

语音识别技术发展及应用

微信智聆语音技术中心

new trend
new technology
new application

Cloud + community
Developer conference

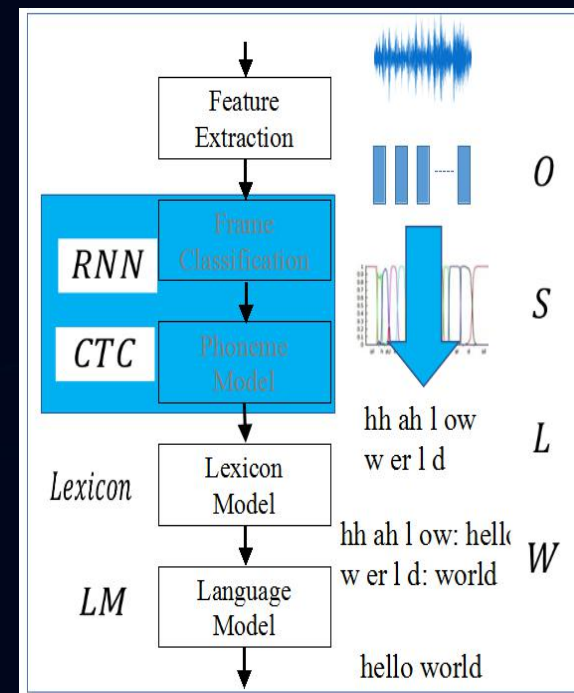
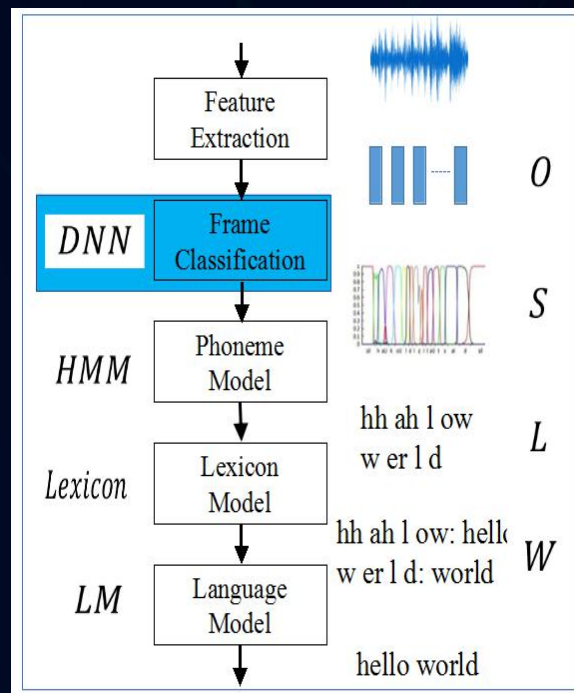
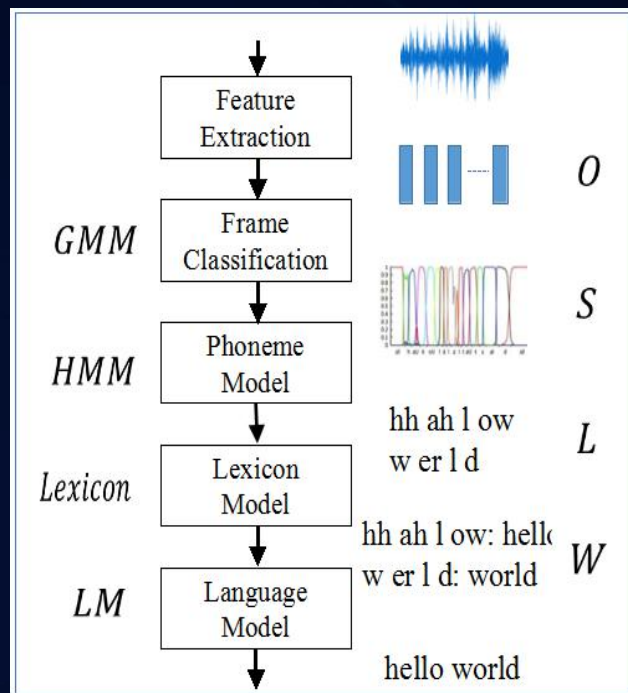
大纲

- 语音识别技术简史
- 微信智聆语音平台
- 微信智聆云端案例

new trend
new technology
new application

Cloud + community
Developer conference

语音识别简史



传统系统

- ◆ 特征提取：MFCC，PLP
- ◆ 声学模型：HMM/GMM模型，输出HMM state似然值
- ◆ 词典模型：查表
- ◆ 语言模型：Ngram模型

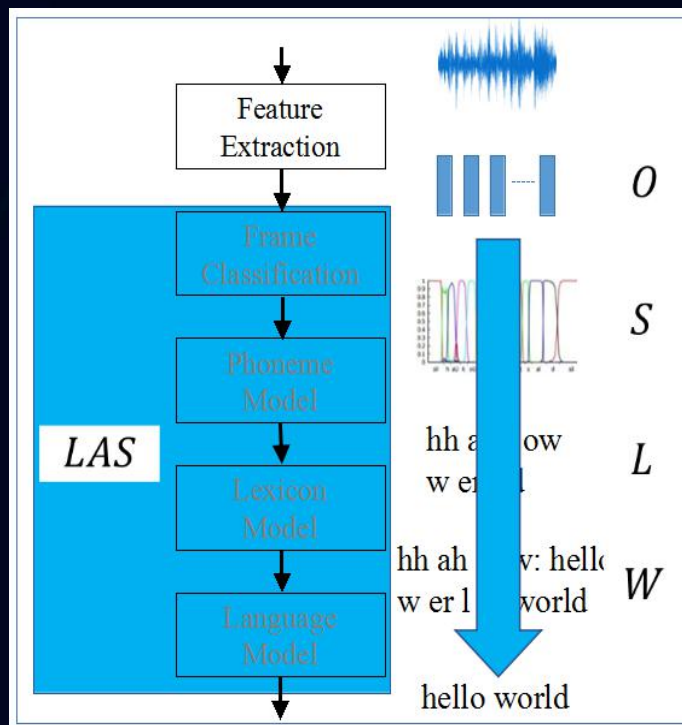
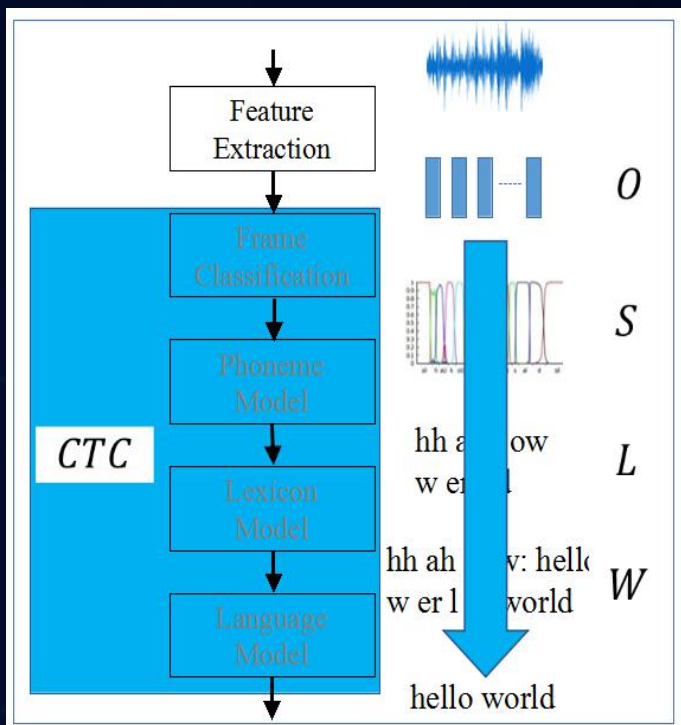
2009年引入深度学习

- ◆ 特征提取：Fbank
- ◆ 声学模型：HMM/DNN，HMM/RNN
输出HMM state后验概率
- ◆ 词典模型：查表
- ◆ 语言模型：Ngram模型

2014年端到端初探

- ◆ 特征提取：Fbank
- ◆ 声学模型：RNN，输出音素后验概率
- ◆ 词典模型：查表
- ◆ 语言模型：Ngram模型

语音识别简史



2017年端到端系统之CTC

- ◆ 端到端模型：RNN，CLDNN；训练准则CTC准则
- ◆ 输出word后验概率
- ◆ 不再需要词典，语言模型

2017年端到端系统之Attention

- ◆ 端到端模型：encoder-decoder attention框架
- ◆ Encoder：输入特征到中间态 h_s
- ◆ Attention：中间态转换 $h_s \rightarrow h_t$
- ◆ Decoder： h_t 转换预测词序列



new trend
new technology
new application

Cloud + community
Developer conference

微信智联语音平台——团队简介

微信智联 (WeChat iHearing)

● 微信智联

腾讯自主研发语音技术品牌

专注于语音识别，语音合成，声纹认证等语音人工智能技术的研发

● 服务对象

C端：各类APP、智能硬件

B端：政府机构、企业单位

● 核心价值观

本着踏实的精神

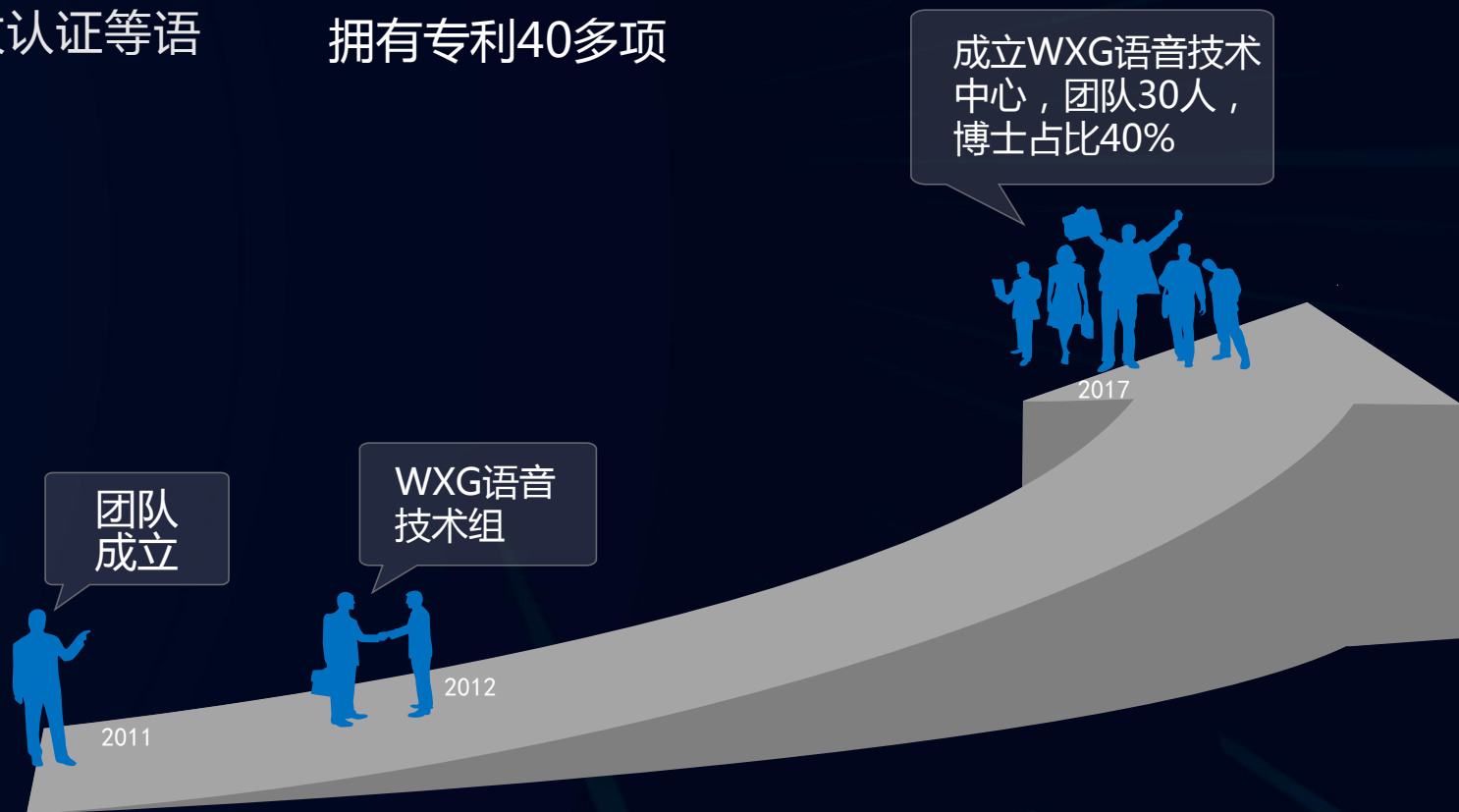
重视细节的优化

落地有用的产品

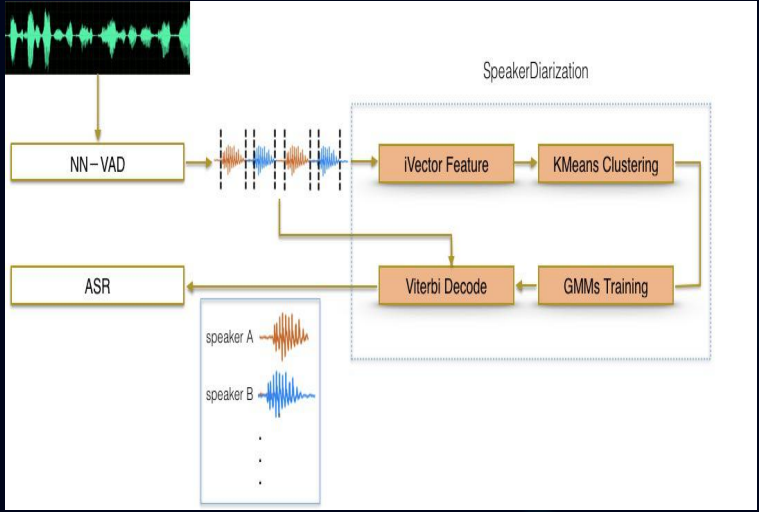
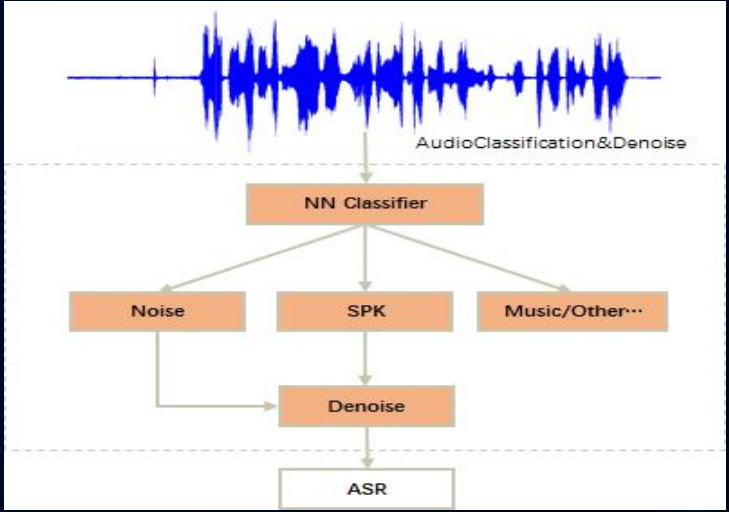
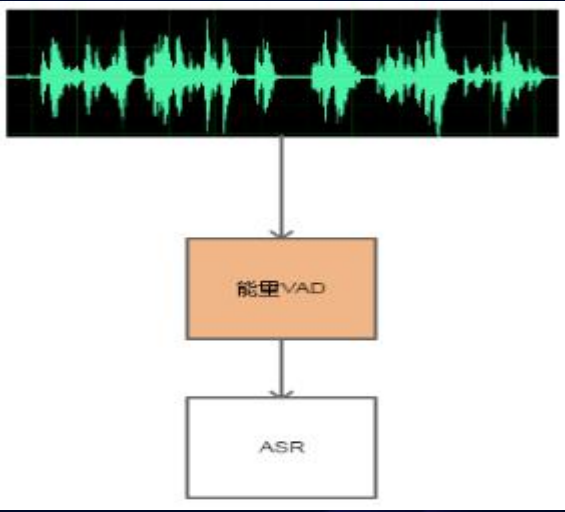
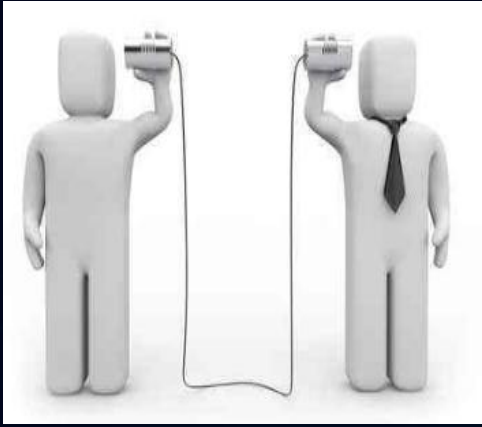
● 论文和专利

发表语音专业会议icassp，interspeech等十余篇

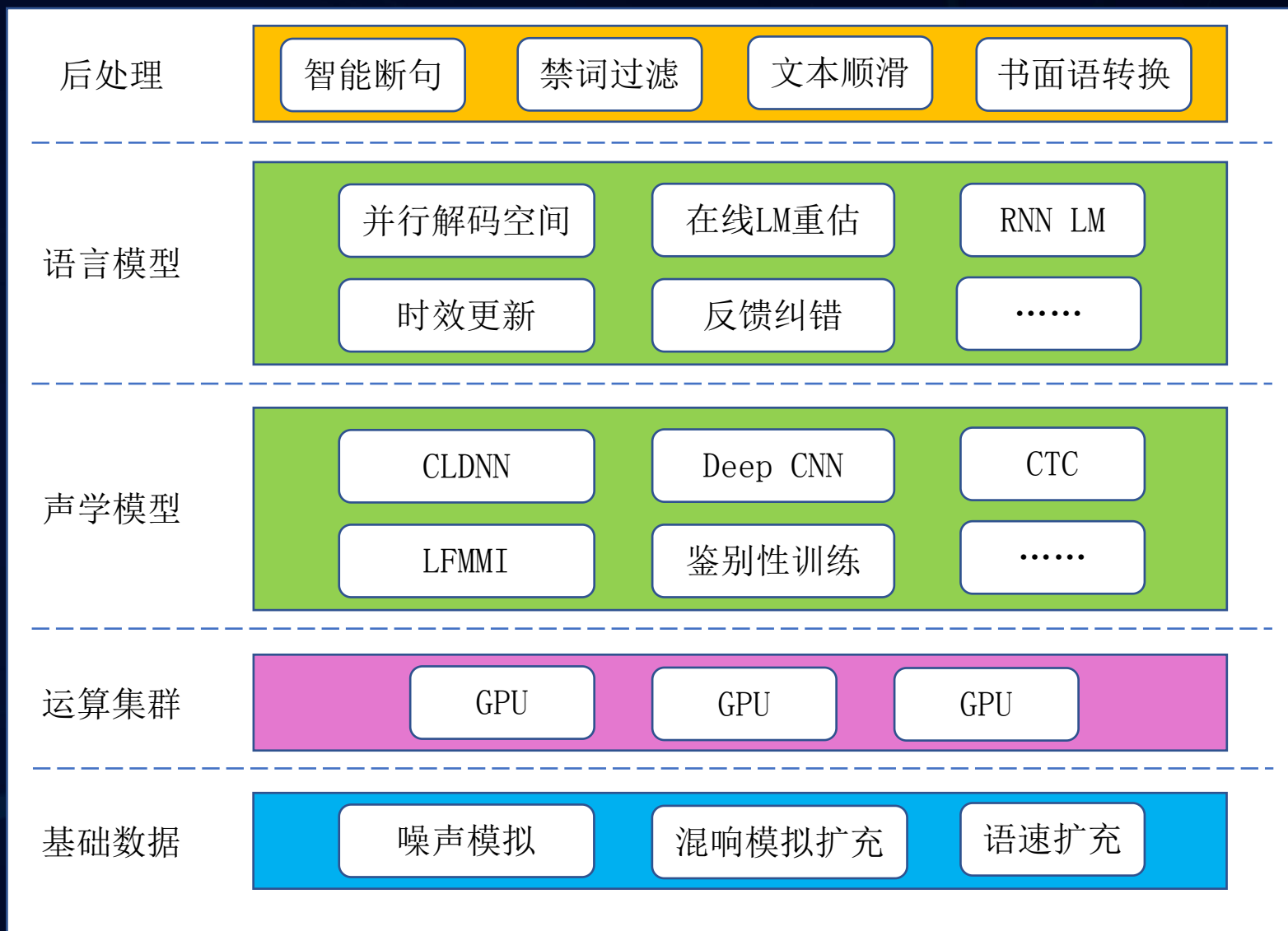
拥有专利40多项



微信智联语音技术——前端处理



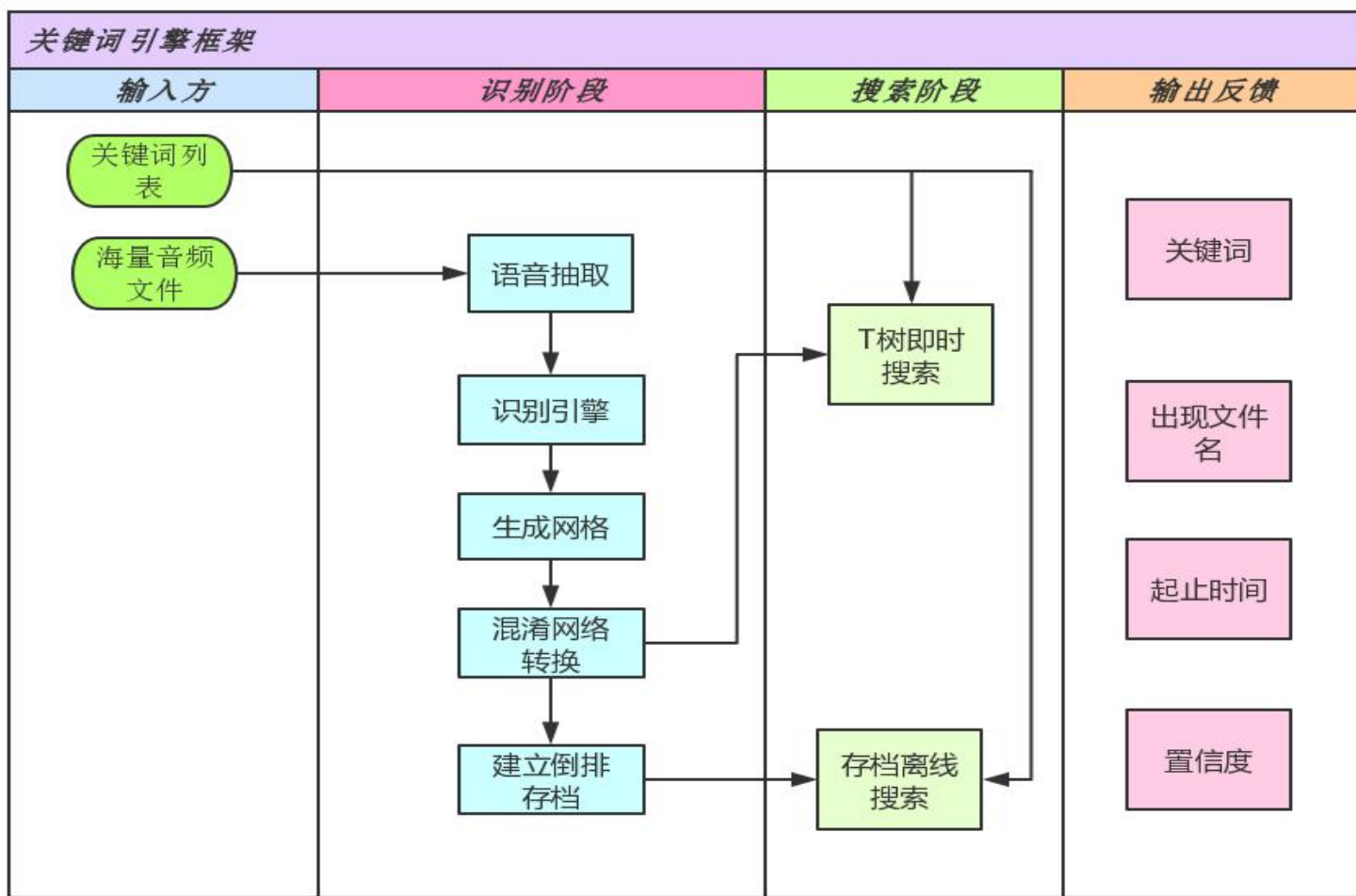
微信智联语音技术——基础语音识别



微信智联语音识别系统特点

- ◆ 支持中文，粤语，英语识别
- ◆ 多种数据argument技巧，扩充训练数据多样性
- ◆ 先进的深度学习架构：CLDNN架构，LFMMI/CTC训练方法
- ◆ 在线大语言模型重估，性能提升同时保证高效解码
- ◆ 多候选以及词级别置信度输出
- ◆ 精准智能断句以及后处理

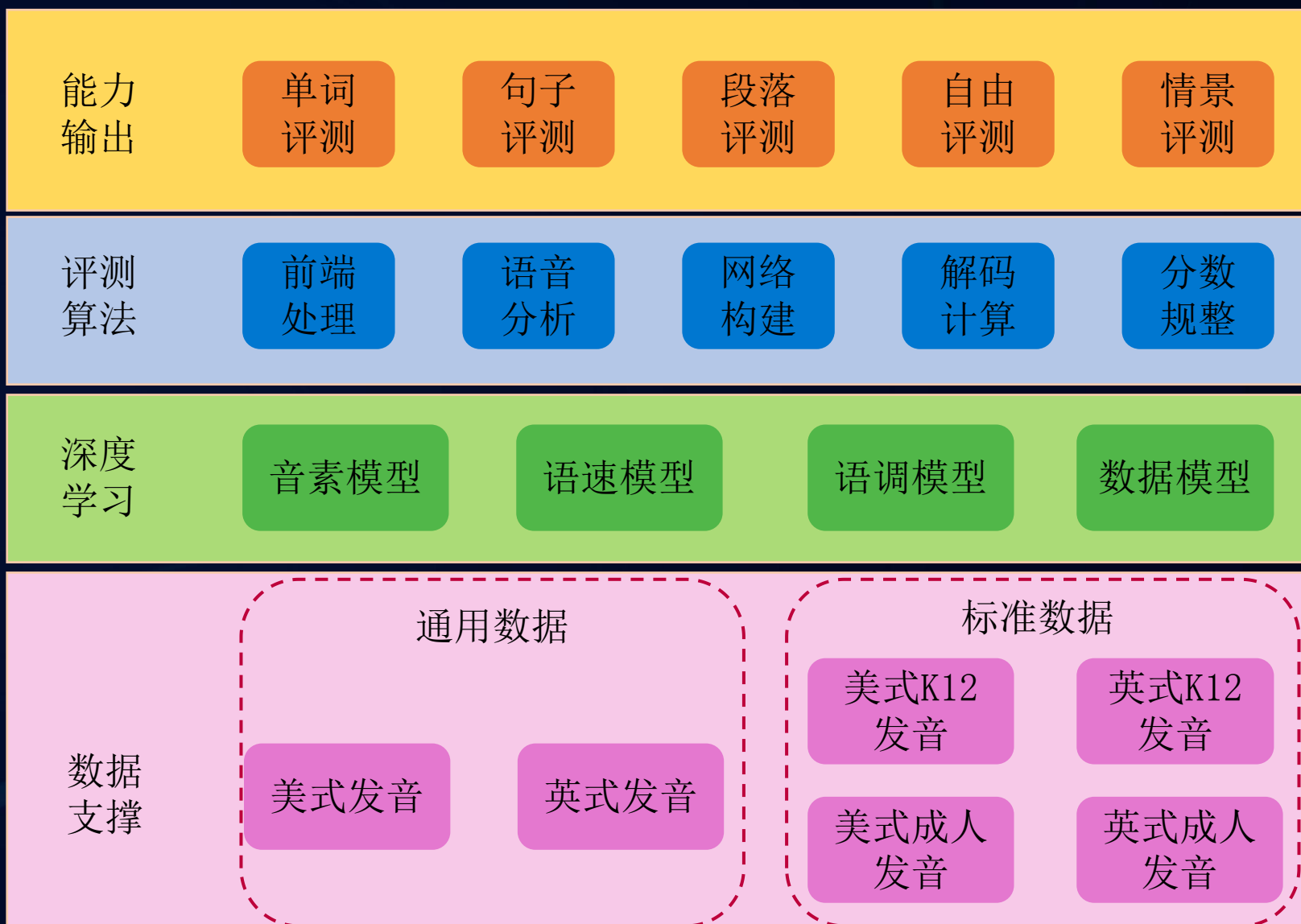
微信智联语音技术——关键词检测



关键词检测技术

- ◆ 资源小，速度快，适用于海量音频数据检索
- ◆ 支持实时检索和历史检索两种模式
- ◆ 等精度召回率达90%

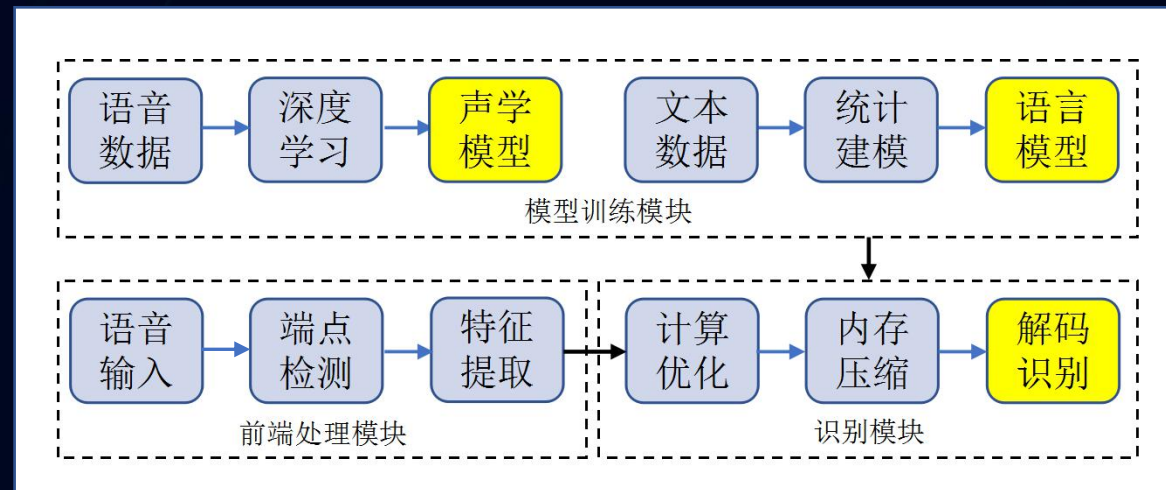
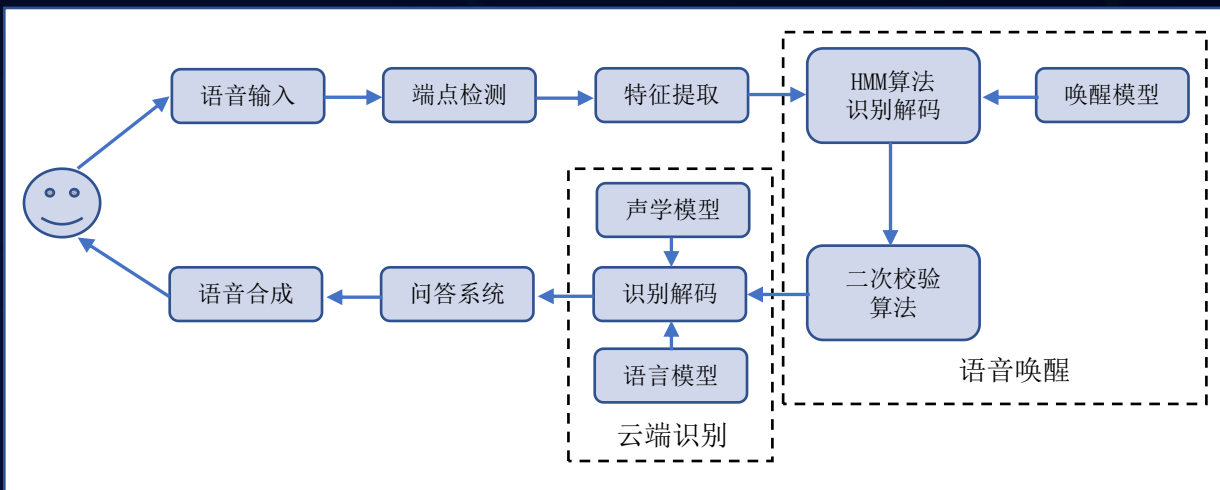
微信智聆语音技术——发音评估



发音评估技术

- ◆ 支持中英文两个语种发音评估
- ◆ 支持单词，句子，段落，自由说等多种模式打分
- ◆ 支持儿童到成年全年龄段打分
- ◆ 发音准确度，完整度，流畅度多维度打分
- ◆ 专家打分相似度90%以上

微信智联语音技术——嵌入式技术



语音唤醒技术

- ◆ 唤醒成功率95%+
- ◆ 误唤醒低于1次/24小时
- ◆ 单词语唤醒资源<900kB
- ◆ 运算 < 90MIPS
- ◆ 模型优化以及多次校验技术

嵌入式语音识别技术

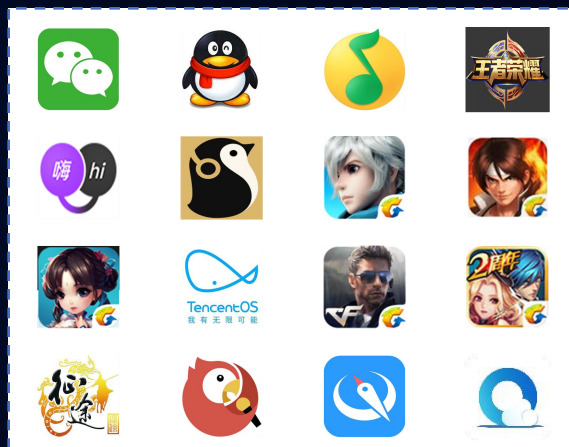
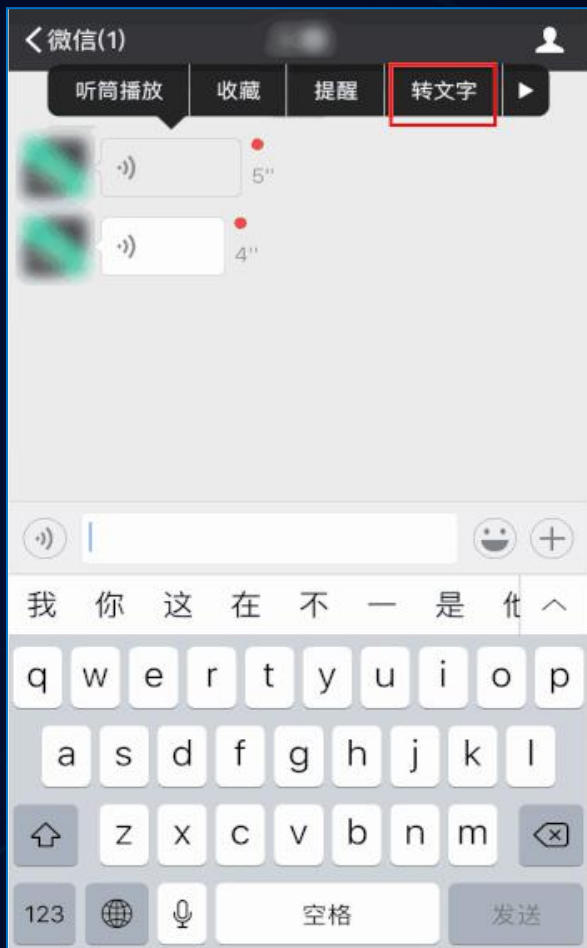
- ◆ 语音输入法场景准确率90%+
- ◆ 离线包资源大小 < 50M
- ◆ 运行时内存占用 < 170M
- ◆ 使用期间CPU占用提升 20%



new trend
new technology
new application

Cloud + community
Developer conference

微信智联云端案例——移动业务线



微信智联云端案例——电话业务线



智能客服电话



金融、交通、教育、保险...

3W+



知乎

珍爱网
zhenai.com

世纪佳缘
jiayuan.com

思坞科技

商客通®

腾讯客服
KF.QQ.COM

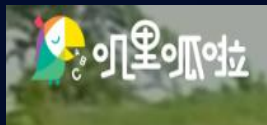
微信智聆云端案例——智聆口语评测



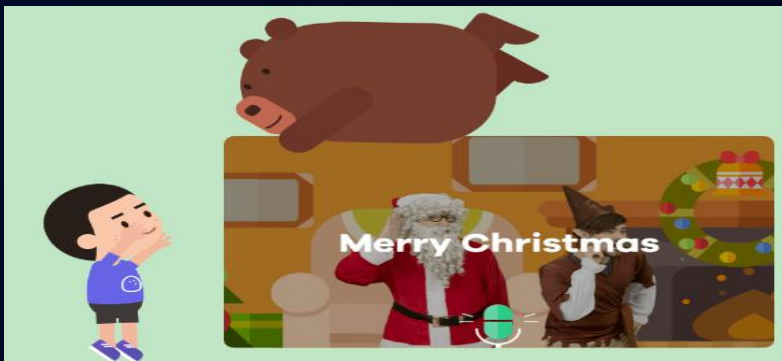
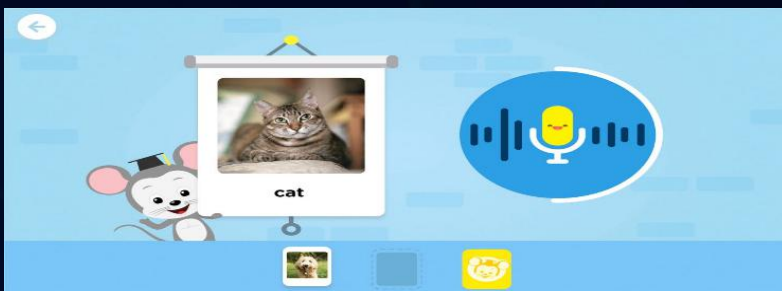
ABCMouse
(儿童教育产品)



腾讯云教育SOE
(Smart oral evaluation)



叽里呱啦
(儿童教育产品)



微信智联云端案例——腾讯同传



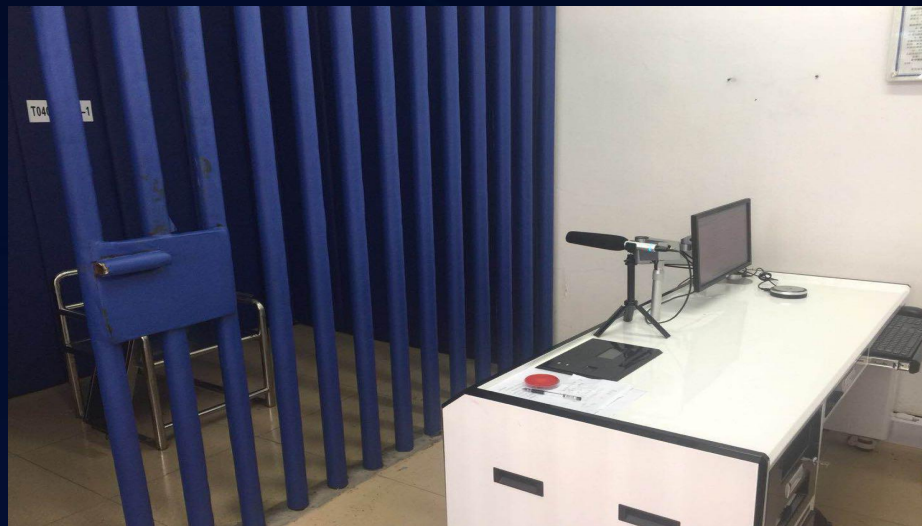
已支持数百场科技、医疗等行业会议
其中国家级、行业级会议100余场。

腾讯同传

传在现场采用的是全部是依靠机器来完成，而不是采用人机协作的方式进行同传。这需要很大的勇气。腾讯同传是我们内部两个团队共同努力的结果。微信智联提供的语音识别技术相当于腾讯同传的耳朵，比如

interpretation is the result of the joint efforts of our two internal teams. The speech recognition technology provided by Wechat iHearing is equivalent to the ears of Tencent simultaneous interpretation.

微信智联云端案例——公检法业务线





new trend
new technology
new application

Cloud + community
Developer conference