要业务涵盖乘用车和商用车整车、关键总成、汽车零部件、汽车装备、汽车金融等相关服务业务，产品覆盖豪华、高端、主流各区隔市场；汽车出口 100 多个国家和地区。

2003 年公司总部从湖北十堰迁至武汉；2017 年东风汽车公司更名为东风汽车集团有限公司；2020 年成立岚图汽车科技有限公司，发展自主高端新能源汽车；2022 年成立猛士汽车科技公司，进入豪华电动越野车领域；2023 年发布主流科技电动品牌东风奕派、国民纯电专业品牌东风纳米。近年来，公司持续完善新能源汽车产业、产业链、产业生态布局，形成全面优于燃油车时代的战略格局。深耕混动、纯电、氢动三条技术路线，推出乘商新能源专属平台和产品，马赫超级混动专用发动机热效率多次刷新行业纪录，达到 47.06%。