

题目编号：XH-202615

# 复杂交互场景的抗干扰语音指令识别技术 比赛方案

## 一、发榜单位

美的集团（上海）有限公司

## 二、题目名称

复杂交互场景的抗干扰语音指令识别技术

## 三、题目介绍

随着智能语音设备的普及和人工智能技术的快速发展，人机交互已经遍及智能家居、车载系统、公共服务等多样化领域，其中基于语音的交互方式为用户提供了更自然便捷的交互体验，推动了智能化应用场景的拓展与深化。然而，在复杂开放场景中，由于多种噪声干扰、多说话人语音重叠等问题，语音识别系统的准确性和鲁棒性往往严重下降。尤其在唤醒词识别与指令理解环节，如何准确识别目标说话人的指令，同时有效拒识非目标说话人，成为提升语音交互系统可靠性和实用性的关键挑战。

因此，本次比赛聚焦语音识别的抗干扰能力提升和性能优化，要求选手设计一套鲁棒的语音指令识别系统，针对唤醒音频的目标发音人提升语音识别率，同时拒识非唤醒发音人，并通过算法和模型结构优化，兼顾识别性能和模型推理效率。具体任务包括：

1. 搭建面向目标发音人的语音识别模型，自动适配不同唤醒音频的信道、音色和语速条件；

2. 系统应具备抗语音干扰能力，能够正确拒识非目标发音人的语音；

3. 针对目标与干扰发音人同时说话的情况，仅识别目标发音人的语音内容；

4. 提升在不同信噪比条件下的语音指令识别率，降低误识别率。

#### **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。

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、科研院所等作为参赛主体提交申报。

#### **五、答题要求**

各参赛队需要设计并实现一套结合唤醒音频的发音人信息，有效识别远场家居控制语音指令，并拒识非唤醒说话人信

息的语音识别系统。

具体包括作品设计报告、测试报告和使用说明等文档，以及作品的源代码、可执行、推理的程序，建议使用 `python` 实现。作品相关文档至少包括以下内容：

①技术设计方案，功能、指标、实现原理、软件流程等；

②测试验证方案，测试数据、结果分析、实现功能、特性等；

③其他，除上述规定文档以外的其他作品相关资料。

## 1. 初赛

组委会提供包含唤醒音频、唤醒文本、待识别音频、待识别文本标签的小规模开发数据集（测试集 A），参赛团队基于给定的数据进行模型开发与优化。组委会将于 4 月通过美的 AI 研究院官方网站 <https://ai.midea.com/> 发布初赛测试集 A，参赛团队自行在所有测试样本上完成推理，并在网站相应页面上传模型和推理脚本，通过邮件形式提供测试集 A 的识别结果、运行耗时、模型参数量等指定信息，邮箱：

`zhangzf148@midea.com`。

测试集包括正样本测试集和拒识测试集。测试数据集，测试数据样例形式如：

```
{  
    "唤醒音频名字": "唤醒音频",
```

```

    "唤醒文本名字": "唤醒标签",
    "识别音频名字": "识别音频",
    "识别文本名字": "识别标签"
}

```

唤醒音频和识别音频来自于同一轮语音交互。对于拒识音频，识别标签为空。

参赛团队根据测试集将推理结果按照以下格式保存为单个JSON 文件进行提交。其中 **id** 为测试音频名字，**content** 为推理结果，**duration** 为 **batch** 为 1 推理音频时测试集总推理时间，单位为秒（s）。

```

{
    "result": {
        "results": [
            {"id": "id1", "content": "xxx", "label": "xxx",
"cer": "xx"},
            {"id": "id2", "content": "xxx", "label": "xxx",
"cer": "xx"}
        ],
        "final_cer": "xx",
        "duration": "t"
    }
}

```

测试集 **B** 不公开。组委会将在比赛提交时间截止后，按照得分排序检查代码和耗时。一旦发现作弊，将取消该团队的所有成绩。

## 2. 总决赛

初赛最终排名以测试集 **B** 成绩和代码核查结果为准，筛选出 10~20 支队伍入围。

入围团队需准备方案介绍 **PPT** 进行现场答辩，内容包含技术方案开发集测试结果。对于入围作品，大赛组委会将综合考虑主客观标准进行判分。

## 六、作品评选标准

初评：根据测试集上的客观指标确定初赛排名，100 分制。

客观评价指标：

### （一）目标发音人识别字错率

1. 定义：在目标发音人语音中，系统错误识别的语音内容占总语音内容的比例

2. 计算公式： $CER = \text{错误识别的字数} / \text{目标发音人总字数} * 100\%$

3. 说明：这是衡量系统对目标发音人语音内容识别能力的核心指标

4. 权重 40%

## （二）拒识率

1. 定义：系统正确拒识别非目标发音人语音的比例

2. 计算公式： $RR = \text{正确拒识别非目标语音数} / \text{总非目标语音数} * 100\%$

3. 说明：该指标评估系统对非目标发音人语音的抗干扰能力，拒识率越高，系统越能避免误识别

4. 权重 40%

## （三）模型推理效率

1. 定义：系统在给定硬件条件下的推理速度和资源消耗，例如 L20 上推理 10s 音频的时间和资源消耗。

2. 测试方法：测量每条语音的平均推理时间和模型的内存占用

3. 说明：该指标评估系统的实际部署性能，推理速度越快、资源占用越低越好

4. 权重 20%，其中推理时间和内存占用指标各占 10%

## （四）信噪比鲁棒性

1. 定义：系统在不同信噪比（SNR）条件下的语音识别性能变化

2. 测试方法：在三种信噪比（如 5db, 0db, -5db）下测试 CER，并计算平均值

3. 说明：该指标评估系统在复杂噪声环境下的鲁棒性

## （五）重叠语音下的目标发音人的识别率

1. 定义：在目标与干扰人同时说话的音频上，仅针对目标发音人语音内容计算 CER

2. 测试方法：目标语音和干扰语音在时间上存在一定的重叠

3. 说明：该指标评估系统在多说话人语音重叠场景下的识别能力

测试集包括正样本测试集和拒识测试集。核心指标为正样本测试集的目标发音人识别字错率、拒识测试集非目标发音人的拒识率以及模型的推理效率，占比权重分别为 40%、40%和 20%。4 和 5 体现测试集数据分布，信噪比分布从-5db 到 5db，说话人重叠率从 0 到 100%，同一时间段最多有 2 个说话人。除了人声干扰，还存在类似空调噪声等非人声干扰。。

决赛参赛队按照顺序进入答辩环节，评委对参赛作品：

设计新颖性（20 分）；性能调优达成度（30 分）、应用价值和落地可行性（20 分）、方案设计完整性和扩展性（20 分）和参赛队答辩表述能力等综合因素（10 分），以 100 分制进行打分，根据分数决定决赛名次。在作品答辩时需要向评委说明作品设计方案、作品实现技术等内容。

## **七、作品提交时间**

2026 年 5 月至 9 月上旬，各参赛团队选择榜单中的题目开展研发攻关，各组织协调机构应组织学生参赛，安排专业人员给予指导，为参赛团队提供支持保障。

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20日前，由发榜单位完成初审，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月，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](http://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**

请参赛团队在规定时期之前将具体作品打包压缩，发送至发榜单位的指定邮箱：[zhangzfl48@midea.com](mailto:zhangzfl48@midea.com)。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。以上材料无需在“挑战杯”官网提交。

## **九、赛事保障**

1. 对于参加本项目的参赛团队，本单位为参赛者攻关答题提供远程技术答疑，技术配套基于作品提交阶段而定。

2. 为此次比赛本单位专门组建专家指导团队，团队成员由美的 AI 研究院资深专家和高级工程师担任。

3. 本单位在参赛团队完成相关审核等程序后可提供参观调研的机会。

## **十、设奖情况及奖励措施**

### **1. 设奖情况**

奖项设置主要面向学生赛道。

根据赛事安排，发榜题目原则上评出 1 个“擂主”（从特等奖中评选），评出特等奖 5 个，一、二、三等奖各 5 个，最终授奖数量可视作品申报数量和质量情况动态调整。

### **2. 奖励措施**

本单位将结合项目实际，拟奖励擂主税后 8 万元/个（叠加特等奖激励共税后 10 万元）；奖励特等奖每支队伍税后 2 万元；奖励一等奖每支队伍税后 1 万元；奖励二等奖每支队伍税后 0.5 万元；奖励三等奖每支队伍税后 0.2 万元。

比赛中表现优异的获奖选手，将有机会进入美的集团人才储备池，各个获奖团队主要负责人将并优先获得实习及就业的

推荐机会。

实际发奖数量将依据作品提交整体情况及赛事组委会的评审结果来确定。

### 3. 奖金发放方式

比赛名次公布后一个月内，美的比赛专班人员会与团队组长/代表取得联系，填写奖金申请表。待获奖团队代表提供信息后1个季度内，将奖金一次性发放至获奖团队组长/代表的银行卡中。

## 十一、比赛专班联系方式

### 1. 专家指导团队

顾问专家：张老师，联系电话：18930400730

顾问专家：张老师，联系电话：15851842545

负责比赛期间技术指导保障。

### 2. 赛事服务团队

联络专员：张老师，联系电话：13661421940

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

“科技尽善，生活尽美”——美的集团是一家集智能家居、工业技术、楼宇科技、库卡、新能源、美的医疗、安得智联七大业务板块为一体的全球化科技集团，形成美的、小天鹅、华凌、东芝、COLMO、TEKA、美芝、威灵、高创、合康、科陆、Clivet、Arbonia、菱王、KUKA、万东和安得等多品牌组合，每年为全球超过 5 亿用户及各领域重要客户提供产品和服务。目前，美的在全球拥有超 400 家子公司、38 个研发中心和 635 个主要生产基地，业务覆盖 200 多个国家和地区。

2024 年，美的集团营业总收入 4091 亿元，同比增长 9.5%；归母净利润 385 亿元，同比增长 14.3%。2025 年前三季度，营业总收入 3647.2 亿元，同比增长 13.8%；归母净利润 378.8 亿元，同比增长 19.5%。美的现有员工约 19.8 万人，其中海外员工超 4 万人、研发人员超 2.3 万人。2025 年，美的集团位列《财富》世界 500 强第 246 位，2024 年位列 Brand Finance 全球最有价值科技品牌 100 强第 36 位。美的集团积极践行社会责任，近十年纳税超 1500 亿元。在国内国际双循环、国产替代和产业升级的背景下，美的集团聚焦七大业务板块，坚持“科技领先、用户直达、数智驱动、全球突破”四大战略主轴，持续加大在数字化、IoT、AI 和全球化方面的投入，近十年研发投入超 1000 亿元，实现 ToC 与 ToB 业务并重发展，致力于成为全球智能家居的领先者、智能制造的赋能者。