

2024第九届信也科技杯 全球AI算法大赛

智辨真言 · 数启未来

2024 FinVolution Global Data Science Competition
Deepfake Speech Detection Challenge

基于短时傅里叶变换的高效语音鉴伪方案

短时傅里叶变换与Flash Attention的结合

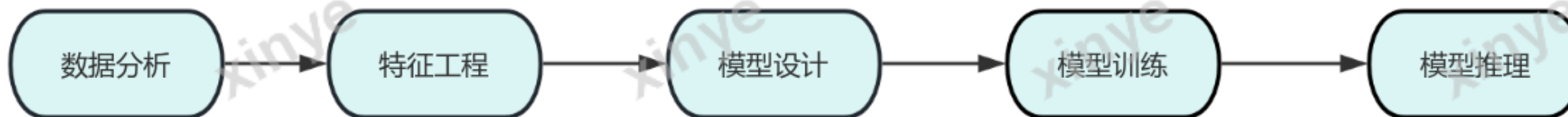
分享人：唐永威 来自：初淅沥以萧飒 团队

目录

1. 团队介绍
2. 方案简介
3. 特征工程
4. 模型设计
5. 模型训练&推理

方案简介

方案流程



特征工程



短时傅里叶变换



NLP: 序列特征

每一帧的傅里叶变换视为单个token的embedding。



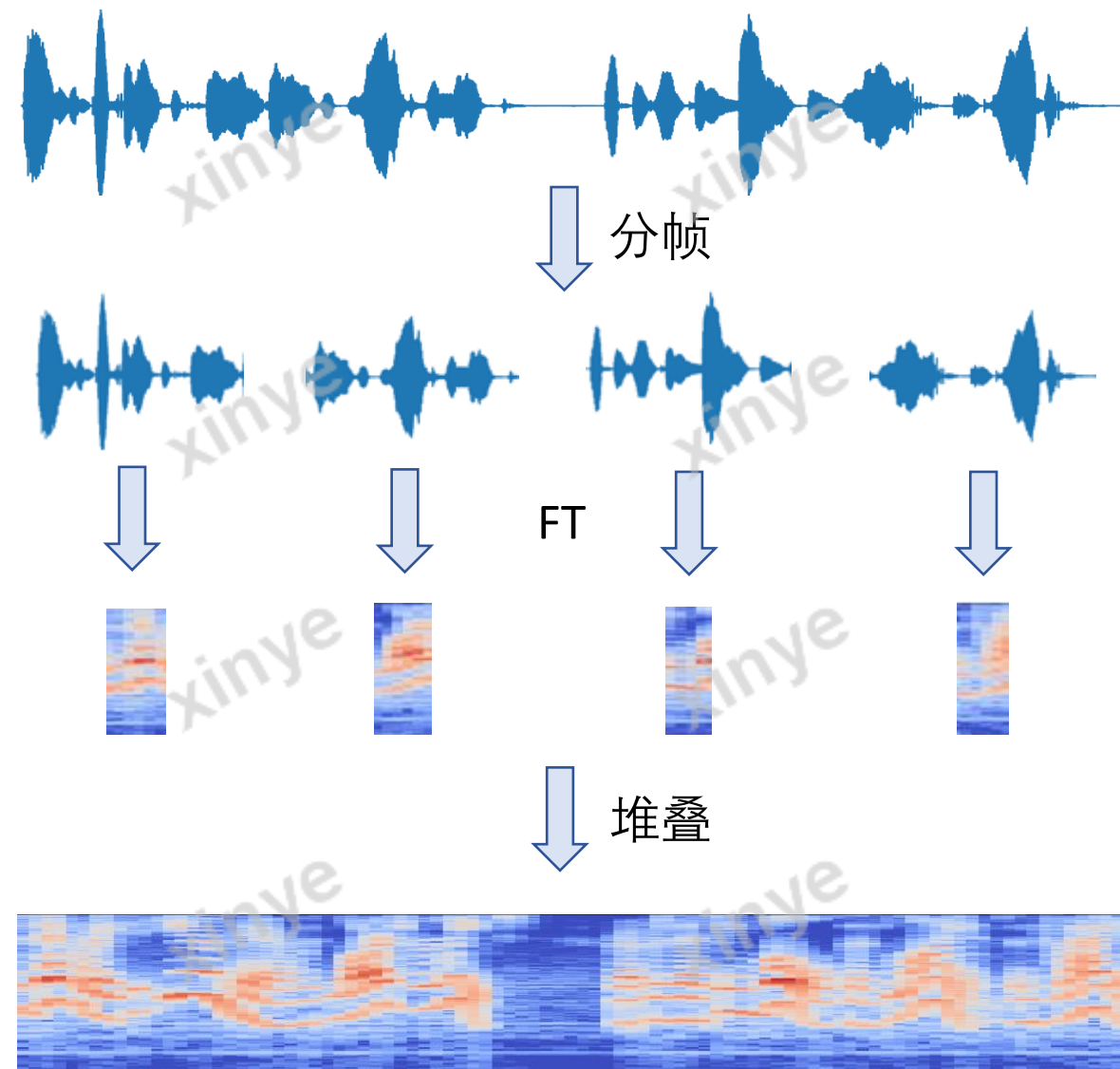
CV: 图像特征

所有帧的傅里叶变换合并到一起，将帧数视为宽、傅里叶变换视为高、整体视为一通道的图像。



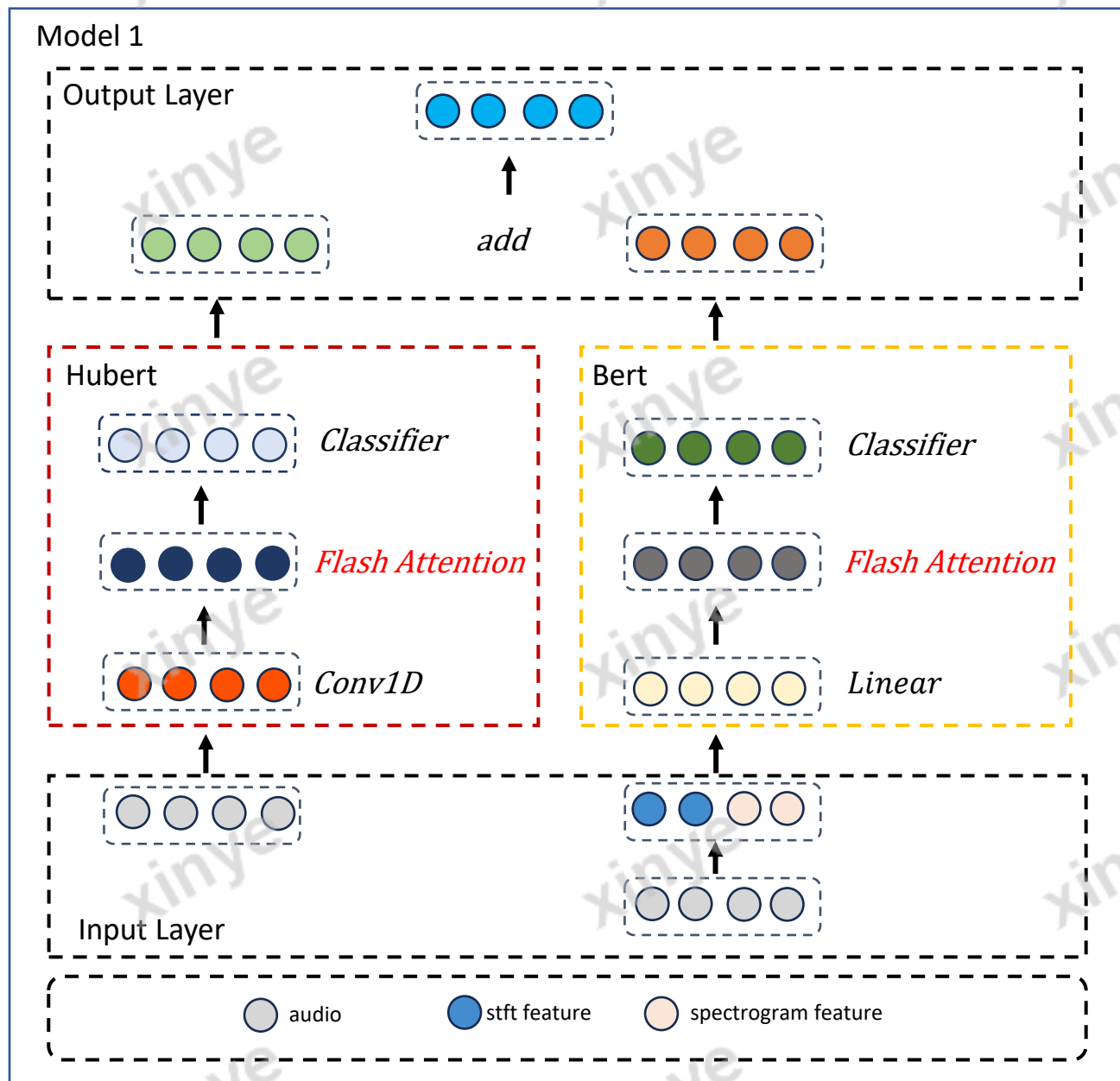
提升显著

初赛3个百分位的提升，复赛2个百分位的提升，两份不同的测试集均能提升，展现了较好的鲁棒性。

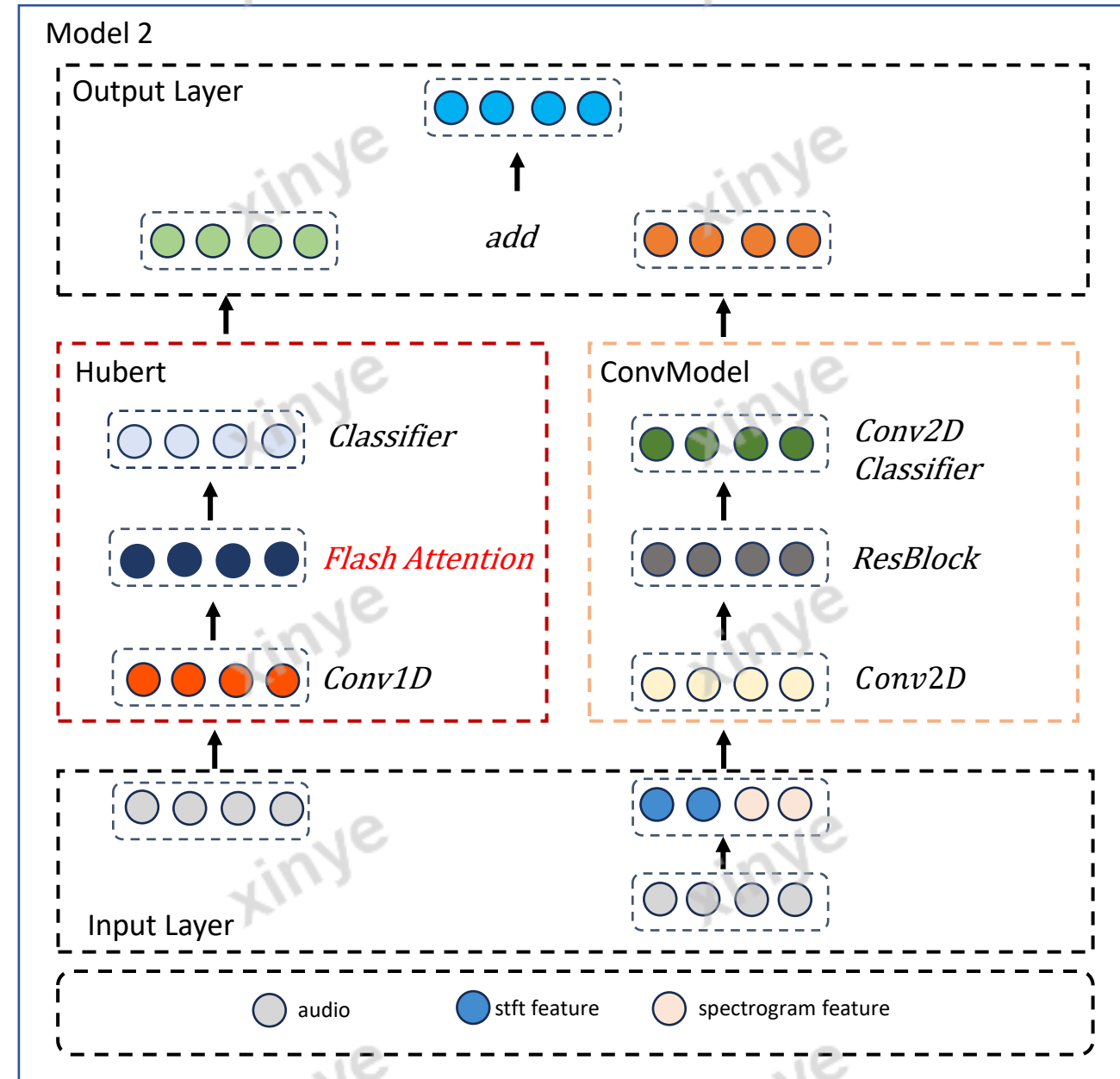


模型设计

0.786

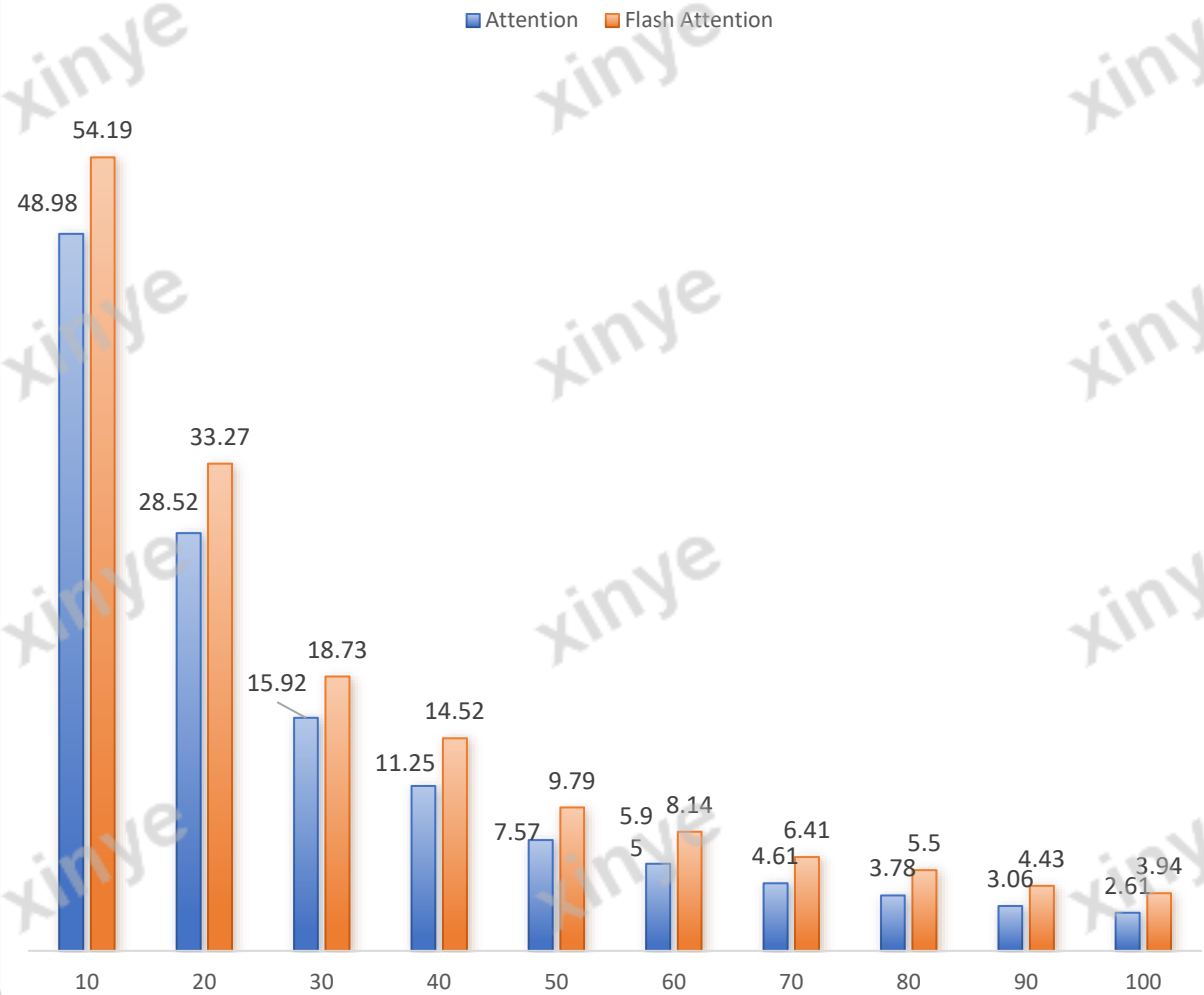


0.788

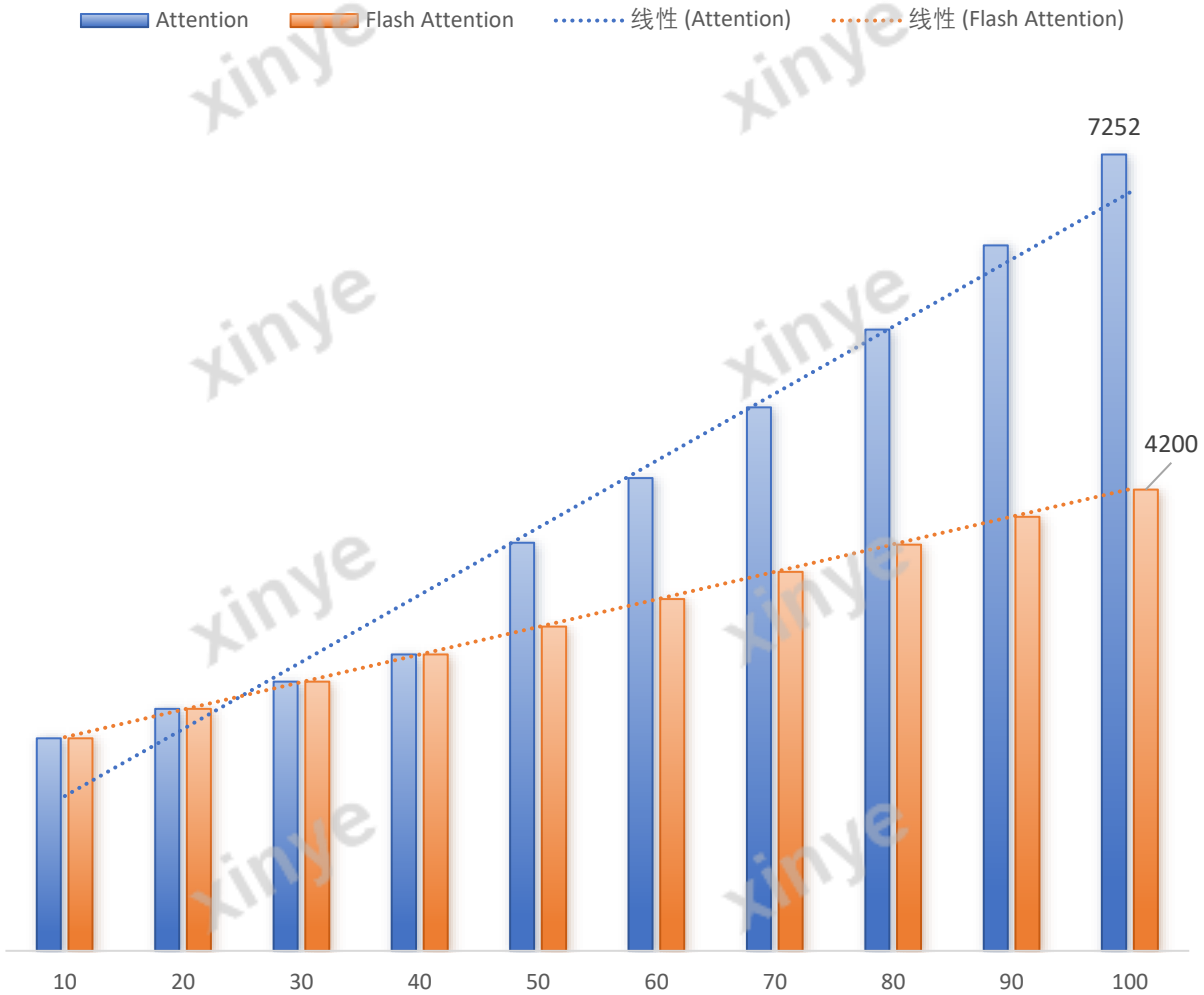


Flash Attention

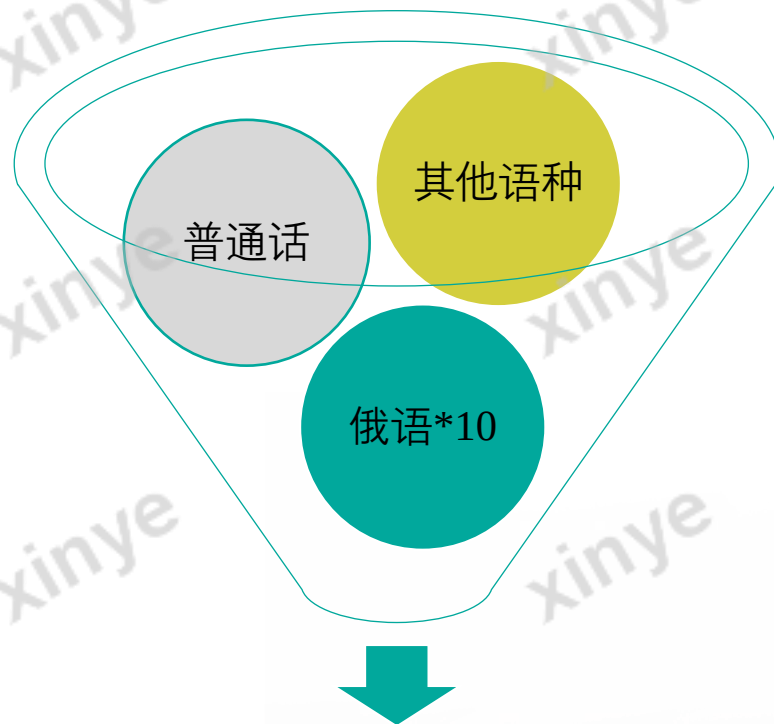
速度对比



显存对比



模型训练&推理



2-4段随机拼接

训练

训练集

公开数据

随机拼接

循环填充

训练

模型文件

推理

循环填充

概率

THANK YOU

2024 FinVolution Global Data Science Competition
Deepfake Speech Detection Challenge

2024第九届信也科技杯
全球AI算法大赛